More 😄 , Less :) The Competition for Paralinguistic Function in Microblog Writing

Umashanthi Pavalanathan and Jacob Eisenstein
School of Interactive Computing
Georgia Institute of Technology
Atlanta, GA 30308
{umashanthi + jacobe}@gatech.edu

Abstract

Many non-standard elements of 'netspeak' writing can be viewed as efforts to replicate the linguistic role played by nonverbal modalities in speech, conveying contextual information such as affect and interpersonal stance. Recently, a new non-standard communicative tool has emerged in online writing: emojis. These unicode characters contain a standardized set of pictographs, some of which are visually similar to well-known emoticons. Do emojis play the same linguistic role as emoticons and other ASCII-based writing innovations? If so, might the introduction of emojis eventually displace the earlier, user-created forms of contextual expression? Using a matching approach to causal statistical inference, we show that as social media users adopt emojis, they dramatically reduce their use of emoticons, suggesting that these linguistic resources compete for the same communicative function. Furthermore, we demonstrate that the adoption of emojis leads to a corresponding **increase** in the use of standard spellings, suggesting that **all** forms of non-standard writing are losing out in a competition with emojis. Finally, we identify specific textual features that make some emoticons especially likely to be replaced by emojis.

Introduction

Figure 1 shows four tweets from Britney Spears, an American celebrity. In all of these tweets, Spears uses a variety of non-standard orthographies to present her personality and convey her love to her fans and friends. However, there is a change in her writing style: the ASCII-based non-standard variants disappear after she started using ***emojis***, a new set of unicode pictographs (Figure 2). In some cases, there is a direct replacement for ASCII-based non-standard variants: the heart-shaped emoticon *<3* is replaced by the similar-looking ❤️ emoji, and the smiling wink emoticon *;)* is replaced by the 😉 emoji. Even non-standard spellings such as A*hhhh,* for which there is no obvious emoji replacement, disappear after Spears begins to use emojis. This change in Britney Spears's writing style corresponds to a change in the technology itself: the introduction of emojis in social media platforms.

[Insert Figure 1 here]

[Insert Figure 2 here]

This example is emblematic of a more general question: what happens to social media language when there are changes in the communicative affordances of the underlying technology? Prior work has explored large scale adoption of new conventions in social media, such as re-tweeting practices on Twitter, and its variation based on users' social and informational goals (boyd *et al.*, 2010; Kooti *et al.*, 2012). In contrast, our focus is not on the social transmission of new conventions, but on how these innovations trigger consequences that cascade throughout the linguistic system. Indeed, while online language is understood as part of an integrated socio-technical system (Cherny, 1999), the mechanics of the interaction between the technical and linguistic components of this system are still only loosely understood. We investigate this intersection in the present work; specifically, we focus on the effects of the introduction of emojis on other stylistic components of online writing in Twitter.

The use of non-standard writing style in online interactions such as social media is often viewed as an attempt to replicate nonverbal elements of communication. These nonverbal elements are also known as paralinguistic cues (Carey, 1980), and are used to modify meaning, convey emotion and interpersonal attitudes, or reveal contextual information. In social media, examples of such paralinguistic cues abound: expressive lengthening and capitalization (*e.g., SOOOO* instead of *so*) to emphasize the point of the author or to indicate a different tone; phonetic spelling (*e.g., aight* instead of *alright*) to indicate relaxed pronunciation (Eisenstein, 2015); emoticons (*e.g., :-(*) to convey emotional affect in an utterance that would otherwise read neutrally. The examples in Figure 1 suggest that emojis can be viewed as another item in this arsenal of paralinguistic cues. From a functionalist perspective, we may view emojis as **competing** with these other non-standard forms for the right to convey non-propositional, interactional meanings.

The stakes of this competition are high because emojis differ from other non-standard forms in one crucial way: unlike emoticons, phonetic spellings, and non-standard orthographies, emojis are not the result of individual innovations propagating through a social network. Rather, they are created and adopted through a standardization process that is run by the Unicode Consortium, a non-profit organization whose membership is made up mostly from large software and technology companies. In sociolinguistic terms, emojis are “change from above”, while other paralinguistic cues are “change from below” (Labov, 2006). As emojis have been introduced into a Twitter environment that was already rife with non-standard forms, we find these two forces for change to be in direct conflict.

To shed light on the impact of the introduction of emojis into the system of paralinguistic communication in online writing, we quantitatively evaluate the following two hypotheses:

Hypothesis I: Adopting emojis causes writers to use fewer emoticons. (This would suggest that emojis and emoticons are competing for a shared paralinguistic function.)

Hypothesis II: Adopting emojis causes writers to use fewer ASCII-based non-standard orthographies overall. (This would suggest that emojis are displacing all forms of non-standard orthography.)

Since emoticons are one form of ASCII-based non-standard orthography, we can consider Hypothesis I as a special case of Hypothesis II.

To test these hypotheses, a simple correlation analysis is not sufficient. Consider the temporal trends of emoji and emoticon usage shown in Figure 3: emoji usage is increasing while emoticon usage is steadily decreasing. However, correlation between these trends does not imply that the increase in emoji usage **causes** a corresponding decrease in emoticon usage. In fact, from this evidence, we cannot even be sure that any individuals are changing the way that they write: an alternative explanation is that users who prefer emoticons are becoming less active over time, while new users tend to prefer emojis. Therefore, to make any causal claims, we need to examine the changes happening at the level of individual users, before and after the introduction of emojis.

[Insert Figure 3 here]

We use a large scale causal inference analysis to test our hypotheses. Specifically, we approximate a randomized experiment using observational data from Twitter. We consider the users who were early adopters of emojis as the treatment group. Using a matching approach to causal estimation, we select a corresponding control group of users with similar pre-treatment characteristics. We then measure the effects of emoji adoption on writing style by comparing the frequency of various linguistic features for these users before and after their first use of emojis.

Our results show that Twitter users who adopt emojis tend to decrease their usage of emoticons, suggesting that emojis and emoticons are competing the same paralinguistic role. We also find that adoption of emojis leads to an **increase** in the use of standard spellings, suggesting that emojis are replacing all forms of ASCII-based non-standard orthographies. These findings lend support to both of our hypotheses and suggest a movement away from ASCII-based non-standard orthographies. Further, we find that emoticons with particular stylistic features, such as horizontal orientation (e.g., *O_O* and -_-), are more likely to be displaced than other emoticons.

Background and Related Work

Emojis are "picture characters," including not only faces, but also pictographs representing objects, actions, concepts, and ideas. Figure 2 shows examples of emoji characters used in Twitter, and Figure 1 shows examples in real tweets. In contrast to emoticons, which are created from ASCII character sequences, emojis are represented by unicode characters, and are continuously increasing in number with the introduction of new characters in each new unicode version.[1] In mid-2015, new emoji characters were introduced to represent people with different skin tones and hair colors;[2] this diversity may help to show attributes of authors such as identity.

Emojis originated in Japanese mobile phones in the late 1990s. They have recently become popular worldwide in text messaging and social media, thanks in part to the adoption of smartphones supporting input and rendering of emoji characters. The increasing popularity of emojis has led to the selection of one of the pictographs, 😂 (widely known as the "face with tears of joy" emoji), as the *Oxford Dictionaries Word of the Year 2015.*[3] Twitter introduced emojis to its web interface in early 2014, following emoji support on Twitter's apps for Android and iOS.[4] Emojis are also becoming widely popular in marketing, and recently Twitter started to support custom emojis for brands.[5]

Non-standard Orthography in Online Writing

Crystal (2006) overviews non-standard writing conventions emerging in computer mediated communication (CMC), such as repeated letters (*e.g., sooo cooolll*), repeated punctuation marks (*e.g., yes!!!!*), and use of asterisks (*e.g., the *real* question*). One of the roles of such non-standard orthography is to express paralinguistic information, in the absence of nonverbal cues present in speech. Examples include the usage of repeated exclamation marks at the end of an utterance to emphasize the point of the author, asterisks surrounding words to indicate that those words contained within them are to be heard with a different tone than the rest of the utterance, and a series of periods to indicate pause (Carey, 1980). Kalman and Gergle (2014) studied how non-standard character repetition is used to enrich communication using a large dataset of e-mails. Emoticons are another form of non-standard orthography, which are particularly relevant to emojis due to their pictographic nature; we discuss them in the next subsection. In our study we contrast these forms of paralinguistic communication with emojis.

Emoticons

There is a long history of research on the role of emoticons in textual communication. While some researchers view emoticons as signals of emotional expression (Rezabek and Cochenour, 1998; Wolf, 2000; Crystal, 2006), later research has shown that the usage of emoticons in CMC goes beyond the expression of emotions, and that the emotional interpretation of emoticons depend on properties of the communicative context, the author, and the reader (Derks *et al.*, 2007; Schnoebelen, 2012; Park *et al.*, 2013). Dresner and Herring (2010) identify three broad linguistic functions of emoticons: (1) as emotion indicators, mapped directly onto facial expressions (*e.g.*, happy or sad), (2) as non-emotional meaning, mapped conventionally onto facial expressions (*e.g.*, joking), and (3) as an indication of the speaker's communicative intention (*e.g.*, ending an utterance with a smiley to mitigate the commitment of the author’s claims or opinions). Emojis seem to be able to play similar roles (Figure 1), which is why we contrast them with emoticons in this study.

Emojis

With the increased popularity of emojis in CMC, researchers have started to explore the role of emojis in textual communication. Stark and Crawford (2015) provide details of the origin of emojis and examine emojis as historical, social, and cultural objects. Kelly and Watts (2015) interviewed a culturally diverse set of 20 participants about how they use emojis in mediated textual communication with close personal ties.

As with emoticons, emojis are sometimes viewed as a vehicle for expressing sentiment and affect: for example, Novak *et al.* (2015) developed a sentiment lexicon for emojis.[6] However, as with emoticons, the communicative role of emojis is considerably broader and more nuanced than a direct expression of emotion. For example, the interviews performed by Kelly and Watts (2015) revealed that emojis are used for a range of purposes beyond affect, such as maintaining a conversational connection, permitting a playful interaction, and creating a shared and secret uniqueness within a particular relationship. Furthermore, the interpretation of emotions may be complicated by differences across viewing platforms (Miller *et al.*, 2016).

The previous literature has begun to establish the communicative role of emojis. The emphasis on interactional meaning supports our view that emojis are in competition with other forms of non-standard orthography, particularly emoticons. Our study is the first to quantify this competition and show that adoption of emojis leads to a corresponding decrease in the use of other non-standard forms.

Causal Inference and Social Media Analysis

Our study is built on the use of causal inference for analyzing observational data. This statistical framework is widely used in fields such as epidemiology and political science, but until recently it has been rare in social media research (Muchnik *et al.*, 2013; King *et al.*, 2014). In general, observational studies of causal phenomena are susceptible to confounds because subjects are not randomly assigned to treatment and control groups as in randomized experiments. For example, the Twitter client from which users post tweets (*e.g.*, web browser, smartphone app) is a potential confound, because different Twitter clients have varied text-input capabilities, which could affect both the treatment (adoption of emoji) and the outcome (differences in linguistic style). Several statistical techniques have been developed to mitigate these confounds in observational studies of causal phenomena, including ***matching*** (Rosenbaum and Rubin, 1983; Ho *et al.*, 2007) and ***stratification*** (Frangakis and Rubin, 2002).

There has been some recent work employing causal inference techniques such as matching in large scale quantitative studies using observational social media data. For example, Reis and Culotta (2015) used matching to create treatment and control groups of users to understand the effects of exercise on mental health; De Choudhury *et al.* (2016) used matched samples to examine dietary choices and nutritional challenges in food deserts using Instagram posts; Cheng *et al.* (2015) used matching on users to study antisocial behavior in online forums. We apply causal inference approaches to the analysis of linguistic style for the first time.

Dataset

We gathered a corpus of tweets from February 2014 to August 2015, using Twitter's streaming API. After streaming a large number of tweets, we selected a set of user accounts for the study population. We then extracted counts of standard and non-standard linguistic features from the tweets of these individuals. In the remainder of this section, we describe these steps in detail.

Selecting the Study Population

As we are focusing on the writing styles of individual Twitter users, it is important to identify personal Twitter accounts by eliminating spam and automated accounts (Yardi *et al.*, 2009) as well as accounts belonging to organizations, which might be maintained by more than one individual. Recently, McCorriston *et al.* (2015) have developed a classifier to identify personal accounts on Twitter; however, for a reasonable classifier accuracy it requires a minimum of 200 tweets per user account, which is hard to acquire due to Twitter API's rate limits. Therefore, we use a set of filtering criteria to eliminate non-personal accounts.

We removed Twitter accounts with more than 5,000 followers or followees and more than 200 tweets on average each month. While popular and high-volume users are of potential interest, particularly in the spread of linguistic styles, these individuals are by definition atypical. To filter out tweets that are not originally posted by users, we removed retweets (repetitions of previously posted messages of others) by excluding messages which contain the "retweeted_status" metadata or the "RT" token. We included only users who have written at least five original tweets on average each month, and removed users who have written more than 10 percent of their tweets in any language other than English.

To quantify the writing style of users we use two sets of ***lexicons***: ***emoticon tokens*** and ***standard tokens***. (Tokenization was performed using a version of the Twokenize program,[7] which we modified to handle emojis.) We define ***lexicon usage rate*** as the ratio of the number of tokens that appear in a lexicon to the total number of tokens. The lexicons are described in the remainder of this section.

Extracting Emoji Tokens

To extract emoji characters from tweets, we converted the messages into unicode representation and used regular expressions to extract unicode characters in the ranges of the "Emoji & Pictographs" category of unicode symbols (other categories include non-Roman characters such as different numbering systems and mathematical symbols). Using this method we identified 1,235 unique emoji characters in a random sample of tweets spanning a period of more than a year (February 2014 to August 2015). Figure 3(a) shows the percentage of emoji character tokens (*i.e.*, $\frac{\text{no. of emoji tokens}}{\text{no. of total tokens}}$) over time in a sample of tweets collected over 19 months.

Extracting Emoticon Tokens

As there is no comprehensive list of Twitter emoticons (and new emoticons get introduced over time), we followed a data-driven approach to identify emoticons. We retrieved an initial set of emoticon-like tokens by using several high-recall, low-precision regular expression heuristics, *e.g.*, two or more ASCII characters with at least one non-alphanumeric character. We then manually annotated all the items that made up 95 percent cumulative frequency of emoticon-like tokens, by looking at their usage in random examples of tweets. For this purpose, we regarded pictographic tokens with more than one ASCII character as emoticons.

After removing tokens that are not used as emoticons, there were 44 and 52 unique emoticons extracted from tweets of March 2014 and March 2015, respectively. In both cases, the twenty most frequent emoticons made up 90 percent of all emoticon tokens. Table 1 shows the extracted emoticon tokens and their cumulative frequencies. Figure 3(b) shows the percentage of emoticon character tokens over time in a sample of tweets collected over 19 months.

[Insert Table 1 here]

Table 2 shows some basic usage statistics of emojis and emoticons in our sample. Notable differences include: emojis appear at the beginning of tweets more often than emoticons (12.40% of tweets begin with an emoji, but only 0.77% of tweets begin with an emoticon); emojis are more likely to appear as the sole token in a tweet (1.01% of tweets contain only an emoji, but 0.17% of tweets contain only an emoticons); and emojis appear more often in messages with user mentions and hashtags. These differences indicate that the linguistic role of emoticons and emojis cannot be identical, since their distributional properties are very different. But nonetheless, the overall **function** (Eggins, 2004) might be similar enough for emojis to replace emoticons in many cases, and it is what we test in our analysis.

[Insert Table 2 here]

Building a Standard Token List

To test the hypothesis that the rise of emojis leads to language becoming more standard (Hypothesis II), we require a list of words that are widely used in online writing and are considered to be standard. Existing dictionaries, such as the standard Unix dictionary, do not have enough coverage of entity names and new words. Examples of standard tokens that do not appear in such dictionaries include: entity names such as *Kardashian* (Kim Kardashian: TV personality), *Minaj* (Nicki Minaj: a musician), and *Genisys* (Terminator Genisys: a science fiction action film) as well as new words such as *emoji*, *instagram*, *crowdfunding*, and *captcha*. Therefore, to incorporate entity names and new words, we augmented standard dictionaries with a vocabulary extracted from English Wikipedia. We regard the vocabulary from Wikipedia as standard because Wikipedia has an explicit editorial practice to avoid words that are not widely accepted in writing.[8] From a dump of English Wikipedia articles (retrieved in October 2015) we extracted terms which appeared in more than 50 different articles[9] (Table 3). The resulting word list contains 621,968 unique tokens.

[Insert Table 3 here]

Study Design

We formulate our research question of whether the introduction and adoption of emojis leads to a corresponding decrease in non-standard orthography on Twitter as a causal inference problem. The causal inference questions are: (1) whether the rise in emojis causes a decline in emoticon usage rate and (2) whether the rise in emojis causes an increase in standard token usage rate.

Consider an idealized experimental setting: we would start with a set of individuals who have never heard of emojis, randomly sample treatment and control groups (which would be matched on all characteristics because they were randomly selected), and then experimentally intervene by introducing the treatment group to emojis. The treatment effect would be the difference in the lexicon usage rates between the two groups after the treatment.

We approximate this randomized experimental setup using observational data from Twitter. We used a matching design to reduce the misestimation of the treatment effect when using observational data. Twitter introduced emojis to its web interface in April 2014, following emoji support on Twitter's apps for Android and iOS.[5] Therefore, we considered the month of March 2014 as the ***pre-treatment*** period and the month of March 2015 as the ***post-treatment*** period.

[Insert Figure4 here]

For the ***treatment group***, we selected authors who had not used any emoji characters in March 2014, but used at least five emoji characters in March 2015. We chose authors who had not used any emoji characters in both March 2014 and March 2015 into the ***control group*** (Figure 3). Figure 5 and Figure 7 show the distribution of emoji usage rates in both the treatment and control groups, before and after treatment. Note that these figures are nearly but not exactly the same; we use two different study populations for the two analyses, due to the matching criteria explained later. Further, the dramatic differences in the post-treatment emoji usage rate for the treatment and control groups are by construction, since the control group is designed to consist of individuals who do not adopt emojis.

[Insert Figure 5 here]

[Insert Figure 6 here]

[Insert Figure 7 here]

[Insert Figure 8 here]

Confounding Factors

Because we approximate a randomized experiment using observational data, experimental units (*i.e.*, Twitter authors) are not randomly assigned to the treatment and control groups. If there are confounding factors that influence both the treatment and outcome, then the treatment effect estimation will be incorrect.

For example, the client from which users post tweets might correlate with both word choice and use of emojis. If a client added an autocorrect feature in between the pre-treatment and post-treatment periods and if this client also facilitated the use of emojis, this would create the spurious impression that emojis cause better spelling. Therefore, pre-treatment differences between the control and treatment groups can cause substantial estimation problems. While we cannot identify all possible confounding factors, we focus on pre-treatment differences in ***lexicon usage,*** and on the ***client used to post tweets***.

Lexicon Usage

Different authors might have different writing styles in terms of the frequencies of lexical categories they use in their tweets: for example, some authors prefer a more standard style, others will prefer to be more expressive, etc. These styles – or some underlying root cause – might be correlated with the decision to adopt emojis, which is the treatment in our analysis. Therefore, pre-treatment lexicon usage is a confounding factor in our causal analysis.

To overcome this confound, we match each author from the treatment group with an author from the control group who has comparable lexicon usage. As tweets are restricted to 140-characters and each token competes for that space, we match on the lexicon usage rate at the token level (*i.e.*, $\frac{\text{no. of lexicon tokens}}{\text{no. of total tokens}}$), rather than at the message level. By matching on the lexicon usage rate we ensure that the treatment and control authors have similar writing style before treatment. Similar to lexicon usage rate, we define ***emoticon usage rate*** as $\frac{\text{no. of emoticon tokens}}{\text{no. of total tokens}}$, ***standard token usage rate*** as $\frac{\text{no. of standard tokens}}{\text{no. of total tokens}}$, and ***emoji usage rate*** as $\frac{\text{no. of emoji tokens}}{\text{no. of total tokens}}$.

Clients Used to Post Tweets

Different clients (*e.g.*, Windows PC vs. iPhone) have varied emoji and emoticon input capabilities. In addition, the autocorrect feature also varies for different clients. Prior work (Gouws *et al.*, 2011; Perreault and Ruths, 2011) found that the tweets posted using different clients have different stylistic and topical tendencies. To minimize the confounds due to different Twitter clients, we match each author in the treatment with an author in the control who is using the same client.

Table 4 shows a list of frequent clients of tweets in our dataset, measured at the tweet level from the "source" field in the tweet metadata. The top seven clients cover a cumulative share of nearly 75 percent of all tweets. When examining the author-level client usage, nearly 60 percent of authors used a single client, 30 percent of authors used two clients, and the remaining 10 percent used three or more different clients. The majority of the authors using two clients used either one of the popular mobile platforms (*i.e.*, “Twitter for iPhone” or “Twitter for Android”) and “web”.[10] Because some authors’ tweets originate from multiple clients, during matching we considered only the authors who have used a single client more than 80 percent of the time.

[Insert Table 4 here]

Matching

Through matching we can ensure that both the treatment and control units have similar characteristics other than the dependent variable: for a given treated unit with a particular set of confounding factors, we would look for a control unit with as similar as possible confounding factors. Our matching procedure is outlined in Figure 9.

[Insert Figure 9 here]

Matching to test Hypothesis I

In analysis-I, to test the hypothesis that the rise in emojis leads to a corresponding decline in emoticon usage, we matched authors in the treatment and control groups based on their emoticon usage rate and on their preferred clients. This matching procedure yielded 3,112 treatment-control author pairs for analysis-I. Figure 6 (left) shows the distribution of emoticon usage rates in both the treatment and control groups, before treatment; by design, they are nearly identical.

Matching to test Hypothesis II

In analysis-II, to test the hypothesis that the rise in emojis leads to language becoming more standard, we matched authors in the treatment and control groups based on their standard token usage rate and on their preferred clients. This matching procedure yielded 3,115 treatment-control author pairs for analysis-II. Figure 8 (left) shows the distribution of standard token usage rates in both the treatment and control groups, before treatment; again, by design these distributions are nearly identical.

Estimation of Treatment Effects

When computing the treatment effect, we need to account for the control group's outcome as well. This is because the overall lexicon usage rate is changing with time (see Figure 3), and hence there might be changes in the lexicon usage rates of the control group even without any intervention. For this purpose we used a ***difference-in-differences*** method (Angrist and Pischke, 2008), which calculates the effect of a treatment by comparing the average change in the outcome variable after treatment for the treatment group to that of the control group.

We computed the treatment effect in terms of the following quantities:

Lexicon usage rates

$X^{t}_{i,\,pre}$ pre-treatment lexicon usage rates for author *i*, who is in the treatment group *t*

$X^{c}_{j,\,pre}$ pre-treatment lexicon usage rates for author *j*, who is in the control group *c*

$X^{t}_{i,\,post}$ post-treatment lexicon usage rates for author *i*, who is in the treatment group *t*

$X^{c}_{j,\,post}$ post-treatment lexicon usage rates for author *j*, who is in the control group *c*

Differences

Y^{t}_{i} the difference between the post-treatment and pre-treatment lexicon usage rates for author *i*, who is in the treatment group *t*

$Y^{t}_{i} = X^{t}_{i,\,post} - X^{t}_{i,\,pre}$

Y^{c}_{j} the difference between the post-treatment and pre-treatment lexicon usage rates for author *j*, who is in the control group *c*

$Y^{c}_{j} = X^{c}_{j,\,post} - X^{c}_{j,\,pre}$

$\overline{Y}^{t}$ the average difference between post-treatment and pre-treatment lexicon usage rates for the treatment group which has n_t authors

$\overline{Y}^{\,t} = \frac{1}{n_t}\sum_{i \in T} Y_i^{\,t}$

$\overline{Y}^{\,c}$ the average difference between post-treatment and pre-treatment lexicon usage rates for the control group which has n_c authors

$\overline{Y}^{\,c} = \frac{1}{n_c}\sum_{j \in C} Y_j^{\,c}$

We can then define the average treatment effect as,

$ATE = \bar{Y}^{\,t} - \bar{Y}^{\,c}$.

Our null hypothesis in both the analyses is that there is no treatment effect, so that any observed differences must be due to chance variation in a finite sample.

Causal Inference Analysis

Analysis-I: Effects of Emoji Adoption on Emoticon Usage

In the first analysis (analysis-I) we test the hypothesis that using emojis causes individuals to use fewer emoticons. The distribution of authors' emoticon usage before treatment is shown in Figure 6 (left). By design, both the treatment and control groups have similar distributions of emoticon usage before treatment, with an average emoticon usage rate of 0.40% – that is, 0.40% of the tokens from tweets of both the treatment and control authors are emoticons.

There is a decrease in emoticon usage for the treatment group after treatment, as in the distribution shown in Figure 6 (right). After the treatment, the treatment group has an average emoticon token usage rate of 0.12%, while the control group has an average rate of 0.31% of emoticon tokens.

The average treatment effect is a 0.19% decrease in emoticon symbols per token; after treatment, the treatment group is more than 2.5 times less likely to use emoticons than the control group. This difference is statistically significant by a paired t-test ($t = -9.612$), at $p < 10^{-21}$. Results from this analysis are summarized in Table 5.

[Insert Table 5 here]

While all authors use fewer emoticons in the post-treatment period, the decrease was substantially larger for those authors who began using emojis in the time between the pre-treatment and post-treatment periods. Still, the reader may wonder why emoticon usage decreases even for the control group. One possible explanation might be the effects of ***alignment*** or ***accommodation*,** whereby conversational participants tend to mirror the language of their interlocutors (Giles and Ogay, 2007; Danescu-Niculescu-Mizil *et al.*, 2011). If Twitter users are accommodating to each other in general, then even users who never use emojis will reduce their emoticon usage to align with the increasing number of emoji users. Testing this hypothesis is left for future work.

Analysis-II: Effects of Emoji Adoption in Standard Word Usage

Emoticons are one form of ASCII-based non-standard orthography and are visually similar to emojis; as noted in the introduction, there are many other ways of expressing paralinguistic cues. In Analysis-II, we ask whether these forms of non-standard orthography are also disrupted by emojis.

Rather than trying to account for every form of non-standard orthography, we focus on the complement: we compute the frequency of **standard** tokens, and test whether this varies with the usage of emojis. The distribution of authors' standard token usage rates before treatment is shown in Figure 8 (left). By design, both the treatment and control groups have a similar distribution of standard token usage rates before treatment, with an average standard token usage rate of 85.2% (*i.e.*, 85.2% of tokens from tweets of authors in both the treatment and control are from our standard token list).

As shown in Figure 8 (right), there is an increase in standard token usage rates for the treatment group after the treatment intervention. During the post-treatment period, the treatment group has an average standard token usage rate of 86.2%, while the control group has an average rate of 85.7%. The difference between standard token usage rates of the treatment and control groups is statistically significant by a paired t-test ($t = 2.755$, $p \approx 1 \times 10^{-3}$), with an average treatment effect of 0.45 percent increase in standard terms per token. This difference in post-treatment standard token usage rates between the treatment and control group shows that while all authors use more standard tokens in the post treatment period, the increase was substantially larger for those authors who began using emojis. Results from this analysis are summarized in Table 5.

For both analysis-I and analysis-II, we performed further **post hoc** analysis to examine whether our findings are sensitive to the specific pre-treament (March 2014) and post-treament (March 2015) periods we chose. We computed the average difference in the respective lexicon usage rates in two periods adjacent to the pre-treatment (March 2014 vs. April 2014) and post-treatment (February 2015 vs. March 2015) periods. We found that these difference are not statistically significant, which suggests that the treatment effects we measure are indeed due to the adoption of emojis.

Are Specific Styles of Emoticons More Threatened by Emojis?

Our results show that the Twitter authors who adopt emojis tend to reduce their usage of emoticons. If emojis are competing with emoticons for a shared paralinguistic function, then this decline in emoticon usage by emoji adopters raises the question of whether emojis replace a particular type of non-standard orthography more often than others. Motivated by prior work on the ***stylistic variation*** of Twitter emoticons (Schnoebelen, 2012), we analyze the relationship between the stylistic characteristics of different emoticons and the changes in their usage by emoji adopters.

Mixed-Effects Binomial Regression Model

To analyze the relationship between the stylistic characteristics of different emoticons and the changes in their usage by emoji adopters, we used a mixed-effects binomial regression model. Mixed-effects models (Pinheiro and Bates, 2006) are recent extensions of regression models, consisting of two parts: ***fixed effects*** and ***random effects***. Fixed effects predictors are the independent variables as in standard regression. Random effects account for any systematic variation within individual experimental units, such as specific users or words.

We model the counts of messages with emoticons (from the treatment group in both the pre-treatment and post-treatment periods) as a binomial regression model. Binomial regression models are a special case of generalized linear models (Gelman and Hill, 2006), in which the observations are counts drawn from a binomial distribution, with the parameters of the binomial set to be a function of the covariates.

$$
\begin{aligned}
C_{e,t} &\sim \text{Binomial}(\theta_{e,t}, N_t) \\
\theta_{e,t} &= \sigma(\beta_t \phi_e + b_t)
\end{aligned}
$$

In this model, $C_{e,t}$ is the count of messages with emoticon type e in time period t; ϕ_e is a vector of predictor features for emoticon type e; N_t is the total count of messages in time period t. The frequency parameter $\theta_{e,t}$ is computed from the logistic transformation of the inner product of weights β_t and predictors ϕ_e, plus an offset b_t. The reason we choose a binomial regression model over a linear regression model is because the observed data is **counts** of messages; a linear regression model would inappropriately assign probability to nonzero and fractional counts. Our model also accounts for changes in overall usage by incorporating the total count of messages in a time period, N_t, as a parameter.

[Insert Table 5 here]

As shown in Table 6, predictors of our model include indicator variables for the treatment period (POST-TREATMENT) and several stylistic features of emoticons, motivated by prior work (Schnoebelen, 2012). For example, NOSE is a predictor to indicate whether an emoticon has a nose (*e.g., :-)* and *:-P*) or not (*e.g., :)* and *:P*). A negative coefficient for this predictor would suggest that emoticons with this feature are more likely to be replaced by emojis.

Because we measure the count of messages with different emoticons in two different time periods and there might be variability in this count across different emoticons in each period, we included random intercepts for each emoticon. Inclusion of these random intercepts increase the robustness of our analysis by preventing any single emoticon from dominating the estimation of overall stylistic effects.

Results

Results of the regression analysis are shown in Table 7. The regression coefficient for the post-treatment indicator variable, POST-TREATMENT, is negative and highly significant. This is expected because we observe a decline in the overall emoticon usage rate for the treatment authors in the post-treatment period.

[Insert Table 7 here]

ORIENTATION, the predictor for emoticons which have horizontal orientation (*e.g., O_O* and -_-), and WINK-EYES, the predictor for emoticons with wink eyes (*e.g., ;-)* and *;P*), both have significant negative coefficients. This indicates that the emoticons with these features are especially impacted by the adoption of emojis. There are a few possible explanations. One possibility is that emojis are particularly well-suited for conveying the paralinguistic functions of these emoticons; for example, the

WINK-EYES emoticons often convey sarcasm or irony, and there may be emojis that are especially suited for this function. However, this would not explain the decrease in frequency of horizonal emoticons, which can express a range of different paralinguistic cues. Horizontal emoticons are very similar to ***kaomojis*** (Bedrick *et al.*, 2012), which are emoji-like characters widely used in East Asia. Since emojis also originated in East Asia (Stark and Crawford, 2015), it may be that the substitution of emojis for emoticons is particularly prevalent among individuals with affinity for East Asian cultural practices (recall that our dataset is restricted to individuals who tweet primarily in English). Further investigation is left for future work.

Discussion and Future Work

Twitter users who adopt emojis tend to reduce their usage of emoticons, in comparison to the matched users who do not use emojis. This result supports our first hypothesis: the increasing frequency in emojis leads to a corresponding decline in emoticons.

Of course, since Twitter has a restriction on the number of characters, in some sense all features compete for space, but not necessarily for the same linguistic functions. Nonetheless, the overwhelming majority of Twitter messages are not near the character limit (Eisenstein, 2013), indicating that this is unlikely to be the main reason for the decrease in emoticon characters. Given evidence that emoticons and emojis both perform a paralinguistic function (Dresner and Herring, 2010; Kelly and Watts, 2015), it seems more likely that the replacement of emoticons by emojis reflects the ongoing success of emojis in a competition for this role. Our analysis of the changes in the usage of specific emoticons by emoji adopters indicates that emoticons with horizontal orientation and “winking” eyes are especially impacted by the introduction of emojis. These emoticons may play roles that are especially well-suited for emojis, or they may be used by individuals who are especially likely to adopt emojis; teasing apart these two explanations is a project for future work.

Results from the second analysis show that the users who adopt emojis tend to use standard tokens at an increased rate after emoji adoption. This suggests that emojis are in competition with the entire range of non-standard orthographies, including expressive lengthening (*e.g., coooooolll!!!*), non-standard words (*e.g., gud*), and abbreviations (*e.g., lol*). If these trends continue, emojis may be a socio-technical solution to the “problem” of non-standard orthography, representing a rare case in which a institutionally-driven change from above succeeds in displacing bottom up linguistic variation.

Because we approximate a randomized experiment using observational data, there are a number of challenges in making causal claims. Although we control for the key confounding factors that affect both the treatment and outcome, there is always the possibility of other confounding attributes of which we are unaware. This quantitative approach could be paired with a qualitative ethnographic investigation to learn about user attitudes towards emojis and non-standard orthography. Another limitation of the current study is that it is restricted to Twitter. Analyzing emoji usage in other social media sites such as Facebook, Tumblr, and Instagram would reveal if similar patterns emerge across platforms.

Aside from addressing these limitations, there are several other directions for future work. We view the increasingly widespread integration of pictographic content with text to be a particularly interesting linguistic development of the CMC period. The current study provides evidence the linguistic function of emojis by examining their competition with emoticons. More fine-grained linguistic analysis is needed to determine how emojis function and interact with text in various syntactic and pragmatic contexts. We are also interested in analysis across users: it is possible that more committed Internet users may use emojis in a more structured way, particularly in regard to ordering and combination of multiple emojis. Such research may be considered as first steps towards a systematic understanding of how the pictographic and textual modalities work together to create meaning in contemporary computer-mediated communication.

Figures and Tables

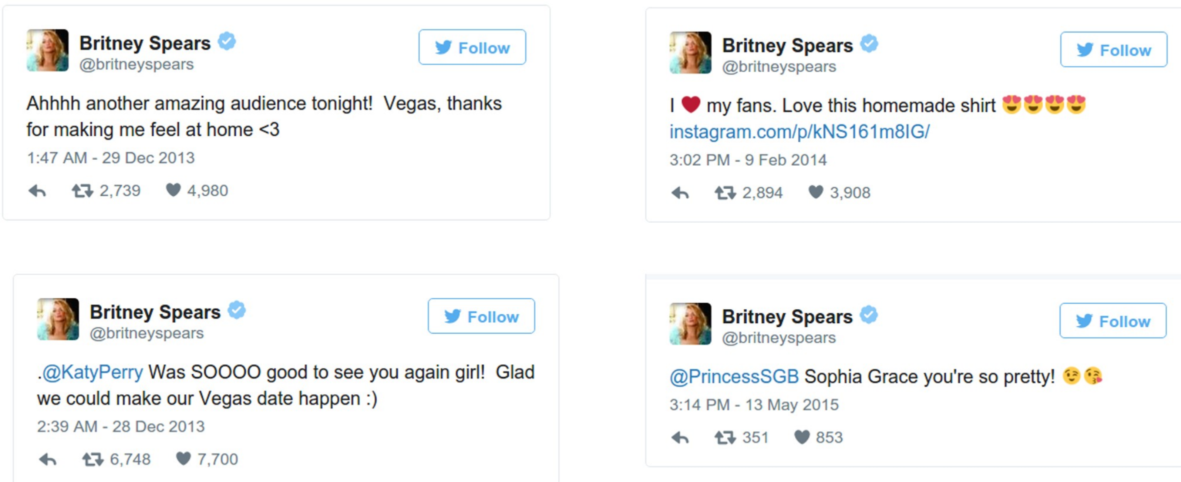
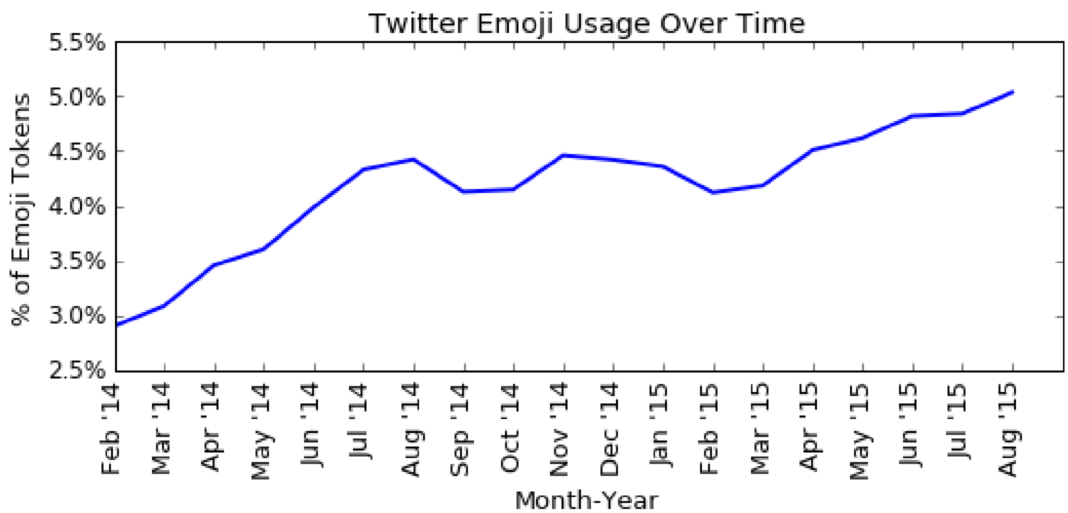
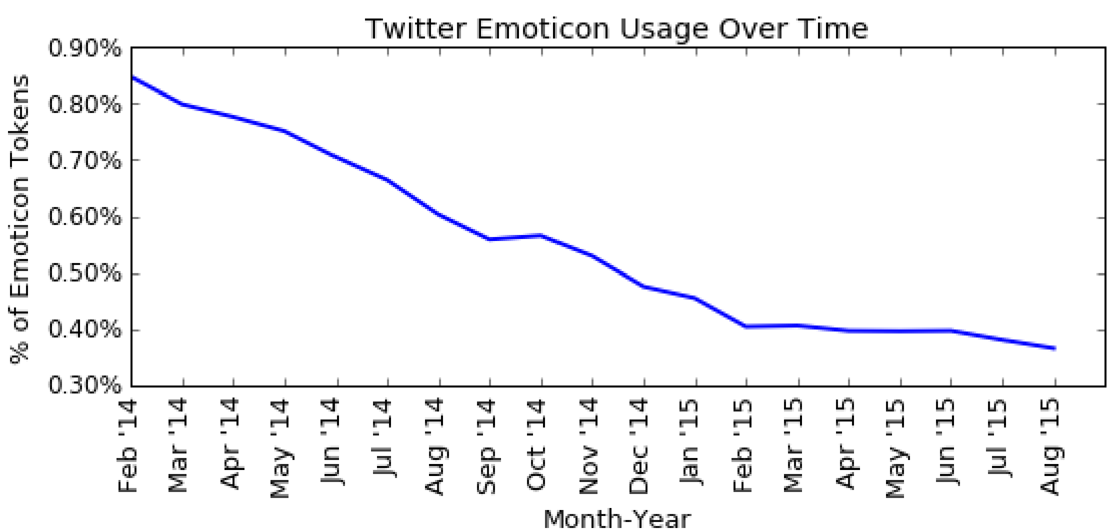


Figure 1: Example tweets from the musician Britney Spears.

Figure 2: Examples of emoji characters used in Twitter (created using http://www.iemoji.com).



(a)



(b)

Figure 3: Temporal trend of percentage of (a) Emoji and (b) Emoticon tokens in tweets from our sample.

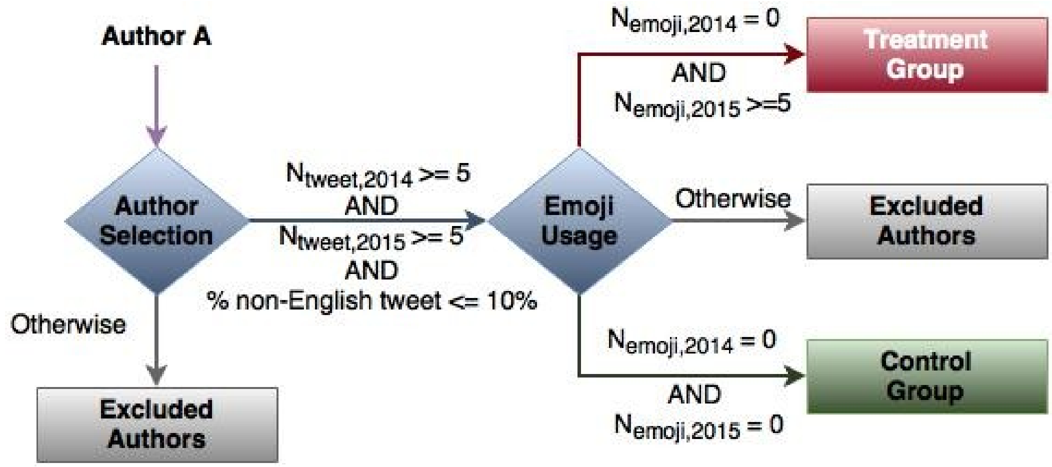


Figure 4: Selection of authors into treatment and control groups.

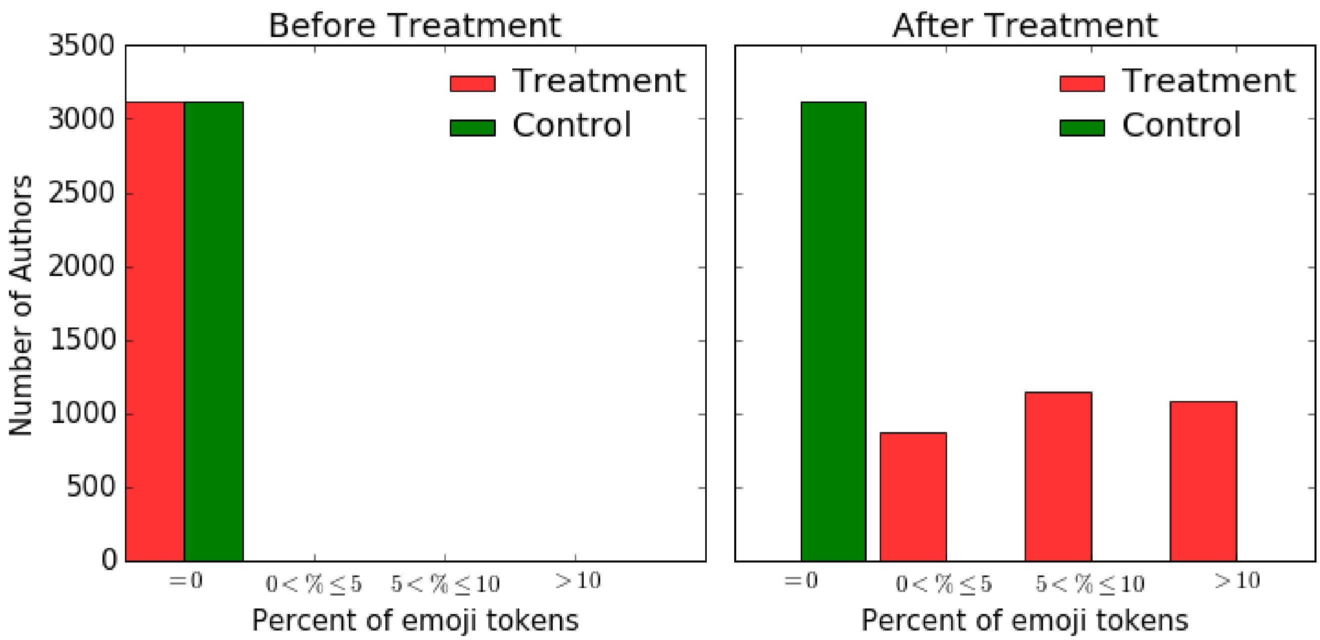


Figure 5: Analysis-I: Emoji usage of treatment and control groups.

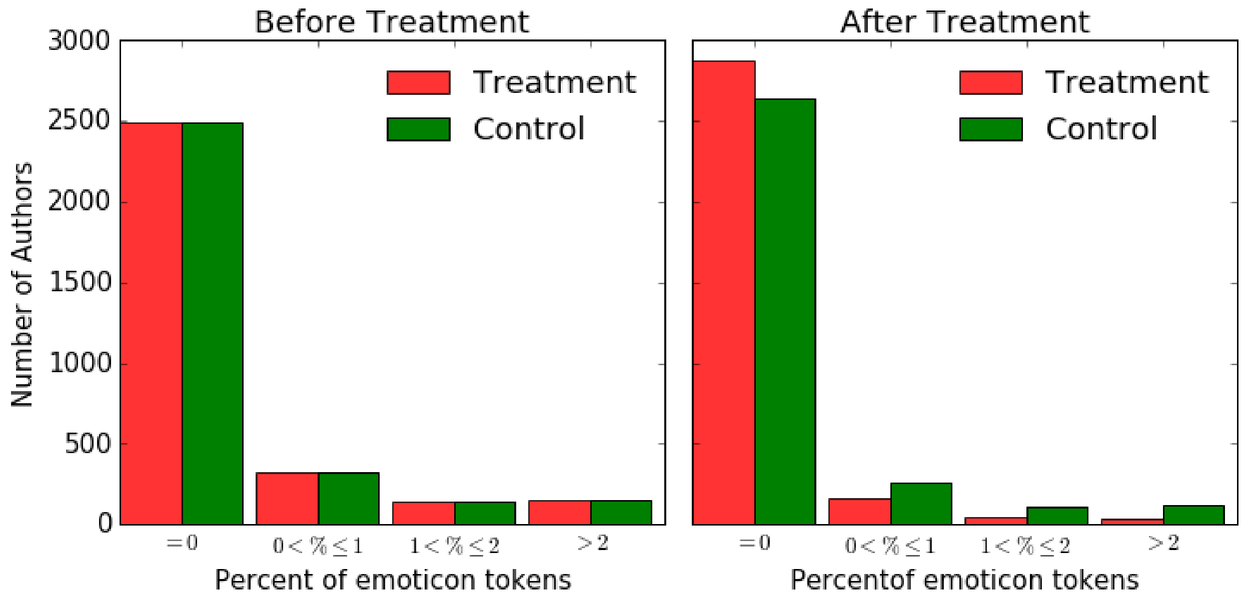


Figure 6: Analysis-I results: Emoticon usage of treatment and control groups.

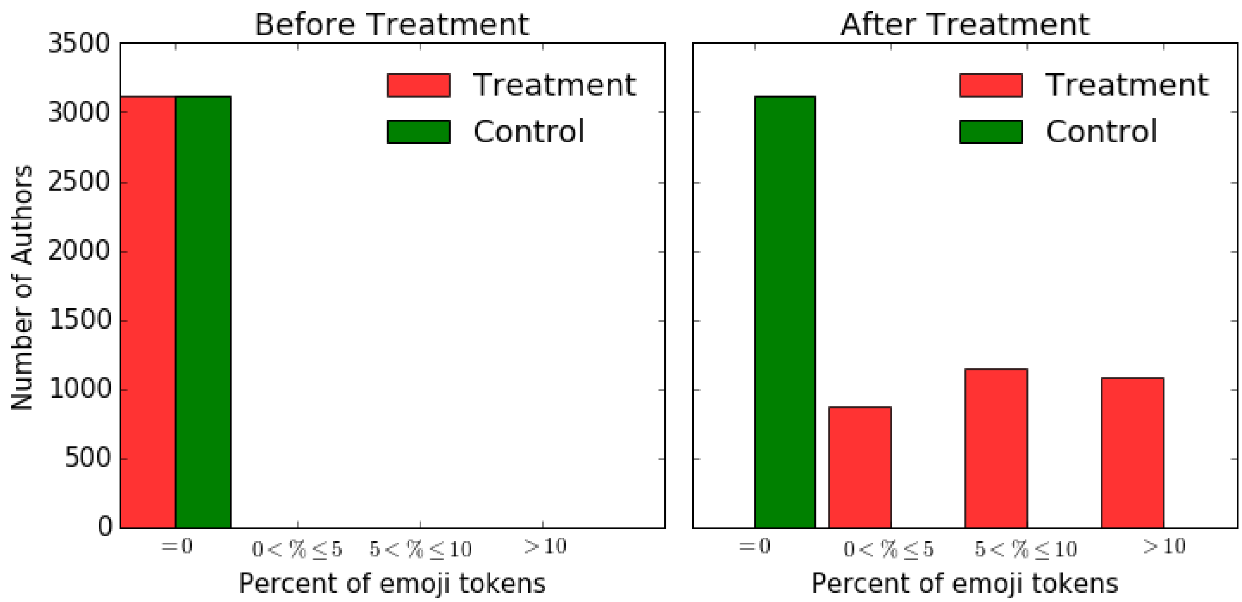


Figure 7: Analysis-II: Emoji usage of treatment and control groups.

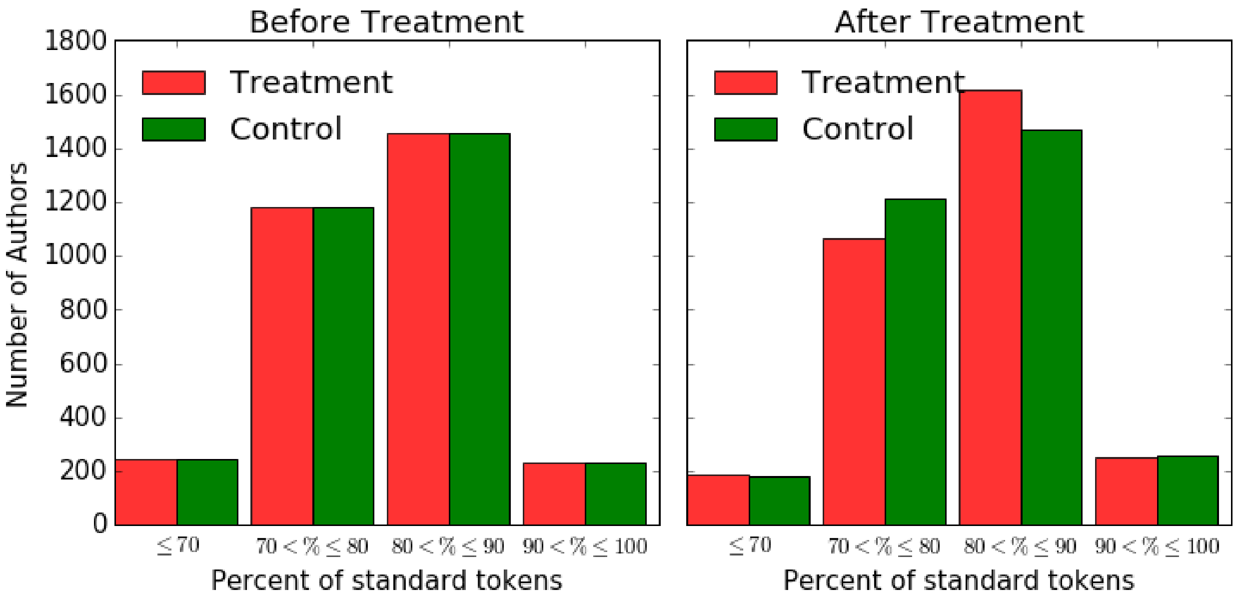


Figure 8: Analysis-II results: Standard token usage of treatment and control groups.

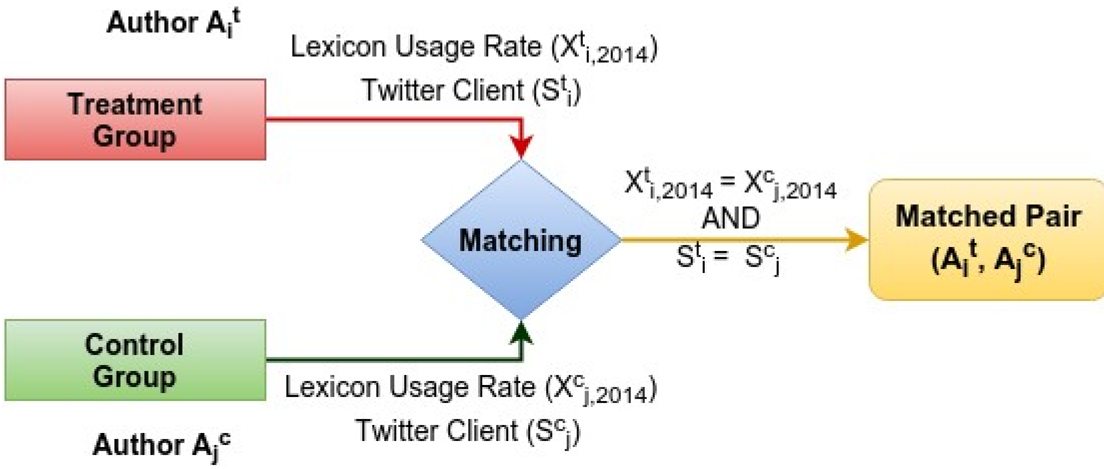


Figure 9: Matching procedure used to select treatment and control author pairs.

March 2014		March 2015	
Emoticon	Cumulative Frequency	Emoticon	Cumulative Frequency
:)	40.60	:)	40.31
:(	53.15	:(	51.85
;)	61.23	:D	59.36
:-)	68.20	;)	66.37
:D	74.97	:-)	72.50
:/	78.55	:/	77.64
:P	80.60	:'(	79.75
;-)	82.30	:P	81.77
:p	83.57	;-)	83.54
-_-	84.81	:-(	84.57
:-(	85.92	:p	85.56
:o	86.51	^_^	86.32
:O	87.03	=)	86.96
=)	87.50	-_-	87.53
^_^	87.91	*^*	88.09
:-D	88.32	:-D	88.63
:-/	88.66	:O	89.17
;D	88.99	:o	89.71
-_-	89.29	(-:	90.03
(-:	89.51	:-/	90.34

Table 1: Twenty most frequently-used emoticons and their cumulative frequencies.

Measure	**Emoticon**	**Emoji**
percent of messages with feature	2.05%	13.85%
percent of authors who have used the feature at least once	49.16%	64.46%
Of the messages with emoticons/emojis, percent with ...		
feature at the beginning	0.77%	12.40%
feature at the end	47.51%	54.15%
feature as sole token	0.16%	0.33%
Tokens of only this feature type	0.17%	1.01%
Of the messages with mentions/hashtags, percent of messages with the feature when ...		
@-mention		
at the beginning	5.14%	17.10%
not at the beginning	3.75%	27.12%
hashtag		
at the beginning	0.66%	4.10%
not at the beginning	1.15%	8.21%

Table 2: Emoticon and emoji usage statistics.

Source	**No. of unique terms**
Redhat Unix Dictionary	479,829
Ubuntu Unix Dictionary	99,171
English Wikipedia Corpus	266,695*
Combined List	621,968

Table 3: Standard word list statistics. (*Terms which appeared in more than 50 different articles.)

Tweet client	Percent of tweets
Twitter for iPhone	36.11
web	15.91
Twitter for Android	12.32
twitterfeed	3.32
TweetDeck	2.76
Twitter for iPad	2.04
TweetAdder v4	1.52

Table 4: Different clients from which users posted tweets and their tweet-level usage. These top seven clients cover a share of nearly 75 percent cumulative frequency of all clients.

		Analysis-I	**Analysis-II**
		Average *percent* of emoticon token rate	Average *percent* of standard token rate
Pre- treatment Period	Treatment Group	0.40%	85.20%
	Control Group	0.40%	85.17%
	t-statistics	-0.027	0.191
	p-value	9.79×10^{-1}	8.48×10^{-1}
Post-treatment Period	Treatment Group	0.12%	86.15%
	Control Group	0.31%	85.66%
	t-statistics	-9.612	2.755
	p-value	1.01×10^{-21}	5.89×10^{-3}
Average Treatment Effect		0.19%↓	0.45%↑

Table 5: Causal inference results for Analysis-I and Analysis-II.

Feature	**Description**	**Example**
POST-TREATMENT	Indicator variable for treatment period	
NOSE	Emoticon with a nose	*:-) :-/*
ORIENTATION	Horizontal facial orientation	*O_O -_-*
EQUAL-EYES	Eyes represented with an equal sign	*=) =D*
WINK-EYES	Wink eyes	*;-) ;D*
TONGUE	Mouth with a tongue	*:P ;P*
SLANT	Slant mouth	*:/ :-*
SMILE	Smiling mouth	*:) ;)*
D-MOUTH	Laughing mouth	*:D ;-D*
O-MOUTH	'O' mouth	*:O :-O*
TEARS	Emoticon with tears	*:') :'(*
REVERSED	Reversed order of symbols	*(-: (:*

Table 6: Stylistic features of emoticons, motivated by the analysis in Schnoebelen (2012).

Feature	Estimate	p-value
INTERCEPT	-7.4013	$<2e^{-16}$ ***
POST-TREATMENT	-0.4672	$3e^{-5}$ ***
NOSE	-0.1198	0.4169
ORIENTATION	-1.1698	0.0009 ***
EQUAL-EYES	-0.3620	0.3882
WINK-EYES	-0.5275	0.0006 ***
TONGUE	-0.2383	0.4167
SLANT	0.3697	0.0748 .
SMILE	-0.2020	0.1164
D-MOUTH	0.0092	0.9663
O-MOUTH	0.4227	0.3569
TEARS	-1.3384	0.2014
REVERSED	0.0847	0.9189
Significance codes (one-tailed p-values): *** : 0.001, ** : 0.01 , * : 0.05 , . :0.1 ,`` '' 1		

Table 7: Mixed-effect binomial regression results. We used the glmer method available in the lme4 package of R.[11]

About the Authors

Umashanthi Pavalanathan is a PhD student in the School of Interactive Computing at the Georgia Institute of Technology. Her area of research is computational analysis of language in online social media, with a focus on language variation and change. E-mail: umashanthi@gatech.edu

Jacob Eisenstein is an Assistant Professor in the School of Interactive Computing at the Georgia Institute of Technology. He studies sociolinguistic variation in online social media, with an emphasis on computational methodologies motivated by statistical machine learning. Jacob is also active in the field of Natural Language Processing, designing software algorithms to automatically identify structure and meaning in text and speech. He is a recipient of the CAREER award from the National Science Foundation, and the George M. Sprowls Award for his doctoral thesis at the Massachusetts Institute of Technology. E-mail: jacobe@gatech.edu

1 http://www.unicode.org/reports/tr51/index.html#Selection_Factors

2 http://www.unicode.org/reports/tr51/index.html#Diversity

3 http://blog.oxforddictionaries.com/2015/11/word-of-the-year-2015-emoji/

4 https://twitter.com/support/status/451399393850052610

5 http://www.nytimes.com/2016/03/07/business/media/picture-this-marketers-let-emojis-do-the-talking.html?_r=2

6 http://kt.ijs.si/data/Emoji_sentiment_ranking/

7 http://www.cs.cmu.edu/~ark/TweetNLP/

8 https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style#Vocabulary

9 We experimented with different threshold values for the number of unique articles and the overall results of our analyses hold the same for different values.

10 Note that these usage distributions are different from the findings reported by Perreault and Ruths (2011), who used a tweet corpus collected in 2011. The increased prevalence of mobile devices might be a reason for authors using multiple devices to post tweets.

11 https://cran.r-project.org/web/packages/lme4/lme4.pdf

References

Joshua D Angrist and Jörn-Steffen Pischke, 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

Steven Bedrick, Russell Beckley, Brian Roark and Richard Sproat, 2012. “Robust kaomoji detection in twitter,” In: *Proceedings of the Second Workshop on Language in Social Media*, Association for Computational Linguistics, pp. 56–64, at http://aclweb.org/anthology/W/W12/W12-2107.pdf, accessed 03 August 2016.

danah boyd, Scott Golder and Gilad Lotan, 2010. “Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter,” In: *43rd Hawaii International Conference on System Sciences (HISS)*, IEEE, pp. 1–10. doi: http://dx.doi.org/10.1109/HICSS.2010.412, accessed 03 August 2016.

John Carey, 1980. “Paralanguage in computer mediated communication,” In: *Proceedings of the 18th Annual Meeting on Association for Computational Linguistics*, Association for Computational Linguistics, pp. 67–69. doi: http://dx.doi.org/10.3115/981436.981458, accessed 03 August 2016.

Justin Cheng, Cristian Danescu-Niculescu-Mizil and Jure Leskovec, 2015. “Antisocial Behavior in Online Discussion Communities,” In: *Proceedings of the International Conference on Web and social media (ICWSM)*, at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10469, accessed 03 August 2016.

Lynn Cherny, 1999. *Conversation and community: chat in a virtual world*, CSLI publications.

David Crystal, 2001. *Language and the Internet*, Cambridge University Press.

Cristian Danescu-Niculescu-Mizil, Michael Gamon and Susan Dumais, 2011. “Mark my words!: linguistic style accommodation in social media,” In: *Proceedings of the 20th International Conference on World Wide Web*, ACM, pp. 745–754. doi: http://dx.doi.org/10.1145/1963405.1963509, accessed 03

August 2016.

Munmun De Choudhury, Sanket Sharma and Emre Kiciman, 2016. "Characterizing Dietary Choices, Nutrition, and Language in Food Deserts via Social Media ," In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social computing*, ACM, pp. 1157–1170. doi: http://dx.doi.org/10.1145/2818048.2819956, accessed 03 August 2016.

Daantje Derks, Arjan ER Bos and Jasper Von Grumbkow, 2007. "Emoticons and social interaction on the Internet: the importance of social context," *Computers in Human Behavior*, volume 23, number1, pp. 842–849. doi: http://dx.doi.org/10.1016/j.chb.2004.11.013, accessed 03 August 2016.

Eli Dresner and Susan C Herring, 2010. "Functions of the Nonverbal in CMC: Emoticons and Illocutionary Force," *Communication Theory*, volume 20, number3, pp. 249–268. doi: http://dx.doi.org/10.1111/j.1468-2885.2010.01362.x, accessed 03 August 2016.

Suzanne Eggins, 2004. *Introduction to Systemic Functional Linguistics: 2nd edition*, Bloomsbury Academic.

Jacob Eisenstein, 2015. "Systematic patterning in phonologically-motivated orthographic variation," *Journal of Sociolinguistics*, volume 19, number2, pp. 161–188. doi: http://dx.doi.org/10.1111/josl.12119, accessed 03 August 2016.

Jacob Eisenstein, 2013. "What to do about bad language on the internet," In: *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL)*: Association for Computational Linguistics, pp. 359–369, at http://www.anthology.aclweb.org/N/N13/N13-1037.pdf, accessed 03 August 2016.

Constantine E Frangakis and Donald B Rubin, 2002. "Principal Stratification in Causal Inference," *Biometrics*, volume 58, number1, pp. 21–29. doi: http:/dx.doi.org/10.1111/j.0006-341X.2002.00021.x, accessed 03 August 2016.

Andrew Gelman and Jennifer Hill, 2006. *Data Analysis using Regression and Multilevel/Hierarchical Models*, Cambridge University Press.

Howard Giles and Tania Ogay, 2007. “Communication Accommodation Theory,” *Explaining Communication: Contemporary Theories and Exemplars*, pp. 293–310.

Stephan Gouws, Donald Metzler, Congxing Cai and Eduard Hovy, 2011. “Contextual bearing on linguistic variation in social media,” In: *Proceedings of the Workshop on Languages in Social Media*, Association for Computational Linguistics, pp. 20–29, at http://dl.acm.org/citation.cfm?id=2021109.2021113, accessed 03 August 2016.

Daniel E. Ho, Kosuke Imai, Gary King and Elizabeth A. Stuart, 2011. “MatchIt: Nonparametric Preprocessing for Parametric Causal Inference,” *Journal of Statistical Software*, volume 42, number8, pp. 1–28, at http://www.jstatsoft.org/v42/i08/, accessed 03 August 2016.

Yoram M Kalman and Darren Gergle, 2014. “Letter Repetitions in Computer-Mediated Communication: a unique link between spoken and online anguage,” *Computers in Human Behavior*, volume 34, pp. 187–193. doi: http://dx.doi.org/10.1016/j.chb.2014.01.047, accessed 03 August 2016.

Ryan Kelly and Leon Watts, 2015. “Characterising the inventive appropriation of emoji as relationally meaningful in mediated close personal relationships,” *Experiences of Technology Appropriation: Unanticipated Users, Usage, Circumstances, and Design*, at http://opus.bath.ac.uk/46780/, accessed 03 August 2016.

Gary King, Jennifer Pan and Margaret E Roberts, 2014. “Reverse-engineering censorship in China: Randomized experimentation and participant observation,” *Science*, volume 345, number 6199. doi: http://dx.doi.org/10.1126/science.1251722, accessed 03 August 2016.

Farshad Kooti, Haeryun Yang, Meeyoung Cha, P Krishna Gummadi and Winter A Mason, 2012. “The Emergence of Conventions in Online Social Networks,” In: *Proceedings of the International*

Conference on Web and Social Media (ICWSM), at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4661, accessed 03 August 2016.

William Labov, 2006. *The social stratification of English in New York City*, Cambridge, U.K.: Cambridge University Press.

James McCorriston, David Jurgens and Derek Ruths, 2015. "Organizations Are Users Too: Characterizing and Detecting the Presence of Organizations on Twitter," In: *Proceedings of the International Conference on Web and Social Media (ICWSM)*, at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10578, accessed 03 August 2016.

Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen and Brent Hecht, 2016. "Blissfully Happy" or "Ready toFight": Varying Interpretations of Emoji," In: *Proceedings of the International Conference on Web and Social Media (ICWSM)*, at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13167, accessed 03 August 2016.

Lev Muchnik, Sinan Aral and Sean J Taylor, 2013. "Social Influence Bias: A Randomized Experiment," *Science*, volume 341, number6146, pp. 647–651. doi: http://dx.doi.org/10.1126/science.1240466, accessed 03 August 2016.

Petra Kralj Novak, Jasmina Smailović, Borut Sluban and Igor Mozetič, 2015. "Sentiment of Emojis," *PLOS ONE*, volume 10, number12, pp. doi: http://dx.doi.org/10.1371/journal.pone.0144296, accessed 03 August 2016.

Jaram Park, Vladimir Barash, Clay Fink and Meeyoung Cha, 2013. "Emoticon Style: Interpreting Differences in Emoticons Across Cultures." In: *Proceedings of the International Conference on Web and Social Media (ICWSM)*, at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6132, accessed 03 August 2016.

Mathieu Perreault and Derek Ruths, 2011. "The Effect of Mobile Platforms on Twitter Content

Generation.” In: *Proceedings of the International Conference on Web and Social Media (ICWSM)*, at http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2798, accessed 03 August 2016.

Jose Pinheiro and Douglas Bates, 2006. *Mixed-effects models in s and s-pLUS*, Springer Science & Business Media. doi: https://dx.doi.org/10.1007/b98882, accessed 03 August 2016.

Virgile Landeiro Dos Reis and Aron Culotta, 2015. “Using matched samples to estimate the effects of exercise on mental health from Twitter,” In: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, pp. 182–18, at http://dl.acm.org/citation.cfm?id=2887007.2887033, accessed 03 August 2016.

Landra Rezabek and John Cochenour, 1998. “Visual Cues in Computer-Mediated Communication: Supplementing Text with Emoticons,” *Journal of Visual Literacy*, volume 18, number2, pp. 201–215. doi: http://10.1080/23796529.1998.11674539, accessed 03 August 2016.

Paul R Rosenbaum and Donald B Rubin, 1983. “The central role of the propensity score in observational studies for causal effects,” *Biometrika*, volume 70, number1, pp. 41–55. doi: http://10.1093/biomet/70.1.41, accessed 03 August 2016.

Tyler Schnoebelen, 2012. “Do you smile with your nose? stylistic variation in Twitter emoticons,” *University of Pennsylvania Working Papers in Linguistics*, volume 18, number2, pp. 14, at http://repository.upenn.edu/pwpl/vol18/iss2/14, accessed 03 August 2016.

Luke Stark and Kate Crawford, 2015. “The conservatism of Emoji: work, affect, and communication,” *Social Media+ Society*, volume 1, number2. doi: http://dx.doi.org/10.1177/2056305115604853, accessed 03 August 2016.

Alecia Wolf, 2000. “Emotional Expression Online: Gender Differences in Emoticon Use,” *CyberPsychology & Behavior*, volume 3, number5, pp. 827–833, http://dx.doi.org/10.1089/10949310050191809, accessed 03 August 2016.

Sarita Yardi, Daniel Romero, Grant Schoenebeck and others, 2009. “Detecting spam in a Twitter network,” First Monday, volume 15, number. doi: http://dx.doi.org/10.5210/fm.v15i1.2793, accessed 03 August 2016.