Hiding Behind the Words of Others: Does Redundant Word Choice Reflect Suppressed Individuality When Tweeting in the First Person Singular?

Chapter · January 2019

DOI: 10.1007/978-3-319-94709-9_59

CITATIONS
0

READS
68

4 authors:

Eliza Barach
University at Albany, The State University of New York
9 PUBLICATIONS **9** CITATIONS
SEE PROFILE

Samira Shaikh
University of North Carolina at Charlotte
50 PUBLICATIONS **198** CITATIONS
SEE PROFILE

Vidhushini Srinivasan
University of North Carolina at Charlotte
3 PUBLICATIONS **2** CITATIONS
SEE PROFILE

Laurie Beth Feldman
University at Albany, The State University of New York
116 PUBLICATIONS **3,616** CITATIONS
SEE PROFILE

Hiding Behind the Words of Others: Does Redundant Word Choice Reflect Suppressed Individuality When Tweeting in the First Person Singular?

Eliza Barach[1(✉)], Samira Shaikh[2], Vidhushini Srinivasan[2], and Laurie Beth Feldman[1,3]

[1] State University of New York – University at Albany, 1400 Washington Avenue, Albany, NY 12222, USA
{ebarach, lfeldman}@albany.edu
[2] University of North Carolina at Charlotte, Charlotte, NC 28223, USA
{sshaikh2, vsriniv6}@uncc.edu
[3] Haskins Laboratories, 300 George Street, New Haven, CT 06211, USA

Abstract. We analyze naturally occurring social media data that derive from Twitter messages posted over a 24-h period in immediate reaction to the Paris terrorist attacks in November 2015. We separately examine patterns for tweets with first-person singular pronouns (I) and first-person plural pronouns (WE), the corresponding variations in valence, arousal, proportion of words in various LIWC categories, and diversity of word choices within those categories. Negatively valenced word choices revealed greater mean differences between I and WE than did positively valenced words. Novel was that tweets with I exhibited a more uniform distribution of word choices and greater linguistic alignment, for most of the LIWC categories and for both positively and negatively valenced word choices, relative to tweets in WE. Greater diversity differences associated with pronoun choice when valence is negative than when it is positive suggest less self-disclosure when tweeting with first-person singular than plural pronouns.

Keywords: Social media · Linguistic alignment · Pronoun choice
Valence · Self-disclosure

1 Introduction

Language is the primary medium for human social interactions. With the advancement of technology and new methods to generate and analyze data, investigations into spontaneous language production have become useful for understanding how people's word choices in their daily interactions reflect who they are as well as what they are doing [1]. Communication by Social Media, which is now one of the most common internet-based activities [2], has become a prolific source of text to analyze. Specifically, the platform Twitter, in which conversations or "tweets" are posted publicly, is one of the largest available repositories of naturally occurring language [3]. As a

J. I. Kantola et al. (Eds.): AHFE 2018, AISC 783, pp. 603–614, 2019.
https://doi.org/10.1007/978-3-319-94709-9_59

microblog, Twitter provides a social network structure and an avenue for the flow of information, where users can tweet updates and "follow" other users so that other users' tweets are updated in his or her feed [4] (i.e. generating a compilation of all the tweets from users the individual follows). Moreover, Twitter's format encourages users to disclose personal details about their daily life, share and seek information [5] communicate with many other users who are not necessarily confined to a narrow group of "friends" [6]. Online communication can be triggered spontaneously, or be elicited by specific events [7], such as catastrophic world events (e.g. terrorist attacks) [8].

Using Twitter data as a corpus from which to investigate conversational characteristics of users has already proven itself to be a valuable scientific endeavor [8]. For example, previous research has revealed differences in patterns of word choice among users responding to the Charlie Hebdo terrorist attack in Paris in 2015 [8]. Indeed, we know that word choice and in particular pronoun choice, can be revealing about an individual's emotional state; across multiple studies research has shown that increased use of first-person singular (I) is associated with increases in negatively valenced word choice [8] as well as negative affective states [1, 9].

While previous work has focused on the counts and proportion of words within different categories [10, 11] as well as the pattern of word choices that co-occur with particular pronouns [8], little work has addressed how diversity/uniformity of word choice varies with pronoun choice. Generally, research has shown that social exchanges reflect a pattern of accommodation in communication, whereby users in remote as well as face-to-face conversations tend to converge to one another's communicative behavior [3]. This coordination among speakers occurs at a variety of linguistic levels including word choice, syntactic structures, utterance length, pitch and gesture [3]. In one case, analyses of communication accommodation (also called linguistic alignment) among Twitter users, revealed convergence, such that users chose words from the same categories as defined by the Linguistic Inquiry Word Count (LIWC) [12]. While the research showed significant alignment among users (i.e. conversations between individuals contained approximately equal number of words from the same category (e.g., A*rticle*) even though those numbers were atypical relative to a baseline) [12], their measure of alignment did not consider the frequency of specific word choices. Instead, alignment focused only on the number of word forms that belong to a particular category. Further, categories included in the LIWC-based analyses were limited to "non-topical style dimensions" having few to no content words that are free to vary across topics. For example, analysis on the prevalence of particular *positive emotion* (e.g., love, nice, sweet) and *negative emotion* (e.g., hurt, ugly, nasty) words as well as *cognitive processe*s (e.g., cause, know, ought) or *social process* (e.g., mate, talk, they) words were not a priority. Thus, alignment focused primarily on counts of non-content words with only secondary attention to sentiment.

In the present study we describe social dynamics within a virtual group formed in response to a particular global crisis. We analyze naturally occurring social media data that derive from Twitter messages in immediate reaction to the Paris terrorist attacks in November 2015. Our Twitter corpus is defined by the hashtags users append to their tweets thereby dramatically restricting the topic domain. Our primary measure of behavior in this virtual group context is based on variation in (content) word choice. In our previous work, we have documented patterns based on an entropy based measure

of lexical diversity in virtual groups that respond to a real life event [8, 13] as well as in more permanent groups who share a goal or interest that persists over time [14]. Our method differs from most current approaches that describe patterns of vocabulary usage with respect to deviations of a group mean frequency from a baseline mean.

Frequency is typically based on number of instances of a word from a particular category without regard to whether one word occurs multiple times or multiple words occur only once. For example, ongoing emotional distress (viz., depression) has been documented to increase the incidence of first-person singular pronouns (I) as well as the prevalence of negatively valenced words in essays by depressed writers relative to those without distress [15]. Similarly, analyses of diary entries associated with coping with the 2001 NYC attack show that changes in word frequency over a two-week period were most dramatic in those who often used critical terms (e.g., Osama, World Trade Center) in their writing and were thus characterized as most preoccupied with the event [16]. Additionally, diary entries exhibited a shift in pronoun choice, such that the use of the first-person plural pronouns (WE) dramatically increased and correspondingly the use of first-person singular pronouns (I) decreased [16]. This shift from the use of first-person singular pronouns to first-person plural pronouns suggests a change in focus from attending to themselves to thinking about friends, family and others within their group [1]. Indeed, this shift is consistent with the idea that the use of first-person singular pronouns is related to the independent or individualist self and the use of first-person plural pronouns is related to the interdependent or collectivist self [17–19]. Moreover, the mere presence of pronouns is associated with differences in word choice. Specifically, the use of more strongly valenced words increases with pronouns, relative to when pronouns are absent [8], which denotes greater emotion. Interestingly concreteness, which is often interpreted as a measure of psychological distance [20], also varies along with pronoun choice, where first-person plural pronouns tend to accompany less concrete word choices, than first-person singular word choices [8]. However, these previous analyses were based on mean token frequency and failed to consider diversity in word choice among those who discuss the same event. Reduced lexical diversity when the topic is held constant can signal communication coordination and linguistic alignment among speakers and can thus be revealing about social identity in the context of a group [21, 22].

One factor that is likely to influence the potential for alignment is the tendency to talk about oneself (self-disclosure) because the degree of self-disclosure tends to be associated with degree of interaction in support seeking posts [23]. In the present study, we link self-disclosure to lexical diversity with the understanding that the tendency for linguistic alignment and reduced diversity varies systematically depending on pronoun choice and the corresponding perspectives of the self (i.e. I, individualist self vs. WE, collectivist self). Thus, the current study analyzes tweets in terms of the uniformity or diversity of word choice as it co-varies with pronoun choice when immediately responding to a terrorist attack.

Some have reported an increase in the use of the first-person plural when responding to a terrorist attack, interpreting it as a need to be part of a group in the face of an outgroup threat [24]. However, these findings are over an extended window of time (sometimes a two-week period), which may allow for the development of an online collectivist community, whereas a more restricted window of time may not.

Moreover, the greater use of I in association with more negative affective states or of WE in association with more positive affective states, would be consistent with previous research in which responses to a catastrophic event exhibited a brief drop in the total number of *positive emotion* words [16] as well as a tendency for negatively valenced and *negative emotion* words when using first-person singular pronouns [1, 8, 9, 15]. Thus, we expect to replicate the pattern of greater or more negatively valenced word choices and less positively valenced word choices when tweeting in the pronoun I relative, to tweeting in the pronoun WE.

To our knowledge, this is the first study that will address linguistic alignment based on first-person pronoun choice (i.e. is there a difference in uniformity of concomitant word choice when tweeting in I vs. WE). We extend analyses of linguistic alignment by examining word choice uniformity (lexical diversity), rather than LIWC category usage based on counts. We analyze LIWC word categories to see whether linguistic alignment generalizes across categories of content words and how that may vary with pronoun choice in response to a terrorist attack. Using a measure of lexical diversity rather than word count, we examine reduced lexical diversity and greater alignment of word choice for tweets with the pronoun I relative to WE. If this is a general case of alignment then we hypothesize that this pattern should be evident for both positively valenced and negatively valenced word choices, as well as for word choices in the LIWC categories. If the pressure of self-disclosure also plays a role, because attention to the self is exaggerated when using the I pronoun [1], lexical diversity based on our entropy measure of alignment may be reduced more among words whose valence is negative than whose valence is positive.

2 Data Collection and Analysis

The data we collected consisted of all the tweets that contained at least one of 14 hashtags pertaining to the November 2015 Paris attacks collected via Twitter streaming API. We then defined a subset category of hashtags so that our final data set always contained at least one hashtag that pertained to Paris (see Table 1). The corpus consisted of 43,851 tweets were shared in one day immediately following the Paris attacks.

Table 1. Examples of some of the hashtags in our tweet collection

#franceattack	#IStandWithFrance	#Paris
#franceshooting	#IStandWithParis	#ParisAttack
#FranceSolidarity	#PrayersForParis	#PeaceForParis

To determine whether lexical variation and linguistic alignment is associated with pronoun choice in our data, we classified tweets with the relevant hashtags as to whether first-person singular pronouns (I, me, my, mine) were present or whether first-person plural pronouns (we, our, ours, us) were present. Two sets of data pertaining to the November Paris attacks were then developed based on pronoun choice: the I set and the WE set. Separating tweets that use first-person singular pronouns (I) and tweets

with first-person plural pronouns (WE) allowed us to differentiate between the collectivist self and individualist self [17–19, 25], and are henceforth designated as I and WE. Moreover, based on an extension of Pan et al., [23] tweets in I that presumably reflect the individualist self and tweets in WE, that presumably reflect the collectivist self, differ with respect to the potential for self-disclosure. In order to ascertain whether pronoun choice was systematic within an individual, for each user ID, we tracked the number of tweets in the corpus that contained an I pronoun and the number of tweets that contained a WE pronoun. We determined that 1% of tweeters were affiliated with both groups and 99% were affiliated with only one.

Each of the two sets of data (I and WE) was pre-processed as follows: We first removed stop-words (non-content words such as the, a etc.) from the tweets, and then converted all words to their lower-case equivalents and stemmed them so that inflected forms (terrorist, terrorists) would be counted together. Then for all words with valences more extreme than 25% (viz., 25% most positive and 25% most negative) in Warriner and Kuperman [26], we calculated the frequencies with which each word appeared in each pronoun set (I, WE). We then retrieved the valence (the pleasantness of a given word) and arousal scores (intensity of evoked emotion, e.g., terrified > grief) of extreme words. We calculated weighted means for arousal (and valence) for each word using the following formula: multiplying the arousal (valence) of a word by the number of times that word appeared in the corpus and dividing by sum of the frequencies of all words. Means are summarized in Table 2 and I and WE word clouds with font size proportional to relative token count are depicted in Fig. 1.

Table 2. Weighted valence and arousal means and valence entropy

Pronoun	Polarity	Valence	Arousal	Entropy	CI
I	NEG	2.36	5.89	5.019	5.0186–5.0195
WE	NEG	2.52	5.67	6.193	6.1929–6.1940
I	POS	7.06	4.31	5.374	5.3737–5.3743
WE	POS	7.10	4.22	6.114	6.1140–6.1146

To examine the variation in the types of word choices for each data set (I and WE), we analyzed the type and token counts of words associated with the different categories in the LIWC data base [10], with the requirement that the cumulative frequency (the sum of the number of times each word appeared) in each LIWC category exceeded 1000. This resulted in the further analysis of the following LIWC categories: *affiliation* (e.g., help, friend, encourage), *anger* (e.g., hate, kill, annoyed), *cognitive processes* (e.g., know, wish, thought), *death* (e.g., kill, war, murder), *drives* (e.g., friend, win, attack), *family* (e.g. daughter, dad, aunt), *insight* (e.g., think, know, believe), *power* (e.g., strong, superior, bully), *positive emotion* (e.g., love, safe, support), *negative emotion* (e.g., attack, terror, suffer), *religion* (e.g., pray, jihad, soul), *risk* (e.g., safety, danger, threat), *sadness* (e.g., tragedy, cry, sad), and *social processes* (talk, give, sympathy). There was a total of 7373 tweets that included a variant of the I pronoun and 6440 tweets that included a variant of the WE pronoun; more tweets that use I fails to replicate previous studies collected in response to crises events. However, our time

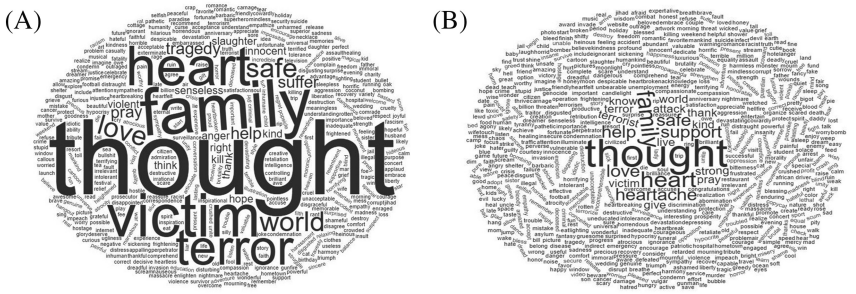

Fig. 1. (A) word cloud showing the most common words that occurred with first-person singular pronouns (I), (B) word cloud showing the most common words that occurred with first-person plural pronouns (WE)

window is limited to 24-h, which may be too short to capture the development of an online collectivist community and may also have consequences for the types of effects we can detect. LIWC category percentages were computed by dividing the total frequency of words in each category for each pronoun (I and WE) by the total number of tweets that included either the I pronoun or the WE pronoun (See Fig. 2 for LIWC percentages).

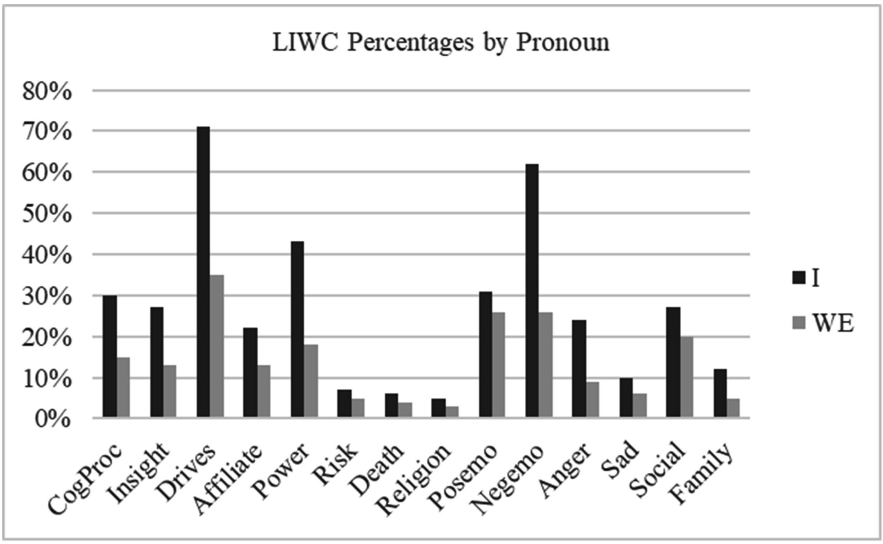

Fig. 2. LIWC percentages

To ascertain the lexical redundancy (or reduced word choice variation) of the extremely valenced words as well as the words in the LIWC categories for tweets in each pronoun, we calculated entropy. The concept of entropy is derived from

information theory [27]. Entropy is defined as a unit-less measure of uncertainty such that the more unpredictable the pattern is, for example here, in terms of number of different words and the frequency of each, the higher the entropy. We calculated entropy separately for words with the 25% most extreme negative and positive values so that we could compare variation in word choice in tweets with I and WE pronouns for valence. Lesser variation as revealed by greater alignment would be consistent with greater mutual influence; greater variation when positive could be symptomatic of more independence between tweeters. The criterion for entropy analyses of a LIWC category for each pronoun set required that at least 100 different words appear within a category. This resulted in the inclusion of the following LIWC categories: *cognitive processes*, *drives*, *power*, *positive emotion*, *negative emotion*. Again, lesser variation amongst the individual categories suggests greater shared influence (linguistic alignment) and greater variation would signal greater independence (less linguistic alignment). Greater differences between negative and positive valence for tweets associated with I than with WE would be consistent with a role for self-disclosure.

We determined entropy scores by summing the total number of extremely valenced words and dividing the number of occurrences for each word by the total number of emotionally valenced words to get relative frequency. Those numbers were logged and then we calculated the product of relative frequency by its log and changed its sign. Because entropy is sensitive to unequal sample sizes, we used the ChaoJust adjustment which takes into account this sample issue. In the end, low entropy scores indicate greater redundancy and higher scores indicate less redundancy. Entropy scores on words with extreme valence are listed in Table 2. This same process was used for calculating entropy among the words in each of the included LIWC categories (see Table 3 for LIWC entropy scores).

Table 3. LIWC token and type count, entropy scores, and CIs

	Token (Type) count		Percentage		Entropy	
LIWC category	I	WE	I	WE	I (CI)	WE (CI)
Cognitive processes	4108 (64)	2110 (56)	30%	15%	2.34 (2.3375–2.3396)	3.14 (3.1350–3.1388)
Drives	9820 (117)	4835 (137)	71%	35%	3.86 (3.8632–3.8637)	4.99 (4.9867–4.9878)
Power	5876 (50)	2509 (58)	43%	18%	2.86 (2.8613–2.8621)	4.07 (4.0678–4.0693)
Positive emotion	4266 (120)	3546 (120)	31%	26%	4.44 (4.397–4.441)	4.65 (4.6486–4.6502)
Negative emotion	8585 (154)	3594 (140)	62%	26%	4.38 (4.3820–4.3829)	5.49 (5.4858–5.4870)
Social processes	3784 (64)	2762 (51)	27%	20%	3.00 (2.9991–3.001)	3.86 (3.8593–3.8608)

3 Results

3.1 Valence and Arousal and Entropy

Mean valence and arousal scores were calculated for each of the sets of data associated with the November Paris attack corpora. We further divided each data set based on polarity, such that negative words and positive words were in separate analyses. As valence scores depart from a score of 1.00, they become more positive. As arousal scores depart from 1.00 emotions become more intense. Mean valence and arousal scores for each pronoun are summarized in Table 2 and word clouds for each pronoun based on relative frequency for co-occurring word choices are depicted in Fig. 1. Greater uniformity in font size is consistent with higher entropy and greater lexical diversity.

Mean negative valence for tweets with I (2.36) was more negative than WE (2.52), which replicates our previous findings [8]. In contrast, mean positive valence for tweets with I (7.06) was slightly less positive than WE (7.10). In summary, tweeting in I was associated with more negatively valenced word choices and less positively valenced word choices than tweeting in WE. Results support the pattern of greater negativity or negative emotional states when communicating with first-person singular than plural pronouns [16]. Valence influenced the degree of arousal for word choices such that negatively valenced arousal was consistently greater than positively valenced arousal, irrespective of corresponding pronoun choice. Moreover, regardless of valence, mean arousal scores indicated that words that co-occurred with I (Negative arousal 5.89; Positive arousal 4.31) tended to be more extreme in arousal than words that co-occurred with WE (Negative arousal 5.67; Positive arousal 4.22), suggesting greater emotional intensity when tweeting in I. In summary, tweets using I tended to be more extreme in valence as well as in emotional intensity (arousal) than tweets using WE. Additionally, negatively valenced words showed greater differences between I and WE and bigger differences in arousal than did positively valenced words.

Entropy scores capture diversity of word choice and were determined by summing the total number of extremely valenced words and dividing the number of occurrences for each word by the total number of emotionally valenced words to get relative frequency. Those numbers were logged and then the product of relative frequency by its log was calculated, followed by a change in its sign. In the end, low entropy scores indicate greater redundancy (more alignment) and higher scores indicate less redundancy (less alignment). Valence entropy scores are listed in Table 2.

Regardless of valence, corresponding word choices for tweets in I exhibited lower entropy scores, than tweets in WE (I NEG 5.019, I POS 5.374; WE NEG 6.193, WE POS 6.114), suggesting greater linguistic alignment. Valence entropy differences for tweets in I, also showed more redundancy and greater linguistic alignment for negatively valenced word choices than positively valenced word choices. This asymmetry is consistent with a disinclination for self-disclosure. In contrast valence entropy differences for tweets in WE revealed a relatively weak albeit opposite pattern where less redundancy and less linguistic alignment was associated with positively valenced word choices than negatively valenced word choices.

In summary and consistent with the tendency to hide behind the words of others, tweeting in I, and presumably tweeting as the individualist self, is associated with greater linguistic alignment and a disinclination for self-disclosure relative to tweeting in WE and presumably as the collectivist self. Further, valence entropy differences varied with pronoun choice, such that negatively valenced word choices that occurred with I were more uniform and exhibited greater linguistic alignment, relative to positively valenced word choices that occurred with I. With respect to word choices for WE, negatively valenced word choices were only slightly more diverse and exhibited less linguistic alignment overall, relative to positively valenced word choices.

3.2 LIWC and Entropy

For each pronoun set (I, WE), we examined the proportion of accompanying word choices within each category from the LIWC in which the total number of occurrences exceeded 1000 (Percentages by pronoun are listed in Fig. 2.). The categories meeting this requirement were *affiliation, anger*, *cognitive processes*, *death*, *drives*, *family*, *insight*, *power*, *positive emotion*, *negative emotion*, *religion*, *risk*, *sadness*, and *social processes*.

For all eligible categories, there was a greater percentage of words that co-occurred with tweets in I relative to tweets in WE. The greatest differences can be seen between I and WE in the following categories: *cognitive processes* (I 30%; WE 15%), insight (a sub category of *cognitive processes*; I 27%; WE 13%), *drives* (I 71%; WE 35%), *affiliation* (a sub category of *drives*; I 22%; WE 13%), *power* (a sub category of *drives*; I 43%; WE 18%), *negative emotion* (a sub category of *psychological processes*; I 62%; WE 26%), *anger* (a sub category of *negative emotion*; I 24%; WE 9%), *sadness* (a sub category of *negative emotion*; I 10%; WE 6%), and *family* (a sub category of *social processes*; I 12%; WE 5%). Larger differences in the proportion of *negative emotion*, *anger*, and *sadness* words with regard to I and WE, replicate previously reports that negative affective states are often more strongly associated with the use of first-person singular pronouns [16] and when people respond to negative events they tend to use more negative emotion words [28].

We compared entropy analyses by pronoun for the same LIWC categories. These analyses are less influenced by different number of tweets in I and WE than those based on counts. Less variation and the tendency for greater linguistic alignment in I suggests less self-disclosure. Entropy scores for the LIWC categories and CIs are listed in Table 3. Of particular relevance to an interpretation that incorporates self-disclosure are the pronounced differences for *negative emotion* and *drives.*

In summary, all categories exhibited greater linguistic alignment for I tweets than WE tweets despite more tweets in I than WE. Here, we emphasize that the systematic reduction in lexical diversity between I and WE present for *negative emotion* words was weaker for *positive emotion* words. This finding is consistent with a stronger disinclination for self-disclosure for tweets in I, especially when emotions are negative [23]. Unanticipated was that words classifiable as *drives* also exhibited disproportionately more tweets in I than WE along with greater linguistic alignment. We speculate that in short time frames such as our 24-h window for data collection, *drive* word choices could be less diverse than WE, but with more time for the emergence of a

collective, word choice associated with WE might be most amenable to a pattern of decreased lexical diversity.

4 Summary and Future Directions

In this report, we investigate spontaneous responses to a catastrophic event via tweets posted over a period of 24-h following the event. Building from our previous work on tweets produced by a massive-scale network that were associated with a catastrophic event, we extend our focus beyond means for word ratings on the psychological indicators of valence and arousal. Here, we introduce entropy measures to capture linguistic alignment as a function of pronoun use and consider differences in valence with pronoun choice as an indicator of self-disclosure. With respect to means, tweeting in I relative to WE was associated with the use of more emotionally intense negatively valenced words, a greater percentage of *negative emotion* words, *anger*, words and *sadness* words, fewer positively valenced word choices and greater linguistic alignment. Results support the greater inclination to use negativity or negative emotional states when using first-person singular than when using first-person plural pronouns [16].

In general, tweeting in I was more common than tweeting in WE and resulted in slightly higher frequencies of usage for words across many LIWC categories. Not only was tweeting in I associated with more emotionally intense negatively valenced, *negative emotion,* and *drives* word choices, but it also accompanied more redundant vocabulary (reduced lexical diversity) than did tweeting in WE. As a rule, reduced diversity can be a marker for greater linguistic alignment, which is characteristic of more intense or sustained social interaction. This pattern is consistent with the literature linking use of I pronouns to greater potential for interaction with others [23]. The tendency to use more *drive* words and more uniform *drive* word choices when tweeting in I suggests that at short time frames *drive* word choices can be more uniform than WE. Whether at longer time frames, with the emergence of a collective mentality, patterns would change so that word choice associated with WE would show the greater convergence is worthy of future study.

Furthermore, first-person singular pronouns are consistent with an individualist perspective, including attention on the self [1] and in the context of a crisis event, feelings of vulnerability may accompany attention on the self and lead to a disinclination to self-disclose in tweet interactions. At present, the self-disclosure hypothesis rests on the asymmetry between strong lexical uniformity (alignment) of negative but weaker uniformity of positive word use.

Additionally, in the context of a crisis event, emotional contagion and the transfer of emotional states to others is likely to be more negative than positive. At the same time, the tendency to respond more strongly to negative relative to positive stimuli arises across a variety of natural settings and experimental tasks [29], such that negative words, often signal danger or emergency and are thus attended to more closely than positive words [30, 31]. In this framework, increased interaction between tweeters as revealed by a greater use of I pronouns concurrent with generally exaggerated attention to the negative, may induce automatic mimicry [29]. Interpreted as contagion, the

general salience of the negative could manifest itself in word choice when producing a tweet and account for greater lexical uniformity for negative than positive word choice without invoking self-disclosure. Either negative emotion states are more contagious and undergo greater alignment or the inclination to self-disclose negative emotion states is suppressed more when tweeting from an individualist perspective (I). In future work we will explore responses over a time period that extends beyond 24-h as well as the generality of the link between self-disclosure when using first-person singular pronouns and negative emotional contagion as reflected in linguistic alignment. One strategy will be to examine retweets (i.e. re-posting another users' tweet) in response to crisis events to disentangle reactions of self-disclosure from emotional contagion. Retweets are less likely to exhibit comparable patterns for I and WE, and specifically, the disinclination for self-disclosure when valence is negative.

References

1. Chung, C.K., Pennebaker, J.W.: The psychological functions of function words. In: Social Communication, pp. 343–359. Psychology Press, New York (2007)
2. De Bruyn, A., Lilien, G.L.: A multi-stage model of word-of-mouth influence through viral marketing. Int. J. Res. Mark. **25**, 151–163 (2008)
3. Danescu-Niculescu-Mizil, C., Gamon, M., Dumais, S.: Mark my words!: linguistic style accommodation in social media. In: Proceedings of the 20th International Conference on World Wide Web, pp. 745–754. ACM (2011)
4. Purohit, H., Hampton, A., Shalin, V.L., Sheth, A.P., Flach, J., Bhatt, S.: What kind of #conversation is Twitter? Mining #psycholinguistic cues for emergency coordination. Comput. Hum. Behav. **29**, 2438–2447 (2013)
5. Java, A., Song, X., Finin, T., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis, pp. 56–65. ACM (2007)
6. Boyd, D., Golder, S., Lotan, G.: Tweet, tweet, retweet: conversational aspects of retweeting on twitter. In: 2010 43rd Hawaii International Conference on System Sciences (HICSS), pp. 41–49. IEEE (2010)
7. Stefanone, M.A., Lackaff, D., Rosen, D.: Contingencies of self-worth and social-networking-site behavior. Cyberpsychol. Behav. Soc. Netw. **14**, 41–49 (2011)
8. Shaikh, S., Feldman, L.B., Barach, E., Marzouki, Y.: Tweet sentiment analysis with pronoun choice reveals online community dynamics in response to crisis events. In: Advances in Cross-Cultural Decision Making, pp. 345–356. Springer, Cham (2018)
9. Weintraub, W.: Verbal Behavior in Everyday Life. Springer, New York (1989)
10. Pennebaker, J.W., Francis, M.E., Booth, R.J.: Linguistic Inquiry and Word Count (LIWC): LIWC2001. Lawrence Erlbaum, Mahwah (2001)
11. Tausczik, Y.R., Pennebaker, J.W.: The psychological meaning of words: LIWC and computerized text analysis methods. J. Lang. Soc. Psychol. **29**, 24–54 (2009)
12. Doyle, G., Yurovsky, D., Frank, M.C.: A robust framework for estimating linguistic alignment in twitter conversations. In: Proceedings of the 25th International Conference on World Wide Web, pp. 637–648 (2018)
13. Shaikh, S., Lalingkar, P., Barach, E., Feldman, L.B.: Cross-cultural reactions to crisis events via language and emoticon use. In: International Conference on Applied Human Factors and Ergonomics, pp. 23–28. Springer, Cham (2017)

14. Feldman, L.B, Aragon, C., Chen, N.C., Kroll, J.F.: Emoticons in informal text communication: a new window on bilingual alignment. Bilingualism Lang. Cogn. **21**, 209–218 (2017)
15. Rude, S., Gortner, E.M., Pennebaker, J.W.: Language use of depressed and depression-vulnerable college students. Cogn. Emot. **18**, 1121–1133 (2004)
16. Cohn, M., Mehl, M.R., Pennebaker, J.W.: Linguistic markers of psychological change surrounding September 11, 2001. Psychol. Sci. **15**, 687–693 (2004)
17. Na, J., Choi, I.: Culture and first-person pronouns. Pers. Soc. Psychol. Bull. **35**, 1492–1499 (2009)
18. Twenge, J.M., Campbell, W.K., Gentile, B.: Changes in pronoun use in american books and the rise of individualism, 1960–2008. J. Cross Cult. Psychol. **44**, 406–415 (2012)
19. van Baaren, R.B., Maddux, W.W., Chartrand, T.L., de Bouter, C., van Knippenberg, A.: It takes two to mimic: behavioral consequences of self-construals. J. Pers. Soc. Psychol. **84**, 1093–1102 (2003)
20. Trope, Y., Liberman, N.: Construal-level theory of psychological distance. Psychol. Rev. **117**, 1024 (2010)
21. Moscoso del Prado Martín, F.: Grammatical change begins within the word: causal modeling of the co-evolution of Icelandic morphology and syntax. In: Bello, P., Guarini, M., McShane, M., Scasselatti, B. (eds.) Proceedings of the 36th Annual Conference of the Cognitive Science Society, Austin, TX, pp. 2657–2662 (2014)
22. Moscoso del Prado Martín, F.: Vocabulary, grammar, sex and aging. Cogn. Sci. (2016)
23. Pan, W., Feng, B., Skye Wingate, V.: What you say is what you get: how self-disclosure in support seeking affects language use in support provision in online support forums. J. Lang. Soc. Psychol. **37**, 3–27 (2017)
24. Pennebaker, J.W.: The Secret Life of Pronouns. Bloomsbury Press, New York (2013)
25. Brewer, M.B., Gardner, W.: Who is this "We"? Levels of collective identity and self representations. J. Pers. Soc. Psychol. **71**, 83–93 (1996)
26. Warriner, A., Kuperman, V., Brysbaert, M.: Norms of valence, arousal, and dominance for 13,915 English lemmas. Behav. Res. Methods **45**, 1191–1207 (2013)
27. Shannon, C.E.: A note on the concept of entropy. Bell Syst. Tech. J. **27**, 397–423 (1948)
28. Kahn, J.H., Tobin, R.M., Massey, A.E., Anderson, J.: Measuring emotional expression with the linguistic inquiry and word count. Am. J. Psychol. **120**, 263–286 (2007)
29. Rosekrans, M.A.: Imitation in children as a function of perceived similarity to a social model and vicarious reinforcement. J. Pers. Soc. Psychol. **7**, 307–315 (1967)
30. Kuperman, V., Estes, Z., Brysbaert, M., Warriner, A.B.: Emotion and language: valence and arousal affect word recognition. J. Exp. Psychol. Gen. **143**, 1065–1081 (2014)
31. Pratto, F., John, O.P.: Automatic vigilance: the attention-grabbing power of negative social information. J. Pers. Soc. Psychol. **61**, 380–391 (1991)