Cross-Cultural Reactions to Crisis Events via Language and Emoticon Use

Conference Paper *in* Advances in Intelligent Systems and Computing · July 2018

DOI: 10.1007/978-3-319-60747-4_3

4 authors:

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

Prasanna Lalingkar
University of North Carolina at Charlotte
1 PUBLICATION **1** CITATION
SEE PROFILE

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

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

Cross-Cultural Reactions to Crisis Events via Language and Emoticon Use

Samira Shaikh[1], Prasanna Lalingkar[1], Eliza Barach[3] and Laurie Feldman[2, 3]
[1]University of North Carolina at Charlotte
[2]Haskins Laboratory, New Haven, Connecticut
[3]State University of New York – University at Albany

Abstract. We analyze language and emoticon use on a social media platform to describe large-scale human reaction to a crisis event. We focus on a targeted corpus of 2 million tweets defined by a set of hashtags and posted on the social media platform Twitter. We analyze these data under the framework of Construal Level Theory and compute lexical diversity and concreteness values of words across subsets of the data that differ with respect to geographical distance from the event. We find that word count and lexical variation among concrete words (but not average concreteness) decreased with increasing geographical distance. In addition, we investigate the use of non-verbal signals through the use of emoticons and emojis in these subsets. Overall, our findings make significant contributions in quantifying and contrasting cross-cultural communication with respect to large-scale human response to a crisis event, specifically a terrorist attack. The results presented here are novel in that they demonstrate what we learn from large-scale nonverbal as well as verbal communication analyzed in the framework of the Construal Level Theory.

Keywords: Human Factors · Social Media · Natural Language · Construal

1 Introduction

Social media permits a sense of immediacy whose impact cannot be overestimated (Giddens, 2013) and many have argued that it changes our concepts of time and space and how we experience them (Panayiota Tsatsou, 2009). Construal Level Theory (CLT) (Trope and Liberman, 2010) posits that people experience psychologically distal events by forming more abstract mental construals of such events than they do for events that are psychologically proximal. These mental construals manifest themselves in the language people use to describe the events, specifically in concreteness values (defined as the degree to which a concept denoted by the word refers to a perceptible entity). For instance, people may experience geographically distant, and hence psychologically distal, events by forming mental construals of such events at higher levels of abstraction than events that are geographically proximal (Henderson et al., 2006). As an example, the word *heart* has a high concreteness value and may be used to describe a geographically proximal event, whereas the word *love*, which has a low concreteness (and hence high abstractness) value, may be used to describe a geographically distant event.

In this article, we test our hypotheses that people will use different vocabulary to describe the same crisis event, depending upon their own geographical distance from that event. We focus on natural language productions on social media platforms immediately following a terrorist attack. We use our analyses to challenge the claim that the sense of immediacy engendered by social media compresses our experience of time and space in ways that are evident in our spontaneous written language production.

2 Data Collection and Method

We focus on a targeted corpus gathered using a specific set of hashtags. , composed of 2 million tweets posted on the social media platform Twitter. The tweets were posted in a 24-hour period following the terrorist attacks in Paris, France in November 2015. The unique hashtags across the corpus are listed in Table 1.

Table 1. List of unique hashtags in our tweet collection

Hashtags			
#ParisShooting	#ParisShootings	#parisattack	#parisattacks

Table 2. Proportion of Tweets containing emoticons and/or emojis in each subset

Within-France		**Outside-France (randomly chosen)**	
Total Tweets	54744	Total Tweets	54744
% of Tweets with Emoticons and/or emojis	11%	% of Tweets with Emoticons and/or emojis	7%

In order to test our hypotheses on a large corpus of linguistic constructions, we created two subsets of tweets from the original set of 2 million – each subset containing 54744 tweets in English. One subset contained tweets that were posted from within France (determined using metadata associated with each tweet, specifically the geotags and geolocations). We call this subset *within-France*. We created a second subset of 54744 English tweets chosen randomly from those tweets out of the original 2 million that were posted outside of France (we call this subset *outside-France*).

For each subset, we determined the number tweets that contain non-verbal symbols, *viz*. emoticons and emojis (11% in the *within-France* subset and 7% in the *outside-France* subset). These proportions are shown in Table 2.

For each subset, we also identified the top 10 most frequently occurring emoticons/emojis. We show these in Table 3 below.

Table 3. Top 10 most frequently occurring emoticons/emojis in both subsets

Rank	Within-France	Outside-France
1	red circle	☹
2	☹	man kissing woman
3	folded hands	folded hands
4	man kissing woman	broken heart
5	:/	pensive face
6	=(	red circle
7	camera	:/
8	pensive face	=(
9	broken heart	crying face
10	france flag	disappointed face

We note that some emoticons and emojis are common within the top ten most frequent symbols in the *within-France* and *outside-France* subsets. Interestingly, the French flag appears in the within-France set (at rank 10) suggesting patriotic feelings emergent in the tweets posted within France after the crisis tweet.

3 Analysis of Data

Our hypothesis is that *within-France* tweets will contain more linguistic expressions with higher concreteness (lower abstractness) ratings than those that originate elsewhere, since these expressions describe events with closer geographical distance.

We use an existing concreteness lexicon (Brysbaert, Warriner and Kuperman, 2014) to identify words in tweets that have extreme concreteness scores (>=4, on a scale of 1-5) and extreme abstractness scores (<=2). We compute the following measures of extreme concreteness and abstractness: 1) means of each tweet 2) counts of extremely scored words in tweets and 3) vocabulary richness/lexical variation measured by entropy of word choice. We shall describe the lexical variation analysis in detail next.

We computed lexical diversity from information entropy (Shannon, 1948) of the frequency distribution of individual words and used it as a measure of vocabulary richness words (defined as unique character sequences delimited by spaces or punctuation; not including the emoticons themselves). For each word, we computed its frequency and then converted that to a percent of the total word frequency (relative fre-

quency). We expressed each value in log base 2 and computed the negative product of the log times the relative frequency and summed those values. Higher values calculated using this measure indicate greater uncertainty and greater lexical diversity. Low values indicate a more redundant word choice. We conducted separate analyses for the most extremely rated (top 25%) concrete and the most extremely rated (top 25%) of abstract words.

4 Results and Discussion

4.1 Analysis of Lexical Variation

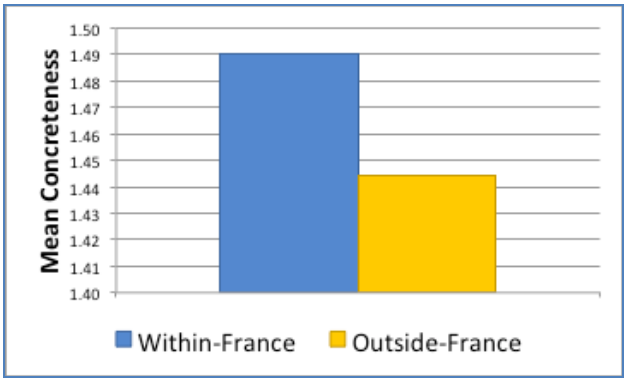
For concrete words, lexical variation was found to be greater inside France than out 15.23 vs. 15.02), while lexical variation for abstract words across tweets was less *within-France* than *outside-France* (15.19 and 15.34). This suggests that *outside-France* tweeters tended to use and reuse a small set of very concrete words compared to their use of abstract words.

Table 4. Lexical Diversity of words with extreme concreteness values (high and low) in within-France and outside-France subsets

	High Concreteness	**High Abstractness**
Within-France	15.23	15.19
Outside-France	15.02	15.34

However, means (SD) revealed that tweets in France were more abstract (1.44 (.08)) than those out-side (1.50 (.04)), which suggests that tweeters from *outside-France* selected a larger, more varied but less abstract set of words. Figure 1 shows the mean concreteness of tweets posted *within-France* and *outside-France*.

Figure 1. Mean concreteness of words in tweets posted within and outside France



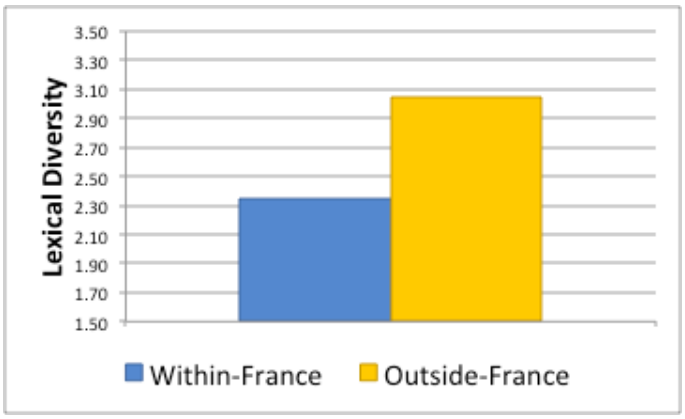
Here we find that variability and means for concreteness and abstractness worked systematically in complementary directions. Our results are consistent with the Construal Level Theory in that lexical variation among concrete words (but not average concreteness) decreased with increasing physical distance from the event. Similarly, with respect to extremely abstract words, within France tweets were less abstract and less lexically diverse.

4.2 Analysis of Lexical Variation with Emoticons/Emojis

Following Feldman et al. (2016) who argue that emoticons in text function like gestures or paralinguistic elements in language, we compared tweets with and without emoticons with respect to lexical variation.

Of tweets *within-France*, 11% included emoticons and/or emojis, compared to 7% of *outside-France* tweets. The emoticons :(, :/ and =(constituted 82% of emoticons in *within-France* tweets but only 65% of *outside-France* tweets. Entropy analysis of emoticon use in tweets revealed that there is less variation in the distribution of emoticons in tweets that originate in France compared to those that originated elsewhere (2.35 vs. 3.05 resp.). These findings are summarized in Figure 2.

Figure 2. Mean entropy of words in tweets with emoticons/emojis posted within and outside France



We also analyzed the words that co-occur with the most frequent emoticons under the assumption those words reveal what the user is experiencing in response to our target crisis event (see also Barbieri et al., 2016). In the context of a ☹, *within-France* tweets were less likely to mention the words 'prayers', 'thought' and 'thanks' and more likely to mention 'crazy' and 'crying' than *outside-France* tweets. Mention of 'thoughts' and 'prayers' were also more characteristic of *outside-France* than of *within-France* tweets that accompany a ☹.

5 Summary and Future Work

Social psychologists have always been interested in the behavior of groups and insights from contagion behavior show that the behavior of an individual is influenced not only by the people he/she knows but also by the behavior of the people that person knows. With the increasing popularity of social media platforms such as twitter, geographically distant individuals are now able to connect and respond to crises events such as terrorist attacks, as was the case in our work.

We analyze natural language productions on social media posted immediately (within 24-hours) of a crisis event that occurred in France in November 2015. We compute measures of lexical diversity to test our hypotheses that tweets posted from places that are geographically distant relative to the event will reflect the use of more abstract language when compared to tweets posted geographically closer to the event (i.e. within France). We find that our results are consistent with the Construal Level Theory and that lexical variation among concrete words decreased with increasing physical distance from the event.

Interestingly, geographic distance to the focal event plays a role in tweeting styles. Potentially this difference could be attributed to the development of in-group and out-group, such as the within-France tweeters and outside-France tweeters, with former as the in-group and the latter as the out-group.

In future work, we wish to extend our analyses to other facets of psychological distance, including social and temporal distance. We also aim to replicate our analyses on other large corpora, which may or may not be centered around a crisis event.

References

1. Barbieri, Fr., Ronzano, F., & Saggion, H. (2016). "What does this Emoji Mean? A Vector Space Skip-Gram Model for Twitter Emojis." Language Resources and Evaluation conference, LREC, Portoroz, Slovenia
2. Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. Behavior research methods, 46(3), 904-911.
3. del Prado Martín, F. M. (2011). Macroscopic thermodynamics of reaction times. Journal of Mathematical Psychology, 55(4), 302-319.
4. Feldman, L. B. , Aragon, C., Chen, N-C. and Kroll, J. F. (2016). Emoticons in text may function like gestures in spoken or signed communication. *Behavioral and Brain Science*. DOI:, http://journals.cambridge.org/
5. Giddens, A. (2013). *The consequences of modernity*. John Wiley & Sons. pp. 36-39
6. Henderson, M. D., Fujita, K., Trope, Y., & Liberman, N. (2006). Transcending the" here": the effect of spatial distance on social judgment. *Journal of personality and social psychology*, *91*(5), 845.
7. Shannon, C. E. (1948). A note on the concept of entropy. *Bell System Tech. J*, *27*, 379-423.
8. Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. Psychological review, 117(2), 440
9. Panayiota Tsatsou, D. (2009). Reconceptualising time and space in the era of electronic media and communications. *PLATFORM: Journal of media and communication*, *1*, 11-32.