Language Limitations in Rumor Research? Comparing French and English Tweets Sent During the 2015 Paris Attacks

Tom Wilson
Human Centered Design and Engineering
University of Washington
tomwi@uw.edu

Stephanie A. Stanek
Human Centered Design and Engineering
University of Washington
stanek14@uw.edu

Emma S. Spiro
Information School
Department of Sociology
University of Washington
espiro@uw.edu

Kate Starbird
Human Centered Design and Engineering
University of Washington
kstarbi@uw.edu

ABSTRACT

The ubiquity of social media facilitates widespread participation in crises. As individuals converge online to understand a developing situation, rumors can emerge. Little is currently known about how online rumoring behavior varies by language. Exploring a rumor from the 2015 Paris Attacks, we investigate Twitter rumoring behaviors across two languages: French, the primary language of the affected population; and English, the dominant language of Internet communication. We utilize mixed methods to qualitatively code and quantitatively analyze rumoring behaviors across French and English language tweets. Most interestingly, temporal engagement in the rumor varies across languages, but proportions of tweets affirming and denying a rumor are very similar. Analyzing tweet deletions and retweet counts, we find slight (but not significant) differences between languages. This work offers insight into potential limitations of previous research of online rumoring, which often focused exclusively on English language content, and demonstrates the importance of considering language in future work.

Keywords

social media, rumoring, language, crisis informatics, information diffusion.

INTRODUCTION

The convergence of people and information at physical sites of crisis is an established phenomenon (Kaufman et al. 1958; Kendra and Wachtendorf 2003; Palen et al. 2009). Although this activity was once limited to those geographically proximal to the event, the ubiquity of social media means that participation in a crisis is now widespread (Hughes et al. 2008). Individuals and organizations from both inside and outside the physically affected area can converge on social media and participate in *collective sensemaking*—the process whereby people attempt to fill information gaps and understand a developing situation (Starbird et al. 2016), a consequence of which can be the emergence of rumors (Shibutani 1966; Starbird et al. 2016)— the circulation of unverified information (Bordia and Difonzo 2004).

Prior work has explored various aspects of online rumoring during crises, particularly on Twitter (e.g. Huang et al. 2015; Maddock et al. 2015; Spiro et al. 2012; Starbird et al. 2015, 2016; Zeng et al. 2016). The majority of this prior research however, focuses on Twitter communications in English—which only constitute around 50% of all tweets (Hong et al. 2011; Poblete et al. 2011). Less understood is how online rumoring behavior on

Twitter may vary by language, and the challenges associated with conducting online rumor research across languages. This is despite the Twitter network being structured by language (Hale 2014), with the majority of a country's Twitter content written in its own dominant language (Mocanu et al. 2013). As social media platforms are now established communication channels during crisis events (Starbird et al. 2015), the intersection of language and online rumoring behavior certainly warrants further attention.

This study focuses on tweets posted during and immediately after the November 2015 Paris Attacks. We explore rumoring behaviors across two languages: French (FR), the primary language of the affected population; and English (EN), the dominant language of Internet communication. Broadly, we seek to understand *how individuals' rumoring behavior on Twitter varies according to their proximity to an emerging crisis event.* In particular, we are curious whether proximity to a crisis, for which language serves as an imperfect but useful proxy, intersects with individuals' rumoring behavior. Considering the FR tweets as *proximal* to the crisis, and the EN tweets representing the broader audience, we specifically analyze four aspects of Twitter rumoring behavior across the languages: 1) tweet deletion and account protection since the event; 2) rumor spreading (affirmation) and correction (denial); 3) temporal engagement in the rumor, and; 4) retweeting of the rumor.

We adopt a mixed methods approach to qualitatively code and quantitatively analyze a false rumor that the *Les Halles* shopping center in central Paris was also the site of a shooting during the November 2015 Paris Attacks. Our study reveals that FR tweets are more likely than EN tweets to have been removed from public view—i.e. deleted or protected—since the event. We note similarities in rumoring and correction behaviors, but differences in the temporal engagement across languages. Although we discover variation in retweeting of rumor tweets, we do not find the differences in our sample statistically significant.

This study is a work in progress insofar as it represents an initial exploration into online rumoring across languages during a crisis event. The findings presented here are drawn from a specific rumor—the *Les Halles* rumor that emerged during the Paris attacks. This rumor provided a distinctive opportunity to investigate rumoring across languages because the name of a specific location, *Les Halles,* was used in both French and English—a convenience not often present in studies of emergent events on social media. Our study makes two main contributions. First, our analyses provide insight into the potential shortcomings (or not) of prior research into online rumoring, which are often limited to English language without considering how rumoring behaviors may vary across languages. Second, we initiate discourse into research that spans language divides, and raise the issue of considering language in studies of social media.

RELATED WORK

Language Use on Twitter

Several studies have examined language use on Twitter. A large-scale analysis of the linguistic geography of Twitter discovered that most content produced within a country was written in its dominant language (Mocanu et al. 2013). (Hale 2014) also found Twitter relationships to be structured by language, with users more likely to retweet content written in the same language as their own. In a cross-language study of Twitter usage, (Hong et al. 2011) found that 51% of tweets were English, and more than 100 languages were used. Among the ten most-used languages, differences in the use of Twitter features (e.g. hashtags, retweets, URLs, mentions, replies) were apparent. A similar study of Twitter usage by country revealed differences in the use of language itself and Twitter features (Poblete et al. 2011). Each of these studies is concerned with general Twitter usage, not Twitter usage during crisis events.

In the crisis context, (Thomson et al. 2012) conducted cross-language comparisons of online anonymity and information source credibility on Twitter in the wake of the 2011 Fukishima nuclear crisis in Japan. In this study, individuals identified as proximal to the event demonstrated an increased tendency to share information from credible sources, suggesting that during crises, physical proximity plays a role in information sharing behaviors on Twitter.

Social Media Use and Online Rumoring During Crisis Events

Crisis researchers have identified several roles of social media platforms in the crisis context: for social convergence (Hughes et al. 2008), for information sharing (Palen and Liu 2007), for situational awareness (Vieweg et al. 2010), and for self-organizing digital volunteers (Starbird and Palen 2011). Examining the Twitter retweet function, (Starbird and Palen 2010) found it was used as an informal recommendation system during mass emergencies, with individuals identified as geographically local to the event more likely to retweet local event-related information. The role of proximity was also explored in (Huang et al. 2015). This qualitative study of individuals who sought information on social media during the Boston Marathon Bombing found that

information seeking and sharing behavior is likely influenced by emotional and physical proximity to the crisis.

Research into online rumoring during crisis events, particularly on Twitter, has led to the identification of multidimensional signatures to characterize types of rumors (Maddock et al. 2015); models to interpret rumor transmission rates (Zeng et al. 2016); efforts to understand the effect of situated proximity on the spread of rumors (Huang et al. 2015); and the role of expressed uncertainty in collective sensemaking (Starbird et al. 2016). In these studies, the influence that language may or may not have on rumoring is not explored, therefore the applicability of the findings to other languages remains unclear.

THE NOVEMBER 2015 PARIS ATTACKS: EVENT BACKGROUND (ALL TIMES IN CET)

On Friday November 13 2015, a series of coordinated terrorist attacks occurred in Paris, France. Nine men attacked six locations across the city, including the *Stade de France* and the *Bataclan* Theatre. Beginning at 21:20, the attackers detonated suicide explosions and fired upon crowds of people at popular restaurants and nightspots. An assault to free the hostages ended the attacks at 00:23 on November 14 2015. During the attacks, 130 people were killed and 368 injured, making it the deadliest attack on France since the Second World War. The dispersed nature of the event meant the situation was extremely confusing, with attacks and response efforts simultaneously occurring at multiple sites across the city. As events across the city unfolded, a rumor emerged of a shooting at the *Les Halles* shopping center in central Paris. The rumor later turned out to be false, and was attributed to the police response to attacks in adjacent districts. Twitter activity surrounding the *Les Halles* rumor is the focus of this study.

DATA

Data were collected using the Twitter Streaming API. Three separate collections were initiated to pick up Twitter correspondence (tweets) during the Paris Attacks. The collections went online at 21:39, 22:33, and 22:47 CET on November 13 2015. We used search terms that would collect EN and FR tweets, including: *paris, paris shooting, paris hostages, paris theatre, parisattacks, paris attack, ParisFusillades, Les Halles, shopping centre*, and *PorteOuverte*. As the event unfolded, tweet volumes were high, leading to data collection being rate limited at ~50 tweets per second per collection.

For the subsequent analysis we selected all original tweets—non-retweets—that contained the regular expression term "*halles*" and that Twitter identified as English or French. We limited our sample to tweets posted after 22:45 CET, a point in time when our Twitter collections and the tweet volumes had stabilized, and before 08:00 CET on November 16 2015, when the volume and frequency of tweets relevant to the *Les Halles* rumor in our corpus was very low. The final dataset contains 264 EN tweets sent by 227 distinct users, and 593 FR tweets sent by 551 distinct users. To place this in the context of our collections, during the same time period we collected almost 3 million original tweets related to the Paris attacks in general, of which approximately 1.8 million are EN and approximately 0.4 million are FR.

METHODS AND FINDINGS

We employed mixed methods to qualitatively code and quantitatively analyze the *Les Halles* rumor. Our analysis focused on cross-language comparison along four dimensions: 1) tweet deletion and account protection; 2) rumor spreading and correction; 3) temporal engagement in the *Les Halles* rumor; and 4) retweeting of rumor tweets. We present the methods and findings of each of these dimensions in the following sub-sections.

Tweet Deletion and Account Protection Across Languages

Investigation of tweet deletion and account protection behavior originated from a legal and ethical responsibility to remove tweets that had been deleted or protected (made private) by their author (Maddock n.d.). On July 7, 2016 we determined the current availability of our *Les* Halles tweets by passing the tweetIDs to the Twitter API. If tweets were still available (i.e. not deleted, suspended, or protected), we would receive the tweet and its associated metadata in response. If a tweet was unavailable, the Twitter API returned one of the following error codes: `144 No status found with that ID`—meaning the user has since deleted the tweet; `63 User has been suspended`—meaning the account that authored the tweet has since been suspended; and `179 Sorry, you are not authorized to see this status`—implying the tweet's author has since protected their account and tweet(s). An important distinction between the mechanism of tweet deletion and tweet protection is that individual tweets can be deleted but not protected—protection is done at the account level and therefore when a person protects their account, all their tweets are protected.

Upon completion of this process, 415 FR and 199 EN tweets were still available (i.e. public), and subject to

further analysis. Table 1 summarizes the findings related to tweet deletion and account protection. 30% of FR tweets and 24.6% of EN tweets containing "*halles*" are unavailable. A greater proportion of FR tweets were deleted by the user: 17.4% FR compared to 13.3% EN. The largest difference across languages is tweet protection; 3.71% of FR tweets, but just 0.75% of EN tweets, originate from accounts that were public during the event but have since been protected i.e. made private.

Lan	Available	Unavailable		
		Deleted	Protected	Suspended
EN	199 (184) 75.38%	35 (29) 13.26%	2 (2) 0.75%	28 (15) 10.61%
FR	415 (401) 69.98%	103 (98) 17.37%	22 (18) 3.71%	53 (42) 8.94%

Table 1. Breakdown of tweet availability status. The numbers in parentheses (xx) denote the number of distinct users responsible for the tweets.

We used Chi-squared tests of independence to explore the association between language and deletion/protection behavior. Though FR tweets were more likely to have been deleted than EN tweets, the difference was not statistically significant (X^2 (1, N = 752) = 2.61, p=0.11). However, the difference for protected tweets was statistically significant (X^2 (1, N = 638) = 6.20, p=0.01), indicating that FR tweets were more likely to have originated from an account that has changed to protected since the event. This shows that those using the primary language of the affected population to engage in the *Les Halles* rumor were more likely to have protected their account since the Paris Attacks.

Rumoring and Corrections Across Languages

Comparison of rumoring behavior across languages first required translation of the French tweets into English so that they could be qualitatively coded. We recruited native French translators from the freelance marketplace UpWork (http://www.upwork.com). From the 36 responses from our job posting on the site, we selected three translators based upon their translation experience, client satisfaction rating, and familiarity with social media communications. Two translators were asked to independently translate the 415 available FR tweets and explain acronyms, slang, and subtleties in the language. Two authors of this paper verified cross-translator agreement. In 66 cases the tweet translations did not convey the same meaning, and the third translator was asked to independently translate those tweets. Selected translations were based upon a simple majority, whereby two of three translators were in agreement.

To code the rumor we adopted a similar coding scheme and process to (Andrews et al. 2016; Starbird et al. 2015), focusing on the distinction between passing along vs. correcting a rumor. Two coders previously trained to classify rumor tweets independently coded the available EN and FR tweets using a mutually exclusive coding scheme of five categories: *Affirm*—the tweet affirms or supports the [*Les Halles*] rumor; *Deny*—the tweet denies or corrects all or part of the rumor; *Neutral*—the tweet does not affirm or deny the rumor (used sparingly); *Unrelated*—part or all of the tweet is unrelated to the rumor; *Uncodable*—part or all of the tweet is unintelligible. Tweets lacking between-coder agreement were discussed until a consensus was reached.

	AFFIRM & DENY		AFFIRM			DENY		
Lan	Rumor tweets	Distinct users (tweet/user)	No. of tweets (% of rumor)	Distinct users (tweet/user)	RT mean (sum)	No. of tweets (% of rumor)	Distinct users (tweet/user)	RT mean (sum)
EN	137	134 (1.02)	124 (90.51%)	122 (1.02)	6.08 (754)	13 (9.49%)	12 (1.08)	27.46 (357)
FR	279	270 (1.03)	253 (90.68%)	247 (1.02)	22.37 (5659)	26 (9.32%)	26 (1.00)	10.12 (263)

Table 2. Breakdown of Les Halles rumor tweets illustrating rumoring, correction, and retweeting behavior across languages.

Table 2 summarizes the findings and illustrates the similarities in rumoring behavior across languages. 68.8% of

available EN and 67.2% of available FR tweets that contained the term "*halles*" are rumor-related tweets that affirm or deny the *Les Halles* rumor. Furthermore, the percentage of affirmations and denials for each language are very similar: 90.5% of EN and 90.7% of FR rumor tweets affirm the rumor; and 9.5% of EN and 9.3% of FR rumor tweets deny the rumor. There is no statistically significant difference between the languages in terms of the proportion of affirmations to denials (X^2 (1, N = 416) = 0.003, p=0.95).

Temporal Signatures of Rumor Engagement Across Languages

We visualized engagement in the *Les Halles* rumor across languages through the temporal signature—the volume of tweets per 10-minute interval (Figure 1). This time series shows EN tweets reaching a higher initial peak than FR tweets (~30 tweets per 10 minutes for EN tweets vs. ~20 tweets per 10 minutes for FR tweets). However, the FR signature demonstrates a more sustained engagement in the rumor, with a sharp resurgence several hours later, and continuing at a few tweets per ten-minute interval for several hours after the initial peak. Interestingly, the shapes of the signatures are quite different in places—i.e. spikes in volume in the two languages do not always align, suggesting somewhat divergent conversations.

A notable spike in FR engagement is apparent around 18:00 CET on November 15. We traced this activity to the release of a firecracker at a memorial event held for the victims of the Paris Attacks, which took place close to the Les Halles shopping center. French tweets describe a *movement de foule*—a stampede—as people attempted to flee from what they feared was another attack. This sub-rumor was very localized; there were 24 FR tweets posted during a 1-hour period, but just 2 EN tweets during that time.

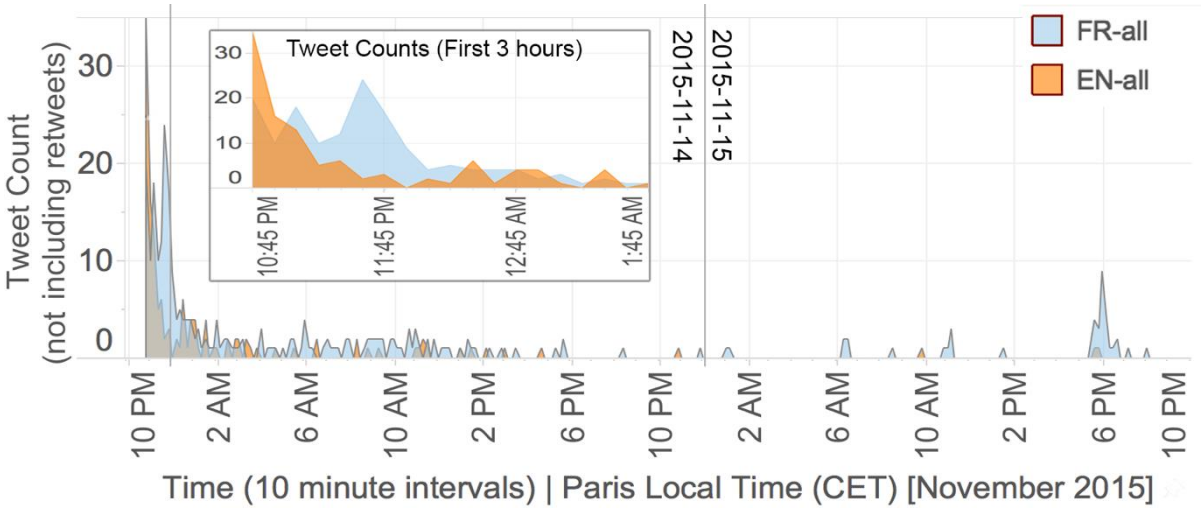

Figure 1. Time-series of rumor engagement across languages.

Rumor Retweeting Across Languages

We obtained the retweet (RT) counts for each rumor tweet from the Twitter API on July 7, 2016, approximately 8 months after the Paris Attacks. The sum and mean RT values were calculated for each of the EN and FR rumor tweets (Table 2). On average, FR affirm tweets were retweeted 22.4 times, whereas EN affirm tweets were retweeted 6.1 times, on average. Conversely, EN denials were retweeted more than FR denials. On average, EN deny tweets were retweeted 27.5 times, compared to 10.1 times for FR deny tweets. Like many measures of social media activity however, our RT data are highly skewed and zero-inflated. Although the raw counts suggest a tendency for the data, T-tests to compare the log-transformed RT counts of EN and FR affirm tweets and EN and FR deny tweets did not identify statistically significant differences.

Data Limitations

Collecting Twitter data during developing crisis event is inherently challenging. Our collections for this event are imperfect: Due to the emergent nature we missed the onset of the rumor, and high Twitter volumes led to rate limiting. Comparing RT counts in our collection to RT counts reported by Twitter, we note our collection was more comprehensive for FR tweets, likely due to rate limits across the three collections, capturing 44.90% of FR RTs and 29.68% of EN RTs. We also note that aside from language other differences likely exist among the Twitter users, although in this study we did not access users' profile information.

DISCUSSION AND CONCLUSION

The *Les Halles* rumor at the center of this study presented a rare opportunity to investigate rumoring across languages because the topic at the center of the rumor is also a specific location—the *Les Halles* shopping center— which was used in both French and English tweets. Collecting data related to crisis-related rumors is challenging, and encountering rumors with characteristics such as the *Les Halles* rumor provided a practically infrequent opportunity to explore rumoring in this context across languages.

Many studies of social media phenomena, particularly in the crisis context, focus on a single language—e.g. the primary language of the affected population, or that of the researchers conducting the analysis. In this study, we were able to span the language divide and attempt to tease out some of the potential limitations of this approach. We explore how language, which may function as an imperfect proxy for proximity, intersects with tweet deletion and protection; rumoring and correction; temporal engagement; and retweeting behavior of a specific rumor on Twitter, during a major crisis event of global interest. Our analyses consider these aspects of Twitter rumoring behavior from two perspectives: people using the primary language of the affected population (French); and people using the dominant language of the Internet and the wider audience (English). Although the data we examined were similar across several dimensions (e.g. proportion of affirming vs. denying tweets), we identified a few dimensions where there did appear to be a difference between languages (e.g. temporal engagement, account protection), underscoring the importance of considering language in future studies of social media.

Most important, as it sheds light on the validity of previous research on rumor corrections (Huang et al. 2015; Maddock et al. 2015; Starbird et al. 2015; Zeng et al. 2016), we find no significant difference in the proportion of rumor affirmations versus rumor denials across the languages. However, the temporal signatures of rumor engagement reveal slightly different patterns: English tweets appear earlier in this rumor but their volume quickly declines whereas the French 'local' conversation is more sustained. This suggests that those using the language of the affected population—who are perhaps more proximal to the event—were more involved in the online collective sensemaking process over time, and responsible for more of the original content, contributing eye-witness observations and knowledge to the conversation (Starbird et al. 2010). We also noticed a resurgence in the *Les Halles* rumor tweets on November 15, which was specific to the local language of the affected population—French—but did not appear in English—i.e. this related rumor did not propagate in English, our proxy for the wider audience. This may suggest some fatigue from the wider audience (Bica et al. 2017) not directly impacted by the event, while those proximal to the event were still very much affected and sensitive to developments.

Our findings also suggest differences between tweet deletion and account protection behavior. French tweets are more likely to have been deleted (though this difference was not statistically significant), and more likely to have originated from accounts that had been changed to protected after the event. Prior work in the crisis context found that deleting tweets is a method of self-correcting misinformation posted or shared [5]. The difference in deletion rate could therefore indicate self-correction was more prevalent among those using the language of the affected population. For protected accounts, one explanation for the higher numbers in FR is that local people who had protected accounts prior to the event changed their status (temporarily) to public during the event, which enabled them to participate in collective sensemaking and communicate with a broader audience. From our work we do not know *when* exactly tweets were deleted or account protected, but this opens up an interesting avenue for future research: Analyzing tweet deletions/account protections over time and across languages in the days, weeks, and months after a crisis event. Such a study would provide further insight into the role of language and proximity on self-correction and related behavior.

Although we believe that meaningful differences between languages regarding how rumor tweets are retweeted exist, we were unable to detect a statistically significant difference with the data used in this study. Future work is needed to further explore differences in the retweeting of tweets affirming and denying a rumor across languages. This exploration will require more data—possibly from several crisis events during which a rumor or rumors with characteristics similar to the *Les* Halles rumor emerge. In addition to analyzing retweet counts at a specific point in time, as we did here, exploring the retweet counts of affirm/deny tweets over time will provide a much richer picture of the spread of misinformation during emerging events.

This study provides an initial exploration into online rumoring across languages, focusing on a specific rumor that emerged during the November 2015 Paris Attacks. Our findings offer a preliminary indication that some aspects of rumoring behavior (e.g. corrections) do not vary substantially by language, but that other aspects (e.g. temporal engagement, tweet deletion/account protection) do differ in meaningful ways. Certainly, language is an aspect that requires further consideration when conducting studies of social media.

ACKNOWLEDGEMENTS

This research is a collaboration between the emCOMP lab and DataLab at the University of Washington and was supported by NSF Grants 1420255, 1541688, 1640940, and PAC121028. We also wish to thank the UW SoMe Lab and Conrad Nied for providing infrastructure support, as well as the many students who provided significant assistance to this project.

REFERENCES

Andrews, C., Fichet, E., Ding, Y., Spiro, E. S., and Starbird, K. 2016. Keeping Up with the Tweet-dashians: The Impact of 'Official' Accounts on Online Rumoring. *Cscw* 452–465.

Bica, M., Palen, L., & Bopp, C. 2017. Visual Representations of Disaster. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1262-1276). ACM.

Bordia, P., and Difonzo, N. 2004. Problem Solving in Social Interactions on the Internet: Rumor As Social Cognition. *Social Psychology Quarterly* **67**(1) 33–49.

Hale, S. 2014. Global connectivity and multilinguals in the Twitter network. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14* 833–842.

Hong, L., Convertino, G., and Chi, E. H. 2011. Language Matters in Twitter : A Large Scale Study Characterizing the Top Languages in Twitter Characterizing Differences across Languages Including URLs and Hashtags. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (1) 518–521.

Huang, Y. L., Starbird, K., Orand, M., Stanek, S. A., and Pedersen, H. T. 2015. Connected through crisis: emotional proximity and the spread of misinformation online. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)* 969–980.

Hughes, A. L., Palen, L., Sutton, J., Liu, S. B., and Vieweg, S. 2008. "Site-Seeing" in Disaster : An Examination of On-Line Social Convergence. *5th International ISCRAM Conference* (May) 44–54.

Kaufman, H. F., Clifford, R. a., Killian, L. M., Fritz, C. E., and Mathewson, J. H. 1958. Convergence Behavior in Disasters: A Problem in Social Control. *American Sociological Review* **23**(9) 102.

Kendra, J. M., and Wachtendorf, T. 2003. RECONSIDERING CONVERGENCE AND CONVERGER LEGITIMACY IN RESPONSE TO THE WORLD TRADE CENTER DISASTER. *Research in Social Problems and Public Policy* 97–122.

Maddock, J. (n.d.). Using Historical Twitter Data for Research : Ethical Challenges of Tweet Deletions.

Maddock, J., Starbird, K., Al-Hassani, H. J., Sandoval, D. E., Orand, M., and Mason, R. M. 2015. Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* 228–241.

Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., and Vespignani, A. 2013. The Twitter of Babel: Mapping World Languages through Microblogging Platforms. *PLoS ONE* **8**(4)

Palen, L., and Liu, S. B. 2007. Citizen communications in crisis: anticipating a future of ICT-supported public participation. *Natural Hazards* 727–736.

Palen, L., Vieweg, S., Liu, S. B., and Hughes, A. L. 2009. Crisis in a networked world features of computer-mediated communication in the April 16, 2007, Virginia Tech Event. *Social Science Computer Review* **27**(4) 467–480.

Poblete, B., Garcia, R., Mendoza, M., and Jaimes, A. 2011. Do All Birds Tweet the Same ? Characterizing Twitter Around the World Categories and Subject Descriptors. *Society* 1025–1030.

Shibutani, T. 1966. *Improvised news: A sociological study of rumor*, Ardent Media.

Spiro, E., Fitzhugh, S., and Sutton, J. 2012. Rumoring during extreme events: A case study of Deepwater Horizon 2010. *Proceedings of the 3rd …* 275–283.

Starbird, K., and Palen, L. 2010. Pass it on?: Retweeting in mass emergency. *Proceedings of the 7th International ISCRAM Conference* (December 2004) 1–10.

Starbird, K., and Palen, L. 2011. Voluntweeters: Self-Organizing by Digital Volunteers in Times of Crisis. *Proceedings of the Conference on Human Factors in Computing Systems (CHI)* 1071–1080.

Starbird, K., Palen, L., Hughes, A. L., and Vieweg, S. 2010. Chatter on the red: what hazards threat reveals about the social life of microblogged information. *CSCW '10 Proceedings of the 2010 ACM conference on Computer supported cooperative work* 241–250.

Starbird, K., Spiro, E., Edwards, I., Zhou, K., Maddock, J., and Narasimhan, S. 2016. Could This Be True?: I Think So! Expressed Uncertainty in Online Rumoring. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* 360–371.

Starbird, K., Spiro, E. S., Arif, A., Chou, F., Narasimhan, S., Maddock, J. I. M., Shanahan, K., and Robinson, J. 2015. Expressed Uncertainty and Denials as Signals of Online Rumoring. *Collective Intelligence* (December 2014) 1–4.

Thomson, R., Ito, N., Suda, H., Lin, F., Liu, Y., Hayasaka, R., Isochi, R., and Wang, Z. 2012. Trusting Tweets : The Fukushima Disaster and Information Source Credibility on Twitter. *Iscram* (April) 1–10.

Vieweg, S., Hughes, A. L., Starbird, K., and Palen, L. 2010. Microblogging During Two Natural Hazards Events : What Twitter May Contribute to Situational Awareness. 1079–1088.

Zeng, L., Starbird, K., and Spiro, E. S. 2016. Rumors at the Speed of Light? Modeling the Rate of Rumor Transmission During Crisis. *49th Hawaii International Conference on System Sciences (HICSS)* 1969–1978.