2014 47th Hawaii International Conference on System Science

Cues to Deception in Social Media Communications

Erica J. Briscoe
Georgia Tech Research Institute
erica.briscoe@gtri.gatech.edu

D. Scott Appling
Georgia Tech Research Institute
scott.appling@gtri.gatech.edu

Heather Hayes
Georgia Tech Research Institute
heather.hayes@gtri.gatech.edu

Abstract

With the increasing reliance on social media as a dominant communication medium for current news and personal communications, communicators are capable of executing deception with relative ease. While past-related research has investigated written deception in traditional forms of computer mediated communication (e.g. email), we are interested determining if those same indicators hold in social media-like communication and if new, social-media specific linguistic cues to deception exist. Our contribution is two-fold: 1) we present results on human subjects experimentation to confirm existing and new linguistic cues to deception; 2) we present results on classifying deception from training machine learning classifiers using our best features to achieve an average 90% accuracy in cross fold validation.

1. Introduction

The instantaneous nature of social media communication has drastically changed the way many individuals receive social information, assess news about the world, and discuss their interpretations of everyday life and events. With new, pervasive technology, most anyone can address large numbers of people instantly. Less benign, is the recently increased ability for individuals to engage in targeted messaging with malicious intent, as through social engineering [1]. Inherent to this increased access is a large influence potential, where deception is a common utilized method [2]. In line with previous research [3,4,5], our interest focuses on deception that involves a sender crafting text-based messages in an attempt to affect the beliefs of the message receiver through the use of deceit.

Breaking news stories and world events – for example, the Arab Spring [6] – are heavily represented in social media, making them susceptible topics for influence attempts via deception. Whether to intentionally cause harm or not, deceptive messages are often taken at face value, as exampled by the 2012 storm 'Sandy', where false tweets about the Wall Street stock exchange flooding were widely reported at journalistic fact and later confirmed to have been an intentional false report [7]. The style of communications that take place in social media are evolving from those in traditional computer-mediated communication (CMC) and we expect that the indicators of deception found in CMC [5] are evolving as well, especially considering both users' expectations and the additional constraints of using social media (e.g., the 142 character limitation in Twitter).

Well-crafted deceptive messaging is difficult to detect, a difficulty compounded by the fact that people are generally naïve believers of information that they receive. Through studying modern forms of communication, as that found in social media, we can, though, begin to develop an understanding of how users' expectations lead them to detect deception and how deception strategies are exhibited through linguistic cues. That cues to deception exist was first proposed by Ekman [8]. We also follow Zuckerman et al.'s [9] theory that no one behavior is consistently indicative of lying, rather, we can look for the psychological process that may occur to a greater or lesser extent, when people are being deceptive. Further, we use Zukerman et al.'s [9] conception of the cognitive aspects of deception, that characterized deception as requiring complicating cognitive processes that, subsequently, have indicators.

In this paper we report a study of cues to deception and automated methods for deception detection in social media by application and extension of deception research using human subjects' experimentation. We conduct in-laboratory experiments using an experimental social media platform, "FaceFriend", to: 1) determine which linguistic cues are implicitly reflective of deception in social media-type platforms 2) Evaluate if individuals utilize linguistic cues to determine the presence of deception and 3) build and measure the performance of deception detection and deception strategy classifiers on deceptive text.

978-1-4799-2504-9/14 $31.00 © 2014 IEEE
DOI 10.1109/HICSS.2014.186

The paper is organized as follows: in the next section we review the literature on the use deception cues in computer mediated communication and related disciplines; then, we describe two human subjects experiments and discuss the results of statistical significance testing and the results of supervised learning experiments; lastly, we present conclusions and future work.

2. Communication in Social Media

Communications that take place in social media are evolving from other forms of traditional CMC, resulting in various differences and expectations about linguistic properties of this type. For example, Pérez-Sabater's [10] work found that email communications are more formal, where Facebook communications lack structural features such as opening and closings. These differences likely also reflect the differences in audience between email and other specifically social-network oriented means of communication. Email is most often directed at a specific person with the understanding that no one but the recipient will be reading the message. Modern social network-oriented communication are often used under the assumption that many people are potential recipient (or at least observers) of a message.

In addition to basic text-only interaction, many online social media offer new communication affordances like emoticons, embedded pictures or videos, and links to related external content. Though new technology often provides increased ability, the interaction design of many social networking sites also often provides constraints on communication, such as the 140 character limit on Twitter. These interface designs and differences in audience are likely to affect users' creation and perception of content [11].

3. Cues to Deception in Communication

A variety of cues may indicate deceptive messaging [3,4,12,13], ranging from those deriving from face-to-face communication to less dynamic forms, such as those in text-based messaging. Table one provides examples of both face-to-face and linguistic cues as determined by DePaulo, et al. [12]. For example [14] presented a framework for detecting deceit using a dynamic Bayesian model of eye movement that achieved 82% classification accuracy. Our work concentrates solely on linguistic cues, with the goal of understanding how the increasing reliance on new communication media may provide similar indicators of deception.

Table 1. Examples of Face-to-Face, Textual Cues, and Speech Cues from DePaulo, et al. [7] that can be used to study deception in social media communication artifacts.

Cue Type
Face to Face
Eye contact
Gaze aversion
Shrugs
Posture shifts
Computer Vision
Textual Cues
Self-references
Negative statements / complaints
Generalizing terms
Speech Cues
Frequency
Pitch
Amplitude (loudness)

Our work builds off of the exploration of linguistic cues, notably investigated by Burgoon et al. [15] and as Zhou et al. [13] in CMC. This work tested and identified several linguistic cues that are potential indicators of deception (e.g. sentence length, pausality, emotiveness). Using these linguistic cues, previous work [13], [4], investigated the accuracy of machine learning methods (such as decision trees and neural networks) to make deception classifications over text. Neural networks provided, roughly, a 79 percent accuracy on a small data set using cross fold validation. Their features were based on the levels of linguistic cues as determined from the messages comprising their training set.

Our work extends this line of study into communication over social network platforms by performing experiments using a social media style website to replicate conditions similar to those on real social media websites. To understand if the cues determined by [3, 12, 13] are consistent with the use of the type of communication typical to social media, we conducted a human-subjects experiment wherein participants performed two tasks relevant to deception within a larger study of social media usage using a familiar style social networking platform.

2. Experiment 1: Linguistic Cues to Perception Detection

2.1 Participants

Sixty-one participants were paid for their participation in the study, ranging in age from 18-58 with 32 females and 29 males. All participants indicated frequent usage of social media and were fluent English speakers in a pre-experiment survey.

2.2 Method

Subjects were seated at a computer station and asked to use a mock online social media site (*Facefriend*). The mock site was intended to resemble a popular social networking site (see Figure 1). Participants completed three pre-experiment surveys, including a personality test and a survey regarding their social media and email usage (another survey not relevant to these results was also conducted).

Each subject was displayed two series of six screenshots that depicted conversations on the *FaceFriend* platform. Each conversation involved two people (with representative icons) and consisted of 3 statements by each person, for 6 total in each conversation (see Figure 1). The manipulation of the linguistic cues between the two series was identical, but the order of the statements was changed to control for order effects. Subjects were told that the context of the conversations surrounded a decision making task where conversation participants had to decide the optimal ordering of a set of items that would be needed to survive a crash landing in an isolated desert environment (derived from the Desert Survival Problem, [15]). Subjects were verbally instructed that it was likely that some number of the conversation participants had been lying during the conversations in order to win the game. The subjects were then allowed unlimited time to read and analyze the conversations and asked to identify any individuals (via a list of screen names and checkboxes) they thought were being deceptive in their communications. The task was not forced-choice; the subjects could pick from 0-12 conversation participants in each series. After identifying the individuals, the subjects were prompted to explain each of their choices in a free-form response.

The conversations displayed to participants were manipulated so as to control for the specific linguistic properties under evaluation (listed in Table 2) as derived from [12] and [8] (we liken our sentiment measure to previous measures of 'emotiveness'). Each conversation (of the six) in the series was created so that on the conversation participants exhibited a specific level of the cue (low or high),

Table 2. A description of the manipulated linguistic cues in Experiment 1. Asterisks indicate novel cues introduced in this study, others were derived from [12].

Cue	Description	Cue Levels
Sentence Length	The average number of non-punctuation tokens used in each sentence per conversation turn.	High – 8 Words
		Neutral – 6 Words
		Low – 4 Words
Sentence Complexity	The complexity of the conversation turn measured as the average Flesch-Kincaid reading score.	High – 5 Prepositions / Complex Phrases
		Medium – 3 Prepositions / Complex Phrases
		Low – 1 Preposition / Complex Phrase
Sentiment	The average sentiment as measured by the AFINN valence dictionary [21] on a range of [-5,+5] across all sentences in a conversation turn.	High - 3 Positive Sentiment-laden Words
		Neutral – 0 Sentiment-laden Words
		Low - 3 Negative Sentiment-laden words
Txt-Informality*	The ratio of the number of suggested spelling corrections to the total number of correctly spelled words in a conversational turn.	High – 6 informal tokens
		Neutral – 4 informal tokens
		Low – 2 informal tokens
Emoticon Usage*	The magnitude of emotions used in a conversational turn.	High – emoticons present
		Low – No emoticons present

while the other participant exhibited the 'neutral' level on average. For example, for the Sentence Length cue, the first commenter on the thread in the presented screenshot would have an average sentence length of 8, corresponding to the 'high' level. The other commenter in the screenshot would have an average sentence length of 6, corresponding to the neutral condition.

The cues were derived from studies that previously found them to be indicators of deception in computer mediated communication. In addition to these historically relevant cues, we manipulated and tested two new cues that were hypothesized to be relevant to deception, *emoticon usage* and *txt-informality*. Emoticon usage was included so as to represent how people may non-verbally convey mood in social media communications. *Txt-informality* is an augmentation of the concept of informality previously tested in [13] and based on new sociolinguistic study of modern communication [16]. While the original metric measured the ratio of the number of typos to correctly spelled words, we expand this to include the common social media usage of known slang and the intentional shortening of words (e.g. kewl, omg!). A description of the manipulated levels of these cues in the experiment is listed in Table 2.

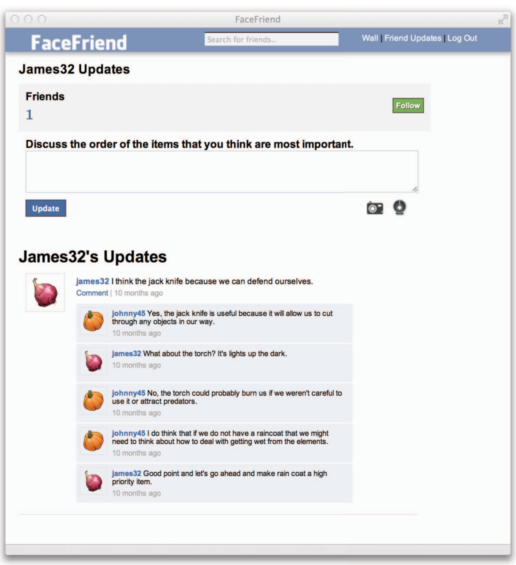

Figure 1. An example of a conversation on the mock social media platform FaceFriend.

2.2 Results

A Chi-square analysis was used to determine the linguistic cues associated with the perception of deception. Overall, the manipulated cues and the perception of deception (as indicated by the subjects' choices of the deceptive participants in the conversations) were not independent of one another (χ^2=81.65; p<.01). No order effects were found – the results were consistent between the two series of conversations. Additional chi-square tests were conducted for each of the five cues.

Sentence complexity and deception perception were not independent of one another (χ^2 =33.08; p<.01), where higher complexity results in lower likelihood of a commenter being identified as deceptive. Sentiment and perception of deception were also related (χ^2 =14.11; p<.01). Here, extreme sentiment results in lower likelihood of a commenter being perceived a liar. Emoticon usage and perception deception were also related (χ^2=23.65; p<.01). A higher use of emoticons results in lower likelihood of a conversational person being perceived a liar.

No relationship was determined between the number of words per sentence or informality and the perception of deception.

3. Experiment 1 Discussion

The results of Experiment 1 determined that higher sentence complexity, emoticon usage, and more extreme sentiment resulted in a lower perception of deception. When comparing this result to previous work in CMC deception, the authors in [17] collected their data using an interview/interviewee task that required some participants to deceive other participants (interviewers) about the theft of a wallet; instead of using social media communication traces they make use of human prepared transcripts of experiment sessions. They found that shorter sentences (i.e. lower average word length) tended to be employed more by deceivers. Our results, however, show that participants do not perceive deception as a function of sentence length. These authors also found that their deceivers tended to employ less complexity during verbal generation, which corresponded to the finding that participants tended to perceive statements with lower complexity as being deceptive and thus would generally be more likely to correctly detect deception on this dimension.

Previous work investigating emotion in CMC, [18], found that, similar to face to face communication, expressing emotion was likely to serve as a 'tool' to bond and to strengthen social relationships. Thus, the presence of emotional indicators may convince a recipient that the sender is attempting to convey closeness and therefore less likely to be deceptive.

In a previous study on deception cues in CMC, [13], the authors found deceptive messages were less complex, though they report a conflicting finding using a different medium (email). This result supports the idea that deception in social media communication is likely to be specific to conditions under which the messages are produced (i.e., informal and text-based).

The results of Experiment 1, compared to previous findings in CMC, suggested that the cues relevant to the perception of deception may be the same as those exhibited in deceptive messaging. Thus, the results of Experiment 1 led to an experiment in which subjects were tasked with active deceptive messaging, as described in Experiment 2, below.

3. Experiment 2: Deceptive Communication

In this experiment, the exhibition of linguistic cues in deceptive communication was investigated to 1) identify which cues are created when a communicator aims to deceive 2) identify the deception strategies individuals employ and 3) build a social media-style communication dataset for doing better deception classification with supervised machine learning methods.

3.1 Participants

The sixty-one participants for Experiment 2 were the same as in Experiment 1.

3.2 Method

Subjects were provided with a written scenario that they were instructed to read. The task involved the 'subarctic survival problem' [18] a group decision-making activity extremely similar to the 'desert survival problem' that was the topic of discussion in Experiment 1. The subjects were instructed to communicate over the *FaceFriend* social media platform to decide the best ranking of items to take from a crashed plane site in an Arctic environment. During their communication, the subjects were instructed to advocate for a specific ranked list of items, also provided, and to do so using deception, when possible. Pilot studies indicated that creating the deception messaging was sometimes difficult for the subjects. To alleviate differences in creative ability, subjects were also provided expert explanations for the advantages provided by each item. For example, the list of items included a 'can of Crisco', which the expert noted could be used to reflect light and signal aircraft. The subjects were allowed to use this information to create similar deceptive explanations. Subjects were told that the other member of the conversation would have no explicit indication of potential deception.

After being explained the task, subjects were asked to log in to the *FaceFriend* social media website to participate in the online interaction with the other party. Unbeknownst to the subjects, the person that they interacted with was a member of the research team (a confederate) who was sitting in another room. During conversation, this confederate responded to the subject's comments using, unless impractical, a list of pre-conceived statements. When impractical to use a pre-conceived response, the confederate would use as short as possible responses (e.g., 'I think we should use the newspaper') to control for confederate-created variance across conversations. After 12 minutes the interaction was stopped and subjects were asked to log out of *FaceFriend*.

Following the interactive session, subjects were shown the list of statements they made during the task and asked to identify which ones were deceptive with a user interface. Upon indicating a particular statement was deceptive, they were prompted to categorize it as a particular deception strategy using a drop-down menu. Strategy choices included Exaggeration [19], Misleading, Omission, Falsification [8], and Other – which allowed free response space for elaboration.

3.2 Results

In order to appropriate analyze the linguistic cues, the participants' statements were first filtered to exclude those that did not contain specific mentions of items used in the ranking task. This was performed so that short responses, such as 'No', were not included. This subset consisted of 254 statements, 132 deceptive and 122 truthful. The same linguistic cues, namely, sentence length, complexity, sentiment, emoticon usage, and txt-informality, from the first experiment were then calculated.

To determine if any of these cues were significantly applied either less or more in deceptive statements, an across-subjects t-test was run on the levels of each cue for the set of deceptive and non-deceptive statements. No significant difference in the levels of Complexity or Sentiment was found between the two sets. There was significant difference for Sentence Length (p<.01) and for Informality (p<.05).

An additional analysis investigated whether the same cues that the subjects chose in Experiment 1 were similarly exhibited in those subjects' deceptive statements. To evaluate this, we ran an across-subjects correlation between the cue levels identified as deceptive in Experiment 1 with those exhibited in Experiment 2. No correlation was found significant at p<.01 level.

3.3. Experiment 2 Discussion

In terms of the cues that are manifest during deception, our results show that only Sentence Length and Informality (as defined in Table 2) occur at significantly different levels within deceptive messaging. The mean value of Sentence Length was high in deceptive messages (7.85 words per sentence in truthful statements; 9.46 in deceptive ones). Txt-informality was lower in deceptive messages (mean of .1376 for deceptive and .1865 for truthful statements). These results do not correspond to those found by Burgoon et al. [17] who found that sentence length decreased and informality increased in deceptive content. This effect may reflect the changing behaviors of users of social media, for example, that are trending towards shorter communications [20].

Interestingly, the levels of cues that were reflected in the subjects' choices of deceptive communication in Experiment 1 were not correlated with those exhibited in Experiment 2. This result suggests that subjects are not aware or not regulating the linguistic cues that they produce or utilize in decisions about message content.

Table 3 presents the counts for each type of deception strategy that the subjects reported using. The strategies are fairly equally used, which supports the generalization of the linguistic cues across deception types.

Table 3. The aggregate count of each type of deception strategy employed by subjects as identified for each statement labeled as deceptive by the subject.

Exaggeration	52
Misleading	56
Omission	38
Falsification	43
Other	17

4. Supervised Learning Experiment

Using the set of linguistic cues (described below) as features, supervised learning classifiers were used to determine whether or not a given statement was truthful or deceptive. The classifier was trained by sweeping the parameter space of settings for each classifier to determine the optimal values to use for higher classification accuracy during cross fold validation.

Three types of features were used. The first set were the cues described in Table 2, the second set were the linguistic cue levels relative to cue levels observed on average by each individual, and third, a set of cues based on the global properties of the dataset, described below. The use of relative cues captures the variance between deceptive and non-deceptive statements at both the local (compared to the surrounding conversation) and global (compared to all conversations) level.

4.1.1. Linguistic Cues. These cues were average sentence length, sentence complexity, sentiment, informality, and emoticon usage.

4.1.2. Locally Relative Linguistic Cues. These cues were average sentence length relative to the average sentence length of all known statements by the individual, sentence complexity relative to the average complexity of all known statements by the individual, relative sentiment to the average sentiment by the individual, relative informality to the average informality of the individual, and relative emoticon usage to the average emoticon usage by the individual.

4.1.3. Globally Relative Linguistic Cues. These cues are the same as the locally relative cues, but compared to the global average of that cue over the entire corpus.

4.2 Deception Classification Results

We performed supervised learning experiments with the previously described features using a parameter optimization with a grid search to find the best parameter configurations for each classifier. Table 4 describes the results of stratified 10-fold cross validation using our best feature set across four different classifiers. The average accuracy was 90%.

Table 4. Stratified 10-fold cross-validated accuracy results for four different supervised learning approaches using best features.

Fold	Random Forest	Gradient Boosting	Support Vector Machines	Percep-tron
1	84.6%	88.4%	88.4%	88.4%
2	84.6%	88%	88.4%	88.4%
3	88.4%	88.4%	88.4%	88.4%
4	92.3%	88.4%	88.4%	88.4%
5	88%	92%	92%	92%
6	92%	92%	92%	92%
7	92%	92%	92%	88%
8	84%	92%	92%	92%
9	92%	92%	92%	88%
10	88%	92%	92%	40%
Average	89% (+/- 2%)	**91%** (+/- 1%)	**91%** (+/- 1%)	85% (+/- 0.07)

4.3 Discussion of Classification

The results of cross validation indicate that deception was successfully classified with very high accuracy. We observe that our classifiers were trained on a roughly equal amount of truthful and deceptive statements from participants, though we expect that the proportions of truthful and deceptive statements that occur in the real world are much less than a 1:1 ratio. Another realistic consideration is that the real-world data would exhibit much more noise than that recorded in a controlled experiment.

Other feature combinations, including ones with sentiment, approached similar accuracy levels but did not produce the highest overall results. The quality of the automated sentiment analysis approach used in the feature extract process may have failed to capture more complicated sentiment and experimentation will continue to explore different sentiment analysis techniques [21, 22].

5. General Discussion

In this study we investigated both how specific linguistic cues are utilized in the perception of deception and how those same cues are manifest in deceptive messaging. This work is intended to extend of previous work (e.g., [7], [9]) that investigated deception in both face-to-face and CMC. Because of the changing face of communication, in light of the ubiquitous use of technology and massive increase in the use of social media, many researchers believe that not only are the means that people use to communicate are rapidly changing, but also the characteristics of the communications themselves (e.g., [23]). The differences between the results of our study and those in previous CMC studies of deception also support this change in communication style and communicator expectations. These findings are supported by the work of Yates' [24], whose hypothesizes that "it is not technology which determines the form and content of CMC but the set of cultural/literacy practices which the users bring to the medium". An important consideration with this study, and those like it, is that because of the newness of the medium, social media communication is rapidly evolving and not yet settled (or 'conventionalized') into its final form, which may take some time [10].

6. Conclusions and Future Work

In this study we have confirmed the presence of several linguistic cues to deception in social media type communications. In doing so, we have introduced novel features that were demonstrated within a supervised machine learning classification task.

In light of the new simultaneous research supporting the differences between native/non-native message senders and communication accommodation theory in social media [10], [25], future work controlling for these phenomena may reveal additional evidence for social media specific linguistic cues. Additional focus will also be given to identifying ground truth data using emoticons and other kinds of social media-specific phenomena (e.g., embedded links, videos) to understand their role in deceptive communication.

Additionally, we are using the dataset described here to understand how personality may be a predictor of the type of deception strategy chosen, as well as identifying cues specific to particular deception strategies. We also intend to extend our features for classification to include discourse cues as [26], which will allow us to account for the context in which statements are made.

10. References

[1] Workman, Michael. "Gaining access with social engineering: An empirical study of the threat." Information Systems Security 16, no. 6 (2007): 315-331.

[2] David B. Buller and Judee K. Burgoon. "Deception: Strategic and nonstrategic communication." *Strategic interpersonal communication* (1994): 191-223.

[3] Lina Zhou, Douglas P. Twitchell, Tiantian Qin, Judee K. Burgoon, and Jay F. Nunamaker Jr. "An exploratory study into deception detection in text-based computer-mediated communication." In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*, pp. 10-pp. IEEE, 2003.

[4] Zhou, Lina, Judee K. Burgoon, Jay F. Nunamaker, and Doug Twitchell. "Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications." *Group decision and negotiation* 13, no. 1 (2004): 81-106.

[5] Judee K. Burgoon,, G. A. Stoner, Joseph A. Bonito, and Norah E. Dunbar. "Trust and deception in mediated communication." In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, pp. 11-pp. IEEE, 2003.

[6] Lora Weiss, Erica Briscoe, Heather Hayes, Olga Kemenova, Sim Harbert, Fuxin Li, Guy Lebanon, Chris Stewart, Darby Miller Steiger, and Dan Foy. "A Comparative Study of Social Media and Traditional Polling in the Egyptian Uprising of 2011." In *Social Computing, Behavioral-Cultural Modeling and Prediction*, pp. 303-310. Springer Berlin Heidelberg, 2013.

[7] Hill, Kashmir. "Hurricane Sandy, @ComfortablySmug, and The Flood of Social Media Misinformation". October 11th, 2012. Retrieved from: http://www.forbes.com/sites/kashmirhill/2012/10/30/hurricane-sandy-and-the-flood-of-social-media-misinformation/

[8] Paul Ekman. *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. WW Norton & Company, 2009.

[9] Zuckerman, Miron, Bella M. DePaulo, and Robert Rosenthal. "Verbal and nonverbal communication of deception." Advances in experimental social psychology 14, no. 1 (1981): 59.

[10] Carmen Pérez-Sabater. "The Linguistics of Social Networking: A Study of Writing Conventions on Facebook." *Linguistik online* 56, no. 6/12 (2012): 82.

[11] Shi, Pan, Heng Xu, and Yunan Chen. "Using contextual integrity to examine interpersonal information boundary on social network sites." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 35-38. ACM, 2013.

[12] Bella M. DePaulo, James J. Lindsay, Brian E. Malone, Laura Muhlenbruck, Kelly Charlton, and Harris Cooper. "Cues to deception." *Psychological bulletin* 129, no. 1 (2003): 74.

[13] Lina, Zhou, Judee K. Burgoon, Douglas P. Twitchell, Tiantian Qin, and Jay F. Nunamaker Jr. "A comparison of classification methods for predicting deception in computer-mediated communication." *Journal of Management Information Systems* 20, no. 4 (2004): 139-166.

[14] Nisha Bhaskaran, Nwogu Ifeoma, Mark G. Frank, and Venu Govindaraju. "Deceit detection via online behavioral learning." In *Proceedings of the 2011 ACM Symposium on Applied Computing*, pp. 29-30. ACM, 2011.

[15] J. C. Lafferty, P. M. Eady, and J. Elmers. "The desert survival problem." *Experimental Learning Methods* (1974).

[16] Thurlow, Crispin, and Alex Brown. "Generation Txt? The sociolinguistics of young people's text-messaging." *Discourse analysis online* 1, no. 1 (2003): 30.

[17] Burgoon, Judee K., J. P. Blair, Tiantian Qin, and Jay F. Nunamaker Jr. "Detecting deception through linguistic analysis." In Intelligence and Security Informatics, pp. 91-101. Springer Berlin Heidelberg, 2003.

[18] Eady, P.M., and Lafferty, J.C. The subarctic survival situation. Plymouth, MI: Experimental Learning Methods, 1969.

[19] William C. Mann, and Jörn Kreutel. "Speech acts and recognition of insincerity." *on the Semantics and Pragmatics of Dialogue* 19 (2004): 64.

[20] Daantje Derks, Agneta H. Fischer, and Arjan ER Bos. "The role of emotion in computer-mediated communication: A review." *Computers in Human Behavior* 24, no. 3 (2008): 766-785.

[21] Nielsen Finn Årup. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages 718 in CEUR Workshop Proceedings : 93-98. 2011. http://arxiv.org/abs/1103.2903

[22] Bo Pang and Lillian Lee. "Opinion mining and sentiment analysis." *Foundations and trends in information retrieval* 2, no. 1-2 (2008): 1-135.

[23] Tagliamonte, Sali A., and Derek Denis. "Linguistic ruin? LOL! Instant messaging and teen language." *American Speech* 83, no. 1 (2008): 3-34.

[24] Simeon J. Yates. "Computer-Mediated Communication. The Future of the Letter." *BARTON. D.; HALL, N.(Eds.) Letter writing as a social practice. Amsterdam/Philadelphia: John Benjamins* (2000): 233-251.

[25] Danescu-Niculescu-Mizil, Cristian, Michael Gamon, and Susan Dumais. "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.

[26] Lina Zhou and Yu-wei Sung. "Discourse Cues to Online Deception." In The *Quality in Government and Business Symposium*, at *the 44th Annual Hawaii International Conference on System Sciences* p. 1-10. 2010.