Linguistic Cues Predict Fraudulent Events in a Corporate Social Network

Max Louwerse (mlouwerse@memphis.edu)
Department of Psychology / Institute for Intelligent Systems
University of Memphis
Memphis, TN 38152 USA

King-Ip Lin (davidlin@memphis.edu)
Department of Computer Science / Institute for Intelligent Systems
University of Memphis
Memphis, TN 38152 USA

Amanda Drescher (adreschr@memphis.edu)
Department of Psychology / Institute for Intelligent Systems
University of Memphis
Memphis, TN 38152 USA

Gün Semin (G.R.Semin@uu.nl)
Faculty of Social Sciences,
Utrecht University
3584 CS Utrecht, The Netherlands

Abstract

There is an increase in deception studies investigating which non-linguistic and linguistic cues best predict deception. Even though these studies have shown participants consistently use specific cues to deception when they are asked to deceive somebody in a particular situation, it is less clear how these findings translate to non-experimental settings, for instance, do these cues also apply in cases of global deception in social networks. This paper investigated whether fraudulent events can be related to linguistic cues of deception within records of a large corporate social network. Specifically, we investigated the Enron email dataset using a model of interpersonal language use. Results suggest that during times of fraud, emails were composed with higher degrees of abstractness.

Keywords: deception, social cognition, computer mediated communication, corpus linguistics.

Introduction

Humans lie because it helps them manipulate the impressions people have of them. Apologizing for being late (even though you could have been on time), telling a police officer you really thought the speed limit was 40 (even though you knew it was 35), and thanking the waitress for guiding you to your table (even though you had waited for 20 minutes and she just did her job), all help to establish an interpersonal glue between you and your social environment. We tell many lies, on average one or two a day (DePaulo & Kashy, 1998).

Of course, there are gradations in the acceptability of twisting the truth. Some lies are blatant transgressions with potentially far reaching consequences, such as cases related to fraud, others are harmless and would have very little or no consequences. Most research in the cognitive sciences on deception centers on lies with little consequences. In fact, very little research has been done on cases of deception with far reaching consequences, for the liar or the recipient of the lie.

Liars leave non-linguistic and linguistic footprints in their attempts to hide the truth, both in cases of blatant and not so blatant half-truths (DePaulo, et al., 2003). Several experiments have investigated these footprints using a paradigm whereby a participant in a deception condition is asked to tell a lie and/or to tell the truth. For instance, Newman, Pennebaker, Berry and Richards (2003) conducted a study in which they asked pro- (and anti-) abortion participants to produce both pro- and anti-abortion stories. They found that deceptive communication had fewer first-person singular pronouns, fewer third-person pronouns, more negative emotion words (e.g., *hate*, *anger*, *enemy*), fewer exclusive words (e.g., *but*, *except*), and more motion verbs (e.g., *walk*, *move*, *go*). Apparently liars wanted to dissociate themselves from their words (fewer first person pronouns), and made an attempt to create a story that seemed less complex (fewer exclusive words) and more concrete (more action words).

Hancock, Curry, Goorha, & Woodworth (2008) came to a very similar conclusion. They investigated deception in asynchronous computer-mediated communication. Participants were asked to write stories on five different topics. Half of the participants were asked to not tell the truth. Hancock et al. (2008) found that lies consisted of fewer words, more questions, fewer first person pronouns and more words pertaining to senses (e.g., *see*, *listen*) than truthful discussions.

Both Newman et al. (2003) and Hancock et al. (2008) found pronoun use, lowered word quantity, emotion words and lower cognitive complexity to be linguistic cues affiliated with deception. Both the experimental design and the findings of these two studies are prototypical for much of the empirical work on deception.

DePaulo et al. (2003) conducted a meta-analysis of experimental literature that investigated cues to deception. They reviewed 116 studies that looked into deceptive cues when people told lies. Results showed, for instance, that liars raised their chins more, pressed their lips more, and had larger pupil dilations than truth tellers. Moreover, lies had more verbal and vocal uncertainty, less verbal and vocal immediacy, were more ambivalent, less plausible and had less logical structure, with less contextual embedding.

However, DePaulo, et al. (2003) warned that these (and other) deception cues were moderated by motivation and transgressions. That is, when participants were more motivated to succeed and when the lies were about transgressions, the deception cues were more pronounced. These moderators are important to note. In fact, it is worth pointing out that the deception studies DePaulo et al. reviewed typically consisted of college students (87.1%), who lied to strangers (88.80%), with lies about transgressions (85.34%).

Indeed, the cues found in the studies DePaulo et al. (2003) used in their meta-analysis are extremely helpful to gaining further insight into deception. In these cases of deception researchers can compare the repertoires of deception cues that humans can use in their lying acts. At the same time, these cues come from unidirectional individual cases in which the participant is asked to act out a lie. It might well be the case that in ecologically situated settings no cues, or different cues, may be observed.

Furthermore, lies often do not impact only the liar. Instead, important cases of lying involve more than a single individual who is aware of the lie. Such instances, where a group of people become part of a collective deception are of a more global nature affecting a social network of people, whereby the individual feelings of guilt and shame are reduced due to a diffusion of responsibility. Examples of deception within a social network include cases of false bookkeeping, mislabeling of accounts, and corruption (Clinard & Yeager, 2006).

Knowing whether (and which) cues to deception can be found in social networks might not answer the question what deception cues humans will use, but it does answer the question whether (and which) deception cues humans generally use. Moreover, such an investigation would be informative in identifying deception strategies in cases of fraud detection or counterintelligence.

This study investigated whether deception in corporate social networks could be detected using linguistic cues.

Enron Email Dataset

The ideal corpus for a study on deception in corporate social networks is the Enron email dataset (Klimt & Yang, 2004). This dataset consists of email messages from various Enron executives/employees obtained from the accounts of 150 executives.

Enron Corporation is most famous for the elaborate network of accounting fraud spread throughout the organization. The company formed in 1985 through the merger of Houston Natural Gas (HNG) and InterNorth Inc. After years of extensive reorganization and rebranding by CEO Kenneth Lay, Enron formed into one of the world's leading natural gas, electricity, and communication companies. Despite its six-year title within Fortune magazine as "America's Most Innovative Company," Enron's network of accounting fraud prompted an SEC inquiry that ultimately lead to the dissolution of the accounting firm Arthur Andersen and a declaration of bankruptcy by Enron Corporation in 2001.

The Enron email dataset is extremely useful for the purposes of this study. First, the dataset is highly diverse, consisting of over 20,000 different senders. Second, the emails cover a relatively large time span (1999-2001). Most importantly perhaps, there is detailed information available on Enron Corporation, its rise and fall and its fraudulent activities (Diesner, Frantz, & Carley, 2005).

While the advantage of this corpus lies in its ecological validity as well as its diversity in senders, receivers, and topics, the disadvantage is that it is very difficult to determine which emails are deceptive and which emails are not. That is, even though Enron as a whole has been known for its deception, that deception cannot be uniquely attributed to specific people or specific topics. As a result, the best way to identify deception is to use those time stamps during which it was clear – in hindsight – that fraudulent activities took place.

There are a number of studies that have analyzed the Enron dataset. Most of these studies looked at the dynamics of the structure and properties of the organizational communication network (Diesner, et al., 2005). Very few studies have looked at deceptive cues in this email corpus. Keila and Skillicorn (2005) is an exception. They used the four deception categories mentioned earlier (first person pronouns, exclusive words, negative emotion words, and action verbs) to categorize the corpus into emails of interest (which were labeled as unusual and deceptive if they showed evidence of the four categories). Keila and Skillicorn's analysis used singular value decomposition (SVD) as the primary analysis technique and successfully showed how emails can be clustered on the basis of the four deception categories. Importantly, Keila and Skillicorn did not test whether these linguistic cues predicted deception.

The current paper tested exactly this question: can a relation be found between linguistic cues in the Enron email data set and fraudulent events? Because we are dealing with interpersonal communication, we investigated this question using the Linguistic Category Model (LCM).

Linguistic Category Model

There is a range of algorithms we could apply to a corpus like the Enron email dataset (Jurafsky & Martin, 2008). However, because we are dealing with a large number of emails sent by different people on a variety of topics covering a time span of many months, it is desirable to use an algorithm based on a model of interpersonal communication. There are very few computational models

Table 1. Overview categories in the Linguistic Category Model (LCM).

Verbs in this category:	DAV	IAV	SAV	SV	ADJ
Refer to a particular activity.	+	-	-	-	-
Refer to a physically invariant feature of the action.	+	-	-	-	-
Refer to a general class of behaviors.	-	+	-	-	-
Have an action with a clear beginning and end.	+	+	-	-	-
Have associated semantic valence, positive or negative.	-	+	+	-	-
Refer to a single behavioral event.	+	+	+	-	-
Refer to a specific object.	+	+	+	+	-
Refer to a specific situation.	+	+	+	-	-
Refer to a specific context.	-	-	-	-	-
Require context for sentence comprehension.	+	-	-	-	-
Express the emotional consequence of an action.	-	-	+	-	-
Refer to mental and emotional states.	-	-	-	+	-
Readily take progressive forms.	+	+	+	-	-
Are freely used in imperatives.	+	+	+	-	-
Require interpretation beyond description.	-	+	+	+	+

available in the field of social cognition (Newman, et al., 2003).

One successful model of interpersonal language is the Linguistic Category Model (LCM, Semin, 2000; Semin & Fiedler, 1988, 1991). The model consists of a classification of interpersonal (transitive) verbs that are used to describe actions or psychological states and adjectives that are employed to characterize persons. This classification gives insight into the meanings of verbs and adjectives that people use when they communicate about actors and their social events. The model makes a distinction between five different categories of interpersonal terms (Semin & Fiedler, 1991):

(a) Descriptive Action Verbs (DAV) refer to single, specific action with a clear beginning and end, such as *hit*, *yell*, and *walk*.
(b) Interpretative Action Verbs (IAV) refer to different actions with a clear beginning and end, but do not share a physical invariant feature, such as *help*, *tease*, *avoid*.
(c) State Action Verbs (SAV) refer to behavioral events, but refer to the emotional consequence of an action rather than the action itself, such as *surprise*, *amaze*, *anger*.
(d) State Verbs (SV) refer to enduring cognitive or emotional states with no clear beginning or end, such as *hunger, trust, understand.*
(e) Adjectives (ADJ) refer to a characteristic or feature qualifying a person or concept, such as *distraught, optimal*.

These five categories can be seen as a continuum from concreteness (DAV) to abstractness (ADJ). The distinction between the categories is obtained on the basis of a number of conventional grammatical tests and semantic contrasts (Miller & Johnson-Laird, 1976). An overview of the five categories is presented in Table 1.

Several studies have shown that the LCM can adequately capture differences in interpersonal language use predicted by theories in social psychology (see Stapel and Semin, 2007).

Semin and Fiedler (1991) proposed an aggregate of the five categories in the form of an abstractness score. This score was formed by the following straightforward formula:

Abstractness score =

$$\frac{DAV + (2 \times (IAV + SAV)) + (3 \times SV) + (4 \times ADJ)}{DAV + IAV + SAV + SV + ADJ}$$

Semin and Fiedler (1991) make the important claim that items scoring high on abstractness (i.e., through abstractness score, or a high frequency of abstract categories, such as adjectives):

1) generate much disagreement;
2) are difficult to verify; and
3) are low in informativeness of the situation.

These claims are relevant for the purposes of the current paper. We hypothesize that when fraudulent events take place it is more likely that the language used is difficult to verify, is low in informativeness of the situation, and is likely to be subject to disagreement (because it is harder to verify and is low in informativeness). In short, we predict that fraudulent events relate to higher abstractness scores in interpersonal communication.

In the computational implementation of the LCM model we identified all verbs and adjectives that matched the criteria identified by Semin and Fiedler (1988; 1991). This set of words was then sent through the CELEX database (Baayen, Piepenbrock & Van Rijn, 1993) to obtain derivations and inflections. The final LCM result was a list of 31,444 words in total, classified in five categories: DAV (17,884), IAV (9,224), SAV (1,533), SV (433), and ADJ (2,370). In addition, adjectives were broken down by the same categorical separations as the verb categories: DA-ADJ (467), IAV-ADJ (1,564), SAV-ADJ (220), SV-ADJ (119).

Table 2. Overview of Enron Corporation events used in Study 1 and 2. Superscripts mark multiple events.

Variable	Description of Variable	Month and Year
Layoffs	Employees within Enron Corp. were laid off.	12/01
CEO	Indicating involvement of the CEO within any coded event.	3/00, 8/00, 11/00, 1/01-4/01, 8/01[6], 10/01[3], 11/01
Fraudulent Paperwork Filed Signed	Filing and/or signing of fraudulent paperwork (by the CEO or COO.)	3/00[2], 8/00
Fraudulent Comments	Enron made fraudulent comments, to the employees and/or investors.	1/01[2], 9/01[2]
Discussion of Ethics	A discussion of ethics occurred between Enron executives or between the CEO and employees	7/00, 3/01, 5/01, 8/01[2], 9/01, 10/01
Selling Enron Shares	Selling of Enron stock by high-level executives occurs.	11/00, 5/01, 6/01, 7/01[2], 8/01[2], 9/01[2]
Rolling Blackouts Initiated	Intentional initiation of rolling blackouts in California.	1/01
Meetings with Nat'l Political Figures	High-level Enron executives met with national political figures incl. the Secr. of the Treasury and the Secr. of Commerce	2/01, 3/01, 4/01, 8/01, 10/01[4], 11/01
Financial Support of Political Candidate	High-level Enron executives (CEO & President) provided financial support for a newly elected national political figure.	1/01
Profit Announced	Profits were announced for the quarter.	4/01
Loss announced	Losses were announced for the quarter.	10/01
SEC Inquiry Developments	Beginning of the SEC inquiry and the point at which the SEC inquiry became a formal investigation.	10/01[2]
Shredding Occurs	Shredding of Enron documents in Enron and/or Arthur Andersen accounting firm.	10/01[2]
Shredding Stopped	Shredding of Enron documents stopped in Enron and/or Arthur Andersen.	10/01, 11/01
Fraud Announced	Enron admitted to having overstated the company's profits	11/01
Bankruptcy Filed	Bankruptcy was filed.	12/01

The content of each of the 255,637 messages was extracted, and the frequency of words in each of the five LCM categories was determined. These frequencies were normalized to account for the number of words in an email.

Sixteen events related to the rise and fall of Enron Corporation, and occurring during the time of the emails, were identified. These events are given in Table 2. Note that some events are directly related to fraudulent activities (e.g., Fraudulent paperwork filed signed; Fraudulent comments; Shredding occurs) and others indirectly (Selling Enron shares; Rolling blackouts initiated; Financial support of political candidate). These events were dummy-coded using a 1 for the presence and a 0 for the absence of an event in the month and year (Cohen, Cohen, West, & Aiken, 2003). This resulted in a database of the sender, the normalized frequency of the LCM categories in each email, and the events linked to the time the email was received.

A mixed-effect regression model analysis was conducted on the normalized frequency of LCM categories, with events as fixed factors, and email sender and email date (year and quarter) as random factors (Louwerse & Jeuniaux, 2010). The model was fitted using the restricted maximum likelihood estimation (REML) for the continuous variable (the normalized frequency of the LCM category). F-test denominator degrees of freedom were estimated using the Kenward-Roger's degrees of freedom adjustment to reduce the chances of Type I error. It is important to point out that mixed effect regression models are very robust with regards to unequal cell sizes, which are a necessary consequence of this dataset.

Given the sheer size of the LCM wordlist, the diversity of topics, senders, and dates (the latter two controlled for in the mixed effect regression model) it is surprising to find any fraudulent event being predicted by the data. Nevertheless, as Table 3 shows, several events can be successfully related to linguistic cues. Recall that, according to the LCM, emails scoring high on abstractness are difficult to verify and are low in informativeness of the situation. Table 3 supports this idea. For instance, during the times that shredding occurred, shredding stopped, and fraud was announced, emails scored higher on abstractness.

Moreover, the most abstract category according to the LCM model is the adjectives. Discussion of ethics, financial support of political candidate, shredding occurs, shredding stopped, fraud announced, and bankruptcy filed, all predicted a higher frequency of adjectives.

Even though these results generate new research questions, there is evidence that the LCM model allows for predicting fraudulent events. Earlier in this study, however, we reviewed studies that found categories such as pronoun use, word quantity, emotion words and cognitive complexity to be affiliated with linguistic cues to deception. Although we do not have access to the exact linguistic cues of some of these categories, we can create an algorithm that approximates these cues. This is what was done in a second study.

Study 2

In the second study we used some of the categories that Newman et al. (2003) and Hancock et al. (2008) reported to be linguistic cues to deception in their experiments: first person pronouns, third person pronouns, causal adverbs, negation (both analytic and synthetic negation), the connective "but", and the length of the email in number of words.

As in Study 1, each of these seven categories was compared with the dummy-coded events in Table 2 using a mixed-effect regression model, thereby controlling for sender and date of the emails.

Table 4 shows the results of this analysis. Events such as fraud announced, bankruptcy filed, fraudulent paperwork filed/signed, and layoffs were related to first person pronouns in emails. However, this relation was in the opposite direction of the one found by Newman et al. (2003) and Hancock et al. (2008).

Fraudulent comments, meetings with national political figures, SEC inquiry developments, and stopping of shredding were related to a higher frequency of negations (analytic negations). It is also noteworthy in these findings that negations were predicted by the stopping of shredding, but not by the occurring of shredding.

Overall, these findings are less uniform than the findings presented in Study 1. This lack of uniformity may be due to the incompleteness with respect to several of the linguistic cues assessed by Newman et al. (2003) and Hancock et al. (2008). Furthermore, the dataset analyzed here did not necessarily represent individual views on situations, unlike the situational data analyzed by Newman et al. (2003) and Hancock et al. (2008). Despite these discrepancies, the findings of the second study are helpful, as a tool of comparison to those in the first study.

Discussion and Conclusion

This study investigated whether linguistic cues can be linked to fraudulent events in a corporate social network. Various studies have looked at linguistic cues to deception. However, unlike the study presented here, these studies used carefully controlled experiments in which participants were asked to give their individual views to a receiver. Most notably, participants were placed in a lying or truthful condition. These studies provide an excellent insight in ways to deceive others, but it is at least an empirical question whether the same linguistic cues can predict deception in more ecologically valid situations. Moreover, it is worth determining whether linguistic cues of deception can be identified in large social networks.

The results of the two studies presented here on the Enron email dataset, a large record of a corporate social network, suggest that abstractness of an email is most indicative of fraudulent events.

By no means are we arguing that by using the LCM model we can predict whether an email consists of fraudulent information or not. At the same time, our results suggest that during times of fraudulent activities messages are sent out with a higher level of abstractness than during times such fraudulent activities are absent or less prevalent.

The work presented here can be extended along a number of dimensions. First, it might well be possible that the LCM categories used here allow for a different abstractness formula that better predicts the result.

To our knowledge this is the first study that has analyzed the impact of fraudulent events on the interpersonal language use of a large social network. Even though the results invite further research, the findings presented here are encouraging, and provide valuable information to the field of deception and interpersonal language use.

Table 3. Significant results mixed effects regression analysis LCM categories. Pluses mark positive relations, minuses negative relations (++ $p < .01$, + $p < .05$, - $p < .05$, -- $p < .01$)

	DAV	IAV	SAV	SV	DA-ADJ	IA-ADJ	SA-ADJ	S-ADJ	ADJ	Abstract-ness
Layoffs	++		++		++	-				
CEO					+			+		-
Fraudulent Paperwork Filed Signed		++								
Fraudulent Comments			-		--					
Discussion of Ethics					--				+	
Selling Enron Shares	-				--					
Rolling Blackouts Initiated					+					
Meetings with Nat'l Political Figures		-		++						
Financial Support of Pol. Candidate			--						++	
Profit Announced					--					
Loss Announced			+		++			+		
SEC Inquiry Developments					++	+		+		++
Shredding Occurs			++		++	+			++	+
Shredding Stopped		+	++		++			+	++	++
Fraud Announced			++		++				++	++
Bankruptcy Filed	++		++		++	-			++	

Table 4. Significant results mixed effects regression analysis various linguistic categories. Pluses mark positive relations, minuses negative relations (++ $p < .01$, + $p < .05$, - $p < .05$, -- $p < .01$)

	1st pers. pronoun	3rd pers. pronoun	causal adverbs	analytic negation	synthetic negation	*but*	word count
Layoffs	+	--		--			-
CEO							
Fraudulent Paperwork Filed Signed	+					-	
Fraudulent Comments				+			
Discussion of Ethics							
Selling Enron Shares			+				++
Rolling Blackouts Initiated							
Meetings with Nat'l Political Figures				+			
Financial Support of Pol. Candidate							
Profit Announced							++
Loss Announced		+					+
SEC Inquiry Developments				++			
Shredding Occurs							+
Shredding Stopped				++			
Fraud Announced	++						
Bankruptcy Filed	++				++		

Acknowledgments

This research was in part supported by grant NIH 1RC1LM010442-01 and NSF 0904909. The usual exculpations apply.

References

Baayen, R. H., R. Piepenbrock, and H. van Rijn (Eds.) (1993). *The CELEX Lexical Database (CD-ROM)*. University of Pennsylvania, Philadelphia (PA): Linguistic Data Consortium.

Clinard, M. & Yeager, P. (2006). *Corporate crime*. New York: Free Press.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, W, The Enron e-mail dataset. http://www.cs.cmu.edu/~enron/. Last accessed 2/5/2010

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74*, 63–79.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74-118.

Diesner, J., Frantz, T., Carley, K.M. (2005). Communication networks from the Enron email corpus "It's always about the people. Enron is no different". *Computational and Mathematical Organization Theory, 11*, 201-228.

Hancock, J.T., Curry, L., Goorha, S., & Woodworth, M.T. (2008). On lying and being lied to: A linguistic analysis of deception. *Discourse Processes, 45*, 1-23.

Jurafsky, D., & Martin, J.H. (2008). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. New Jersey: Prentice-Hall.

Keila, P.S. & Skillicorn, D.B. (2005). Detecting unusual email communication. *Proceedings of the 2005 conference of the Centre for Advanced Studies on Collaborative research (CASCON 2005)*, 238–246.

Klimt, B. & Yang, Y. (2004). The Enron corpus: A new dataset for email classification research. *Proceedings of the Fifteenth European Conference on Machine Learning*, pp. 217–225.

Louwerse, M.M., & Jeuniaux, P. (2010). The Linguistic and Embodied Nature of Conceptual Processing. *Cognition, 114*, 96-104.

Miller, G. A. & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. N. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*, 665–675.

Semin, G. R. & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology, 54*, 558-568.

Semin, G. R. & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. *European Review of Social Psychology*, 1-30.

Semin, G. R. (2000). Communication: Language as an implementational device for cognition. *European Journal of Social Psychology, 30*, 595-612.

Stapel, D. & Semin, G. R. (2007). The magic spell of language. Linguistic categories and their perceptual consequences. *Journal of Personality and Social Psychology*, 93, 23-33.