

Linguistic deception theory:  
Is the world of  
e-discovery ready?



# Executive summary

Academic research has shown that applying linguistic cues of deception to e-discovery analyses may help identify situations where e-mail authors are writing deceptively or avoiding certain words to evade suspicion. Yet researchers have generally been limited to using historical e-mails or artificial data sets developed specifically for the research. Can computer applications automate the detection of e-mails likely to be deceptive? Taking the next steps will require new research using real-life e-mails.

### E-discovery: current practices and challenges

Given the high volume of data that needs to be searched when a legal matter or allegation of misconduct arises, legal counsel and forensic specialists have tried to find more efficient ways of combing through e-mails and other electronic records to identify relevant information. Out of this mission, the e-discovery field emerged.

E-discovery, the technology-enhanced analysis of electronic records, can take many forms. For instance, The Sedona Conference® Working Group on electronic document retention and production suggests that “selective use of keyword searches can be a reasonable approach when dealing with large amounts of electronic data. Examples of search terms include the names of key personnel, date ranges, and terminology related to a specific event. It is also possible to use technology to search for ‘concepts,’ which can be based on ontologies, taxonomies, or data clustering approaches.”<sup>1</sup>

---

With today’s e-discovery techniques, many potentially significant e-mails may be missed because people discussing illegal activity tend to be deceptive in their writing.

However, while today’s e-discovery techniques have come a long way, they still may miss many potentially significant e-mails. Why? Because people discussing illegal activity tend to be deceptive in their writing — and e-discovery analysis techniques may not overcome the misleading nature of their language.

What can be done to help improve today’s e-discovery capabilities to address deception in e-mails?

### Computational linguistics: A cross-disciplinary approach

Part of the solution may lie in recent research in the computational linguistics field on deception. This area of study combines linguistics with computer science and potentially could augment the analysis of keywords, concepts, and circumstantial data. Several types of analysis can be classified as linguistics-based cues,<sup>2</sup> including:

- Analysis of word-type frequency to help identify whether certain kinds of words are used more or less frequently in deceptive communications
- Identification of deceptive word substitutions — for example, attempting to detect whether code words are being used in communications

1 The Sedona Principles, Second Edition (2007), Cmt. 11.a. In this citation, the term “ontologies” is used with a specific computer science meaning. It refers to semantic data models used to integrate heterogeneous databases (see <http://tomgruber.org/writing/ontology-definition-2007.htm> for more discussion).

2 Zhou et al. (2004)

- Analysis of the length of messages in electronic communications to help identify whether there is a correlation between length and deception

---

Pennebaker’s findings suggest that individuals who try to deceive generally use fewer first-person pronouns and exclusive words and more negative-emotion words and action verbs.

For years, linguists have tried to automate the study of the emotional, cognitive, structural, and process components of verbal and written communications. Some of the seminal work on identifying deceptive writing in computer-mediated communication has been carried out by David Skillicorn at Queen’s University and Jeff Hancock at Cornell University. Both have benefited from research by James Pennebaker at The University of Texas.<sup>3</sup>

### Automating deceptive-text detection

From 1992 to 1997, Pennebaker’s group developed a text analysis software application, Linguistic Inquiry and Word Count (LIWC), that gained strong acceptance among linguists who study deception. As part of his research, Pennebaker studied ways to automate the detection of deceptive text in a supervised environment. He used linguistic cues to compare documents preidentified as being deceptive with those known to be truthful. Pennebaker’s team uncovered differences in word usage between the two document sets that suggested a word-use model to identify text likely to be deceptive.

For example, individuals who try to deceive generally use fewer first-person pronouns (such as I, me, and my) and exclusive words (such as except, but, and without) and more negative-emotion words (such as hate, worthless, and sad) and action verbs (such as walk, move, and go).<sup>4</sup>

- Fewer first-person pronouns may indicate the authors’ attempts to dissociate themselves from their words.
- Fewer exclusive words may indicate the telling of a less complex story whose details are easier to create and consistently remember.
- More action verbs may be the result of attempts to reduce the number of exclusive words or distract from a lack of subtlety through plenty of action.

3 Newman et al. (2003)

4 Newman et al. (2003)

- The increased frequency of negative-emotion words may indicate an internal conflict between self-image and the deception.

Skillicorn's work built on these deception cues and combined Pennebaker's findings with computer science tools such as Singular Value Decomposition (SVD). SVD is used in many existing e-discovery tools to find clusters of concepts within sets of documents — for example, grouping e-mail messages about baseball. To date, however, there seem to be no commercial e-discovery tools that use SVD to study linguistic deception cues in e-mail.

### Combing through Enron's e-mails

Keila and Skillicorn (2005) tested Pennebaker's deception cues on a set of more than 500,000 e-mails from former Enron employees, which the Federal Energy Regulatory Commission made public. The team used a subset of the words from the LIWC program and distilled observations of Pennebaker's linguistic cues to rank e-mails based on how well they fit a profile of deception<sup>5</sup>. Their findings suggest that Pennebaker's deception cues "do capture deceptive emails well in the Enron email dataset, although with some adaptation to account for the fact that these emails are (intended to be) written in a business context. ... Some appropriate fraction of the most likely emails can then be selected for further analysis."<sup>6</sup> These might include e-mail messages from senior executives or individuals of interest based on other facets of an investigation, which could be cross-referenced with the findings of this linguistic analysis.

### Spotting substitutions

Skillicorn (2005) and Fong, Skillicorn, and Roussinov (2006) also tested an automated process to detect situations in which a word was replaced to deceive an outside reader. For instance, Skillicorn explains that terrorist communications might replace "attack" with "wedding." This replacement maintains the syntax of a sentence because weddings happen at a particular place and time and require a group to coordinate travel and meeting. The researchers hypothesized that they could identify word substitutions by analyzing word frequency.<sup>7</sup>

<sup>5</sup> Keila and Skillicorn noted that the overall use of first-person pronouns was rare in the Enron e-mail data set, presumed to be a particular characteristic of business communications, and adjusted their deception profile to account for this attribute.

<sup>6</sup> Keila and Skillicorn (2005)

<sup>7</sup> Skillicorn (2005)

## Skillicorn's team claimed that Pennebaker's deception cues "do capture deceptive e-mails well in the Enron e-mail dataset, although with some adaptation to account for the fact that these e-mails are (intended to be) written in a business context."

Based on this, they developed a number of measures to identify if a word is out of context, an indication of word substitution. As the only large, contemporary set of real-life corporate e-mails available to researchers, the Enron e-mails again proved to be the most useful data for study. The researchers altered half of the selected pool of Enron e-mail messages with substitutions of the first noun in each selected e-mail, yet kept the syntax of the sentences intact. Skillicorn reported that in applying his technique he observed a 95 percent detection rate with an 11 percent false-positive rate for identifying word substitution.

### Scrutinizing style

In a different twist, Hancock, Curry, Goorha, and Woodworth (2004) studied changes in the linguistic style of both senders and receivers of e-mails and instant messages. Researchers asked the senders to be deceptive when communicating about two out of four discussion topics. The study revealed two key findings:

- Senders of deceptive messages used more words overall, made more references to others, and included more sense-based descriptions (such as seeing and touching) than when they were telling the truth.
- Receivers of deceptive messages used more words and sense terms, asked more questions, and used shorter sentences than when they were being told the truth.

### Studying e-mails as they're written

The team of Zhou, Burgoon, Nunamaker, and Twitchell (2004) noted that "most experimental data in the prior research were collected via interview, interrogation, observation, or analysis of written statements of a specific past event." Their focus instead was to test a set of linguistic cues on e-mail communications where there is an opportunity for the writers "to create and revise their messages so as to make them as persuasive as possible."

The researchers structured a scenario in which pairs of students communicated exclusively by e-mail. One student in each pair was instructed to use deception to persuade the other. The study found that deceptive senders use higher quantities of words, more expressivity, less formality, and more uncertainty in the language of their messages<sup>8</sup> than both the receivers of their messages and truthful message-senders. They also appear to display less immediacy, complexity, diversity, and specificity of language.

#### The need for research using real-life data

As described above, academic studies show intriguing potential for application to e-discovery. However, they are largely untested using real-life data. This is why computational linguistic research should be advanced to the next level — more tests using actual e-mail sets from corporations. The need is vital. Two key shortcomings exist in current e-discovery techniques:

- There may be e-mails of interest containing deceptive language that cannot be reliably identified through the use of keyword searches.
- The authors of deceptive e-mails often avoid certain words or use code words to evade detection.

Research using real-life data sets could determine whether linguistic cues help find instances of deception that can be further analyzed to determine their real meaning, that is, to help determine whether the deception is relevant to the investigative puzzle. A comparison of the tests might look at a series of past instances in which an investigative team of lawyers and forensic specialists found deceptive smoking gun e-mails through a combination of e-discovery, review by knowledgeable parties, investigative interviews, and other efforts. Using the same pool of e-mails obtained during those investigations, along with control sets of e-mails, the researchers could apply linguistic analysis and then draw conclusions as to whether the linguistic analysis found the same smoking gun e-mails, found other e-mails of interest, and found these e-mails more quickly and cost-effectively.

---

## Ultimately, new research should determine whether linguistic cues more effectively or efficiently identify instances of deception that deserve further analysis.

Overall, research using real-life data might yield insights on the following questions:

- Do linguistic analysis techniques produce results consistent with findings obtained using traditional e-discovery tools? If not, why?
- What commonalities exist between electronic communications identified as significant using

traditional e-discovery techniques and those using the linguistic cues discussed in this paper?

- Does linguistic analysis add value to the investigative process?
- What are the pros and cons of reconfiguring existing e-discovery tools to conduct linguistic analyses? Are there comparative advantages or disadvantages to developing greenfield software solutions?

#### Conclusion and next steps

The research featured in this paper shows enough promise to warrant testing the various techniques on real-life e-mails. This would help build stronger empirical evidence for the effectiveness of analyzing linguistic deception cues. To date, Skillicorn's use of Pennebaker's deception cues, in combination with computer science approaches, provides the most relevant findings for e-discovery because the study used a set of e-mails from ex-Enron employees and management. We'd like to see real-life e-mail data tested using Hancock's innovations that study the changes in linguistic patterns both by deceptive authors and the truthful recipients of those messages.

Now the question remains: How do you move from academia's existing body of work toward analyses that can be applied to corporate investigations and litigation discovery?

Private-sector e-discovery specialists maintain terabyte upon terabyte of electronic communications as part of their engagements. However, e-discovery research that uses this data must not violate laws, regulations, and engagement requirements relating to client confidentiality and data privacy and protection.

How might researchers work within these bounds while advancing the science of e-discovery? Several considerations include:

- Developing a stronger alliance between the legal community, private-sector e-discovery specialists, and researchers so acceptable standards for conducting the research might be established and live data sets might be made available to the researchers
- Indexing all words in a live data set, removing the proper nouns, and replacing these proper nouns with mundane and trackable terms or codes
- Exploring the use of data sets from companies similar to Enron — defunct, yet with large, available e-mail data sets

It will be important to test the effectiveness and efficiency of new linguistic approaches to e-discovery compared to existing techniques. We believe that if further research supports confidence in linguistic analysis of deception, subsequent work may focus on practical applications, such as new configurations of existing tools or greenfield software applications. E-discovery's growing importance as an investigative resource suggests that investment in such innovations could be worthwhile — and harness the potentially powerful combination of linguistic cues and computer science to identify deceptive e-mails.

<sup>8</sup> For definitions of these categories, see Zhou et al. (2005).

Daniel Mosher is a senior manager in the Forensic & Dispute Services practice of Deloitte. He focuses on corporate investigations of fraud and corruption allegations.

The views expressed in this article are those of the author and do not necessarily reflect the views of Deloitte Financial Advisory Services LLP.

This publication contains general information only, and Deloitte Financial Advisory Services LLP is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect

your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte Financial Advisory Services LLP, its affiliates and related entities shall not be responsible for any loss sustained by any person who relies on this publication.

#### Contact information

**Daniel Mosher**  
Senior Manager  
Forensic & Dispute Services  
Deloitte Services LP  
+1 203 563 2771  
dmosher@deloitte.com

#### References

- **Buller, D. B., J. K. Burgoon, A. Buslig, and J. Roiger (1996).**  
"Testing Interpersonal Deception Theory: The Language of Interpersonal Deception," *Communication Theory*, 6, 268–289.
- **Dulaney, E. F. (1982).**  
"Changes in Language Behavior as a Function of Veracity," *Human Communication Research*, 9, 75–82.
- **Fong, Sze Wang, David Skillicorn, Dmitri Roussinov (2006).**  
"Measures to Detect Word Substitution in Intercepted Communication," in the proceedings of IEEE International Conference on Intelligence and Security Informatics (ISI2006), May 23–24, 2006.
- **Hancock, J. T., L. Curry, S. Goorha, and M. T. Woodworth (2004).**  
"Lies in Conversation: An Examination of Deception Using Automated Linguistic Analysis," *Proceedings, Annual Conference of the Cognitive Science Society*, 26, 534–540.
- **Keila, P. S. and D. B. Skillicorn (2005).**  
"Detecting Unusual and Deceptive Communication in E-mail," external technical report, School of Computing, Queen's University, Kingston, Ontario, Canada, 2005; see <http://www.cs.queensu.ca/TechReports/Reports/2005-498.pdf>
- **Knapp, M. L. and M. A. Comadena (1979).**  
"Telling It Like It Isn't: A Review of Theory and Research on Deceptive Communications," *Human Communication Research*, 5, 270–285.
- **Knapp, M. L., R. P. Hart, and H. S. Dennis (1974).**  
"An Exploration of Deception as a Communication Construct," *Human Communication Research*, 1, 15–29.
- **Mehrabian, A. (1971).**  
"Nonverbal Betrayal of Feeling," *Journal of Experimental Research in Personality*, 5, 64–73.
- **Newman, M. L, J. W. Pennebaker, D. S. Berry, and J. M. Richards (2003).**  
"Lying Words: Predicting Deception from Linguistic Style," *Personality and Social Psychology Bulletin*, 29, 665–675.
- **Skillicorn, David (2005).**  
"Beyond Keyword Filtering for Message and Conversation Detection," in the proceedings of IEEE International Conference on Intelligence and Security Informatics (ISI2005), pages 231–243, May 2005  
*The Sedona Principles, Second Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production* (The Sedona Conference® Working Group Series, 2007).
- **Zhou, Lina, Judee K. Burgoon, Jay F. Nunamaker, and Doug Twitchell (2004).**  
"Automating Linguistics-Based Cues for Detecting Deception in Text-based Asynchronous Computer-Mediated Communication," *Group Decision and Negotiation*, 13: 81–106.