December 2006

An Analysis of Text-Based Deception Detection Tools

Christie Fuller
Oklahoma State University

David Biros
Oklahoma State University

Douglas Twitchell
Illinois State University

Judee Burgoon
University of Arizona

Mark Adkins
University of Arizona

An Analysis of Text-Based Deception Detection Tools

Christie M. Fuller
MSIS Deparment, Oklahoma State University
christie.fuller@okstate.edu

David P. Biros
MSIS Deparment, Oklahoma State University
david. biros@okstate.edu

Douglas P. Twitchell
Illinois State University
dtwitch@ilstu.edu

Judee K. Burgoon
CMI, University of Arizona
jburgoon@cmi.arizona.edu

Mark Adkins
CMI, University of Arizona
madkins@cmi.arizona.edu

ABSTRACT

The quality of information can be degraded when individuals attempt to deceive others through information manipulation. This can be very influential in a text-based domain. In recent years, tools have been developed that, while not initially designed for this domain, have been adapted successfully for use in identifying deception in text-based communication. These text analysis tools, which utilize features such as parsing and categorizing, are emerging as accurate tools to identify cues that may be useful in distinguishing deceptive from truthful communications. These deception detection tools have been applied to problems such as security screening, criminal incident statements, and evaluation of online communication patterns. This paper provides a comparative analysis of the features and capabilities of two of the more promising tools and identifies how their use might fit within existing theoretical constructs.

Keywords

Deception, Cues, Deception Detection, Linguistic Analysis, Text

INTRODUCTION

Deception has previously been defined as "a message knowingly transmitted by a sender to foster a false belief or conclusion by the receiver" (Buller and Burgoon, 1996). Though research in the detection of deceptive communication has been ongoing for some time, humans have not proven to be very capable lie detectors (Vrij, Edward, Roberts and Bull, 2000). Only a few groups of professionals, such as secret service agents, have been reported to exceed chance levels (Ekman, O'Sullivan and Frank, 1999). Based on a synthesis of research from 200 documents and 23,500 judges, Bond and DePaulo (2006) concluded that average detection accuracy is only 54%.

In light of the difficulty human detectors have in recognizing deception, several methods have been developed to assist or replace humans in deception detection. These include the polygraph, Statement Validity Assessment, and Reality Monitoring (Vrij et al., 2000). Statement Validity Assessment is composed of three elements, one of which is criteria-based content analysis (CBCA). CBCA is used to systematically assess a statement. Reality monitoring is based on the notion that memories derived from real versus imagined events differ on several characteristics. A number of criteria are available to evaluate a statement based on these characteristics (Vrij, 2000). Though these detection alternatives do exist, they may be intrusive or fail to provide immediate feedback. Another alternative, computerized voice stress analysis, is less invasive, but not feasible in many situations (Twitchell, Jensen, Burgoon and Nunamaker, 2004).

One promising aid in deception detection is linguistic analysis (Qin, Burgoon, and Nunamaker, 2004; Zhou, Burgoon, Nunamaker and Twitchell, 2004), which has great applicability, given the rise in text-based communication in everyday life (Zhou, Burgoon and Twitchell, 2003) and the difficulty with which people recognize verbal forms of deceit. For example, one study showed that people lie in 14% of emails and 21% of instant messages (Hancock, Thom-Santelli and Ritchie, 2004). Another study by George and Keane (2006) examining deceptive resumes found that respondents identified less than a third of the deceptions in text. A particularly challenging context for detecting deceit from text is that of criminal investigators, who routinely must make decisions regarding the veracity of statements made by persons of interest. This suggests a need for methods of deception detection designed for analyzing text. Automated tools that enhance decision making processes while remaining unobtrusive would be invaluable. It is the decision support function that such tools could provide that makes the current study relevant to the information systems domain.

This paper will offer an analysis of the features of two prevailing tools currently being used to detect deception in verbal communication. Two main steps in the overall process of automated linguistic analysis when applied to the domain of deception detection are (1) identifying and extracting the cues to be used for deception detection and (2) classifying text as deceptive or truthful based on those cues (Adkins, Twitchell, Burgoon and Nunamaker, 2004). The focus of this study is on the first step in this process.

LITERATURE REVIEW

There is a growing body of literature in deception detection (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton and Cooper, 2003; Vrij, 2000; 2005), much of which is focused on human interaction (Buller and Burgoon, 1996; DePaulo et al., 2003; McCornack, 1992). The dynamics of face-to-face (FtF) deception have been well investigated, but text-based deception is still relatively new. Prevailing theories of deception include the Four Factor Theory (Zuckerman and Driver, 1985), Cue Leakage Theory (Ekman, 1985; Ekman and Friesen, 1969), Reality Monitoring (Johnson and Raye, 1981), Interpersonal Deception Theory (Buller and Burgoon, 1996), and Information Manipulation Theory (McCornack, 1992). Additionally, known cues to deception have recently been summarized in the self-presentation perspective of deception (DePaulo et al., 2003).

Many deception detection studies derive from the Cue Leakage Theory, focusing on identifying deceptive cues that a sender might leak out to a receiver. Many of these cues are physical in nature, including eye contact, hand gestures, and facial expressions (Ekman and Friesen, 1969; Ekman and Friesen, 1974). However, recent studies have started to focus on deception detection based on verbal cues (Zhou et al., 2003a; Zhou et al., 2004a). Interpersonal deception theory (IDT) and information manipulation theory (IMT) are based on principles of interpersonal communication and consider the interaction of the deceiver and the receiver in the deceptive interaction. While these theories have been used primarily in studying face-to-face communication, their consideration of the strategic relationship between the participants in the communication process lends themselves to adaptation to the text environment. They support the argument that deception will be evident in the quantity, quality, clarity, relevance, and personalization of deceptive messages—all features that can be captured with linguistic cues.

There is as yet no uniform way to categorize text-based deception cues, but Zhou and colleagues (Zhou, Burgoon, Twitchell, Qin and Nunamaker, 2004) advanced one classification scheme. They organized indicators into *quantity, specificity, affect, expressivity, diversity, complexity, uncertainty, informality, and nonimmediacy*. Quantity matches the quantity dimension in IDT and IMT. It refers to how many words, verbs, sentences, and the like are present, i.e., it reflects the length of the utterance. Specificity is quality-related in reflecting the amount of actual details present. Also somewhat related to quality are the affective tone and amount of expressiveness through use of adjectives and adverbs. Complexity and diversity of the vocabulary and syntax speak to the clarity of the message. Terms expressing uncertainty or ambiguity reflect ways to avoid giving relevant answers, and informal and nonimmediate language are means to distance speakers from their messages or responsibility for any actions in question.

Some cues may prove to be better discriminators than others depending on the medium and context. A recent study by Burgoon, Qin and Twitchell (2006) identified 17 different linguistic cues that may prove to be good discriminators for deception in written communication. Table 1 lists these variables and shows whether the mean for the variable was found to be significantly greater for truthful or deceptive statements.

Recently, computer based tools have emerged to aid human examiners in linguistic cue analysis. These include the Agent 99 Analyzer (A99A) and Linguistic Inquiry and Word Count (LIWC) software. Both tools offer the ability to detect deception in text. This study focuses on testing the tools on a standard set of truthful and deceptive statements. The data utilized for this study are a set of real world statements involving high stakes situations.

CATEGORY	VARIABLE	Deceptive > Truthful	Truthful > Deceptive
Quantity	Word count	*	
	Verb count	*	
	Sentence count	*	
Specificity	Modifier count	*	
	Affect ratio	*	
	Sensory ratio		*
Diversity	Lexical diversity		*
	Redundancy	*	
	Content word diversity		*
Personalization	Non-self references	*	
	2nd person pronouns	*	
	Other References	*	
	Group pronouns	*	
Non-immediacy	Immediacy terms	*	
	Spatial far terms		*
	Temporal nonimmediacy	*	
	Passive voice		*

Note: Non-self references refers to a composite variable formed from 2nd person pronouns, other references and group pronouns. Immediacy Terms is a composite variable formed from Spatial Far terms, Spatial Close and Temporal Non-immediacy.

Table 1: Pilot Study Indicators of Deception

AUTOMATED DECEPTION DETECTION TOOLS

Agent 99 Analyzer

Agent99 is a suite of tools developed at the University of Arizona for aiding deception detection, including deception detection in text (Zhou et al., 2004a; Zhou et al., 2004b; Zhou, Twitchell, Qin, Burgoon and Nunamaker, 2003) and video (Meservy, Jensen, Kruse, Burgoon, Nunamaker, Twitchell, Tsechpenakis and Metaxas, 2005) and deception detection training (Cao, Crews, Lin, Burgoon and Nunamaker, 2003). One of the tools included in the suite is the Agent99 Analyzer (A99A), which was built for detecting deception in text. It was built using the open-source General Architecture for Text Engineering (GATE) (Cunningham, 2002). As implied by its name, GATE is an architecture or platform for creating and running a wide variety of text engineering software.

A99A utilizes GATE for two reasons. First, GATE's architecture is based on modularity. With a small amount of programming, deception cues are easily added to the architecture and depicted in the graphical user interface as *Processing Resources*. GATE allows the user to graphically choose which cues or combination of cues to run on the text. The text, in turn, is also managed graphically with each suspect statement modeled as a *Language Resource*, which can be grouped into corpora for processing. Second, GATE comes with a number of built-in text analysis tools that are suitable for use with deception detection. The most important of these to deception is the part-of-speech tagger (Cunningham, Maynard, Bontcheva, Tablan, Ursu, Dimitrov, Dowman, Aswani and Roberts, 2005), which allows the computation of many of the deception cues including verb count, modifier count, content word diversity, non-self references, second person pronouns, other references, group pronouns, and passive voice. Additionally, GATE comes with other processing resources that split text into sentences and individual words.

One advantage of the part-of-speech tagger that comes with GATE is its statistical nature (Hepple, 2000). The tagger generates rules based on probabilities that are gathered through the use of a large, manually-tagged corpus of text. The statistical process gives the tagger the ability to robustly handle such things as previously unseen words and misspelled words. All words are given a part-of-speech based on the tagger's best guess, with an accuracy of about 97% (Hepple, 2000). While not always correct, the tagger should be accurate enough for the uncertain task of detecting deception.

Utilizing variables reflecting quantity, complexity, uncertainty, non-immediacy, expressivity, diversity, informality, specificity, and affect, the effect of modality in deception has been investigated (Qin, Burgoon, Blair and Nunamaker, 2005), as well as the effects of time and sequence on deceptive responses (Burgoon et al., 2006). An additional feature of A99A is its ability to generate output from GATE that can be used in classification models, such as neural networks, decision trees, discriminant analysis and logistic regression to automatically determine whether statements or messages are deceptive or truthful. After the models are built, they can be examined to identify important cues and the direction of those cues (Zhou et al., 2004b).

LIWC

LIWC (Pennebaker and Francis, 2001) processes text based on four main dimensions: standard linguistic dimensions, psychological processes, relativity, and personal concerns. Within each of these dimensions, a number of variables are represented. For example, the psychological processes dimension contains variable sets representing affective and emotional processes, cognitive processes, sensory and perceptual processes, and social processes. In total, the default dictionary, which includes a total of 2300 words and word stems, serves as the basis for 74 output variables. With a few exceptions, the output variables represent a percentage of the total words belonging to a particular category. One interesting variable of LIWC is percentage of words found in the LIWC dictionary. This can perhaps be viewed as a measure of how much of the statement the tool is able to process. LIWC was initially created to identify basic emotional cognitive and emotional dimensions and has since been expanded and refined. The current version of the software captures about 80% of words used in writing and speech, as measured across 43 studies. The user may also include additional dictionaries. For this study, only default dictionaries of LIWC were utilized.

Newman, Pennebaker, Berry, and Richards (2003) proposed that the language dimensions of self-references, negative emotions, and cognitive complexity could be associated with deception. The use of motion and exclusive words were proposed as indicators of cognitive complexity. The study found that the other references variable was also a predictor of deception. Based on the work of Newman et al. (2003), Bond and Lee (2005) used LIWC to code the statements of prisoners. In addition to the categories studied by Newman et al., Bond and Lee also used LIWC to code Reality Monitoring Terms.

Hancock and colleagues (Hancock, Curry, Goorha and Woodworth, 2004) have also examined the use of automated linguistic analysis in deception. Their research, which draws on Interpersonal Deception Theory (Buller and Burgoon, 1996) and the Self-Presentation Perspective (DePaulo et al., 2003; Vrij, 2000), hypothesized differences in word counts, pronoun usage, words related to feelings and senses and exclusive words will differentiate deceptive and truthful communications. The study used LIWC to analyze eight variables in the four categories described above. A repeated measures GLM design was used to determine which variables differed between deceptive and truthful communications, and also to examine differences between sender and receivers. Deceptive senders used more words, more "other" pronouns such as "he", "she" and "they", and more sensory terms.

METHODOLOGY

The current study used a sample of criminal incident statements collected from a military base in the Midwest United States. The statements were actual statements written by suspects or witnesses involved in criminal incidents on the military base. Most of the statements were found to be truthful. However, in some instances, military law enforcement investigators learned of additional information suggesting the suspect or witness statements were deceptive. When the investigators question the suspects or witness about the incident, they "come clean" and admit they lied on their statements. This serves as ground truth with respect to the veracity of the incident statements. For this study, 30 deceptive and 30 truthful statements were analyzed. The written statements were transcribed into text files for use by the linguistic analysis tools, A99A and LIWC. Statement transcription followed a standardized process and attempted to capture the statements exactly as written, matching grammar, punctuation, capitalization, and so forth. Though the User Manual for LIWC directs the user to correct any misspellings and grammatical errors in transcribed data, for this project, the original transcription was maintained to allow for a more direct comparison of A99A and LIWC. LIWC does not count misspelled words so this may result in fewer words counted overall and some words placed in inappropriate categories. However; this is not expected to have a large effect on results and it was deemed to be more important to maintain consistency between programs.

For A99A, the variables listed in Table 1 were available based on previous research. Approximate matches for the corresponding variables in LIWC were identified by reviewing the LIWC User Manual for descriptions of the variables and sample words belonging to the categories. Table 2 below shows the approximate variable matches in each program.

A99A Variables	LIWC Variables
Word Count	Word Count
Affect Ratio	Affect
Sensory Ratio	Sensory and Perceptual Processes
Lexical Diversity	Unique Words
Non-self References	Other References
2nd Person Pronouns	Total Second Person
Other References	Total Third Person
Group Pronouns	1st Person Plural
Spatial Far+ Spatial Close Terms	Space

Table 2: A99A and LIWC variable matches

No match could be found for verb count, sentence count, redundancy, modifier count, or passive voice. The modal verbs variable was available in both A99A and LIWC and therefore was available to represent uncertainty. Type token ratio was the closest match to both lexical diversity and content word diversity. As it was a closer match to lexical diversity, only this comparison was made and content word diversity was dropped as a variable for comparison. Additionally, immediacy terms could not be directly compared, as the available LIWC variable for temporal non-immediacy appears to also subsume temporal-immediacy. Therefore, only spatial terms were evaluated, and this was an evaluation of spatial far and close terms, based on the availability of the space variable in LIWC.

RESULTS & ANALYSIS

Using both A99A and LIWC, the prepared text files were analyzed to calculate the relevant values for the desired variables. For each program, these results were then separately analyzed to determine which variables could be used to distinguish truthful and deceptive statements. For the variables calculated using A99A, significant differences were found between mean values of the variables for truthful and deceptive statements for all variables except affect ratio and modal verbs. Similarly, for the variables calculated using LIWC, significant differences were found between mean values of the variables for truthful and deceptive statements for all variables except affect ratio and modal verbs. These results show that both programs found significant differences in the mean values for truthful and deceptive statements for the same set of variables. The direction of the differences between variables was as expected and consistent with previous results, shown in Table 1, for all variables, except for the LIWC Sensory and Perceptual Processes variable. While the corresponding Sensory ratio variable was significantly greater in truthful than deceptive statements when calculated by A99A as expected, LIWC found a significantly greater mean for deceptive than truthful statements.

These results lend credibility to the use of these tools in deception detection and other text analysis tasks. The similar results achieved with each tool suggest that cues which have been appropriately defined can be automated to assist investigators. These results might also allow us to draw limited comparisons between different studies using different tools when the variables are defined similarly for both tools. For most of the variables analyzed in this study, the definitions of the variables are relatively straightforward. For example, the list of third person pronouns is fairly well-defined. The results are mixed for less obvious variables such as affect and spatial terms.

A99A Variable	**LIWC Variable(s)**
Word Count	**Word Count**
Affect Ratio	Affect
Sensory Ratio	**Sensory and Perceptual processes**
Lexical Diversity	**Unique Words**
Non-self References	**Other References**
Second Person Pronouns	**Total Second Person**
Other References	**Total Third Person**
Group Pronouns	**1st Person Plural**
Spatial terms	**Space**
Modal Verbs	Modal Verbs
Note: Bold indicates significant difference in mean for variable between truthful and deceptive statements for respective program at 0.05 level.	

Table 3: Results

Despite these promising findings on most variables, the tools failed to detect significant differences on variables previously suggested to be useful as predictors of deception in text, such as affect and modal verbs (Zhou et al., 2004a). It may be that the type of statement being analyzed reduced the presence of affective terms such as "good" or "bad" or produced the same amount in both truthful and deceptive statements. Alternatively, the lack of significance in either program may have been the result of looking at this variable at an aggregate level. Some previous studies have separated this variable into more than one variable (Hancock, Curry, Goorha and Woodworth, 2005; Zhou et al., 2004a), Given that modal verbs have shown to be effective discriminators in other studies, the nonsignificant results on this indicator, like affect, are an argument favoring a multi-indicator model in which only some of the potential indicators are likely to be present in a given statement. Also not to be discounted as an explanation for the nonsignificant findings on these cues is sample size. Only 60 statements were used in this study, which may not be adequate to find significant differences on all cues.

CONCLUSION

The use of linguistic-based cues to aid in deception detection has been attracting increased research interest. As criminal case loads involving deception continue to challenge investigators, new decision support tools are need to help them determine the veracity of person-of-interest statements. This study has examined two tools that have been used previously in such studies. Using a sample of real word data, significant differences were found on eight of ten variables analyzed for both programs. The direction of significance varied from expectations for only one variable, LIWC Sensory and Perceptual Processes variable. The variables utilized in this study were those suggested as appropriate for this context and medium. The list of variables used was based on previous studies utilizing A99A. Some variables were excluded due to a lack of a matching variable in LIWC. Many of these matches could not be made because of LIWC's lack of a part-of-speech tagger, though LIWC has many additional output variables which were not utilized. The consistency in results between the programs suggests opportunities for expanding this comparison to other variables appropriate for deception detection in other domains and contexts. Further, there may be additional prospects for integrating and expanding the dictionaries used by either program. Future studies might extend these comparisons to utilizing the output of each program as the input for classification models. This might further test the capabilities of each program to distinguish between truthful and deceptive statements.

REFERENCES

1. Adkins, M., Twitchell, D.P., Burgoon, J.K., and Nunamaker, J.F., Jr. (2004) Advances in automated deception detection in text-based computer-mediated communication, in Dawn A. Trevisani and Alex F. Sisti (Eds.) *Proceedings of SPIE -- Volume 5423 Enabling Technologies for Simulation Science VIII*, August 2004, Orlando, FL, USA, 122-129.
2. Bond, C. F., and DePaulo, B. M.(2006), Accuracy of deception judgments, *Personality and Social Psychology Reports* (in press).

3. Buller, D.B., and Burgoon, J.K. (1996) Interpersonal deception theory, *Communication Theory*, 6, 3, 203-242.
4. Burgoon, J., Qin, T., and Twitchell, D. (2006) The dynamic nature of deceptive verbal communication, *Journal of Language & Social Psychology*, 25, 1, 1-22.
5. Cao, J., Crews, J.M., Lin, M., Burgoon, J., and Nunamaker, J.F., Jr. (2003) Designing Agent99 Trainer: A learner-centered, web-based training system for deception detection, in H. Chen, R. Miranda, D. Zeng, T. Madhusudan, C. Demchak, and J. Schroeder (Eds.) *Proceedings of the First NSF/NIJ Symposium on Intelligence and Security Informatics (ISI 2003), Lecture Notes in Computer Science (LNCS 2665)*,June 2-3, 2003, Tucson, AZ, Springer-Verlag, 358-365.
6. Cunningham, H. (2002) GATE, a general architecture for text engineering, *Computers and the Humanities*, 36, 2, 223-254.
7. Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V., Ursu, C., Dimitrov, M., Dowman, M., Aswani, N., and Roberts, I. (2005) Developing Language Processing Components with GATE Version 3 (a User Guide) http://gate.ac.uk/sale/tao/index.html#x1-1710008.4,".
8. DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., and Cooper, H. (2003) Cues to deception, *Psychological Bulletin*, 129, 1, 74-118.
9. Ekman, P. (1985) Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage, WW Norton & Company, New York.
10. Ekman, P., and Friesen, W.V. (1969) Nonverbal leakage and clues to deception, *Psychiatry: Journal for the Study of Interpersonal Processes*, 32, 88-105.
11. Ekman, P., and Friesen, W.V. (1974) Detecting deception from the body or face, *Journal of Personality and Social Psychology*, 29, 3, 288-298.
12. Ekman, P., O'Sullivan, M., and Frank, M.G. (1999) A few can catch a liar, *Psychological Science*, 10, 3, 263-265.
13. George, J.F, and Keane, B.T., (2006) Deception detection by third party observers, Deception Detection Symposium, 39th Annual Hawaii International Conference on System Sciences.
14. Hancock, J., Curry, L., Goorha, S., and Woodworth, M. (2004) Lies in conversation: An examination of deception using automated linguistic analysis, in Kenneth Forbus, Dedre Gentner and Terry Regier (Eds.) *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society*, August 4-7, Chicago, Illinois, 534-540.
15. Hancock, J., Thom-Santelli, J., and Ritchie, T. (2004) Deception and design: the impact of communication technology on lying behavior, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, April 24-29, Vienna, Austria, ACM Press, 129-134.
16. Hancock, J.T., Curry, L., Goorha, S., and Woodworth, M. (2005) Automated linguistic analysis of deceptive and truthful synchronous computer-mediated communication, in *Proceedings of the Thirty-Eighth Annual Hawaii International Conference on System Sciences (HICSS '05)*, January 3-6, 2005, 22c-31c.
17. Hepple, M. (2000) Independence and commitment: Assumptions for rapid training and execution of rule-based POS taggers, in *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, October 2000, Hong Kong, 278-285.
18. Johnson, M.K., and Raye, C.L. (1981) Reality monitoring, *Psychological Review*, 88, 1, 67-85.
19. McCornack, S.A. (1992) Information manipulation theory, *Communication Monographs*, 59, 1, 1-16.
20. Meservy, T.O., Jensen, M.L., Kruse, J., Burgoon, J.K., Nunamaker, J.F., Jr., Twitchell, D.P., Tsechpenakis, G., and Metaxas, D.N. (2005) Deception detection through automatic, unobtrusive analysis of nonverbal behavior, *IEEE Intelligent Systems*, 20, 5, 36-43.
21. Newman, M.L., Pennebaker, J.W., Berry, D.S., and Richards, J.M. (2003) Lying words: predicting deception from linguistic styles, *Personality and Social Psychology Bulletin*, 29, 5, 665-675.
22. Pennebaker, J.W., and Francis, M.E.(2001) Linguistic inquiry and word count: LIWC 2001, Erlbaum Publishers, Mahwah, NJ.
23. Qin, T.T., Burgoon, J., Blair, J.P., and Nunamaker, J.F., Jr. (2005) Modality effects in deception detection and applications in automatic-deception-detection, *Proceedings of the Thirty-Eighth Annual Hawaii International Conference on System Sciences (HICSS '05)*, January 3-6, 2005, 23b.

24. Qin, T.T., Burgoon, J., and Nunamaker, J.F., Jr. (2004) An exploratory study on promising cues in deception detection and application of decision tree in *Proceedings of the Thirty-Seventh Annual Hawaii International Conference on System Sciences (HICSS '04.)*

25. Twitchell, Jensen, M.L., Burgoon, J.K., and Nunamaker, J.F., Jr. (2004) Detecting deception in secondary screening interviews using linguistic analysis, in *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*, October 3-6, 2004, 118-123.

26. Vrij, A. (2000) Detecting lies and deceit: The psychology of lying and the implications for professional practice, John Wiley & Sons, New York.

27. Vrij, A. (2005) Criteria-based content analysis: A qualitative review of the first 37 studies, *Psychology Public Policy and Law*, 11, 1, 3-41.

28. Vrij, A., Edward, K., Roberts, K.P., and Bull, R. (2000) Detecting deceit via analysis of verbal and nonverbal behavior, *Journal of Nonverbal Behavior*, 24, 4, 239-263.

29. Zhou, L., Burgoon, J., and Twitchell, D. P. (2003) A longitudinal analysis of language behavior of deception in e-mail, in H. Chen, R. Miranda, D. Zeng, T. Madhusudan, C. Demchak, and J. Schroeder (Eds.) *Proceedings of the First NSF/NIJ Symposium on Intelligence and Security Informatics (ISI 2003), Lecture Notes in Computer Science (LNCS 2665),*June 2-3, 2003, Tucson, AZ, Springer-Verlag, 102-110.

30. Zhou, L., Burgoon, J.K., Nunamaker, J.F., Jr.., and Twitchell, D.P. (2004) Automated linguistics based cues for detecting deception in text-based asynchronous computer-mediated communication: An empirical investigation, *Group Decision and Negotiation*, 13, 1, 81-106.

31. Zhou, L., Burgoon, J.K., Twitchell, D.P., Qin, T.T., and Nunamaker, J.F., Jr.- (2004) A comparison of classification methods for predicting deception in computer-mediated communication, *Journal of Management Information Systems,* 20, 4, 139-165.

32. Zhou, L., Twitchell, D.P., Qin, T., Burgoon, J.K., and Nunamaker, J.F., Jr. (2003) An exploratory study into deception detection in text-based computer-mediated communication, in *Proceedings of the Thirty-Sixth Annual Hawaii International Conference on System Sciences (HICSS '03)*, Big Island, Hawaii.

33. Zuckerman, M., and Driver, R.E. (1985) Telling lies: verbal and nonverbal correlates of deception, in: *Multichannel intergration of nonverbal behavior*, Erlbaum, Hillsdale, NJ.