Individual Differences in Deception and Deception Detection

Sarah Ita Levitan, Michelle Levine, Julia Hirschberg
Dept. of Computer Science
Columbia University
New York NY, USA
{sarahita,mlevine,julia}@cs.columbia.edu

Nishmar Cestero
Dept. of Psychology
Boston University
Boston MA, USA
nishmarc@bu.edu

Guozhen An, Andrew Rosenberg
Dept. of Computer Science
Queens College, CUNY
Queens NY, USA
{gan@gc,andrew@cs.qc}.cuny.edu

Abstract— **We are building a new corpus of deceptive and non-deceptive speech, using American English and Mandarin Chinese adult native speakers, to investigate individual and cultural differences in acoustic, prosodic, and lexical cues to deception. Here, we report on the role of personality factors using the NEO-FFI (Neuroticism-Extraversion-Openness Five Factor Inventory), gender, ethnicity and confidence ratings on subjects' ability to deceive and to detect deception in others. We report significant correlations for each factor with one or more aspects of deception. These are important for the study of trust, cognition, and multi-modal information processing.**

Keywords-Deception; cross-linguistic; personality.

I. INTRODUCTION

Finding new methods for detecting deception is a major goal of researchers in psychology and computational linguistics as well as commercial, law enforcement, military, and intelligence agencies. While many new techniques and technologies have been proposed and some have even been fielded, there have been few significant successes. The goal of our research is to develop techniques to identify deceptive communication in spoken dialogue. As part of this investigation, we are focusing on how within-culture and cross-cultural differences between deceivers as well as their common characteristics impact deceptive speech behavior. Our research focuses solely on cues drawn from the speech signal, which have been little studied.

In this paper, we describe results of experiments correlating gender, ethnicity, and personality characteristics from the NEO-FFI Five Factor Analysis [1] with subjects' ability to deceive and to judge deception in others' speech. We also examine the importance of subjects' reported confidence in their judgments in deception production and detection. In Section 2, we describe previous work on cues to deception and deception detection. In Section 3 we discuss our experimental design, data collection and annotation. In Section 4 we describe results of our correlations of personality, gender, ethnicity, and confidence on deception production and detection. We conclude in Section 5 with a discussion of our results.

II. DECEPTION DECTECTION

Previous research on deceptive behavior has studied standard biometric indicators commonly measured in polygraphy (cardiovascular, electrodermal, and respiratory), facial expression, body gestures, brain imaging, body odor, and lexical and acoustic-prosodic information.

Biometric measures are widely acknowledged, even by polygraphers, to be inadequate for deception detection, performing at no better than chance. Useful groundwork has been laid in identifying potential facial expression cues of deception by Ekman et al. [2][3]. However, attempts to identify deception from facial expressions are questioned by some researchers [4]. Moreover these approaches are difficult to automate, requiring delicate image capture technology and laborious human annotation. There have been promising results using automatic capture of body gestures as cues to deception [5][6], but this method requires multiple, high-caliber cameras to capture movements reliably. Similarly, the use of brain imaging technologies for deception detection is still in its infancy [7] and these require the use of MRI techniques, which are not practical for general use. Additional biometric indicators of deception such as body odor are beginning to be investigated [8] but these studies, like brain imaging, are in very early stages.

Some researchers and practitioners have examined language-based cues to deception. These include Statement Analysis [8], SCAN [10][11], and some of the text-based signals identified by John Reid and Associates [12]. These efforts have been popular among law enforcement and military personnel, though little tested scientifically (although Bachenko et al. [13] have partially automated and validated some features used in Statement Analysis). Other lexical cues to deception have been developed and tested empirically by Pennebaker and colleagues [14][15] and by Hancock et al. [16]. There has also been research focused on lexical cues to deception in written online communication [17][18].

Little work has been done on cues to deception drawn from the speech signal. Simple features such as intensity and hypothesized vocal tremors have performed poorly in objective tests [19][10][21][22], although other features examined by Harnsberger et al. [23] and Torres et al. [24] have had more success. In previous work on deception in American speech, Hirschberg et al. [25] developed automatic deception detection procedures trained on spoken cues and tested on unseen data. These procedures have achieved accuracies 20% better than human judges. In the process of identifying common characteristics of deceivers, they also

 ISBN: 978-1-61208-390-2

noticed a range of individual differences in deceptive behavior, e.g., some subjects raised their pitch when lying, while some lowered it significantly; some tended to laugh when deceiving, while others laughed more while telling the truth. They also discovered that human judges' accuracy in judging deception could be predicted from their scores on the NEO-FFI, suggesting that such simple personality tests might also provide useful information in predicting individual differences in deceptive behavior itself [26].

Differences in verbal deceptive behavior in different cultures have been identified by several researchers [27][28]. Studies of deceptive behavior in non-Western cultures have primarily focused on understanding how culture affects *when* people deceive and *what* they consider deception [29][30]. Studies investigating the universality of deceptive behavior have found that, while stereotypes may exist [31] these may not correlate with actual deceptive behavior [32][33] and that culture-specific deception cues do exist [27][28][34].

In the work presented here, we investigate both the ability to deceive and to detect deception considering gender and ethnicity and examining new cues to deception: features extracted from the NEO-FFI personality inventory [1] and subjects' reported confidence in their abilities.

A. Experimental Design

To investigate questions of individual and cross-cultural differences in deception perception and production, we are collecting a large corpus of cross-cultural deceptive and non-deceptive speech. We employ a variant of the 'fake resume' paradigm to elicit both deceptive and non-deceptive speech from native speakers of Standard American English (SAE) and Mandarin Chinese (MC), both speaking in English. Each conversation in the corpus is between a pair of subjects who are not previously acquainted with one another. To date, the corpus includes 134 conversations between 268 subjects.

For the first phase of each session, subjects are separated from one another. Each is told that they will play a lying game with another subject, in which they will alternate between interviewing their partner and being interviewed themselves. As interviewees, they should attempt to successfully deceive the interviewer. As interviewers, they should attempt to determine whether the interviewee is lying or telling the truth. For motivation, they are told that their compensation depends on their ability to deceive while being interviewed, and to judge correctly while interviewing. As interviewer, they receive $1 each time they correctly identify an interviewee's answer as either lie or truth and lose $1 for each incorrect judgment. As interviewee, they earn $1 each time their lie is judged to be true, and lose $1 each time their lie is correctly judged to be a lie by the interviewer.

Subjects are then asked to truthfully complete a 24-item biographical questionnaire. In addition to their true answers, they are told to create a false answer for a random half of the questions. They are given guidelines to ensure that their false answer differed significantly from the truth, to ensure that lying will not be too easy. For example, for the question "Where were you born," the false answer must be a place that the subject has never visited, a false answer to "What is your father's occupation" must be different from their mother's true occupation, and so on. Before the interviews begin, false answers are checked by an experimenter to make sure subjects follow these guidelines. In addition to the biographical questionnaire, each subject completes the NEO-FFI personality inventory [1], which is described below.

While one subject is completing the NEO-FFI inventory, we collect a 3-4 minute baseline sample of speech from the other participant for use in speaker normalization. The experimenter elicits natural speech by asking the subject open-ended questions (e.g., "What do you like best/worst about living in NYC?"). Subjects are instructed to be truthful during this part of the experiment. Once both subjects have completed all the questionnaires and we have collected baseline samples of speech, the lying game begins.

The lying game takes place in a sound booth where the subjects are seated across from each other, separated by a curtain so that there is no visual contact; this is necessary since our focus is on spoken and not visual cues. There are two parts to each session. During the first half, one subject acts as the interviewer while the other answers the biographical questions, lying for half and telling the truth for the other half, based on the modified questionnaire. In the second part of the session, the subjects switch roles. All speech data is collected in a double-walled sound booth in the Columbia Speech Lab and recorded to digital audio tape on two channels using Crown CM311A Differoid head-worn close-talking microphones.

The interviewer is able to ask the questions in any order s/he chooses, and is encouraged to ask follow-up questions to help determine the truth of the interviewee's answers. For each question, the interviewer records his/her judgment, along with a confidence score from 1-5. As the interviewee answers the questions, s/he presses a T or F key on a keyboard (which the interviewer cannot see) for each phrase, logging each segment of speech as true or false. Thus, while the biographical questionnaire provides the 'global' truth value for the answer to the question asked, the key log provides the 'local truth' value for each phrase, which is automatically aligned with each speech segment. At the end of the experiment, subjects complete a brief questionnaire, which includes additional confidence questions.

B. Personality Assessment

The NEO-FFI personality assessment [1] is based on the five-factor model of personality, an empirically-derived and comprehensive taxonomy of personality traits. It was developed by applying factor analysis to thousands of descriptive terms found in a standard English dictionary. It is used to assess the five personality dimensions of:

Openness to Experience. Designed to capture imagination, aesthetic sensitivity, and intellectual curiosity. It is "related to aspects of intelligence, such as divergent thinking, that contribute to creativity" [1]. Those who score low on this dimension prefer the familiar and tend to behave more conventionally. People high in Openness are "willing to entertain novel ideas and unconventional values" [1].

 ISBN: 978-1-61208-390-2

Conscientiousness. Addresses individual differences in self-control, such as the ability to control impulses, but also to plan and carry out tasks. It measures contrasts between determination, organization, and self-discipline and laxness, disorganization, and carelessness.

Extraversion. Meant to capture proclivity for interpersonal interactions, and variation in sociability. It reflects contrasts between those who are reserved vs. outgoing, quiet vs. talkative, and active vs. retiring.

Agreeableness. Measures interpersonal tendencies and is intended to assess an individual's fundamental altruism. Individuals high in Agreeableness are sympathetic to others and expect that others feel similarly.

Neuroticism. Contrasts emotional stability with maladjustment. It is intended to capture differences between those prone to worry vs. calm, emotional vs. unemotional behavior, and vulnerable vs. hardy.

III. ANALYSES AND RESULTS

Although subjects were instructed to lie in response to 12 of the questions, 55 out of 268 subjects did not follow these instructions, and lied in response to more or fewer than 12 questions. The following analyses include 126 pairs, those in which both subjects lied in response to 10-14 of the questions; this restriction ensures that roughly equal amounts of truthful and deceptive speech are available for each subject. This subsample consists of 142 native SAE participants (88 females, 54 males) and 110 native MC participants (69 females, 41 males). Our eventual goal is a corpus balanced for gender and ethnicity, but in this paper we present results only on this sample.

First, we examined how accurately subjects could identify deception in their partners during the lying game. Prior research indicates that human judges perform worse than chance at detecting deception [26][35]. However, in our study subjects correctly identified question responses as truthful or deceptive at a greater than chance level. They were accurate 56.75% of the time (compared to the chance baseline of 49.55%).

To further assess subjects' accuracy, we explored how well subjects detected lies as opposed to truths. To account for the different number of lies across subjects, for each subject we calculated ratio scores for: number of successful global lies to the number of global lies told (*successful lies*); the number of successful lie detections to the number of global lies told (*successful lie detections*); the number of successful truth detections to the number of truths told (*successful truth detections*). Results indicate that people successfully deceived their partner 51.83% of the time. Deceptive answers were correctly identified 48.16% of the time and truthful answers were correctly identified 65.20% of the time.

We investigated whether subjects' ability to detect deception was correlated with their ability to deceive by comparing *successful lies* to *successful lie detections*. Our data indicate that subjects who were better at detecting deceptive answers were also better at deceiving, $r(252) = 0.13$, $p = 0.04$. When separated by gender and native language, it becomes apparent that this correlation is strongest for females, and specifically for SAE females ($r(157) = 0.24$, $p = 0.003$ and $r(88) = 0.29$, $p = 0.005$). We note that, for all subjects, those who were better at detecting deception were also more likely to label their partners' answers as untrue --- whether or not their partner did indeed lie, $r(252) = 0.69$, $p < 0.001$. However, female subjects who were more likely to label their partners' answers lies *were* also better at deceiving, $r(157) = 0.18$, $p = 0.02$.

Next, we examined how individual differences in gender, culture, personality, and confidence ratings interacted with successful deception and deception detection. Independent sample t-tests indicated no effect of subjects' gender or native language on their ability to deceive. In addition, correlational analyses showed no effect of personality factors on subjects' ability to detect deception. This latter finding is in sharp contrast with Enos et al.'s findings for personality differences in success rates of post hoc judges of deception [24] and suggests that personality factors may play a more important role when non-conversational participants rather than those engaged in the conversation are judging deception. In contrast, the personality factor of Extraversion does correlate with subjects' ability to deceive and here we do find cultural and gender differences: MC females' success positively correlates with Extraversion scores ($r(69) = 0.26$, $p = .03$) while SAE males' success *negatively* correlates with their Extraversion scores ($r(54) = -0.36$, $p = .01$). Furthermore, SAE females' deception ability *negatively* correlates with their Conscientiousness scores ($r(86) = -0.22$, $p = .04$).

For confidence ratings, we also find a gender difference: overall, female subjects' ability to detect deception *negatively* correlates with their average confidence in their judgments, $r(157) = -0.20$, $p = 0.01$. This did not hold true for SAE females examined separately although it did for MC females, $r(69) = -0.26$, $p = 0.03$. We hypothesize that interviewers who are less confident in their judgments may ask more follow-up questions and thus obtain more evidence to determine deception. It will be important to look at answer length and number of follow-up questions to test these possibilities. We also found that, for females, average confidence in detecting deception *negatively* correlated with Neuroticism, $r(155) = -0.16$, $p = 0.05$. Not surprisingly, women who are less "neurotic" are more confident in their deception judgments. We will need to check for similar findings for male subjects once we have collected more data.

Finally, we looked at whether the gender and culture of subjects' partners played a role in deception and deception detection. Independent t-tests show no effects so far.

IV. CONCLUSIONS AND FUTURE WORK

Preliminary analysis of a sample of our deceptive speech corpus shows some promising results: We found that subjects who are better at detecting lies are also better at deceiving others, and that this correlation is stronger for females and stronger still for SAE females. While we have not found effects of personality characteristics on our subjects' ability to detect deception, in contrast to Enos et al.

 ISBN: 978-1-61208-390-2

[26], we have found that Extraversion and Conscientiousness scores correlate with ability to deceive, although the direction of this effect differs depending upon gender and ethnicity. We note the difference between the judgment tasks in our experiment vs. [26]'s. Finally, we found that MC women showed a negative correlation between confidence scores and ability to detect deception while SAE women and men in general did not. In addition, for all females, ability to detect deception was negatively correlated with Neuroticism.

We anticipate that the completion of a balanced corpus will clarify and expand some of these findings. We will also finish transcribing our corpus and aligning the transcription with the speech recordings so that we can add acoustic, prosodic, and lexical cues to gender, ethnicity, and personality information for the purpose of building automatic classifiers for deceptive vs. non-deceptive speech.

V. ACKNOWLEDGMENTS

We thank the following students for their contributions to this study: Zoe Baker-Peng, Ling Huang, Melissa Kaufman-Gomez, Yvonne Missry, Elizabeth Petitti, Sarah Roth, Molly Scott, Jennifer Senior, Grace Ulinski, and Christine Wang. This work was partially funded by AFOSR FA9550-11-1-0120.

REFERENCES

[1] P. T. Costa and R. R. McCrae, "Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI)." Odessa, FL: Psychological Assessment Resources, 1992.

[2] P. Ekman, M. Sullivan, W. Friesen, and K. Scherer, "Face, voice, and body in detecting deception," Journal of Nonverbal Behaviour, vol. 15(2), 1991, pp. 125-135..

[3] M. Frank, M. O'Sullivan, and M. Menasco, "Human behavior and deception detection," in J. G. Voeller (Ed.), Wiley Handbook of Science and Technology for Homeland Security, New York: John Wiley & Sons, 2008.

[4] J. Cohn quoted in http://abcnews.go.com/GMA/HealthyLiving/autism-research-benefit-studying-babies-facial-recognition-experts/story?id=9244817&page=3, 12/04/2009.

[5] T. Qin, J. Burgoon, and J. Nunamaker, "An exploratory study on promising cues in deception detection and application of decision tree." Hawaii International Conference on System Sciences, 2004, pp. 23-32.

[6] T. Meservy et al., "Deception Detection through Automatic, Unobtrusive Analysis of Nonverbal Behavior." IEEE Intelligent Systems, 20:5 (September 2005), pp. 36-43.

[7] D. Langleben et al., "Telling truth from lie in individual subjects with fast event-related fMRI." Human Brain Mapping, 2005, 26(4): 262-72.

[8] S. Waterman, "DHS wants to use human body odor as biometric identifier, clue to deception," United Press International (UPI). http://www.upi.com/Top_News/Special/2009/03/09/DHS-wants-to-use-human-body-odor-as-biometric-identifier-clue-to-deception/UPI-20121236627329/, 3/9/2009.

[9] S. Adams, "Statement analysis: What do suspects' words really reveal?" FBI Law Enforcement Bulletin, October 1996.

[10] A. Sapir, "Scientific Content Analysis (SCAN)." Laboratory of Scientific Interrogation. Phoenix, AZ, 1987.

[11] N. Smith, "Reading between the lines: An evaluation of the scientific content analysis technique (SCAN)," Police Research Series. London, UK, 2001.

[12] J. E. Reid and Associates, "The Reid Technique of Interviewing and Interrogation," Reid, John E. and Associates, Inc., 2000.

[13] J. Bachenko, E. Fitzpatrick, and M. Schonwetter, "Verification and implementation of language-based deception indicators in civil and criminal," International Conference on Computational Linguistics, vol. 1, Manchester, 2008, pp. 41-48.

[14] Pennebaker, M. Francis, and R. Booth, "Linguistic Inquiry and Word Count." Erlbaum Publishers, Mahwah, NJ, 2001.

[15] M. Newman, J. Pennebaker, D. Berry, and J. Richards, "Lying words: Predicting deception from linguistic style." Personality and Social Psychology Bulletin, 2003, 29:665-675.

[16] J. Hancock, L. Curry, S. Goorha, and M. Woodworth, "On lying and being lied to: A linguistic analysis of deception." Discourse Processes, vol. 45, pp. 1-23, 2008.

[17] L. Zhou and D. Zhang, "Following Linguistic Footprints: Automatic Deception Detection in Online Communication." Communications of the ACM, 2008, 51(9): 119-122.

[18] R. Mihalcea and C. Strapparava, "The lie detector: explorations in the automatic recognition of deceptive language." In Proceedings of the ACL-IJCNLP 2009. Stroudsburg, PA.

[19] D. Haddad and R. Ratley, "Investigation and evaluation of voice stress analysis technology." http://www.ncjrs.org/pdffiles1/nij/193832.pdf, March 2002.

[20] H. Hollien and J. Harnsberger, "Voice stress analyzer instrumentation evaluation." Technical report, Counterintelligence Field Activity. 2006.

[21] H. Hollien, J. Harnsberger, C. Martin, and K Hollien, "Evaluation of the NITV CVSA." Journal of Forensic Sciences, vol. 53, pp. 183 – 193, 2006.

[22] A. Eriksson and F. Lacerda, "Charlatanry in forensic speech science: A problem to be taken seriously." The International Journal of Speech, Language and the Law, vol. 14:2, pp. 169-193, 2007, http://www.scribd.com/doc/9673590/Eriksson-Lacerda-2007.

[23] J. Harnsberger, H. Hollien, C. Martin, and K. Hollien, "Stress and deception in speech: evaluating layered voice analysis." Journal of Forensic Sciences, vol. 54(3), pp. 642-50, 2009.

[24] J. Torres, E. Moore, and E. Bryant, "A Study of Glottal Waveform Features for Deceptive Speech Classification." ICASSP 2008, Las Vegas.

[25] J. Hirschberg et al., "Distinguishing deceptive from non-deceptive speech." Interspeech 2005. Lisbon.

[26] F. Enos, S. Benus, R. Cautin, M. Graciarena, J. Hirschberg, and E. Shriberg, "Personality factors in human deception detection: Comparing human to machine performance." Interspeech 2006. Pittsburgh.

[27] F. Feldman, "Nonverbal disclosure of deception in urban Koreans." Journal of Cross-Cultural Psychology, vol. 10, 1979, pp. 215-221.

[28] M. Cody, W. Lee, and E. Y. Chao, "Telling lies: Corellates of deception among Chinese." In J.P. Forgas and J. M. Innes, eds. Recent Advances in Social Psychology: An International Perscpective, 1989, pp.359-568.

[29] M. Lapinski and T. Levine, "Culture and information manipulation theory: The effects of self-construal and locus of benefit on information manipulation." Communication studies, vol. 51(1), 2000, pp. 55-73.

[30] J. Seiter, J. Bruschke, and C. Bai, "The acceptability of deception as a function of perceivers' culture, deceiver's

 ISBN: 978-1-61208-390-2

intention, and deceiver-deceived relationship." Western Journal of Communication 66(2), 2002, pp.158-180.

[31] C. Bond and The Global Deception Research Team, "A world of lies." Journal of Cross-Cultural Psychology, vol. 37(1), 2006, pp. 60-74.

[32] A. Vrij and G. Semin, "Lie experts' beliefs about nonverbal indicators of deception." Journal of Nonverbal Behavior, vol. 20(1), 1996, pp. 65-80.

[33] M. Zuckerman, B. DePaulo, and R. Rosenthal, "Verbal and non-verbal communication of deception." In L. Berkowitz (ed.), Advances in Experimental Social Psychology, Academic Press, New York, pp.1-59, 1981.

[34] R. Feldman, L. Jenkins, and O. Popoola, "Detection of deception in adults and children via facial expressions." Child development, 350-355,1979).

[35] M. Aamondt and H. Custer, "Who can best catch a liar?" Forensic Examiner 15(1):6-11. 2006.

 ISBN: 978-1-61208-390-2