Reading between the Lines: Linguistic Cues to Deception in Online Dating Profiles

Catalina L. Toma & Jeffrey T. Hancock
Department of Communication
Cornell University
Ithaca, NY 14850, USA
catalina.toma@cornell.edu, jeff.hancock@cornell.edu

ABSTRACT

This study investigates whether deception in online dating profiles is detectable through a linguistic approach, which assumes that liars nonconsciously produce different word patterns than truth-tellers. We objectively measure deception in online dating profiles and analyze the linguistic composition of the open-ended component of the profile (i.e., "about me" section) using computerized text analysis. Results show that profile deceptions correlate with fewer self-references, increased negations, fewer negative emotion words and fewer overall words used in the textual self-description. Results are discussed in terms of (1) practical implications for detecting deception in online profiles; and (2) theoretical implications regarding the impact of media affordances (i.e., asynchronicity and editability) on the occurrence of linguistic cues to deception.

Author Keywords

Deception, linguistic cues to deception, online dating, social networking sites.

ACM Classification Keywords

J4 Social and behavioral sciences: Psychology

General Terms

Theory

INTRODUCTION

Far from being a crutch for the desperate, online dating is now one of the most frequently used services on the Internet. Despite its popularity, online dating is plagued by concerns about deception, with many users characterizing it as a "leap of faith" and expressing worry that others misrepresent themselves [7]. An important question, then, is whether deception is detectable in online dating profiles *before* meeting potential mates face-to-face. Does the online profile itself provide clues about the veracity of the self-presentation?

Here we propose a linguistic approach to assessing deception in online dating profiles. An emerging body of research [e.g., 2, 3] has shown that liars often use words differently than truth-tellers. This approach uses computerized text analysis to distinguish between deceptive and truthful messages based on these differential word patterns. Since online dating profiles typically contain a textual component where daters describe themselves in their own words, we examine whether this textual self-description provides an indication of the amount of deception present in the profile.

To date, the linguistic approach to deception has been used (1) to differentiate between messages that *contain* deception and messages that do not; and (2) in laboratory settings, where conversation partners composed messages ad hoc. We expand on this paradigm in several important ways. First, we examine whether lies told as part of the overall profile self-presentation (e.g., lies about height or age) rather than lies contained solely in the textual self-description, are related to linguistic changes in this textual self-description.

Second, we examine whether the linguistic approach to deception is useful when liars have ample opportunity to *control* their messages. Due to the technological affordances of asynchronicity and editability [6], online daters have an unlimited amount of time to compose their self-presentation and are also able to revise it in order to make it believable. Are linguistic cues to deception produced even when liars are able to carefully monitor what they say?

Lastly, we examine deceptions that occur in natural environments, where there is a strong incentive to avoid being caught lying. Because lies can have catastrophic consequences for relationship development [7], recent research shows that online daters lie strategically in their profiles. Men tend to misrepresent their height and social status indicators, and women tend to misrepresent their weight and photographs, but these lies are small in magnitude [5]. Do these small and strategic deceptions, composed by people who are highly motivated to avoid detection, result in linguistic cues?

Linguistic cues to deception

The linguistic approach to deception assumes that the emotions and cognitions experienced by liars are reflected through the nonconscious production of certain word types.

Can these linguistic cues be used to differentiate highly deceptive profiles from more honest ones?

Emotional linguistic cues
The act of lying is typically associated with a range of negative emotions, such as anxiety, shame and guilt. These emotions arise because lying is socially undesirable and associated with a sense of failing moral standards [1]. This spectrum of emotions has been shown to manifest itself linguistically in two ways. First, these emotions can reveal themselves directly through an increase in negative emotion words (e.g., "hate," "sorry," "worthless") [2, 3]. Second, they can be revealed through liars' unconscious efforts to distance themselves from them. Psychological distancing is a strategy meant to reduce the discomfort caused by deception, and manifests itself linguistically through a decrease in self-references (e.g., "I," "me") and an increase in negations (e.g., "no," "not," "never") [3]. Given these findings, we hypothesize that:

H1: Highly deceptive online dating profiles will have fewer self-references but more negations and negative emotion words than less deceptive profiles.

Cognitive linguistic cues
The act of lying is also cognitively difficult because it involves fabricating information [1]. This cognitive difficulty can manifest itself linguistically through a decrease in exclusive words and an increase in motion words. Exclusive words, such as "except," "without," "but," differentiate between what belongs to a category and what does not, a cognitively demanding task. By contrast, motion words such as "walk," "move," and "go" represent simple, concrete actions that are easy to string together [2, 3].

An additional marker of cognitive load is the production of fewer overall words [3]. This occurs because the cognitive demands of deception make it easier to manage information and avoid contradictions by saying less. Together, these findings suggest that:

H2: Highly deceptive profiles will have fewer exclusive words and increased motion words, but a lower overall word count than less deceptive profiles.

Effects of media affordances on cues
As discussed earlier, online dating profiles allow users increased control over the production of their messages through the affordances of asynchronicity and editability [5, 6]. These affordances should attenuate the cognitive burden of deception, because they allow liars as much time as they need to construct a believable self-presentation. However, asynchronicity and editability should not affect the negative emotions associated with deception. Thus, we hypothesize that:

H3: Emotionally-related linguistic cues to deception should account for more variance in deception scores than cognitively-related linguistic cues in online dating profiles.

METHOD

Participants and recruitment
Participants were 80 online daters (40 men and 40 women; age $M = 30.55$, $SD = 8.46$, $min = 18$, $max = 53$) with profiles in one of four online dating portals: Match.com, Yahoo Personals, American Singles or Webdate (see [5] for additional information). These services were selected because they are mainstream, widely popular and require users to create detailed self-presentations, which include open-ended self-descriptions.

Participants were recruited through print and online advertisements in the New York City area. The advertisements called for participation in a "self-presentation" study and did not mention deception. Participants signed up through a secure website, where they indicated their username and the service they used. This information was used to limit participants to those over the age of 18 and heterosexual daters. Using these criteria, 251 online daters were invited to participate, 80 of whom came to their appointments and were included in the study.

Procedure
The study was conducted at New School University in New York City. A copy of participants' profile was printed and archived prior to their arrival at the lab. During the research appointment, participants were given their profile print-out and asked to rate the accuracy of their responses on each profile element. Then, participants' height, weight and age were measured by the researcher. Participants were paid $30 for their time.

Measures

Deception index
To assess the extent to which participants lied in their profiles, we used objective measurements of deception, which are preferable to self-report because they are not biased by socially desirable responding (i.e., participants feeling ashamed to admit the true extent of their lies). Three profile elements lend themselves to direct measurement: 1) height, which was measured using a standard measuring tape, 2) weight, which was measured using a standard scale; and 3) age, which was recorded from participants' driver licenses. Absolute deviations from the truth were calculated by subtracting observed measurements from profile statements. These deviations were standardized and then averaged in order to calculate a *deception index*, which we use as an objectively derived measure of the amount of deception present in participants' profiles.

Accuracy of textual self-descriptions
Because it is impossible to objectively assess the accuracy of participants' textual self-description, self-report measures were used for this item. Participants rated the accuracy of the self-description on a scale from 1 (completely inaccurate) to 5 (completely accurate). Accuracy was defined as "the extent to which this information reflects the truth about you now."

Linguistic measures
The textual self-description of every profile was analyzed using LIWC2007 [4]. LIWC is a text analysis software that determines the linguistic composition of transcripts. LIWC compares each word in the transcript with its internal dictionary of 4,500 words and then assigns it to one or several of its 76 word categories. These categories include function words (e.g., articles, negations), psychological processes (e.g., negative emotions, cognitive processes) and personal concerns (e.g., work, home, religion). LIWC has been successfully used to predict numerous psychological outcomes, including deception [e. g., 2, 3].

Each self-description was converted to a text file and run through LIWC. For each self-description, LIWC produced an output indicating the word frequency for each category (e.g., first-person pronouns, negations, negative emotion words, motion words and exclusive words). Word frequencies are expressed as a percentage of the total number of words contained in each file (see Table 1).

RESULTS AND DISCUSSION
LIWC captured 87.89% of the words in daters' textual self-descriptions. On average, the self-descriptions were 156.16 words long ($SD = 118.54$), with no gender differences in length.

Participants rated their self-descriptions as very accurate. On the 1 (extremely inaccurate) to 5 (extremely accurate) scale used, self-descriptions were rated as 4.79 ($SD = 0.41$, $min = 4.00$, $max = 5.00$), suggesting that daters considered them to be almost free of deceptions.

To assess whether profile lies were correlated with linguistic changes in the open-ended self-descriptions, we built separate regression models for the emotional and cognitive indicators of deception, and an overall model containing both types of indicators. None of the indicators were correlated with each other, indicating that multicollinearity was not a problem.

Dimension	Example	Mean	SD
Word count		156.16	118.54
1st-person singular pronouns	I, my, me, mine	7.99%	4.07%
Negations	No, never, not	1.44%	1.36%
Negative emotions	Hate, hurt, ugly	1.15%	1.22%
Exclusive words	But, without,	2.94%	1.97%
Motion words	Walk, move, go	1.69%	1.55%

Table 1. Percentages of words in LIWC categories.

For the emotional indicators, a regression model was built with the deception index as the dependent variable and first-person singular pronouns, negations, and negative emotion words as predictors. The model fit the data well [$F(3, 74) = 5.67$, $p < 0.001$] and explained 15% of the variance in the deception index [$R = 0.43$, $R^2 = 0.19$, $R^2_{adj} = 0.15$]. The standardized coefficients for all the predictor variables were

significant (see Table 2), providing support for H1. However, the coefficient for negative emotion words was in the opposite direction than predicted.

For cognitive indicators, a regression model was built with the deception index as the dependent variable and exclusive words, motion words and word count as predictors. The model did not fit the data well [$F(3, 74) = 1,37$, ns], as neither exclusive words nor motion words were significant predictors. The model was revised by eliminating the non-significant predictors. The revised model was a good fit [$F(1, 76) = 4.19$, $p = 0.04$] and explained 4% of the variance in the dependent variable [$R = 0.23$, $R^2 = 0.05$, $R^2_{adj} = 0.04$] suggesting that word count was the only cognitive-related variable to correlate with deception (see Table 2). H2 was thus only partially supported.

The above models suggest that emotional indicators were more powerful in predicting deception than the cognitive ones. To test this prediction further, a combined model containing both the emotional and cognitive indicators of deception was built. The combined model fit the data well, [$F(4, 73) = 6.72$, $p < 0.001$] and explained 23% of the variance in the deception index [$R = 0.52$, $R^2 = 0.27$, $R^2_{adj} = 0.23$]. The pattern of coefficients from the combined model supports H3. All of the hypothesized emotional cues were significant predictors of the deception index, but the only reliable cognitive cue was word count (see Table 2).

	LIWC category	Std. β	p
Emotional cues			
	I-pronouns	-0.254	0.02
	Negations	0.281	0.01
	Neg. emotions	-0.296	0.008
Cognitive cues			
Original model			
	Word count	-0.228	0.06
	Exclusive words	0.005	0.97
	Motion words	0.024	0.84
Revised model			
	Word count	-0.228	0.04
Overall model			
	Word count	-0.291	0.005
	I-pronouns	-0.279	0.008
	Negations	0.321	0.003
	Neg. emotions	-0.293	0.006

Table 2. Standardized regression coefficients for linguistic indicators of deception.

To summarize, this study investigated whether deception in online dating profiles is associated with linguistic changes in the textual description. We examined whether cues related to the emotional and cognitive aspects of deception would emerge in the unique media context of online dating profiles, where (1) lies tend to occur as part of the overall self-presentation rather than in the textual self-description; (2)

liars have the opportunity to closely monitor their written statements due to the affordances of asynchronicity and editability; and (3) lies are small, because daters are strongly motivated to avoid deception detection. Importantly, online dating lies have been shown to be intentional, and not simply the result of oversights or lack of self-awareness (i.e., daters not knowing their precise height or weight measurements) [5].

Results show that profile lies did correlate with changes in the way online daters wrote about themselves in their open-ended self-descriptions, although the self-descriptions themselves were mostly accurate. Specifically, liars psychologically distanced themselves from their deceptions by producing fewer self-references and more negations. They also wrote shorter self-descriptions, presumably in an effort to avoid contradictions with prior profile statements.

A surprising result was that liars produced fewer, rather than more, negative emotion words. This could be due to the fact that people who lied more were more eager to make a good impression, and thus avoided sounding negative – which is usually a turnoff in dating situations. Future work is needed to clarify the nature of this indicator.

A noteworthy finding is that the linguistic cues identified here accounted for a substantial amount of variance in the deception index (23%) – an effect size that is larger than those observed in many similar studies. Can these linguistic markers be used to classify profiles as deceptive? A logistic regression using the linguistic variables identified above as predictors of high versus low deceptive profiles correctly classified 63% of the profiles (low deceptiveness: 61.5%; high deceptiveness: 65.8%)[χ^2 (4) = 13.55, p = .009]. The model significantly outperforms chance ($p < .01$) and is similar to rates observed in previous studies [2, 3].

While the classification rate is far from perfect, the fact that profiles can be classified along deception better than chance has implications for the many social networking websites (e.g., Facebook, LinkedIn) that have a similar profile structure, with textual self-descriptions and closed-ended questions. These data suggest that it may be possible to detect profile deceptions through analysis of the textual self-description, even though self-presenters are highly motivated to conceal their deceptions.

A theoretical implication of these findings concerns the nature of the linguistic cues to deception. The affordances of the media (i.e., editability and asynchronicity) appeared to attenuate cognitive linguistic cues. Markers of cognitive complexity, such as exclusive words and motion words, which were significant predictors of deception in other studies [2, 3], were not related to deception in this asynchronous context, suggesting that highly controllable media environments, such as online dating, may alleviate the cognitive burden associated with deception. In contrast, emotion-related linguistic cues were unaffected by the media environment, and were in fact the strongest predictors of deception.

A limitation of this study is that its correlational design does not preclude alternative explanations for what caused the occurrence of linguistic cues. For instance, participants who lied more may have had lower self-esteem, which may have led them to write differently. However, the cues found here also emerged in studies using an experimental design, where deception was manipulated [e.g., 2, 3].

To conclude, the present study represents an initial foray in identifying the linguistic correlates of deception in online dating profiles. While any practical application is currently limited, the data suggest the possibility of building a linguistic model that can detect deception in online dating profiles and possibly other kinds of online profiles.

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

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

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

4. Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., & Booth, R.J. (2007). *The development and psychometric properties of LIWC2007*. Austin, TX, LIWC.Net.

5. Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34,* 1023-1036.

6. Walther, J. B. (2007). Selective self-presentation in computer-mediated communication: Hyperpersonal dimensions of technology, language, and cognition. *Computers in Human Behavior, 23*, 2538–2557.

7. Whitty, M. T., & Joinson, A. (2009). *Truth, lies and trust on the Internet*. New York: Psychology Press.