Lying Words: Predicting Deception From Linguistic Styles

Matthew L. Newman
James W. Pennebaker
The University of Texas at Austin

Diane S. Berry
Southern Methodist University

Jane M. Richards
The University of Washington

Telling lies often requires creating a story about an experience or attitude that does not exist. As a result, false stories may be qualitatively different from true stories. The current project investigated the features of linguistic style that distinguish between true and false stories. In an analysis of five independent samples, a computer-based text analysis program correctly classified liars and truth-tellers at a rate of 67% when the topic was constant and a rate of 61% overall. Compared to truth-tellers, liars showed lower cognitive complexity, used fewer self-references and other-references, and used more negative emotion words.

Keywords: *deception; honesty; language; words; pronouns*

In 1994, Susan Smith appeared on television claiming that her two young children had been kidnapped at gunpoint. Eventually, authorities discovered she had drowned her children in the lake and fabricated the kidnapping story to cover her actions. Before Smith was a suspect in the children's deaths, she told reporters, "My children wanted me. They needed me. And now I can't help them" (Kastor, 1994). Normally, relatives will speak of a missing person in the present tense. The fact that Smith used the past tense in this context suggested to trained Federal Bureau of Investigation (FBI) agents that she already viewed them as dead (Adams, 1996).

Although liars have some control over the content of their stories, their underlying state of mind may "leak out" through the way that they tell them—an idea that dates back to Freud (1901). In one of Freud's examples, a doctor was visiting a wealthy patient who was suffering from a long illness. Despite outwardly claiming to have his patient's interests in mind, the doctor remarked, "I hope you will *not* soon leave your bed" (p. 88), revealing his selfish desire to continue treating a wealthy patient.

Telling a false story, by definition, requires describing events that did not happen or attitudes that do not exist. In addition to creating a convincing story, liars also must present it in a style that appears sincere (Friedman & Tucker, 1990). As a result, according to the literature on "reality monitoring," stories based on imagined experiences are qualitatively different from stories based on real experiences (e.g., Johnson & Raye, 1981; Vrij, Edward, Roberts, & Bull, 2000; cf. Undeutsch, 1967). One way to capture the differences between true and false stories is to look at the language people use to tell them.

Social psychology has, with a few notable exceptions (e.g., Feldman Barrett, Williams, & Fong, in press; Fiedler, Semin, & Koppetsch, 1991; Giles & Wiemann, 1993; Leets & Giles, 1997; Ruscher, 2001; Schnake & Ruscher, 1998; Semin & Fiedler, 1988), underappreciated the value of studying people's lan-

Authors' Note: We are indebted to Paul Ekman, Mark Frank, and especially Bella DePaulo for providing transcripts from earlier studies that served as pilot data for our thinking about this article. We are grateful to Ana Arazola, Lorilei Cardenas, Wendy Hiller, Jennifer Mueller, and Danna Upton for their assistance in conducting the research and to Matthias Mehl and Sherlock Campbell for comments on an earlier draft. Parts of this project were funded by a grant from the National Institutes of Health (MH52391). Correspondence should be addressed to James W. Pennebaker, Department of Psychology, The University of Texas at Austin, Austin, TX 78712; e-mail: pennebaker@psy.utexas.edu.

PSPB, Vol. 29 No. 5, May 2003 665-675
DOI: 10.1177/0146167203251529

guage use. A growing body of research suggests that we can learn a great deal about people's underlying thoughts, emotions, and motives by counting and categorizing the words they use to communicate. Of interest, words that reflect how people are expressing themselves can often be more informative than what they are expressing (Pennebaker & King, 1999; Pennebaker, Mehl, & Niederhoffer, in press; cf. Shapiro, 1989). Several features of *linguistic style*, such as pronoun use, emotionally toned words, and prepositions and conjunctions that signal cognitive work, have been linked to a number of behavioral and emotional outcomes. For example, poets who used a high frequency of self-references but a lower frequency of other-references in their poetry were more likely to commit suicide than those who showed the opposite pattern (Stirman & Pennebaker, 2001). Increased use of cognitive words (e.g., *think*, *because*) among college students has been linked to higher grades, better health, and improved immune function (Klein & Boals, 2001; Petrie, Booth, & Pennebaker, 1998).

In the present studies, we took an inductive approach to examining the linguistic manifestations of false stories. First, we used a computerized text analysis program to create empirically derived profiles of deceptive and truthful communications. We then tested the generalizability of these profiles on independent samples. Finally, we compared the predictive ability of these profiles to predictions made by human judges.

The linguistic profiles were created using Linguistic Inquiry and Word Count (LIWC) (Pennebaker, Francis, & Booth, 2001), a text analysis program that analyzes written or spoken samples on a word-by-word basis. Each word is compared against a file of more than 2,000 words, divided into 72 linguistic dimensions. After counting the number of words in each category, the output is given as a percentage of the total words in the text sample. Computerized word count approaches are typically blind to context but have shown promising and reliable results in personality, social, and clinical psychology (Mergenthaler, 1996; Pennebaker et al., 2001; Rosenberg & Tucker, 1979; Stone, Dunphy, Smith, & Ogilvy, 1966). Specifically, the dimensions captured by LIWC have been used in recent studies to predict a number of outcome measures, including social judgments (Berry, Pennebaker, Mueller, & Hiller, 1997), personality (Pennebaker & King, 1999), personality change (Pennebaker & Lay, in press), psychological adjustment (Rude, Gortner, & Pennebaker, 2001), and health (Pennebaker, Mayne, & Francis, 1997).

Our approach to the language of deception has been influenced by analyses of linguistic style when people write or talk about personal topics. Essays that are judged as more personal and honest (or, perhaps, less self-deceptive) have a very different linguistic profile than essays that are viewed as more detached (cf. Pennebaker & Francis, 1996; Graybeal, Seagal, & Pennebaker, in press). Of interest, this linguistic profile also is linked to improvements in the authors' physical health (Campbell & Pennebaker, in press; Pennebaker et al., 1997). This suggests that creating a false story about a personal topic takes work and results in a different pattern of language use. Extending this idea, we would predict that many of these same features would be associated with deception or honesty in communication. Based on this research, at least three language dimensions should be associated with deception: (a) fewer self-references, (b) more negative emotion words, and (c) fewer markers of cognitive complexity.

First, the use of the first-person singular is a subtle proclamation of one's ownership of a statement. Knapp, Hart, and Dennis (1974) hypothesized that liars may avoid statements of ownership either to "dissociate" themselves from their words or due to a lack of personal experience (e.g., Buller, Burgoon, Buslig, & Roiger, 1996; Dulaney, 1982; Knapp & Comadena, 1979; Mehrabian, 1971). Similarly, Wiener and Mehrabian (1968) argued that liars should be more "nonimmediate" than truth-tellers and refer to themselves less often in their stories. Other studies have found that when individuals are made to be self-aware, they are more "honest" with themselves (e.g., Carver & Scheier, 1981; Duval & Wicklund, 1972; Voraurer & Ross, 1999) and self-references increase (e.g., Davis & Brock, 1975). Finally, individuals who respond defensively (i.e., self-deceptively) when discussing personal topics tend to distance themselves from their stories and avoid taking responsibility for their behavior (Feldman Barrett et al., in press; Shapiro, 1989). If this state of mind is reflected in the words people use, then deceptive communications should be characterized by fewer first-person singular pronouns (e.g., *I*, *me*, and *my*).

Second, liars may feel guilty either about lying or about the topic they are discussing (e.g., Ekman, 1985/1992; Knapp & Comadena, 1979; Knapp et al., 1974; Vrij, 2000). Diary studies of small "everyday" lies suggest that people feel discomfort and guilt while lying and immediately afterward (e.g., DePaulo et al., 2003). If this state of mind is reflected in patterns of language use, then deceptive communications should be characterized by more words reflecting negative emotion (e.g., *hate*, *worthless*, *sad*).

Finally, the process of creating a false story should consume cognitive resources (cf. Richards & Gross, 1999, 2000), leading liars to tell less complex stories. From this cognitive perspective, truth-tellers are more

likely to tell about what they did *and* what they did not do. That is, they are making a distinction between what is in the category of their story and what is not. Based on previous emotional writing studies, people using the "honest" style described above show evidence of making distinctions in their stories. Specifically, individuals who use a higher number of "exclusive" words (e.g., *except, but, without*) are generally healthier than those who do not use these words (Pennebaker & King, 1999). In addition, stories that are less complex may focus more on simple, concrete verbs (e.g., "I walked home") rather than evaluations and judgments (e.g., "Usually I take the bus, but it was such a nice day") because the former are more accessible and more readily strung together in a false story. On the surface, this idea is inconsistent with evidence that people automatically associate an object with an attitude or evaluation of that object (e.g., Bargh, Chaiken, Govender, & Pratto, 1992; Fazio, 2001). However, these automatic evaluations are presumed to reflect participants' true attitudes. When people are attempting to construct a false story, we argue that simple, concrete actions are easier to string together than false evaluations. Unpublished data from our labs have shown a negative relationship between cognitive complexity and the use of motion verbs (e.g., *walk, move, go*). Thus, if deceptive communications are less cognitively complex, liars should use more motion verbs and fewer exclusive words.

Measuring and testing several linguistic dimensions at once is an unusual approach to the study of language and deception. Despite the fact that lies in the real world contain a variety of linguistic cues, previous research has typically tested the predictive power of individual linguistic cues (but see Dulaney, 1982; Vrij, 2000; Vrij et al., 2000). This approach does not allow the assessment of a multivariate profile of deception. Where previous work has examined the link between deception and the use of individual cues (e.g., self-references or negative words or cognitive words), the methodology used in the present research allowed us to test whether differential use of multiple cues (e.g., self-references + negative words + cognitive words) is a reliable marker of deception.

The present study tested a number of specific predictions about the linguistic features of deceptive stories. We predicted that the linguistic profile of deception created by LIWC would reflect qualitative differences between truthful and deceptive communications. Specifically, as described above, we hypothesized that liars would use fewer self-references, fewer cognitive complexity words (exclusive words, motion verbs), and more negative emotion words. We also hypothesized that a linguistic profile based on one sample would generalize to an independent sample. Finally, we hypothesized that the profile created by LIWC would be more accurate at detecting deception than untrained human judges.

Method

To maximize the generalizability of our linguistic profile, we asked participants either to lie or to tell the truth about different topics in different contexts. The basic methodology of each of the five studies is described below. Four of the studies (1, 2, 3, and 4) were conducted at Southern Methodist University (SMU) and the fifth study was conducted at the University of Texas at Austin (UT). In all of the studies, the written or transcribed verbal samples of each participant were analyzed using the LIWC program. Information about the demographics and number of writing samples for all studies is listed in Table 1.

EXPERIMENTAL PROCEDURES

Study 1: Videotaped abortion attitudes. The study included 101 undergraduates (54 men, 47 women) at SMU who were videotaped while discussing both their true and false views on abortion. Of the participants, 26 (13 men, 13 women) identified themselves as "pro life" and 76 (42 men, 34 women) identified themselves as "pro choice." A female experimenter ran participants individually. All were asked to state their true opinion regarding abortion and to explain their reasons for favoring that position. All participants also were asked to state that they agreed with the point of view that was actually different from their own and then discuss their reasons for agreeing with that position. Order of true versus deceptive communication was counterbalanced. Participants were told that other people would view their videotapes and attempt to guess their true views. They were encouraged to be as persuasive and believable as possible while delivering both communications. No time limit was set but individuals were encouraged to try to talk for at least 2 min each time.

Study 2: Typed abortion attitudes. The study included 44 undergraduates (18 men, 26 women) at SMU who were asked to type both a truthful and deceptive communication concerning their views on abortion. Overall, 13 participants (8 men, 5 women) identified themselves as pro life and 31 (10 men, 21 women) identified themselves as pro choice. A female experimenter ran participants tested individually in a small laboratory room. All were asked to type directly onto a computer both their true and false attitudes concerning abortion. Participants were told that other people would read their essays and attempt to guess their true views. As in Study 1, order of communication was counterbalanced and participants were strongly encouraged to be as persuasive as possible during each of the 5-min typing sessions.

Study 3: Handwritten abortion attitudes. The study included 55 undergraduates (15 men, 40 women) at SMU, 14 of whom (3 men, 11 women) identified themselves as pro life and 41 (12 men, 29 women) of whom identified themselves as pro choice, who participated in the study. As part of a general project on interpersonal communications, students were given a packet of materials to complete and return to the Psychology Department within 2 weeks. One envelope in the packet contained sheets of lined paper and a written version of the instructions given to participants in Study 2. As before, individuals were asked to provide a truthful and deceptive description of their position on the issue of abortion in a counterbalanced order. Participants were told that other people would read their essays and attempt to guess their true views. Upon completing the packet, participants sealed their writing samples and returned them.

Study 4: Feelings about friends. The study included 27 undergraduates (8 men, 19 women) at SMU who were tested individually within a paradigm developed by DePaulo and Rosenthal (1979) and asked to provide true and false descriptions of people whom they truly like and dislike. In the counterbalanced study, each participant provided four videotaped descriptions about two people they truly liked and two people they truly disliked. The deception manipulation required participants to talk about a person they liked as if they disliked them and about a person they disliked as if they liked them. All four combinations were included as a way to increase the scope of lies that were represented. Participants were asked to talk about each of the four people for 3 min each, during which time they would be videotaped and their heart rate and skin conductance levels would be monitored. In the truthful conditions, participants were encouraged to talk about any characteristics or stories "showing why you truly like (dislike) this person . . . tell us honestly how you feel." In the deceptive conditions, participants were encouraged to "convey a believable, convincing false impression." They were also given the freedom to invent stories or characteristics to help them convey this false impression.

Study 5: Mock crime. The study included 60 Introductory Psychology students (23 men, 37 women) at UT who were asked to sit alone in a small cluttered room for a few minutes. Half were told to simply look around the room and half were told specifically to look inside a book for a dollar bill and "steal it." Participants in both conditions were then told that an experimenter (blind to condition) would enter the room and accuse them of stealing a dollar bill. All participants were explicitly told to deny taking any money. Participants also were told that if they convinced the interrogator of their innocence, they would be given the dollar bill. In reality, the interrogator pretended to believe all participants and all were paid $1 for their participation. During the brief interrogation, participants were taken to another room where (nonoperational) skin conductance leads were attached as part of a bogus polygraph system and a video camera was pointed at the participant. The interrogator asked the participant to (a) indicate whether they had taken any money from the book, (b) describe exactly what they had done while sitting in the room, (c) describe the contents of the room, and (d) describe again exactly what they had done from the minute they entered the room. The interrogation generally lasted less than 2 min.

TABLE 1: Study Demographics

Study	*Description*	N	*Samples/ Person*	*Total Verbal Samples*	*Average Words/ Sample*	*% Female*
1	Video abortion	101	2	202	124	47
2	Typed abortion	44	2	88	290	59
3	Written abortion	55	2	110	170	73
4	Video friends	27	4	108	529	70
5	Video mock crime	60	1	60	156	62
Total		287		568 writing samples		

JUDGES' RATINGS OF COMMUNICATIONS

Human judges at SMU were asked to rate the truthfulness of the 400 communications dealing with abortion attitudes (Studies 1 through 3). The judges were told that they would read transcripts of people presenting a pro-choice or a pro-life position on the issue of abortion rights and that the presenters had been randomly assigned to present one or the other position, as in a debate. Thus, in some cases, the people would be describing feelings that were not truly their own. The judges' task was to guess whether each communication was truthful or deceptive. Judges were specifically instructed that, depending on the particular transcripts that they rated, there could be more actual pro-choice than pro-life opinions, more actual pro-life than pro-choice opinions, or roughly equivalent numbers of each.

For each communication, the judges responded "yes" or "no" to the question, "Is this person expressing his or her true feelings on the matter?" Each communication was evaluated by a total of seven to nine judges. The proportion of judges indicating that they believed each one to be truthful was used as a measure of perceived truthfulness.

DATA PREPARATION PROCEDURES

Each verbal sample was transcribed and entered into a separate text file. Any transcriber comments (e.g., "subject laughs") were removed and misspellings corrected. Each of the 568 text files from the five studies was

analyzed using the LIWC program. Although LIWC can analyze text along more than 72 linguistic dimensions, several categories were excluded from the present analyses. First, variables were excluded from all subsequent analyses if they reflected essay content (e.g., words related to death, religion, occupation, etc.). The logic of this rule was that we sought to develop a linguistic system that would be independent of potential essay content. Second, any linguistic variables that were used at extremely low rates (less then 0.2% of the time) were excluded. Finally, variables were excluded that might be unique to spoken or written transcripts or could be influenced by the interpretations of a transcriber, such as words per sentence and nonfluencies (e.g., *uh*, *umm*). The final list of variables that were employed, then, was reduced to 29 variables (see Table 2).

Finally, before performing any analyses, all of the LIWC categories were standardized within study by converting the percentages to *z* scores. Because base rates of word usage varied depending on both the subject matter (i.e., abortion, friendship, or mock crime) and the mode of communication (i.e., spoken, typed, or handwritten), using standardized percentages allowed us to make comparisons across studies.

Results

The nature of the data allowed us to address a number of questions pertaining to the language of truthfulness versus deception. In the first section of the results, we developed and tested a linguistic profile of deception. In the second section, we compared the accuracy rates of human judges with those of the LIWC profile.

LINGUISTIC PROFILE OF DECEPTION

Analysis strategy. As noted earlier, one of the recurring problems with earlier studies is that researchers have typically correlated use of individual word categories with deception. Such a strategy makes it virtually impossible to form a complete picture of the language of deception. To address this problem in the present research, we used our text analysis program to create a multivariate profile of deception. We first developed a profile for each of the five studies and then developed an overall linguistic profile that combined the results for the individual studies.

Specifically, we analyzed each of the five studies in three steps. We first performed a forward-entry logistic regression,[1] predicting deception based on usage of the 29 LIWC categories in four of the five studies. This logistic regression produced a set of beta weights predicting deception. Second, these beta weights were multiplied by the corresponding LIWC categories in the remaining study and added together to create a prediction equation. These equations formed our operational definition of linguistic profiles. Finally, a second logistic regression was performed, using this equation to predict deception in the remaining study. These three steps were repeated for each of the five studies. In all analyses, deception was coded as a dichotomous variable with truth-telling coded as "1" and lying coded as "0." Thus, categories with coefficients in the negative direction were used at a higher rate by liars and categories with coefficients in the positive direction were used at a higher rate by truth-tellers.

Predicting deception in each study. For each of the five studies, a linguistic profile based on the other four studies was used to predict deception. In the case of Study 1, for example, we examined how well the linguistic profile from Studies 2 through 5 could predict deception in Study 1. A logistic regression predicting deception from the LIWC categories in Studies 2 through 5 revealed a good fit to the data, $\chi^2(4, N = 366) = 34.99$, $p < .001$, that explained 9% of the variance.[2] The coefficients for this model are presented in the first section of Table 3. Based on these four studies, liars tended to use more negative emotion words, fewer sensation words, fewer exclusive words, and more motion verbs. Three of these dimensions—negative emotion words, motion verbs, and exclusive words—were consistent with our predictions. We then multiplied these beta weights from Studies 2 through 5 by the corresponding LIWC categories in Study 1 to create a prediction equation, which was tested in a logistic regression to predict deception in Study 1, $\chi^2(1, N = 202) = 7.58$, $p < .01$ (see Note 3 for a sample equation).[3] As seen in Table 3, this equation explained 4% of the variance and was able to correctly classify 60% of liars and 58% of truth-tellers for an overall accuracy rate of 59%.

We repeated this three-step analysis for each of the other four studies. The initial logistic regressions predicting deception in four studies all revealed good fits to the data (all ps $< .001$). In addition, the separate prediction equations based on the logistic regression procedures were significant in predicting deception in Study 2, $\chi^2(1, N = 88) = 5.49$, $p < .05$, and Study 3, $\chi^2(1, N = 110) = 19.87$, $p < .001$, but not Study 4, $\chi^2(1, N = 108) = 1.18$, $p = .28$, or Study 5, $\chi^2(1, N = 60) = 0.12$, $p = .73$ (see Table 3 for the details of these analyses). The overall classification rate was compared to chance performance (50%) in each of the five studies. The LIWC profiles performed better than chance in Study 1 ($z = 2.25$, $p < .05$), Study 2 ($z = 1.80$, $p < .05$), and Study 3 ($z = 3.40$, $p < .001$) but not Study 4 ($z = .60$, *ns*) or Study 5 ($z = .37$, *ns*).

Predicting deception across five studies. The analyses described so far revealed a relatively consistent linguistic profile of deception, but our goal was to develop a more general picture of the linguistic markers of deception. To do this, we selected the five LIWC categories that were significant predictors in more than one of the analyses

TABLE 2: LIWC Categories Used in the Present Study

Dimension	*Abbreviation*	*Example*	*# Words*	*Mean*
I. Standard linguistic dimensions				
Word Count	WC			238.87
% words captured, dictionary words	Dic			73.67
% words longer than six letters	Sixltr			13.57
Total pronouns	Pronoun	I, our, they, you're	70	12.76
First-person singular	I	I, my, me	9	3.97
Total first person	Self	I, we, me	20	4.72
Total third person	Other	she, their, them	22	4.04
Negations	Negate	no, never, not	31	2.98
Articles	Article	a, an, the	3	7.30
Prepositions	Preps	on, to, from	43	11.93
II. Psychological processes				
Affective or emotional processes	Affect	happy, ugly, bitter	615	3.54
Positive emotions	Posemo	happy, pretty, good	261	2.14
Negative emotions	Negemo	hate, worthless, enemy	345	1.39
Cognitive processes	Cogmech	cause, know, ought	312	8.75
Causation	Cause	because, effect, hence	49	1.39
Insight	Insight	think, know, consider	116	2.16
Discrepancy	Discrep	should, would, could	32	3.93
Tentative	Tentat	maybe, perhaps, guess	79	3.09
Certainty	Certain	always, never	30	1.21
Sensory and perceptual processes	Senses	see, touch, listen	111	1.54
Social processes	Social	talk, us, friend	314	11.01
III. Relativity				
Space	Space	around, over, up	71	2.34
Inclusive	Incl	with, and, include	16	5.96
Exclusive	Excl	but, except, without	19	4.72
Motion verbs	Motion	walk, move, go	73	1.03
Time	Time	hour, day, o'clock	113	2.22
Past tense verb	Past	walked, were, had	144	3.13
Present tense verb	Present	walk, is, be	256	12.50
Future tense verb	Future	will, might, shall	14	1.87

NOTE: LIWC = Linguistic Inquiry and Word Count. # words refers to the number of words per category in the LIWC dictionary. Mean refers to the mean percentage of usage in the present studies. Word count refers to a raw number.

above, reasoning that categories appearing only once might be unique to the mode of communication (i.e., handwritten, typed, or spoken) or the topic (i.e., abortion, feelings about friends, or a mock crime).

These five categories—first-person singular pronouns, third-person pronouns, negative emotion words, exclusive words, and motion verbs—were entered into a simultaneous logistic regression predicting deception in all five of the studies combined. The coefficients for these variables are presented at the bottom of Table 3. We then created a general prediction equation based on these coefficients and entered this into a logistic regression predicting deception in all five studies combined. When all five studies were combined, the general equation explained 8% of the variance and correctly classified 59% of liars and 62% of truth-tellers, $\chi^2(6, N = 568) = 49.82$, $p < .001$, for an overall accuracy rate of 61%. This was significantly better than chance ($z = 5.50$, $p < .001$). To examine the reliability of the different models, we computed Cronbach's alpha on the five prediction equations. The overall alpha was .93, suggesting that many of the linguistic markers of deception are consistent across situations.

Across five studies, deceptive communications were characterized by fewer first-person singular pronouns, fewer third-person pronouns, more negative emotion words, fewer exclusive words, and more motion verbs. Table 4 presents effect sizes and reliability estimates for these linguistic markers. Four of these markers (first-person pronouns, exclusive words, motion verbs, and negative emotion words) were consistent with our predictions. However, the markers were more consistent among the abortion studies, and the general equation explained a notably smaller percentage of the variance in the nonabortion studies. We return to this issue in our discussion.

COMPARING LIWC WITH HUMAN JUDGES

Finally, we investigated how LIWC's ability to identify deception based on linguistic styles would compare to

TABLE 3: Predictors of Deception: Logistic Regression Coefficients Used in Prediction Equations

LIWC Category	β	*% Accuracy (Lie/ Truth/Overall)*	*Adjusted R^2*
Predicting Study 1:			
Studies 2-5 combined		60/58/59*	4
Negative emotion	–.268		
Senses	.270		
Exclusive words	.452		
Motion verbs	–.310		
Predicting Study 2:			
Studies 1 and 3-5 combined		61/57/59*	6
Negative emotion	–.227		
Exclusive words	.286		
Motion verbs	–.358		
Predicting Study 3:			
Studies 1-2 and 4-5 combined		66/69/67**	17
First-person pronouns	.209		
Third-person pronouns	.254		
Exclusive words	.362		
Motion verbs	–.213		
Predicting Study 4:			
Studies 1-3 and 5 combined		52/54/53	1
First-person pronouns	.240		
Articles	–.264		
Negative emotion	–.382		
Exclusive words	.463		
Predicting Study 5:			
Studies 1-4 combined		53/43/48	0
First-person pronouns	.330		
Third-person pronouns	.334		
Exclusive words	.435		
Motion verbs	–.225		
General prediction equation:			
All 5 studies combined		59/62/61**	8
First-person pronouns	.260		
Third-person pronouns	.250		
Negative emotion	–.217		
Exclusive words	.419		
Motion verbs	–.259		

NOTE: LIWC = Linguistic Inquiry and Word Count. For % accuracy, the three percentages listed for each equation are (a) % of liars identified accurately, (b) % of truth-tellers identified accurately, and (c) overall accuracy. For overall accuracy rates, *$p < .05$ and **$p < .001$ when compared to chance performance of 50%. Coefficients in the negative direction mean that liars used the category at a higher rate. All coefficients are significant at $p < .05$ or better.

human judges. Recall that judges rated the perceived truthfulness of all communications dealing with abortion attitudes (Studies 1 through 3; $N = 400$ communications) and the proportion of judges who believed each communication to be truthful had been calculated. We used these data to calculate a "hit rate" for our judges and compared it to LIWC's ability to correctly identify deception. More specifically, a dichotomous classification was made for each communication. If the proportion of judges believing a particular communication to be truthful was greater than 50%, it was defined as "judged truthful." The remaining communications were considered "judged false."

TABLE 4: Effect Sizes and Reliability of Linguistic Predictors Across Five Studies

	Study						
	1	*2*	*3*	*4*	*5*	*Mean* d	*Reliability*
First person	.31	.85	.75	–.02	–.24	.36	.43
Third person	.22	.30	.07	.24	–.21	.16	.28
Negative emotion	–.19	–.27	–.42	.40	–.33	–.15	.36
Exclusive words	.40	1.23	.91	.02	.30	.54	.55
Motion verbs	–.14	.09	–.31	–.40	–.29	–.20	.40

NOTE: Numbers for each study represent Cohen's *d* comparing liars and truth-tellers. Effect sizes in the negative direction mean that liars used the linguistic category at a higher rate. Mean *d* is a weighted mean of these effect sizes. Reliability was calculated using Cronbach's alpha.

We then calculated the proportion of communications that had been correctly identified as truthful or deceptive by the judges and compared them with the LIWC classifications based on the general prediction equation (see above). As seen in Table 5, LIWC correctly classified 67% of the abortion communications and the judges correctly classified 52% of the abortion communications. These proportions were significantly different ($z = 6.25$, $p < .0001$). LIWC performed significantly better than chance ($z = 6.80$, $p < .001$) but the judges did not ($z = .80$, *ns*).

To correct for a potential positivity bias in judges' responses, we conducted a signal detection analysis. We first converted the proportion of correct hits (identifying truthful communications as truthful) and the proportion of false positives (identifying deceptive communications as truthful) to *z* scores. These values of .74 and .71 were converted to *z* scores of $z = .64$ and $z = .55$, respectively. We then calculated d' by subtracting the proportion of false positives from the proportion of correct hits ($d' = .09$, *ns*).

In addition, the two detection strategies showed different patterns of error. The judges were significantly more likely to make "false positive" identifications than "false negative" identifications (71% vs. 27% of judges' errors, respectively; $z = 14.67$, $p < .001$). LIWC, in contrast, was equally likely to make "false positive" identifications and "false negative" identifications (49% vs. 51% of LIWC errors, respectively; $z = .50$, *ns*).

Discussion

Successfully lying to another person usually involves the manipulation of language and the careful construction of a story that will appear truthful. In addition to creating a convincing story, liars also must present it in a style that appears sincere (Friedman & Tucker, 1990).

TABLE 5: Comparison of Human Judges' Ratings With LIWC's Prediction Equations in Three Abortion Studies

	Predicted	
	Deceptive	*Truthful*
LIWC equations		
Actual		
Deceptive (n = 200)	68% (135)	32% (65)
Truthful (n = 200)	34% (69)	66% (131)
Human judges		
Actual		
Deceptive (n = 200)	30% (59)	71% (141)
Truthful (n = 200)	27% (53)	74% (147)

NOTE: LIWC = Linguistic Inquiry and Word Count. N = 400 communications. The overall hit rate was 67% for LIWC and 52% for judges; these were significantly different, $z = 6.25$, $p < .001$. LIWC performed significantly better than chance, $z = 6.80$, $p < .001$, but the judges did not, $z = .80$, *ns*. See the text for an analysis of error rates.

Although liars have some control over the content of their stories, their underlying state of mind may leak out through the style of language used to tell the story. The data presented here provide some insight into the linguistic manifestations of this state of mind.

Specifically, deceptive communication was characterized by the use of fewer first-person singular pronouns (e.g., *I*, *me*, *my*), fewer third-person pronouns (e.g., *he*, *she*, *they*), more negative emotion words (e.g., *hate*, *anger*, *enemy*), fewer exclusive words (e.g., *but*, *except*, *without*), and more motion verbs (e.g., *walk*, *move*, *go*). Four of these linguistic categories—first-person pronouns, negative emotion words, exclusive words, and motion verbs—were consistent with our specific predictions. However, the generalizability of these categories varied depending on the topic. In this discussion, we first examine the meaning of each linguistic predictor and then address the generalization of these predictors.

ELEMENTS OF THE LINGUISTIC PROFILE

First, in the present studies, liars used first-person pronouns at a lower rate than truth-tellers. The lower rate of self-references is consistent with previous literature (but see our discussion below of DePaulo et al., 2003) and is thought to reflect an attempt by liars to "dissociate" themselves from the lie (Dulaney, 1982; Knapp et al., 1974; Mehrabian, 1971; for a review, see Knapp & Comadena, 1979; Vrij, 2000, Chap. 4). Self-references indicate that individuals are being "honest" with themselves (Campbell & Pennebaker, in press; Davis & Brock, 1975; Duval & Wicklund, 1972; Feldman Barrett et al., in press; Shapiro, 1989). Because deceptive stories do not reflect one's true attitudes or experiences, liars may wish to "stand back" (Knapp et al., 1974, p. 26) by investing less of themselves in their words.

Second, liars also used negative emotion words at a higher rate than truth-tellers. Liars may feel guilty, either because of their lie or because of the topic they are lying about (e.g., Vrij, 2000). Because of this tension and guilt, liars may express more negative emotion. In support of this, Knapp et al. (1974) found that liars made disparaging remarks about their communication partner at a much higher rate than truth-tellers. In an early meta-analysis, Zuckerman, DePaulo, and Rosenthal (1981) identified negative statements as a significant marker of deception. The "negative emotion" category in LIWC contains a subcategory of "anxiety" words, and it is possible that anxiety words are more predictive than overall negative emotion. However, in the present studies, anxiety words were one of the categories omitted due to low rate of use.

Third, liars used fewer "exclusive" words than truth-tellers, suggesting lower cognitive complexity. A person who uses words such as *but*, *except*, and *without* is making a distinction between what is in a given category and what is not within a category. Telling a false story is a highly cognitively complicated task. Adding information about what did not happen may require cognitive resources that the typical liar does not possess. Fourth, liars used more "motion" verbs than truth-tellers, also suggesting lower cognitive complexity. Because liars' stories are by definition fabricated, some of their cognitive resources are taken up by the effort of creating a believable story. Motion verbs (e.g., *walk*, *go*, *carry*) provide simple, concrete descriptions and are more readily accessible than words that focus on evaluations and judgments (e.g., *think*, *believe*).

In addition, liars in the present studies unexpectedly used third-person pronouns at a lower rate than truth-tellers. This is inconsistent with previous literature: Liars typically use more other-references than truth-tellers (e.g., Knapp et al., 1974). It is possible that this reflects the subject matter—abortion attitudes—in the majority of the present studies. Talking about abortion necessarily involves talking about women, but this can be done using pronouns (*she*, *her*), more specific nouns (*a woman*, *my sister*), or even proper names. This word use may reflect differences in the underlying psychology between liars and truth-tellers such that people lying about their attitudes added concrete details by referring to specific people instead of using the generic *she*. Although this interpretation is admittedly post hoc, it is consistent with our finding that liars tend to be concrete rather than abstract (see also Knapp et al., 1974).

Previous investigations of the linguistic differences between liars and truth-tellers have yielded mixed results, largely due to substantive differences in the ways linguistic categories have been defined and assessed. In a recent exhaustive meta-analysis, DePaulo et al. (2003) reviewed the combined evidence for 158 different cues to deception. Overall, liars appear to be less forthcom-

ing, less convincing, and more tense than truth-tellers. One major strength of meta-analysis is that it can allow comparisons across different measurement units or techniques. However, this strength can be a problem if the measurement differences actually reflect qualitatively different things.

For example, one of the linguistic cues reviewed by DePaulo et al. (in press) was the use of "self-references." Based on 12 separate estimates, DePaulo et al. computed a combined effect size of $d = .01$, indicating no relationship. Of interest, the variation of effect sizes within the 12 studies is astounding—from –.91 to +.86 (negative numbers indicate that liars used fewer self-references). An examination of these 12 studies reveals very different methods for defining and assessing "self-references," from counting the number of self-reference words (*I, me, my*) (e.g., Bond, Kahler, & Paolicelli, 1985; Buller et al., 1996) to counting the number of times that the speaker was the primary subject of the sentence, as in "she gave me a cookie." These appear to capture rather different things. Research using LIWC's word-count approach suggests that, for example, a person who says "I am not sad" is more likely to become depressed than one who says "I am happy" (Pennebaker et al., 1997). Due to the range in both methodology and findings, the effect size calculated by DePaulo et al. (in press) seems a premature estimate of the link between self-references and deception.

LYING WORDS ACROSS CONTEXT

In the present studies, we predicted that the linguistic markers of deception would for the most part generalize across context, but context did matter to some degree. A model based primarily on talking about abortion attitudes was much more predictive within the same subject matter than across different subject matter. This suggests an interrelationship between the content of a communication and the style of language used to tell it. Although the findings were consistent with our predictions about the linguistic manifestations of (all) false stories, future research is needed to tease apart the linguistic markers of lying from the linguistic markers specific to lying about abortion attitudes. People's opinions on abortion are highly emotional, and this may affect the language used above and beyond the process of creating a false story. However, despite the low predictive power in Studies 4 and 5, there was some consistency in the actual categories related to deception across all five studies (see Tables 3 and 4), and the prediction equations were highly correlated with one another ($\alpha = .93$).

In the real world, context is an important factor in identifying deception. The FBI trains its agents in a technique called statement analysis, which attempts to detect deception based on parts of speech (i.e., linguistic style) rather than the facts of the case or the story as a whole (Adams, 1996). Suspects are first asked to make a written statement. Trained investigators then review this statement, looking for deviations from the expected parts of speech. These deviations from the norm provide agents with topics to explore during interrogation. Adams (1996) gives the example of a man accused of killing his wife. Throughout his statement, he consistently refers to "my wife and I" rather than "we," suggesting distance between the couple. A trained interrogator would then ask the man about his relationship with his wife. If he says they were inseparable, the interrogator would have reason to be suspicious and ask a follow-up question. Thus, linguistic style may be most useful in the hands of a trained expert who knows what to look for and how to use language to reveal inconsistencies (see also Vrij et al., 2000).

Two limitations of the present research deserve attention. First, this particular model is limited to the English language, and possibly to American English. Our argument in the present studies is that liars and truth-tellers will use language in predictably different ways. However, besides the obvious differences in vocabulary, other languages also have different patterns of language use. Thus, the same linguistic markers may not identify deception in other languages. For example, a number of the Romance languages do not require an expressed noun or pronoun; these are often part of the verb. In Spanish, a person may introduce himself by saying "soy Bob" ("I am Bob") without using the pronoun *yo*, meaning *I*. The same is true of Latin, although instances of people lying in Latin are on the decline. Because of these different patterns, the linguistic markers of deception identified in the present study—especially the use of first-person pronouns—may not generalize to other languages. However, it is possible that deception in other languages is characterized by changes in first-person singular verbs, such as a lower rate of *soy* (*I am*). We would suggest that research on linguistic markers of deception in other languages proceed in much the same way as the present study—first creating an empirically derived linguistic profile and then validating this profile on an independent sample.

Second, previous research has suggested that both the motivation to lie and emotional involvement are important moderators of which markers will identify deception. Motivated liars tend to be more tense but slightly more "fluent" in their communications (e.g., Frank & Ekman, 1997; Gustafson & Orne, 1963; for reviews, see DePaulo et al., in press; Ekman, 1985/1992; Vrij, 2000). In the present research, all participants knew that others would try to guess whether they were lying, but external motivation to lie successfully was practically nonexistent. Participants in the mock crime study (Study

5) knew they might get an extra dollar for lying successfully, but the price of a cup of coffee is negligible compared to the things that can be at stake in real-world lies. For example, a person trying to lie about an extramarital affair, or involvement in a tax-fraud scheme, will have very high motivation to lie successfully, and this may have a further influence on the language that is used.

CONCLUSIONS

The research described here took a unique approach to studying language and deception. By using a computer-based text analysis program, we were able to develop a multivariate linguistic profile of deception and then use this to predict deception in an independent sample. This research—along with others from our lab—suggests that researchers should consider noncontent words (also referred to as particles, closed-class words, or function words) in exploring social and personality processes. Markers of linguistic style—articles, pronouns, and prepositions—are, in many respects, as meaningful as specific nouns and verbs in telling us what people are thinking and feeling.

Our data suggest that liars tend to tell stories that are less complex, less self-relevant, and more characterized by negativity. At a broad level, the differences between deceptive and truthful communications identified here are consistent with the idea that liars and truth-tellers communicate in qualitatively different ways (Undeutsch, 1967). Consequently, the present studies suggest that liars can be reliably identified by their words—not by what they say but by how they say it.

NOTES

1. In a logistic regression, forward entry into the model is based on the significance of the Wald statistic for each predictor variable. The variable with the most significant independent contribution to the dependent variable is added first. Variables are added into the model in order of significance until alpha for a single variable exceeds .05.

2. All reports of the percentage of variance explained are based on the Cox and Snell R^2 statistic. The Cox and Snell R^2 is based on the log likelihood for the final model compared to the log likelihood for a baseline model with no independent variables specified. The Cox and Snell R^2 also is adjusted for differences in sample size.

3. Study 1 was predicted with the following equation: (*z* score for negative emotion words * –.268) + (*z* score for sense words * .270) + (*z* score for exclusive words * .452) + (*z* score for motion verbs * –.310).

REFERENCES

Adams, S. H. (1996, October). Statement analysis: What do suspects' words really reveal? *FBI Law Enforcement Bulletin*, pp. 12-20.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*, 893-912.

Berry, D. S., Pennebaker, J. W., Mueller, J. S., & Hiller, W. S. (1997). Linguistic bases of social perception. *Personality and Social Psychology Bulletin, 23*, 526-537.

Bond, C. F., Jr., Kahler, K. N., & Paolicelli, L. M. (1985). The miscommunication of deception: An adaptive perspective. *Journal of Experimental Social Psychology, 21*, 331-345.

Buller, D. B., Burgoon, J. K., Buslig, A., & Roiger, J. (1996). Testing Interpersonal Deception Theory: The language of interpersonal deception. *Communication Theory, 6*, 268-289.

Campbell, R. S., & Pennebaker, J. W. (in press). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science*.

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer-Verlag.

Davis, D., & Brock, T. C. (1975). Use of first person pronouns as a function of increased objective self-awareness and performance feedback. *Journal of Experimental Social Psychology, 11*, 381-388.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74-112.

DePaulo, B. M., & Rosenthal, R. (1979). Telling lies. *Journal of Personality and Social Psychology, 37*, 1713-1722.

Dulaney, E. F. (1982). Changes in language behavior as a function of veracity. *Human Communication Research, 9*, 75-82.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York: Academic Press.

Ekman, P. (1985/1992). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (2nd ed.). New York: Norton.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion, 15*, 115-141.

Feldman Barrett, L., Williams, N. L., & Fong, G. (in press). Defensive verbal behavior assessment. *Personality and Social Psychology Bulletin*.

Fiedler, K., Semin, G. R., & Koppetsch, C. (1991). Language use and attributional biases in close personal relationships. *Personality and Social Psychology Bulletin, 17*, 147-155.

Frank, M. G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. *Journal of Personality and Social Psychology, 72*, 1429-1439.

Freud, S. (1901). *Psychopathology of everyday life*. New York: Basic Books.

Friedman, H. S., & Tucker, J. S. (1990). Language and deception. In H. Giles & W. P. Robinson (Eds.), *Handbook of language and social psychology*. New York: John Wiley.

Giles, H., & Wiemann, J. M. (1993). Social psychological studies of language: Current trends and prospects. *American Behavioral Scientist, 36*, 262-270.

Graybeal, A., Seagal, J. D., & Pennebaker, J. W. (in press). The role of story-making in disclosure writing: The psychometrics of narrative. *Journal of Health Psychology*.

Gustafson, L. A., & Orne, M. T. (1963). Effects of heightened motivation on the detection of deception. *Journal of Applied Psychology, 47*, 408-411.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Bulletin, 88*, 67-85.

Kastor, E. (1994, November 5). The worst fears, the worst reality: For parents, murder case strikes at the heart of darkness. *The Washington Post*, p. A15.

Klein, K., & Boals, A. (2001). Expressive writing can increase working memory capacity. *Journal of Experimental Psychology: General, 130*, 520-533.

Knapp, M. L., & Comadena, M. A. (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., Hart, R. P., & Dennis, H. S. (1974). An exploration of deception as a communication construct. *Human Communication Research, 1*, 15-29.

Leets, L., & Giles, H. (1997). Words as weapons: When do they wound? Investigations of harmful speech. *Human Communication Research, 24*, 260-301.

Mehrabian, A. (1971). Nonverbal betrayal of feeling. *Journal of Experimental Research in Personality, 5*, 64-73.

Mergenthaler, E. (1996). Emotion-abstraction patterns in verbatim protocols: A new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology, 64*, 1306-1315.

Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10*, 601-626.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Lawrence Erlbaum.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 6*, 1296-1312.

Pennebaker, J. W., & Lay, T. C. (in press). Language use and personality during crises: Analyses of Mayor Rudolph Giuliani's press conferences. *Journal of Research in Personality*.

Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*, 863-871.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (in press). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*.

Petrie, K. P., Booth, R. J., & Pennebaker, J. W. (1998). The immunological effects of thought suppression. *Journal of Personality and Social Psychology, 75*, 1264-1272.

Richards, J. M., & Gross, J. J. (1999). Composure at any cost? The cognitive consequences of emotion suppression. *Personality and Social Psychology Bulletin, 25*, 1033-1044.

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: The cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology, 79*, 410-424.

Rosenberg, S. D., & Tucker, G. J. (1979). Verbal behavior and schizophrenia: The semantic dimension. *Archives of General Psychiatry, 36*, 1331-1337.

Rude, S. S., Gortner, E. M., & Pennebaker, J. W. (2001). *Language use of depressed and depression-vulnerable college students*. Manuscript submitted for publication.

Ruscher, J. B. (2001). *Prejudiced communication: A social psychological perspective*. New York: Guilford.

Schnake, S. B., & Ruscher, J. B. (1998). Modern racism as a predictor of the linguistic intergroup bias. *Journal of Language & Social Psychology, 17*, 484-491.

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

Shapiro, D. (1989). *Psychotherapy of neurotic character*. New York: Basic Books.

Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and non-suicidal poets. *Psychosomatic Medicine, 63*, 517-522.

Stone, P. J., Dunphy, D. C., Smith, M. S., & Ogilvy, D. M. (1966). *The General Inquirer: A computer approach to content analysis*. Cambridge, MA: MIT Press.

Undeutsch, U. (1967). *Forensische psychologie* [*Forensic psychology*]. Gottingen, Germany: Verlag fur Psychologie.

Vorauer, J. D., & Ross, M. (1999). Self-awareness and feeling transparent: Failing to suppress one's self. *Journal of Experimental Social Psychology, 35*, 415-440.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. Chichester, UK: Wiley.

Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior, 24*, 239-264.

Wiener, M., & Mehrabian, A. (1968). *Language within language: Immediacy, a channel in verbal communication*. New York: Appleton-Century-Crofts.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1-59). New York: Academic Press.

Received February 25, 2002
Revision accepted June 1, 2002