

GRAMMATICAL DIFFERENCES BETWEEN TRUTHFUL AND DECEPTIVE NARRATIVES

Antonio J. DeCicco and John R. Schafer
Western Illinois University

Law enforcement officers rely on the veracity of written statements from witnesses and suspects to determine guilt or innocence. This study examined the predictive value of grammar structures to differentiate truthful written narratives from deceptive written narratives in the English and Spanish languages. Experiment 1 examined three variables amongst English speakers: total word count, text bridge ratio, and spontaneous negation ratio. Experiment 2 replicated the methodology to assess the efficacy of the three variables in predicting veracity in the Spanish language. Participants in experiment 1 and experiment 2 watched a digital presentation of a person shoplifting an item from a convenience store and wrote truthful and deceptive narratives regarding the shoplifting event. The results of the study showed that deceptive narratives contained significantly fewer words, higher text bridge ratios, and higher spontaneous negation ratios than truthful narratives.

Keywords: detecting deception, deception, lie, truth, paralinguistic, written statements, interviewing, statement analysis

Law enforcement investigators rely on the veracity of written statements from witnesses and suspects to ensure the integrity of the investigation. Analyzing written statements can be the difference between justice and a miscarriage of justice. A common obstacle when analyzing written statements is determining the veracity of the statements. This study examines the words and grammar structures that liars use to evade detection. The research also provides a model to determine veracity of written statements.

Liars typically experience guilt and anxiety, which induce cognitive stress and physiological arousal (Gordon & Fisher, 2002; Hirsch & Wolf, 2001; Knapp, Hart, & Dennis, 1974; Kraut, 1980; Miller & Stiff, 1993; Rudacille, 1994; Walters, 2000; Vrij, 2000; Zuckerman, DePaulo, & Rosenthal, 1981). Physiological arousal causes the unintentional leakage of verbal and nonverbal behaviors that may portend deception (DePaulo, et al., 2003; Ekman, 1992; Miller, 1983; Miller & Stiff, 1993). Observing nonverbal and verbal indicators of deception require the establishment of a behavioral baseline during which there is some certainty that no deception is taking place and against which behaviors observed during deception can be compared.

The research supporting verbal and nonverbal cues to deception is inconclusive and, therefore, not very reliable, especially in practical applications (DePaulo et al., 2003;

Author Note: Correspondence concerning this article should be addressed to Antonio DeCicco, Department of Law Enforcement and Justice Administration, Western Illinois University, Macomb, IL 60103. E-mail: AJ-DeCicco@wiu.edu

Hazlett, 2006; Vrij, 2000). DePaulo et al. (2003) conducted a meta-analysis of 1,338 estimates of 158 nonverbal and paralinguistic cues contained in 120 deception studies. The results of the meta-analysis showed that liars were more nervous than people who told the truth, made more negative impressions than did truthful people, and included fewer announced flaws and more unique content in their communications. The meta-analysis also showed that an increase in the motivation to lie amplified nervousness and inhibition.

Like nonverbal cues to detect deception, paralinguistic cues are poor predictors of deception. Paralinguistic cues include the tonality, inflection, latency, pitch, tension, and duration of speech (Hazlett, 2006). When liars speak, they tend to use a higher voice pitch (DePaulo et al., 2003); however, the change in voice pitch is difficult to detect without the aid of specialized equipment. The remaining paralinguistic deception cues are not good predictors of mendacity (DePaulo et al., 2003).

Content-Based Criteria Analysis (CBCA)

In 1954, at the behest of the West German Supreme Court, psychologists developed Statement Validity Assessment (SVA) to determine the veracity of victims of child abuse. SVA was founded on a hypothesis formulated by Undeutsch, who postulated that statements based on actual memories differ from statements based on fabrication or fantasy (Vrij, 2000). SVA consists of three components: a structured interview, a criteria-based content analysis (CBCA), and an evaluation of the CBCA outcome using the Validity Check-list (Vrij, 2000). Investigators conduct a structured interview of a child using a series of open-ended questions. The transcribed interview is then analyzed using CBCA. CBCA consists of 19 criteria divided into 4 categories: a) general characteristics, which include the criteria logical structure, unstructured production, and quality of details; b) specific content, which includes the criteria contextual embedding, descriptions of interactions; reproduction of conversation; unexpected complications during the incident; unusual details; superfluous details; accurate reports of details, which may have been misunderstood; the relation of external associations; accounts of subjective mental state; and attribution of perpetrator's mental state; c) motivation, which includes the criteria spontaneous corrections; the admission of the lack of memory; doubts raised about one's own testimony; self-depreciation; and the pardon of the perpetrator; and d) offense-specific elements, which include the criterion details characteristic of the offense (Vrij, 2000).

Trained evaluators review statements and judge the presence or absence of the 19 criteria, which are scored on a 3-point scale -- 1 point for the presence of the criteria, 2 points if the criteria are strongly present, and 0 if the criteria are absent. The CBCA is designed to judge statements on a truthful continuum, not a deception continuum. A score of 38 indicates that all the criteria are present, which increases the probability that the statement is true.

The CBCA score then subjected to the Statement Validity Check-list (SVC), which consists of 11 items divided into 4 categories: a) psychological characteristics, which include the criteria of inappropriateness of language and knowledge, inappropriateness of affect, and susceptibility to suggestion; b) interview characteristics, which include the crite-

ria of suggestive, leading or coercive questioning, and overall inadequacy of the interview; c) motivation, which includes the criteria of questionable motives, questionable context of the original disclosure or report, and pressures to report falsely; and d) investigative questions, which include the criteria inconsistency with the laws of nature, inconsistency with other statements, and inconsistency with other evidence.

The CBCA initially was designed to test the truthfulness of statements from children; however, CBCA has been modified and tested on adult populations. The results of this research tend to support the efficacy of CBCA to discriminate between truthful statements and deceptive statements in adult populations (Dana-Kirby, 1997; Landry & Brigham, 1992; Porter & Yuille, 1996; Ruby & Brigham, 1998; Vrij, 2000; Zaparniuk, Yuille, & Taylor, 1995). Vrij (2000) conducted a review of 37 studies testing the reliability of CBCA and found that CBCA was not sufficiently reliable for court proceedings; however, he noted that CBCA could be used as an effective investigative tool.

Theoretical Verbal Analysis

Theoretical Verbal Analysis (TVA) identified 195 criteria distributed over 6 categories that discriminate truthful communications from deceptive communications. The 6 categories are a) autobiographical memory, b) cognitive processes, c) emotional processing and expression, d) defense mechanisms, e) linguistics, and f) self-presentation and attribution. Germane to this study are TVA's truthful and deceptive criteria. The truthful criteria include actions explained in terms of the "big picture;" active language when relating success; the communication of statements in a logical orderly fashion; inclusion of unusual details; focus on events, inclusion of details when relaying a story; the justification of facts with past experiences; the use of active language; the provision of less information about specific interactions with people; the provision of enough information for general understanding; responses that include combinations of action and state verbs; and the use of interpretive action verbs when relaying actions (Connelly et al., 2006).

The deceptive criteria include answers to questions other than the ones asked; vague descriptions of events or situations; the lack of discussion of cause and effect; the frequent use of adjectives such as honest and actually; the use of more state verbs; the infrequent use of adverbs; the interpretation of events through the people involved; a large presence of passive language; less reference to situational, social, or contextual variables shaping behavior; limitations in the use of self-references; no discussion of emotional states; no explanation of behavior; the lack of convincing in relaying outcome; the use of passive language when relaying failure; the use of formal language; statements that are too long or too short based on demands of the interview situation; statements focusing on objective actions not cognitive states; and statements focusing on objective actions and not evaluations of actions (Connelly et al, 2006).

The CBCA and TVA may be effective methods for evaluating the veracity of written communications; however, both systems have several drawbacks. First, CBCA and TVA cannot be used effectively in real-time communications due to the extensive analysis necessary to determine truthfulness or deception. Second, the last two SVA checklist

criteria, consistency with other statements and inconsistency with other evidence, require extensive investigation. Under exigent circumstances when investigators must determine veracity with minimal time and information, CBCA and TVA are not useful techniques to determine veracity due to the lengthy evaluation process. Likewise, TVA requires a lengthy examination of written statements using 31 criteria. Third, in both CBCA and TVA, the criteria are subjective and more than one judge is required to make a determination of veracity. The amount of training necessary to make a judge competent in CBCA or TVA criteria is not standardized. Judges who are not competent in judging CBCA and TVA criteria may make inaccurate decisions as to the veracity of written statements.

Reality Monitoring

Reality Monitoring (RM) was developed to study memory characteristics (Johnson & Raye, 1998). The RM consists of eight criteria: clarity, perceptual information, spatial information, temporal information, affect reconstructability of the story, realism, and cognitive operations. The premise of RM is that real memories are based on perceptual information, contextual information, and affective information (Colwell, 2002; Zhou, Twitchell, Qin, Burgoon & Nunamaker, 2003). In other words, memories contain physical cues such as visual details, smells, and tastes; contextual information, such as when, where, and how an event took place; and affective information, such as personal feelings that occurred during the event. Conversely, fabricated stories derive primarily from cognitive operations and tend to be vague and less detailed (Colwell, 2002; Johnson & Raye, 1998).

Limited experimental studies have been conducted to determine the validity of RM in discerning truthful statements from deceptive statements. Using the RM criteria, Vrij (2000) obtained an accuracy rate of 71% for detecting truthful statements and 74% for identifying deceptive statements. Hofer, Akehurst, and Metzger (cited in Vrij, 2000) obtained an accuracy rate of 61% for detecting truthful statements and an accuracy rate of 70% for detecting deceptive statements. Sporer (1997) obtained an accuracy rate of 75% for detecting truthful statements and an accuracy rate of 65% for detecting deceptive statements.

Linguistic Inquiry and Word Count (LIWC)

Newman, Pennebaker, Berry, and Richards (2003) developed the Linguistic Inquiry and Word Count (LIWC), a linguistic computer software program that analyzes written narratives for mendacity. LIWC is a linguistic database with more than 2,000 words divided into 72 categories used to examine written text for veracity. A developing profile for deceptive statements based on the LIWC suggests that liars tell stories that are less complex, use fewer self-relevant terms, and are more negatively charged. As more studies are conducted using LIWC, a more definitive profile of how liars use language to deceive will become more apparent.

Scientific Content Analysis

Scientific Content Analysis (SCAN) examines written statements based on their internal consistency and does not rely on outside reality to determine veracity as does CBCA and TVA. The SCAN supposes three conditions: the truth, a mistake, and a lie. The truth reflects consistency between the writer's subjective knowledge, the writer's

statement, and outside reality (Sapir, 1996). A mistake reflects consistency between the writer's subjective knowledge and the writer's statement, but is inconsistent with outside reality. A lie reflects inconsistency among the writer's subjective knowledge, the writer's statement, and outside reality. The SCAN does not discern truthful statements from deceptive statements but rather highlights areas in written narratives that may indicate deception and need further clarification.

The efficacy of SCAN has been studied with mixed results (Adams, 1996; Driscoll, 1994; Lesco, 1990; McClish, 2001; Rabon, 1994; Rudacille, 1994;; Smith, 2001; Shearer, 1999). Smith (2001) conducted a study to determine if SCAN is a viable method to distinguish truthful from deceptive written narratives. Smith (2001) concluded that investigators who are experienced in the use of SCAN improved their abilities to detect deception; however, he also found that experienced assessors did not use a standardized method to apply SCAN criteria; thus, subsequently, weakening the causation of the method.

Driscoll (1994) enlisted 30 participants, 25 males and 5 females, to write truthful and deceptive narratives. Driscoll used SCAN to assess the narratives for consistencies that portend deception. The study showed that the SCAN predictor variables: unnecessary connections, violation of the first-person simple past tense formula, failure to deny the accusation, and changes in language, were indicators of deception. Driscoll found the variable failure to deny the accusation as the best predictor of deception.

Chang (2008) examined 125 narratives written by suspects, victims and witnesses from the Lincoln Police Department and Nebraska State Patrol. Chang concluded that six SCAN sequences were the most pertinent in determining veracity of written statements: improper use of pronouns, use of connections, information out of sequence, unimportant information, denial of allegations, and quoted discourse. Chang labeled "use of connections" as a possible way to omit information in written narratives. In furtherance of the current literature, we are interested in the connections people use and whether they use these connections to omit information about an event.

Investigative Discourse Analysis (IDA)

Rabon (1994) extended the principles of SCAN with IDA, a more organized approach to content analysis. The premise of IDA is similar to SCAN in that individual word selections are important and have meaning to the writer. Rabon posited that truthful people select certain words to convey information to the reader, while deceptive people select certain words to influence the perception of the reader rather than convey factual information.

Suiter (2001) compared the efficacy of IDA to CBCA. Suiter asked 46 college freshmen to write both true statements and false statements about events in their days. Suiter subjected these statements to IDA and CBCA criteria. The study determined that false statements use more abjuration terms than truthful statements. Rabon (1994) defined abjuration as words that withdrew the assertion previously made. Abjuration words include, *but, yet, however, although, nevertheless, though, and anyway*. Suiter (2001) found that writers who wrote false statements used abjuration words twice as often as writers who wrote truthful statements, lending some support to the IDA criteria and, by extension, the

SCAN criterion of missing connections. Suiter (2001) also found that the use of the word *when* was prevalent in deceptive statements, indicating a temporal lacuna. Other words that indicated temporal lacunae include *after that*, *afterwards*, *at that time*, *at that point*, *later on*, *by the time*, and *all of a sudden*. Suiter cautioned that a temporal lacuna, in and of itself does not indicate deception. However, this finding does support Rabon's (1994) notion of temporal lacuna and Sapir's (1996) notion of missing connections.

The advantage of studying grammar structures is that in both stressful and non-stressful environments, deceptive people and truthful people use the same grammar rules to construct sentences. The only difference between truthful narratives and deceptive narratives is the omission or obfuscation of the truth. Words are the building blocks for sentences, and grammar rules serve as blueprints for sentence construction. The stability of grammar rules within language, including English and Spanish, provides a more stable platform to study the similarities and differences of grammar structures in truthful and deceptive conditions.

Text Bridges

Text bridges constitute grammatical structures that circumvent withheld information. The proposed study takes a unique approach to detecting deception. Instead of measuring verbal and nonverbal cues triggered by physiological changes, this study examined the grammar structures people use during deception. Most liars do not fabricate entire stories, but rather tell the truth up to the point where they want to conceal information, skip over the information they want to conceal, and continue to tell the truth (Ekman, 1992). At the point liars want to omit information, they use a connector or a "text bridge" to conceal the activity. Grammatically bridging withheld information is similar to bridging a river. A road is laid up to the river's edge and stops. A bridge spans the river, and the road continues on the other side of the river. Bridges come in a variety of designs, but each design must adhere to specific construction standards, or the bridge will collapse. Likewise, sentence construction must follow certain grammatical rules. Deceptive people who lie by omission must use grammar structures that allow them to construct a series of sentences that circumvent or bridge the information that deceptive people desire to withhold. This notion comports with the findings of Sapir (1996), Driscoll (1994), and Chang (2008).

Text bridges include adverbial conjunctives, transitional words, and subordinating words, which are grammatical devices used to smoothly transition from one idea to another idea and from one sentence to another sentence (Forlini et al., 1990; Llorach, 2003). Sentence construction must follow certain grammatical rules. Identifying the grammatical structures isolate the portion of the narrative that contains withheld information. Using grammatical rules to detect withheld information provides a standard measure against which to record changes in truthful and deceptive narratives. A review of English grammar has identified three grammar structures that function as text bridges: adverbial conjunctives, transitional words, and subordinating words (Forlini et al., 1990).

The same categories of text bridges exist in Spanish-Castellano (Llorach, 2003). English sentences divide into four basic types: declarative sentence, the imperative sen-

tence, the interrogative sentence, and the exclamatory sentence. Again, the same exist for Spanish (Llorach, 2003). The current study focused on declarative sentences in English and in Spanish.

A simple declarative sentence contains a subject and a verb, for example, in English one might say, “We went to the beach” or, in Spanish, “Fuimos a la playa.” The two sentences contain a subject and a verb making them simple declarative sentences. A compound declarative sentence contains two or more independent clauses (Forlini et al., 1990; Llorach, 2003), for example, “John read a book, and he wrote an essay” in English or, “John leyó un libro, y escribió un ensayo” in Spanish. The two sentences have two independent clauses connected by “and.” A complex declarative sentence consists of one independent clause and one or more subordinate clauses (Forlini et al., 1990; Llorach, 2003), An example of a complex declarative sentence in English would be, “Tom answered the phone when his wife called” or in Spanish, “Tom contestó el teléfono cuando su esposa llamó.” The word, “when” and “cuando” connect the subordinate clause to the independent clause.

Complex declarative sentences create temporal-spatial lacuna. For example, the complex declarative sentence “When I came home from work, I found my wife dead,” a temporal-spatial lacuna is created between the subordinate clause, “When I came home from work” and the main clause, “I found my wife dead.” The writer does not explain what took place from the time he arrived home until the time he found his wife dead. The text bridge, “when” created this temporal-spatial gap. The use of a text bridge does not indicate deception only that the writer intentionally or unintentionally withheld information.

Subordinate clauses connect unequal but related ideas to form complex sentences (Forlini et al., 1990). Words such as: *after, although, as if, as long as, because, before, even though, if, in order that, since, so, that, than, through, unless, until, when, whenever, where, wherever, and while* create temporal-spatial lacunae, which liars can use to conceal information. Again, truthful people can use subordinating clauses as behavioral contractions to intentionally or unintentionally withhold information because they consider the information mundane or irrelevant to the inquiry.

Transitional and adverbial conjunctives have the same effect as subordinating conjunctives. Transitional words connect themes and ideas or establish relationships (Forlini et al., 1990; Llorach, 2003). In the sentence, “It was sunny on Saturday, so we went to the beach,” the transitional word, “so” connects two ideas. Transitional words consist of words such as: *after, afterward, before, during, earlier, final, first, later, since, meanwhile, then, until, however, in contrast, indeed, instead, nevertheless, on the contrary, yet, on account of, so, therefore, also, besides, and furthermore*, etc. Adverbial conjunctives connect two complete ideas (Forlini et al., 1990; Llorach, 2003). Adverbial conjunctives are words such as: *accordingly, again, also, besides, consequently, finally, furthermore, however, indeed, moreover, nevertheless, otherwise, then, therefore and thus*. These three English and Spanish grammar structures create temporal-spatial lacunae.

Liars Use Fewer Words

Research has shown that liars use fewer words and their statements contain less relevant information (Connelly et al., 2006; Dulaney, 1982; DePaulo et al., 2003; DePaulo, Stone & Lassiter, 1985; Ekman, 1992; Kraut, 1980; Knapp et al., 1974; Miller & Burgoon, 1981; Rabon, 1994; Sapir, 1996; Suiter, 2001; Vrij, 2000; Zuckerman et al., 1981; Zuckerman & Driver, 1985). Liars tend to use fewer words because a story with fewer words is easier to remember (De Paulo et al., 2003; DePaulo et al., 1985; Dulaney, 1982; Knapp et al., 1974; Kraut, 1980; Rabon, 1994; Sapir, 1996; Suiter, 2001; Zuckerman & Driver, 1985). Liars tend to keep descriptors to a minimum because they did not experience the event firsthand and possess fewer facts about the topic of their deception (Connelly et al., 2006; Ekman, 1992; DePaulo et al., 2003; Rabon, 1994; Sapir, 1996; Suiter, 2001; Vrij, 2000). In the event liars are asked to repeat their stories, fewer facts are easier to command than stories filled with a large amount of fabricated details (Vrij, 2000; Ekman, 1992; DePaulo et al., 2003).

Spontaneous Negations

Negations are defined by words such as *no*, *not*, and all contractions of *not* (Adams & Jarvis, 2006). The concept of negations divides into two categories, negations and spontaneous negations. In response to the closed-ended question, "Did you rob the bank?" a deceptive person as well as a truthful person would answer, "No, I did not rob the bank." A negative answer to a direct question is classified as a negation, not a spontaneous negation. When people are presented with open-ended questions, they should relate the actions they took versus the actions they did not take (Sapir, 1996; Rudacille, 1994). For example, when a tourist was asked to list the states he visited during a recent trip to America, he replied, "I visited California, Utah, and Texas, but I never visited New York." The phrase "but I never visited New York" constitutes a spontaneous negation. Of the 47 states the tourist did not visit, he singled out New York as a state he did not visit for a reason. The person making the inquiry must determine the reason why the tourist mentioned New York as a state he did not visit. Commonly used spontaneous negations include "I don't mean to interrupt," "I'm not trying to be obnoxious," and "I don't mean to rain on your parade."

Spontaneous negations serve the same function as text bridges allowing a deceptive person to bridge the information gap. When deceptive people use spontaneous negations, they fail to state what specific actions they did take. Spontaneous negations used during open-ended questions may provide additional cues to differentiate truthful narratives from deceptive narratives, especially when they are used in conjunction with text bridges.

Extant studies identified some potential differences between deceptive and truthful narratives, but they have not definitively defined a model to differentiate the two. The current study examined the efficacy of three variables, total words, text bridge ratio, and spontaneous negation ratio, to collectively predict the veracity of written statements in English and Spanish. We posit the following hypotheses:

Hypothesis 1: Deceptive written narratives contain fewer words than do truthful written narratives.

Hypothesis 2: Text bridge ratios of deceptive written narratives will be higher than text bridge ratios of truthful written narratives.

Hypothesis 3: Deceptive written narratives contain higher spontaneous negation ratios than truthful written narratives.

EXPERIMENT 1

Participants

Participants were drawn from law enforcement and military service populations. Since law enforcement and military personnel must demonstrate minimum literacy skills upon employment, these populations ensured that the participants of this study are able to read and write in English. There were 304 participants, 77% males and 23% females, ranging in age from 20 to 61.

Procedures

Participants were asked to read and sign a consent form. After completing the consent forms, the participants were given an instruction sheet requesting the following demographic information: age, gender, education level, and ethnicity. The instruction sheets also contained instructions for writing the narratives. Participants were divided into two groups, Group I and Group II. Both groups watched a digital video of a shoplifting event that occurred in a convenience store. Both groups were instructed to pretend they were the person depicted in the digital video and write a truthful and deceptive narrative describing their actions in the store. Group I participants were instructed to write their truthful narratives first and then write their deceptive narratives. Group II participants were instructed to write their deceptive narratives first and then write their truthful narratives. To compensate for individualistic writing styles, education levels, age, ethnicity, and gender, each person served as his/her own control. The grammar structures from the truthful and deceptive narrative were compared and objectively measured.

Measures

The independent variable condition has two levels: truthful and deceptive. The dependent variables are total words, text bridge ratio, and spontaneous negation. Due to colinearity between the dependent variables total words and text bridge ratio, the variable spontaneous negation will be expressed as a ratio. The truthful and deceptive narratives were scored as follows: the total number of words was calculated using the word count feature in Microsoft Word; text bridge ratio was calculated by dividing the total number of text bridges by the total number of words; and spontaneous negation ratio was calculated by dividing the total number of spontaneous negations by the total number of words.

Results

An examination of the means indicated that deceptive narratives contained fewer words ($M = 44.11$, $SD = 21.08$), higher text bridge ratios ($M = .035$, $SD = .040$), and higher spontaneous negation ratios ($M = .0191$, $SD = .026$), as compared to the total words in truthful narratives ($M = 59.74$, $SD = 22.34$), the text bridge ratio in truthful narratives ($M =$

.026, $SD = .019$), and the spontaneous negation ratio in truthful narratives ($M = .0009$, $SD = .004$). Means and Standard deviation can be found in table 1.

Table 1
Means and Standard Deviations (N=304)

	Mean	Standard Deviation
<u>Total Words</u>		
Truthful	59.74	22.34
Deceptive	44.11	21.08
<u>Text Bridge Ratio</u>		
Truthful	.026	.019
Deceptive	.035	.040
<u>Spontaneous Negation Ratio</u>		
Truthful	.0009	.004
Deceptive	.0191	.026

A repeated measures ANOVA with a Greenhouse-Geisser correction was conducted to test significant differences between the variables total words, text bridge ratio, and spontaneous negation ratio in truthful and deceptive narratives. The repeated measures ANOVA revealed that there were significant differences between truthful and deceptive narratives (Greenhouse-Geisser = 1092104.89, $F(1) = 3471.19$, $p < .001$).

Moreover, the repeated measures ANOVA showed that deceptive narratives contained significantly fewer words than truthful narratives ($t = 8.87$, $p < .001$); deceptive narratives had significantly higher text bridge ratios than did truthful narratives ($t = -3.53$, $p < .001$); and deceptive narratives contained significantly higher spontaneous negation ratios than did truthful narratives ($t = -11.81$, $p < .001$). These results can be found in table 2.

Table 2.
Repeated Measures ANOVA for the Variables Total Words, Text Bridge Ratio, and Spontaneous Negation Ratio (N=304)

Variables	B	SE	t	p	eta
Total words	15.63	1.76	8.87	.000	.33
Text bridge ratio	-.009	.003	-3.53	.000	.14
Spontaneous negation ratio	-.018	.002	-11.81	.000	.43

A logistic regression was conducted to assess whether the three variables: total words, text bridge ratio, and spontaneous negation ratio significantly predicted the veracity in written statements. When computed together, the three variables significantly predicted veracity in written narratives, Omnibus Tests of Model Coefficients $X^2 = 291.67$, $df = 3$, $N = 304$, $p < .001$. The r square statistic cannot be exactly computed for logistic models, so the Nagelkerke pseudo r squared was computed. The Nagelkerke r square .508 for model containing the three variables total words, text bridge ratio, and spontaneous negation ratio, which indicates that this model explains 51% of the variance. The odds ratio for the three variables total words, text bridge ratio, and spontaneous negation ratio are .97, 1.16, and 4.90 respectively (Table 3). Total words, text bridge ratio, and spontaneous negation ratio correctly predicted truthful narratives 89% of the time and deceptive narratives 67% of the time.

Table 3.

Logistic Regression predicting deceptive written narratives using the variables Total words, Text bridge ratio, and Spontaneous negation ratio (N=304)

Variables	B	SE	Odds ratio	p
Total words	-.032	.005	.97	.000
Text bridge ratio	.155	.048	1.16	.001
Spontaneous negation ratio	1.15	.186	4.90	.000

Discussion

The results of the experiment showed that the combination of the three variables, total words, text bridge ratio, and spontaneous negation ratio, correctly predicted deceptive written narratives 67% of the time and truthful written narratives 89% of the time. The total number of different text bridges used in both truthful and deceptive narratives was 17. The most commonly used text bridges in truthful narratives were *then*, *so*, *after*, *when*, *as*, *while*, *once*, and *next*. The remaining text bridges were used three times or less. The most commonly used text bridges in deceptive narratives were *then*, *so*, *after*, and *when*. The remaining text bridges were used three times or less.

The results were deemed significant in English, but a question arises. Do grammatical differences exist in other languages or are these findings limited to English speakers? To test these questions, we recruited individuals whose native language is Spanish.

EXPERIMENT 2

Participants

Participants were drawn from La Universidad de Veritas, a Costa-Rican university. Participants consisted of faculty, maintenance personnel, and students who were fluent in spoken and written Spanish. There were a total of 103 participants (22 participants were omitted due to failure to follow procedure), 38% of participants were male, and 62%

of participants were female ranging in age from 13-60. Ethnicity in what we deem the Hispanic culture is very different. Ethnicity is not viewed the same way in comparison to Western culture; Hispanic cultures often do not use labels pertaining to ethnicity to label themselves. However, because ethnicity is a common demographic question amongst western cultures, it was included in our preliminary demographic questionnaire. Seventy-nine percent of participants classified themselves as Hispanic, 14% were Mestizo, indigenous to Costa-Rica, 5% were Caucasian, less than 1% were of Asian descent, less than 1% were of African descent, and less than 1% were of “other” ethnicity. Ninety-six percent of participants indicated they had a high school diploma or higher.

Procedures

This experiment examined the predictive value of grammatical structures to differentiate truthful and deceptive narratives written in Spanish by native Spanish speakers from a university located in Costa Rica, who primarily speak the Castellano dialect of the Spanish language. Adopting the same methodology from experiment 1, this experiment examined the variables, total words, text bridge ratio, and spontaneous negation ratio to determine if these variables serve as predictors of veracity in native Spanish speakers.

Results

Repeated one-sample *t*-tests was used to assess the differences between the total number of words, text-bridge ratios, and spontaneous negations in both truthful and deceptive narratives. Means and standard deviations are located in Table 4. Examination of these means suggests that deceptive narratives contain fewer words ($M = 48.43$, $SD=19.07$), higher text bridge ratios ($M = .03$, $SD = .022$), and higher spontaneous negation ratios ($M=.037$, $SD=.028$) than truthful narratives, $M = 51.99$, $SD = 20.36$, $M = .028$, $SD = .024$, and $M = .021$, $SD = .028$, respectively.

Table 4

Means and Standard Deviations (N=81)

	Mean	Standard Deviation
<u>Total Words</u>		
Truthful	51.99	20.36
Deceptive	48.43	19.07
<u>Text Bridge Ratio</u>		
Truthful	.028	.024
Deceptive	.030	.022
<u>Spontaneous Negation Ratio</u>		
Truthful	.021	.023
Deceptive	.037	.028

As seen in Table 5, deceptive narratives ($t(80) = 22.12; p < .0001$) are significantly different than truthful narratives ($t(80) = 22.98; p < .0001$). Text bridges were used at a different rate between truthful narratives ($t(80) = 11.04; p < .0001$) and deceptive narratives ($t(80) = 13.32; p < .0001$). Lastly, spontaneous negations also were used at a different rate in truthful narratives ($t(80) = 8.11; p < .0001$) and deceptive narratives ($t(80) = 11.80; p < .0001$).

Table 5.

Repeated Measures One Sample t-text for the Variables Total Words, Text Bridge Ratio, and Spontaneous Negation Ratio (N= 81)

Variables	Mean	t	p
<u>Deceptive Narrative</u>			
Total Words	48.43	22.12	.000
Text bridge ratio	.030	13.32	.000
Spontaneous negation ratio	.037	11.80	.000
<u>Truthful Narrative</u>			
Total Words	51.99	22.98	.000
Text bridge ratio	.028	11.04	.000
Spontaneous negation ratio	.021	8.11	.000

Discussion

Our speculation that total words, text bridge ratio, and spontaneous negation ratio would be able to delineate between a truthful and deceptive narrative in Spanish was confirmed. Both English speakers and Spanish speakers use text bridges and spontaneous negations at a higher rate when lying than when telling the truth. Also, both studies show liars use fewer words as was indicated in the literature (Vrij, 2008). The most common text bridges that were used in the Spanish written narratives were: *luego, después de, también, cuando, primero, aunque, donde, así (que), hasta, entonces, pues, como no..., antes de, por lo que, eguidamente, sin embargo, mientras, además, para que, por lo tanto, para no, and en la cual*. Please see Table 6 for translations. The two main types of text bridges used were transitional and subordinating conjunctions (Llorach, 2003).

Table 6

Spanish to English Translation of Text-bridges

Spanish	English Translation
Luego	Then/Later
Después (de)	After
Tambien	Also
Pues	Since/then/well
Como no...*	On account of
Antes (de)	Before
Por lo que*	Thus
Seguidamente	Then/the following
Sin embargo	However
Mientras	While
Además	Also/besides
Para que*	So that/in order to
Por lo tanto*	Therefore/hence

Note: * Indicates meaning of word or phrase is based on context as used in a sentence

GENERAL DISCUSSION

The results of this study developed a predictive model for veracity in the English and Spanish languages using the three variables: total words, text bridge ratio, and spontaneous negation ratio. The predictive model is not suggested to be used solely to determine veracity in written narratives. The three variables predict truthful and deceptive narratives at a rate higher than the current model of nonverbal indicia, where the credibility continues to be slightly about the 50th percentile (Connelly et al., 2006; DePaulo et al., 2003; Vrij, 2000).

The variable, total number of words, in a statement, supports the previous literature that liars use fewer words. (Connelly et al., 2006; Dulaney, 1982; DePaulo et al., 2003; DePaulo et al., 1985; Ekman, 1992; Kruat, 1980; Knapp et al., 1974; Miller & Burgoon, 1981; Rabon, 1994; Sapir, 1996; Suiter, 2001; Vrij, 2000; Zuckerman et al. 1981; Zuckerman & Driver, 1985). Again, liars use fewer words so they may be able to recall the information given more readily than if the information was fabricated (Vrij, 2000; Ekman, 1992; DePaulo et al., 2003).

Although all three variables were deemed significant in Spanish, there are differences from the English version of this study. When the experiment was conducted in English, spontaneous negations were the most powerful predictor for veracity in written narratives. In the English version, deceptive narratives were 94% more likely to have a spontaneous negation than in the truthful narrative. In the Spanish results, participants were 30% more likely to use spontaneous negations in their deceptive narratives than in their truthful narra-

tives. The variance is due to the nature of the two idioms. The English language rarely uses negations in oral and written narratives in comparison to the Spanish language. In Spanish, negations are more commonly used. Nonetheless, spontaneous negations were still more prevalent in deceptive narratives written in Spanish than in truthful narratives in Spanish.

Based on the current research, people use spontaneous negations more often because of the psychological changes they undergo when they lie (DePaulo et al., 2003). These psychological changes include: increased skin conductance, increased blood pressure, and increased respiration, which reflect an increase in general anxiety (DePaulo et al., 2003). Therefore, liars use spontaneous negations to combat or inhibit the anticipated increase in anxiety; however, the use of spontaneous negations to decrease anxiety is not supported by any current research.

The use of text bridges had similar results as to the English predictive model version and nonverbal predictive model. In both studies text bridges were used at a higher rate in deceptive narratives than in truthful narratives. Although the rate was found significant, it still only had a predictive value of slightly above the 50th percentile suggesting it is as reliable as the current nonverbal predictive model (Connelly et al., 2006; DePaulo et al., 2003; Vrij, 2000). These results still corroborate Sapir's (1996) notion of missing connections and Rabon (1994) and Suiter's (2001) notions of temporal lacunae. Text bridges signal withheld information. Withheld information does not necessarily mean the information was deceptive, but the writer may think it to be irrelevant to the inquiry, thus unintentionally creating gaps in their written narratives (Suiter, 2001).

LIMITATIONS

The absence of a comparative narrative in practical situations remains to be the principal limitation to the predictive model. Using the predictive model of total words, text bridge ratio, and spontaneous negation ratio is more difficult without comparative values.

Another limitation is that only one portion of the narrative was examined. Written narratives typically contain three components: a prologue, the body (or the incident), and an epilogue. The body or description of the event is typically the most critical component of narratives because the body contains a description of the event or the focus of the inquiry. The variables total words, text bridge ratio, and spontaneous negation ratio may be more effective when used to examine the separate component parts of a narrative body. Using the variables total words, text bridge ratio, and spontaneous negation ratio to analyze the body of a narrative written in isolation may enhance the predictive value of the total number of words, text bridge ratios, and spontaneous negations. The present study consisted of short narratives that only asked the participants to describe the shoplifting event, which would naturally comprise the body of the narrative. Another approach might be to use the prologue and the epilogue as baseline indicators against which to examine the body of the narrative. Additional research is required to confirm the effectiveness of analyzing the body of narratives using total words and text bridge ratio to evaluate veracity.

Experiment 2 is comprised of data collected from la Universidad de Veritas. La Universidad de Veritas is located in San Jose, Costa Rica. In Costa Rica, the dialect most commonly used is the Castellano dialect. The provided sample size does not contain representation of other Spanish dialects. A distinction can be made from one culture to the next, although each may be a descendent of the Spanish language. The way citizens of the United States communicate is different from the way citizens of Britain communicate, even though both languages are labeled, "English." Dialects control the way in which individuals communicate, but we predict, based on Chomsky's theory of consistency in speech (1972), the text bridges used would not vary drastically in different languages or dialects. Instead, one should focus on the fact the writer used a text bridge instead of which actual text bridge word was used.

Cultural barriers seemed to create reluctance to participation; therefore, the sample size was less than initially expected. Research studies are not a common practice in Costa Rica and, because of the customary differences, the cultural barrier hindered the experiment from the targeted sample size of 200 participants.

Future Directions

The current literature on linguistic analysis is limited. Although, the literature that does exist suggests predictive models for deception apprehension can be developed using the stable platform of language. Deception apprehension and the ability to identify truthful indicia are necessary in many settings, especially in legal settings. If the current predictive models can be further developed, they will be able to provide a new technique to aid in the decision of culpability in those legal settings.

Cultural diversity is a prominent aspect of the United States. The term, "melting-pot" is a common description of the United States. The melting-pot includes a large array of languages that are crucial to cultural identity. Spanish-Castellano dialect is just one of the languages that are used most commonly in the United States. Several other idioms also exist in the United States, which mean those working in the legal system must also accommodate those who may speak another language. Legal advocates strive to recreate the event in question, in order to make justified judgments. Oftentimes the most precise way to express occurrences accurately are hindered by cultural barriers through the use of language. Therefore, extending current literature to the extensive quantity of idioms is beneficial in the accuracy of detecting truth from deception.

CONCLUSIONS

The three-variable deception model developed in this study performed as well as and, in some instances, better than the predictive value of nonverbal deception indicia (Connelly et al., 2006; DePaulo et al., 2003; Vrij, 2000). This study demonstrated that studying grammar structures may be an effective method to discern truthful written narratives from deceptive written narratives. The results of this study also provide a new avenue for deception research.

REFERENCES

- Adams, S. H. (1996). Statement analysis: What do suspects' words really reveal? *FBI Law Enforcement Bulletin*, 65, 12-20.
- Adams, S. H., & Jarvis, J. P. (2006). Indicators of veracity and deception: An analysis of written statements made to police. *Language and the Law*, 13, 1-22.
- Chang, G. H. (2008). Effectiveness of content analysis in assessing suspect credibility: Counterterrorism implications. (Unpublished doctoral dissertation). The College at the University of Nebraska, Lincoln, NE.
- Chomsky, N. (1972). *Language and mind*. New York: Harcourt Brace Jovanovich.
- Connelly, S., Allen, M. T., Ruark, G. A., Kligyte, V., Waples, E. P., Leritz, L. E., & Mumford, M. D. (2006). Exploring content coding procedures for assessing truth and deception in verbal statements, year 3, cumulative final report. Norman, OK: University of Oklahoma.
- Colwell, K. W. (2002). The whole truth: Structured interview protocols and reality monitoring criteria. (Unpublished doctored dissertation). Sam Houston State University, Huntsville, TX.
- Dana-Kirby, L. (1997). Discerning truth from deception: Is criteria-based content analysis effective with adult statements? (Unpublished doctored dissertation). University of Oregon, Eugene, OR.
- DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129, 74-118.
- DePaulo, P. J., Stone, J. I., & Lassiter, G. D. (1985). Deceiving and detecting deceit. In B. R. Schlenker (Ed.), *The self and social life* (pp 323-370). New York: McGraw-Hill.
- Driscoll, L. N. (1994). A validity assessment of written statements from suspects in criminal investigations using the SCAN technique. *Police Studies*, 4, 77-88.
- Dulaney, E. F. (1982). Changes in language behavior as a function of veracity. *Human Communication Research*, 9, 75-82.
- Ekman, P. (1992). *Telling lies: Clues to deceit in the marketplace, marriage and politics*. New York: W.W. Norton.
- Forlini, G., Bauer, M. B., Beiner, L., Capo, L., Kenyon, K. M., Shaw, D. H., & Verner, Z. (1990). *Grammar and composition*. Englewood Cliffs, NJ: Prentice Hall
- Gordon, N. J. & Fisher, W. L. (2002). *Effective interviewing & interrogation techniques*. New York: Academic Press.
- Hazlett, G. (2006). Research on detection of deception: What we know vs. what we think we know. In R. Swenson (Ed.), *Educating information - interrogation: Science and art* (pp. 45-62). Washington, DC: National Defense Intelligence College Press.
- Hirsch, A. R., & Wolf, C. J. (2001). Practical methods for detecting mendacity: A case study. *The Journal of the American Academy of Psychiatry and the Law*, 29, 438-444.
- Johnson, M. K., & Raye, C. L. (1998). False memories and confabulation. *Trends in Cognitive Science*, 2, 137-145
- Knapp, M. L., Hart, R. P., & Dennis, H. S. (1974). An exploration of deception as a communication construct. *Human Communication Research*, 1, 15-29.
- Kraut, R. (1980). Humans as lie detectors: Some second thoughts. *Journal of Communication*, 30, 209-216.
- Landry, K., & Brigham, J. C. (1992). The effects of training in criteria-based content analysis on the ability to detect deception in adults. *Law and Human Behavior*, 16, 663-675.
- Llorach, E. A. (2003). *Gramàtica de la lengua Espanola*. Madrid, ES: Editorial Espas Calpe.
- McClish, M. (2001). *I know you are lying*. Winterville, NC: Police Employment.
- Miller, G. R. (1983). Telling like it isn't and not telling it like it is: some thoughts on deceptive communication. In J. I. Sisco (Ed.), *The Jensen lectures: Contemporary communication studies* (pp. 91-116). Tampa, FL: University of South Florida.
- Miller, M. G., & Burgoon, J. K. (1981). Factors affecting assessments of witness credibility. In R. Bray & N. Kerr (Eds.), *The psychology of the courtroom*. New York: Academic Press.
- Miller G. R., & Stiff, J. B. (1993). *Deceptive communication*. Newbury Park, CA: Sage Publications.
- 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.

- Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in one interrogation context. *Law and Human Behavior, 20*, 443-459.
- Rabon, D. (1994). *Investigative discourse analysis*. Durham, NC: Carolina Academic Press.
- Rudacille, W. C. (1994). *Identifying lies in disguise*. Dubuque, IA: Kendall/Hunt.
- Ruby, C. L. & Brigham, J. C. (1998). Can criteria-based content analysis distinguish between truth and false statements of African-American speakers? *Law and Human Behavior, 22*, 369-388
- Sapir, A. (1996). *The L.S.I. course on scientific content analysis SCAN workbook*. Phoenix, AZ: Laboratory for Scientific Interrogation
- Shearer, R. A. (1999). Statement analysis: SCAN or scam? *Skeptical Inquirer, 23*, 40-43.
- Smith, N. (2001). *Reading between the lines: An evaluation of the Scientific Content Analysis Technique (SCAN)*. (Police Research Series.) London: Crown.
- Sporer, S. L. (1997). The less traveled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology, 11*, 373-397.
- Suiter, T. L. (2001). *Linguistic foundations of investigative discourse analysis and content based criteria analysis*. (Unpublished doctoral dissertation). Cornell University, Ithaca, NY.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed) . Chichester, GB: John Wiley & Sons.
- Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. West Sussex, GB: John Wiley & Sons.
- Vrij, A. & Mann, S. (2001). Detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15*, 187-203.
- Walters, S. B. (2000). *The truth about lying: How to spot a lie and protect yourself from deception*. Naperville, IL: Sourcebooks.
- Zaparniuk, J., Yuille, J. C., & Taylor, S. (1995). Assessing the credibility of true and false statements. *International Journal of Law and Psychiatry, 18*, 343-352.
- Zhou, L., Twitchell, T. P., Qin, T., Burgoon, J. K., & Nunamaker, J. F. (2003). *An exploratory study into deception detection in text-based computer-mediated communication*, in *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference*, p. 10, 2003. Hilton Waikoloa: HI
- 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.
- Zuckerman, M., & Driver, R. E. (1985) Telling lies: Verbal and nonverbal correlates of deception. In A. W. Siegman & S. Feldstein (Eds.), *Multichannel interrogations of nonverbal behavior* (p. 129 – 147), Hillsdale, NJ: Erlbaum.

Received: 3/2014

Accepted: 2/2015

DeCicco, A. J., & Schafer, J. R. (2015), Grammatical differences between truthful and deceptive narratives [Electronic Version]. *Applied Psychology in Criminal Justice, 11*(2), 75-92.