Applied Psycholinguistics **31** (2010), 439–462
doi:10.1017/S0142716410000068

# The linguistic correlates of conversational deception: Comparing natural language processing technologies

NICHOLAS D. DURAN, CHARLES HALL, PHILIP M. MCCARTHY, and DANIELLE S. MCNAMARA
*University of Memphis*

Received: June 17, 2008 Accepted for publication: March 21, 2009

ADDRESS FOR CORRESPONDENCE
Nicholas Duran, Department of Psychology, University of Memphis, Memphis, TN 38152. E-mail: nduran@memphis.edu

ABSTRACT
The words people use and the way they use them can reveal a great deal about their mental states when they attempt to deceive. The challenge for researchers is how to reliably distinguish the linguistic features that characterize these hidden states. In this study, we use a natural language processing tool called Coh-Metrix to evaluate deceptive and truthful conversations that occur within a context of computer-mediated communication. Coh-Metrix is unique in that it tracks linguistic features based on cognitive and social factors that are hypothesized to influence deception. The results from Coh-Metrix are compared to linguistic features reported in previous independent research, which used a natural language processing tool called Linguistic Inquiry and Word Count. The comparison reveals converging and contrasting alignment for several linguistic features and establishes new insights on deceptive language and its use in conversation.

Of the spinmeisters, fibbers, or equivocators among us, their success often hinges on the ability to conceal a lie with well-chosen words. However, truth's traces may still lurk amid their verbal eloquence, as subtle linguistic features of language have been shown to reveal inner states of thought and feeling. These features go beyond the literal meaning of words and focus instead on how words are arranged and structured in discourse (Pennebaker, Mehl, & Niederhoffer, 2003). By pursuing these features, some progress has been made in uncovering the linguistic correlates of deception. The gains, though, are not without their unique challenges. Deception is a behavior designed to defeat detection, and thus identifying salient linguistic features of deception may be difficult even for the trained researcher (Vrij, Edward, Roberts, & Bull, 2000). Attempts by human judges to detect deception are fraught with problems of reliability and depth of analysis. One approach to this problem has been to turn to *natural language processing* (NLP) algorithms that incorporate advances in technology and linguistic theory. At the forefront of these technologies is an application called Coh-Metrix[1] (Graesser, McNamara, Louwerse, & Cai, 2004). Coh-Metrix is the largest NLP tool of its kind, with over 700 indices of computed language characteristics that have been validated across a variety of psychological domains (Crossley, Louwerse, McCarthy, & McNamara, 2007; Hempelmann, Rus, Graesser, & McNamara, 2006; McCarthy et al., 2008; McNamara, Louwerse, McCarthy, & Graesser, in press).

In the current study, we take the first steps in using Coh-Metrix to identify features of deception. By doing so, we also address another challenge in the linguistic analysis of deception research. Deception occurs in a variety of settings and for a variety of purposes. Accordingly, the linguistic features relevant to one context do not necessarily hold in another context (Zhou, Burgoon, Nunamaker, & Twitchell, 2004). Moreover, in research conducted thus far, the linguistic features that have been identified for a particular context have not been corroborated, or even extended, with multiple NLP tools. Given the nebulous nature of deception, there is an impetus for researchers to clearly specify the context of the targeted deception, and to use convergent NLP approaches to evaluate the various types of linguistic features. Therefore, to meet these challenges, we build from prior research to compare and establish conceptual validity between NLP tools.

Specifically, we turn to the work of Hancock, Curry, Goorha, and Woodworth (2008), who use an NLP tool called Linguistic Inquiry and Word Count (LIWC; Pennebaker, Francis, & Booth, 2001). In their study, Hancock et al. (2008) collected transcripts of deceptive and truthful conversations that occurred within an instant-messaging (IM) environment. To evaluate their data, those transcripts were submitted to LIWC, a tool that evaluates over 70 dimensions of language. LIWC has gained a tried and trusted reputation for tracking linguistic features that are indicative of social and psychological phenomena, including personality traits (Pennebaker & King, 1999), emotional expression (Kahn, Tobin, Massey, & Anderson, 2007), and mental health (Pennebaker, Mayne, & Francis, 1997). Much like Coh-Metrix, the success of LIWC is aided by its automated and easy to use interface. The two NLP tools also share similarities in their ability to analyze a large number of linguistic features, preeminence in their respective literatures, and accessibility for a general audience. Moreover, the tools have a number of conceptually similar indices (i.e., computational instantiations of linguistic features) that allow for an evaluation of algorithmic validity. By comparing Coh-Metrix with LIWC, we can offer a unique, but complementary analysis that strengthens our investigation into the nature of deceptive language.

To conduct our study, we use the conversational transcripts and LIWC results from Hancock et al. (2008). We do so to provide a basis for comparison within one specific context of deception. According to Zhou (2005), the features of deceptive language vary from context to context, with particular contrast within communication channels (e.g., face to face, telephone, e-mail). Therefore, it becomes necessary to focus on a single communication channel to control for any changes in language use. In Hancock and colleagues' study, the context for deceptive language is expressed as computer-mediated communication (CMC) using IM. These conversations occur as synchronous exchanges between two (or more) interactive participants. In recent years, this CMC channel has received greater attention because of its increased use in business and industrial settings (Juul Andersen, 2005; Quan-Haase, Cothrel, & Wellman, 2005). Because of this increase, it is appropriate to investigate deception in CMC, which is as common, if not more so, than face-to-face conversations (Carlson, George, Burgoon, Adkins, & White, 2004). As in face-to-face conversations, deceivers using IM can monitor the interaction as it occurs, but are not burdened by paralinguistic cues that might otherwise be incriminating. Although this CMC context is growing in popularity and is open to feature deception, there are few studies that explicitly address this communication channel. As such, the Hancock and colleagues' transcripts offer an opportunity to further explore a promising CMC context for deceptive cues.

Another reason to revisit the Hancock et al. (2008) conversational transcripts is to place greater emphasis on the dynamics of deception in real-time conversation. Hancock and colleagues were largely motivated by the research of Burgoon and Buller (1996; Burgoon, Buller, & Floyd, 2001; Burgoon, Buller, Floyd, & Grandpre, 1996), who argue that deception is concomitant to maintaining plausibility in social interaction. Deception often occurs in a dialogue between interlocutors; as such, the linguistic features that identify deceptive competence emerge from the joint contribution of both conversational partners (*sender* and *receiver* of deceptive exchanges). Many researchers claim the mutual influence between conversational partners creates an interdependent relationship in language use (Clark, 1996; Pickering & Garrod, 2004). Hancock and colleagues were particularly interested in whether the receiver engaged in what Niederhoffer and Pennebaker (2002) refer to as *linguistic style matching*, whereby the receiver takes on the linguistic features of the deceptive sender.

The dynamic maintenance of conversational deception also has unique cognitive and social challenges. Although a receiver may be unaware of the veracity of the sender's false statements, the sender must continually stay committed to preserving the receiver's presumption of truth. In doing so, senders must process and comprehend the speech of the receiver while simultaneously planning their own response (Greene, O'Hair, Cody, & Yen, 1985); they must actively monitor the receiver's understanding to establish and maintain conceptual common ground (Clark & Schaefer, 1987); and senders must adjust pragmatic strategies on the fly when discussing different topics. Hancock et al. (2008) and others (e.g., Johnson, Barnhardt, & Zhu, 2004; Zuckerman, DePaulo, & Rosenthal, 1981) have hypothesized that the sender's maintenance of both their own false reality and the receiver's ostensible reality comes at the price of cognitive resources, thereby creating compensatory linguistic behavior on the part of the sender.

Deception in interactive contexts such as conversation also increases the risk of being discovered as a fraud, resulting in *face loss* that is often associated with negative social standing (Brown, 1977). These social factors are embedded in the influences of the culture at large and are inextricably linked to the cognitive demands outlined. Based on these characterizations of conversational deception, we selected sets of Coh-Metrix measures that are operationalized to capture the cognitive and social influences of conversational deception.

In the section that follows, we first review the method Hancock et al. (2008) used for collecting the transcripts of deceptive and truthful conversations. We then present and provide a theoretical rationale for the Coh-Metrix measures chosen for this study. We then compare the data of Hancock and colleagues alongside our expanded approach.

## HANCOCK ET AL.'S CONVERSATIONAL TRANSCRIPTS

### *Participants*

The original cohort of participants from the Hancock et al. (2008) study included 30 male and 36 female upper-level undergraduate students from a private university in the northeastern United States. The 66 participants were randomly paired to create 33 same-sex interlocutor pairs who were unacquainted with each other prior to their participation in the study.

All participants were recruited under the pretense of studying *how unacquainted individuals communicate about various conversation topics*. As such, participants were not aware that deception would be required in the study. Participants' social interaction was also limited by placing each member of a pair in a separate room upon arrival at the laboratory.

### *Procedures*

The experiment was conducted within a text-based, CMC environment. CMC is simply using a computer interface to send message in text, video, or audio format via a computer interface. Participants were led to separate rooms and seated in front of a computer console. The IM software *Netmeeting* was used to collect the written communication of participants. This software allows messages to be sent instantaneously between computers via an Internet connection. Both the sender and receiver of a message enter text into a large interface text window that can be viewed easily. All messages were recorded automatically and stored anonymously.

Still in their separate rooms, participants were randomly assigned the role of *receiver* (the "deceived") or *sender* (the "deceiver") for each dyad. The sender's role was to initiate and maintain a conversation using four simple, icebreaker topics provided by the experimenter. These experimental topics included: *discuss the most significant person in your life*, *talk about a mistake you made recently*, *describe the most unpleasant job you have ever had to do*, and *talk about responsibility*. The four topics were presented to the sender and receiver on a sheet of paper along with the practice topic: *When I am in a large group, I . . .* The practice topic allowed participants to become comfortable with one another in the experimental setting. Along with initiating the conversation, the sender was also responsible for introducing deception to the conversation. Senders were informed that it would be necessary to deceive their partners on two of the topics preselected by the researchers, and to tell the truth on the other two topics. Specifically, they were asked *to NOT tell "the truth, the whole truth, and nothing but the truth."* This broad conceptualization of deception was considered to be the most naturalistic, thus giving senders some flexibility in how they chose to lie. On the sheet of paper with the experimental topics, the two topics that involved NOT telling the "truth, the whole truth, and nothing but the truth" (i.e., to be deceptive) were signaled to the sender with an asterisk. The receivers, blind to the sender's deception, were merely instructed to stay engaged and responsive to the ongoing conversation. The receiver's sheet of paper outlining topic order had no asterisk markers. The presentation of topics, as well as the order of deception, was counterbalanced across all participant pairs.

The online interactions were automatically stored and monitored by the experimenter on a separate, third console. The experimenter's role during the conversational phase was to initiate the conversation and mediate the interaction with the practice topic. Prior to initiation, participants were allowed 5 min to reflect upon the topics, thus allowing senders (i.e., the deceivers) time to prepare the gist of their fabricated responses. There was no time limit to the subsequent conversation, and participants were instructed to stop only when both conversational partners felt they had exhausted the topic matter. After completing all four topics, participants were introduced to each other in person and then fully debriefed.

For preparation of the data, the recorded messages were converted into sender and receiver transcript files according to topic. A total of 264 transcripts were produced, with each dyad generating 8 different transcript files: 4 transcripts of the *sender* dialogue and 4 transcripts of the *receiver* dialogue. Because two of the four topics discussed were considered *deceptive*, there were 4 transcripts labeled *deceptive*: 2 from the sender and 2 from the receiver (recall, however, that the receiver was not aware that the sender was being deceptive). The remaining 4 transcripts were labeled *truthful*.

## LINGUISTIC FEATURES OF DECEPTION

For dependent variables, Hancock et al. (2008) used eight LIWC based linguistic indices. With LIWC, 72 different word characteristics can be tracked per written response. For each of the 72 word characteristics, LIWC provides the percentage of words that adhere to that particular characteristic. Computational algorithms in LIWC compare the content of each transcript to over 2,300 words that have been coded for a variety of psychological and linguistic characteristics; including part of speech, emotional saliency, and cognitive complexity.

In our current study, we used the same transcripts as Hancock et al. (2008) but analyzed them with the Coh-Metrix software. Coh-Metrix was initially developed to explore cognitive constructs of cohesion in written text. Cohesion here refers to the linguistic features that explicitly link words, propositions, and events in a text, which in turn, facilitate a reader's coherent mental representation of a text. To construct a profile of cohesion, Coh-Metrix tracks word-level features that are similar to LIWC, but also incorporates modules and algorithms that assess collocations of words. Coh-Metrix integrates lexicons, syntactic parsers, part of speech classifiers, semantic analysis, and other advanced tools in NLP. Algorithms include referential overlap, proportion of situational dimensions (e.g., causal dependencies), latent semantic similarity, density of connectives, and syntactic complexity. As such, there are over 700 linguistic indices available in Coh-Metrix.

Combinations of these indices have been applied to a wide range of domains, including the validations of coherence breaks in academic texts (Duran, Bellissens, Taylor, & McNamara, 2007; Ozuru, Best, & McNamara, 2004); discriminating low- and high-cohesion versions of academic texts (McNamara et al., in press; McNamara, Ozuru, Graesser, & Louwerse, 2006); identifying shifts in writing style between professional writers, even shifts that occurred during the careers of each respective writer (McCarthy, Lewis, Dufty, & McNamara, 2006); and evaluating the pedagogical importance of authentic and simplified texts for second language acquisition education (Crossley et al., 2007).

The current analysis is the first attempt to use Coh-Metrix to characterize linguistic patterns of conversational deception. However, using over 700 linguistic indices presents two major theoretical problems. One problem is that spurious distinctions are likely to arise when there is an excess of variables. Too many variables can result in a statistical "overfitting," such that small and largely irrelevant differences between deceptive and truthful conditions may be exaggerated. The second problem of using the full set of linguistic indices is the overwhelming task of establishing each index's explanatory power. Before a specific index is used, it should be justified by a general framework of deception; however, no such framework exists (that we are aware of) because deceptive linguistic behavior is highly flexible with different external (e.g., social) and internal (e.g., cognitive) influences (Depaulo et al., 2003). As such, it becomes necessary to first consider the conversational context in which the deception is embedded and only then select linguistic indices that are most relevant to that particular context. For example, it is reasonable to assume that deceptive behavior in a casual conversation will be very different from deceptive behavior in a criminal interrogation. Accordingly, our selection of Coh-Metrix indices was guided by many of the principles of deception established in Hancock et al. (2008) and elsewhere in the deception and communication literature (Burgoon et al., 1996; Zhou, 2005). These principles are based on the cognitive and social influences that are hypothesized to arise during deceptive behavior. Ultimately, we operationalized the linguistic indices in six categorical constructs that will be explained in further detail later in this article: (a) quantity, (b) immediacy, (c) specificity, (d) accessibility, (e) complexity, and (d) redundancy (Table 1).

Each category above is represented by two to three Coh-Metrix indices that were chosen to provide converging validity, one of the explicit goals of our research. At least one of these indices was selected to be conceptually similar to a LIWC index. These similar indices may seem trivially redundant; however, they provide a basis for comparison with Hancock et al. (2008) and for establishing simple measurement reliability. Unfortunately, several categories do not have a representative and/or a conceptually similar LIWC index. These omissions are addressed in turn.

We proceed by briefly explaining the theoretical motivation for each of our six categorical constructs. For each category, we report the results from the Coh-Metrix analysis and interpret the results within a framework of conversational deception. Where possible, we also compare and contrast our results with those of Hancock et al. (2008). As in Hancock and colleagues' work, the Coh-Metrix data is analyzed in a 2 (Message Type: deceptive vs. truthful) $\times$ 2 (Speaker Type:

Table 1. *Categories of deceptive behavior based on linguistic features operationalized by Coh-Metrix*

| Classification | Definition |
|---|---|
| | Quantity |
| Total word count[a] | Total words in text (based on Charniak parser) |
| Words per conversational turn[a] | Mean words per sentence |
| | Immediacy |
| Tentative statements[b] | Modal verbs (e.g., *should, might, may*) |
| Personal pronouns[a] | *I, me, he, they, etc.* |
| | Specificity |
| Spatial[b] | Locational prepositions (e.g., *here*) |
| Temporal[b] | Ratio of temporal elements |
| Questions[a] | Incidence of *wh*-adverbs (e.g., *why, what*) |
| | Accessibility |
| Familiarity of words[b] | Word rating from MRC database |
| Meaningfulness of words[b] | Word rating from MRC database |
| Concreteness of words[b] | Word rating from MRC database |
| | Complexity |
| Negation[a] | Negation connectives (e.g., *did not, except, but*) |
| Sentential complexity[b] | Mean words before main verb of main clause |
| | Redundancy |
| Given information[b] | LSA given-new value |
| Referential overlap[b] | Argument word overlap, adjacent sentences |

*Note:* MRC, Medical Research Council; LSA, latent semantic analysis.
[a]Linguistic cue is an approximate replication of Hancock et al. (2008).
[b]Linguistic cue is novel to the current study.

sender vs. receiver) repeated measures type general linear model procedure. We also provide partial eta squared ($\eta_p^2$) values to assess the strength of any significant effects.

This analytical method allows us to examine the differences not only between deceptive and truthful conversations but also between sender and receiver. As mentioned earlier, the receiver might exhibit a pattern of linguistic style matching with the sender. Alternatively, the sender's behavior may elicit a subtle, but unique pattern of linguistic behavior in the receiver. For these reasons, it is theoretically important to consider the linguistic profiles of both conversational partners in deceptive exchanges.

## COH-METRIX RESULTS AND LIWC COMPARISON

### *Quantity*

In both Hancock et al. (2008) and the current study, the *total word count* and *number of words per conversational turn* were computed and compared between deceptive and truthful conversation transcripts. These indices are theoretically important for assessing the willingness of deceptive senders to proffer information. Senders may use fewer words to minimize the opportunities to incriminate themselves (Colwell, Hiscock, & Memon, 2002). As such, senders' total word count and number of words per conversation turn should be significantly lower in deception than when telling the truth. Conversely, senders want to appear socially involved so as not to violate a social norm of reciprocity that might otherwise raise suspicion (Burgoon et al., 1996). Senders may therefore maintain their word count across truthful and deceptive interactions.

In the current Coh-Metrix analysis, a significant main effect of message type (deceptive vs. truthful) was observed for total word count, $F$ (1, 33) $= 8.87$, $p = .005$, $\eta_p^2 = 0.21$.[2] More words were produced during deceptive conversation ($M = 159.38$, $SE = 9.97$) than truthful conversation ($M = 122.76$, $SE = 9.23$).[3] Senders increased word use from 123.15 words ($SE = 10.21$) in truthful conversations to 158.16 words ($SE = 12.01$) in deceptive conversations. Receivers increased word use from 122.37 words ($SE = 10.39$) in truthful conversations to 160.59 words ($SE = 16.12$) in deceptive conversations. These patterns of results were virtually identical to Hancock et al. (2008), who also found a statistically significant main effect for message type. In neither study was there an effect for speaker type (sender vs. receiver), nor did the total word count between message types differ across speakers (i.e., there was no interaction between message type and speaker type).

The second quantity analysis was on the mean number of words per conversational turn. Using Coh-Metrix, a significant main effect for message type was observed, $F$ (1, 33) $= 3.50$, $p = .05$, $\eta_p^2 = 0.10$. Fewer words were produced per conversational turn in the deceptive conversations ($M = 7.73$, $SE = 0.27$) than per truthful turn ($M = 8.37$, $SE = 0.36$). Senders produced fewer words per conversational turn when deceptive ($M = 7.98$, $SE = 0.42$) compared to telling the truth ($M = 8.19$, $SE = 0.55$), and receivers produced fewer words per conversational turn in the deceptive conversations ($M = 7.48$, $SE = 0.35$) than per truthful turn ($M = 8.55$, $SE = 0.55$). Taking this result in conjunction with the previous total word count results, the Coh-Metrix analysis demonstrates that senders and receivers in deceptive conversations use more words overall, but fewer words per conversational turn. However, this conclusion does not hold for Hancock et al. (2008). In their analysis, they did not find an equivalent decrease in words per conversational turn for senders and receivers in deceptive conversations. Hancock and colleagues instead report a marginally significant interaction (two tailed, $p = .06$), indicating that only receivers used fewer words per conversational turn in deceptive conversations.

The incongruent conclusions between the two computational tools may have resulted from implemented differences of what LIWC and Coh-Metrix consider a word. LIWC simply computes as a word any sequence of alphanumeric characters that is separated by a white space from another sequence. Coh-Metrix, however, computes words on the basis of the Charniak syntactic parser and corresponding word tags. As a result, the definition of a word is more precise. Contractions, for example, are counted in the expanded form (e.g., *don't* > *do not*, *they're* > *they are*). A more relevant difference is that an ellipsis is counted as a distinct pause filler. LIWC would treat *so* . . . as one word, whereas Coh-Metrix would output two words by distinguishing *so* and the ellipsis. This specificity is important for the current study where pause fillers are believed to hold semantic content.

Based on overall word counts between deceptive and truthful conditions, Coh-Metrix counts 2.6 more words on average per deceptive conversational turn and 0.42 more words on average for truthful conversational turn when compared to LIWC's counts. This comparison suggests that Coh-Metrix distinguishes more word types, and that this precision particularly affects the interpretation of the utterance length measurement.

### *Immediacy*

Introducing deception into a conversation always carries the risk of detection. Although the consequences might be no more than slight embarrassment, deceivers may take cautionary measures to distance themselves from their lies, even while engaged in the act of lying. Wiener and Mehrabia (1968) have suggested that deceptive statements are marked by "distancing strategies" that minimize personal involvement with the content of the message. One such distancing strategy is the decreased use of *first person personal pronouns* (e.g., *I*, *me*, *ours*; Newman, Pennebaker, Berry, & Richards, 2003). Related to this decrease, deceptive messages are expected to have a greater number of *second* and *third person pronouns* (e.g., *you*, *s/he*, *it*, *they*) to divert attention from the deceiver.

Another distancing strategy is an increased use of *tentative constructions* with words and phrases like *might*, *would*, *I guess*, *it seems to me*. These are often referred to as *hedges*. Tentative constructions imply a noncommittal to the content of the lie, thereby mitigating negative judgment of personal character or attributions of blame (Vrij & Heaven, 1999).

For the analysis of pronoun use, Hancock et al. (2008) computed the percentage of first, second, and third person pronouns in deceptive and truthful conversations. The researchers found a statistically significant main effect for speaker type (sender vs. receiver) for third person pronouns, as well as an interaction between message type and speaker type for third person pronouns. The main effect provides evidence that senders use more third person pronouns than receivers, but it is more important that the interaction reveals that it is only the deceptive senders who are more likely to discuss others in the third person.

We also used Coh-Metrix to assess pronoun use as a distancing strategy. To do so, we simply computed the percentage of the different pronouns in each conversational transcript. Much like Hancock et al. (2008), we did not find any statistically significant effects for first and second person pronoun use. However, our results for third person pronouns differed from Hancock and colleagues' results. We found a main effect for speaker type and the interaction, with marginally significant values, $F$ (1, 33) = 3.84, $p$ = .06, $\eta_p^2 = 0.10$, and $F$ (1, 33) = 3.20, $p$ = .08, $\eta_p^2 = 0.09$, respectively. Nonetheless, we found the same trend, $F$ (1, 33) = 5.73, $p$ = .02, $\eta_p^2 = 0.15$, showing senders using more third person pronouns during deception ($M$ = 2.93, $SE$ = 0.32) than the truth ($M$ = 1.94, $SE$ = 0.22). The statistical differences here are most likely explained by differences in word count when computing percentages.

In a second immediacy analysis, we used Coh-Metrix to evaluate the distancing strategy of increased tentative construction phrases. There is no equivalent analysis in Hancock et al. (2008). The current approach underscores the advantages of using a syntactic parser and part of speech tagger. With these additional modules, an incidence score (out of 1,000 words) for modal verbs (e.g., *should*, *might*, *may*) can be computed. Despite these noted advantages, the Coh-Metrix index of tentative constructions via modal use was not statistically significant. The Coh-Metrix index may have been too general to make subtle distinctions. Coh-Metrix does not distinguish among different uses of modals. Consequently, all modals were included in the computation, even modals that are nontentative in nature. For example, the root use of *may* and *must* produces a nontentative use in statements like *You must go now* or *You may not*. Taken together, the nonspecific modal index was too general to support the immediacy category.

### *Specificity*

Language has many linguistic features that allow speakers to reconstruct events from memory with certain temporal and spatial characteristics. The reconstructed events are often isomorphic mappings to perceived external events or, as is the case with deception, fabrications generated from internal cognitive processes of imagination and reasoning. As such, the mental representation of each event differs in terms of origin; the event can be initially encoded as a perceptual experience or as a simulation of an imagined experience. According to Reality Monitoring theory (Johnson & Raye, 1981), the temporal and spatial characteristics for each event will differ in terms of specificity. Events that originate in actual perception will have greater temporal and spatial detail than events that originate from internal simulations. To continue with our goal of automatically cataloging the linguistic patterns of deceptive and truthful speech in conversation, we chose two Coh-Metrix indices that capture the linguistic features of temporal and spatial characteristics. The temporal features index tracks words that have a high probability of being embedded in temporal expressions. These words include specifiers (e.g., *next*, *following*), deictics (e.g., *yesterday*, *now*), absolutes (e.g., *1997*, *Monday*), time of day (e.g., *12:00 a.m.*, *noon*), and time periods (e.g., *summer*, *week*). The index is computed as a ratio score that divides the summed occurrence of all temporal words in a conversational transcript by the total number of words in the transcript. For the Coh-Metrix spatial index, the number of locational prepositions (e.g., *here*, *on*, *in*) is counted for each transcript and normalized for differences in transcript length by converting to an incidence score (out of 1,000 words).

There are no equivalent measures for temporal and spatial specificity in Hancock et al. (2008). However, in terms of a *general* specificity, Hancock and colleagues hypothesized that there might be a decrease in general specificity, thus prompting the receiver of a lie to ask more questions for clarification or detail. As such, the number of questions asked by receivers will increase as the sender is lying. To infer an asked question, Hancock and colleagues used LIWC to compute the percentage of sentences ending with question marks. In similar fashion, we used Coh-Metrix to compute a proportion score of *wh*-words (e.g., *why*, *what*) to assess possible changes in receivers question asking behavior.

The first specificity analysis using Coh-Metrix indices of temporal and spatial specificity was not statistically significant. However, for the Coh-Metrix index of general specificity, there was a significant interaction between message type and speaker type for number of *wh*-adverbs used, $F$ (1, 33) = 6.83, $p = .01$, $\eta_p^2 = 0.17$. An analysis of *wh*-adverb use at each level of speaker type for deceptive and truthful messages revealed that senders used fewer *wh*-adverbs, and presumably asked fewer questions when being deceptive ($M = 6.53$, $SE = 0.98$) than when telling the truth ($M = 9.04$, $SE = 1.09$), $F$ (1, 33) = 4.19, $p = .05$, $\eta_p^2 = 0.11$; conversely, receivers used marginally more *wh*-adverbs when being deceived ($M = 10.34$, $SE = 1.23$) than when told the truth ($M = 7.33$, $SE = 1.02$), $F$ (1, 33) = 3.30, $p = .08$, $\eta_p^2 = 0.09$. These patterns of results suggest that receivers ask more questions when being deceived, whereas senders ask fewer questions when being deceptive. In Hancock et al. (2008), they also found the same effect for the receiver, but failed to find a similar effect for the sender. Again, the incongruence might be attributed to differences in the computational approach for operationalizing question use (i.e., proportion of *wh*-adverbs vs. percentage of question marks).

### *Accessibility*

We hypothesized that deceivers would select vocabulary that is easier to retrieve from memory. Based on the seminal work of Paivio (1965) and Underwood and Schulz (1960), word retrieval accessibility is modulated by experiential influences of word meaningfulness, familiarity, and concreteness. Word meaningfulness is operationalized by the number of associations that a word invokes for native English speakers. More associations increase word meaningfulness and the ease of retrieval for that word. Word familiarity is the familiarity of the orthographic form of a word and is typically assessed on a Likert-type scale from 1 to 7. More familiar words are more likely to be retrieved. Finally, word concreteness refers to how easy it is to explicitly ground a word in perceptual experiences. For example, a word like *house* is more easily grounded than an abstract word like *interesting*. As such, concrete words are more easily recalled than abstract words. For word meaningfulness and familiarity, Coh-Metrix provides an average score based on human ratings of over 150,000 words compiled in the Medical Research Council database (Coltheart, 1981). For word concreteness, Coh-Metrix computes abstractness and ambiguity scores by incorporating a module based upon WordNet (Miller, 1995). WordNet is an online lexicon tool that groups words into sets of synonyms that are connected by semantic relations. One such relationship, the hypernym value, refers to the number of levels a word has above it in a conceptual, taxonomic hierarchy. A high hypernym value is a proxy for word concreteness because the word has many distinctive features.

All indices for the accessibility category are computed as incidence scores in Coh-Metrix. There are no equivalent indices for accessibility in Hancock et al. (2008).

There was a statistically significant main effect of message type for word meaningfulness in conversations, $F$ (1, 33) = 7.88, $p$ = .008, $\eta_p^2 = 0.19$. The words used in deceptive conversations were more meaningful ($M$ = 418.47, $SE$ = 1.23) than words used in truthful conditions ($M$ = 412.76, $SE$ = 1.75). Senders' use of meaningful words increased from a rating of 415.21 ($SE$ = 2.30) in truthful conversations to a rating of 418.15 ($SE$ = 1.47) in deceptive conversations. Receivers increased from a rating of 410.31 ($SE$ = 2.60) in truthful conversations to a rating of 418.78 ($SE$ = 2.00) when they were being deceived. No interaction was observed between message type and speaker type.

For the analysis of word concreteness there was a significant interaction between message type and speaker type, $F$ (1, 33) = 5.42, $p$ = .02, $\eta_p^2 = 0.14$. An analysis of word concreteness at each level of speaker type for deceptive and truthful messages suggest that senders use more concrete words when deceptive ($M$ = 340.63, $SE$ = 3.31) than when they are telling the truth ($M$ = 332.99, $SE$ = 2.69), $F$ (1, 33) = 3.25, $p$ = .05, $\eta_p^2 = 0.09$. There was no difference for receivers in deceptive conversations ($M$ = 337.49, $SE$ = 2.34) or truthful conversations ($M$ = 337.77, $SE$ = 3.28).

The third accessibility measure of word familiarity was not statistically significant.

In summary, senders and receivers used more meaningful words when being deceptive, with the deceptive sender specifically using words that are more concrete. As we suggested earlier, these word characteristics facilitate the activation and retrieval of semantic meaning from memory. A consequence of this facilitation is that meaningful and concrete words are more likely to be used if cognitive resources are directed elsewhere (e.g., in concocting a deceptive message during conversation). Thus, the increased use of meaningful and concrete words by deceptive speakers supports our earlier hypothesis that deception places greater demands on cognitive resources.

### *Complexity*

Another linguistic predictor of conversational deception is change in the syntactic complexity of sentential structures. Based on our general hypothesis of cognitive and social demands, deceivers will minimize or compensate for the demand by avoiding sentences with difficult syntactic composition. In Coh-Metrix, a standard index of sentence complexity is the number of words before the main verb of the main clause. It is assumed that as the number of words increases, so does the demand on the speaker's working memory (see Graesser, Zhiqiang, Louwerse, & Daniel, 2006). Assuming that the process of lying would tax a deceiver's memory resources, we can expect a decrease in words before the main verb (i.e., lower complexity) compared to the truth-telling condition.

Alternatively, we could also hypothesize that an increased number of words before the main verb are to be expected in conversational contexts where deceptive messages are created on the fly. An increase in words before the main verb would reveal a *stalling* strategy used to formulate a lie while still staying engaged in the conversation.

Coh-Metrix computes the main verb of each sentence by first automatically parsing each sentence using the Charniak parser (1997, 2000). Each parse generates a syntactic tree that represents the underlying formal grammar. From this formal representation, the main verb of the main clause is identified and preceding words are tallied. The sentential complexity for deception and truth-telling is then assessed by collapsing the sentences of each conversational transcript into a mean score. There is not an equivalent index in Hancock et al. (2008).

A significant main effect of message type was observed for this complexity measure, $F$ (1, 33) = 5.63, $p$ = .02, $\eta_p^2 = 0.15$. More words were used before the main verb in deceptive conversations ($M$ = 7.14, $SE$ = 0.46) than in truthful conversations ($M$ = 5.79, $SE$ = 0.37). Specifically, senders use more words before the main verb when deceptive ($M$ = 6.79, $SE$ = 0.71) than when telling the truth ($M$ = 6.16, $SE$ = 0.61. Likewise, receivers use more words before the main verb ($M$ = 7.50, $SE$ = 0.60) when they are being deceived than in truthful conversations ($M$ = 5.41, $SE$ = 0.43). No interaction was observed between message type and speaker type.

These results suggest that senders and receivers use more syntactically complex sentences in deceptive conversations. Increased sentence complexity does not support the hypothesis that complexity results from working memory demands, but it instead supports the alternative hypothesis that generating deception in spontaneous conversation requires a stalling strategy. For working memory to be the prevailing factor, the sender has to know exactly what they want to say before they say it. It is only under these circumstances that the sender will intentionally minimize the use of words before the main verb. This active strategy of advanced planning is unlikely in the current conversational context.

A second Coh-Metrix index of complexity that is common to LIWC is the number of negation connectives (e.g., *did not*, *is not*, *but*, *except*) that appear in each conversational transcript. Newman et al. (2003) argued that deceptive speakers tend to avoid using negation connectives because they risk presenting incriminating contradictions and muddled detail. Negation connectives require speakers to contrast events that actually occurred with events that did not occur. Although negative connectives help clarify event depictions, the speaker must also recall additional detail from memory. Deceptive speakers must conjure that detail up at that moment. As such, deceptive speakers may have additional challenges because they are "recalling" false details from an already distorted reality—a reality that may be loosely constructed in spontaneous conversation. Thus, the deceiver may sacrifice clarity and use fewer negation connectives to avoid accidental contradictions.

The Coh-Metrix index of negation connectives is a proportion value computed from the Charniak syntactic parser and part of speech taggers. The LIWC index uses the predefined word list and computes the value as a percentage.

This measure of complexity is computed similarly for LIWC and Coh-Metrix and is also assumed to reflect demands on working memory. Our results agree with those reported by Hancock et al. (2008) that there are no statistically significant effects for negation connectives.

### *Redundancy*

In both deceptive and truthful conversations, an important component of event narration is the coherence of statements and ideas. Coherence is a psychological interpretation of comprehension. The greater the coherence, the easier the narration will be to understand (Graesser, McNamara, & Louwerse, 2003). Coherence is modulated by various factors, but a crucial factor is cohesion, which is the explicit language used to connect information and provide conceptual consistency. Most cohesion research suggests that text cohesion influences text comprehension, particularly with texts that consist of formal written monologues (Beck, McKeown, Sinatra, & Loxterman, 1991; McNamara, Kintsch, Songer, & Kintsch, 1996), but little work has been conducted on the relationship between cohesion and coherence in informal spoken dialogue. The question remains as to whether the coherence of a speaker's mental event representation influences the cohesion of their speech. Deceptive speech can potentially address this question because deceivers' mental representations of false events are likely to be less coherent than representations of truthful events. If this is the case, the less coherent deceptive representation may result in less cohesive speech.

It could be possible, however, that incoherent mental representations are not mirrored in speech, but instead the difficulty of remembering and structuring spontaneous deception may promote simpler and more cohesive speech. Characteristics of such language include conceptual redundancy and more accessible words (Duran et al., 2007). We have already demonstrated in this study that deceivers tend to use more accessible words (e.g., high concreteness; high meaningfulness). It may be the case that deceivers also capitalize on conceptual redundancy for greater cohesion.

We evaluated the cohesion of deceptive and truthful conversations with two widely used indices in text analysis that are incorporated in Coh-Metrix: argument overlap (McNamara et al., 2006) and latent semantic analysis (LSA) given-new values (Hempelmann et al., 2005). Both indices are broad indicators of between-sentence conceptual redundancy. This redundancy reinforces information by keeping it focal in a developing narrative. Argument overlap computes explicit overlap between two sentences by tracking the common nouns in either single or plural form. The LSA given-new value also computes overlap between sentences, but it requires more explanation to understand how it works.

The LSA given-new value is based on LSA (Landauer, McNamara, Dennis, & Kintsch, 2007). This measure compares adjacent sentences to determine if the meaning in a target sentence is new (different) or given (redundant) to preceding sentences. Sentence meaning is first computed by representing each word in the sentences as a distributional pattern of frequency occurrences within a large corpus of texts (representation is in vector format). Words that have similar patterns of occurrences are considered similar in meaning. Word similarity vectors are then combined linearly into a composite meaning vector. The target vector is projected into a hyperplane constructed from all preceding composite meaning vectors and based on the target sentences relationship to the hyperplane, a *new-given* value is generated (for more information, see McCarthy et al., in press). High values on both the argument overlap and LSA given-new value suggest high cohesion between sentences. These measures are unique to Coh-Metrix; there is no equivalent in Hancock et al. (2008).

For the first analysis of argument overlap, we did not find any statistically significant effects. However, the more subtle measure, in the LSA given-new value, revealed a statistically significant main effect for message type, $F(1, 33) = 9.32$, $p = .004$, $\eta_p^2 = 0.22$. In the deceptive conversations, there was a higher given-new value ($M = 0.25$, $SE = 0.005$) compared to truthful conversations ($M = 0.23$, $SE = 0.007$). Senders' given-new value was higher when they were deceptive ($M = 0.26$, $SE = 0.007$) compared to when they were telling the truth ($M = 0.24$, $SE = 0.01$). Receivers' given-new value was higher when they were being deceived ($M = 0.25$, $SE = 0.008$) compared to when they were being told the truth ($M = 0.22$, $SE = 0.01$).

These results provide evidence that deceptive conversations contain more given information relative to preceding context. This result should be expected if we consider an important goal for deceivers is to minimize opportunities for self-incrimination. A strategy to avoid self-incrimination may be to reiterate particular topics or themes in the conversation. Deceivers do not reiterate by explicit repetition, as evidenced by the null finding with referential overlap, but by an implicit focus on a few semantic focal points. However, there may be no conscious decision to avoid self-incrimination. Instead, the high LSA given-new value support a hypothesis that redundancy strategies are triggered by differences between memory representations of deceptive and truthful narratives. For example, it is possible that the details of truthful events are more extensively linked in memory than the fabricated details of a lie. As a truthful account unfolds, the activation and recall of remembered details are likely to activate other details in a distributed and global manner; thus, a greater variety of information is available for use. Conversely, the details in a deceptive account are often constructed and cued from the local and developing context. As such, there is less new information activated from memory and deceivers may default to more redundant language.

### *Brief summary of Coh-Metrix analysis*

The overall results of our study demonstrate that the linguistic features that characterize deceptive conversations are substantially different from those that characterize truthful conversations. From the perspective of Coh-Metrix, we can describe deceptive conversations as involving (a) more words overall, but fewer words used per conversational turn; (b) words that are more meaningful; (c) utterances of each conversational turn being more syntactically complex (because of a *stalling* hypothesis); and (d) less unique information introduced during the course of the conversation.

The effects we have discussed thus far changed in the same direction for both sender and receiver. However, other changes in linguistic behavior were specific to either the sender or receiver. For example, personal pronouns and word concreteness increased only for senders while they were being deceptive. Demonstrating

Table 2. *Comparison of similar Coh-Metrix and LIWC index results in conversational transcripts*

| | Coh-Metrix | LIWC |
|---|---|---|
| Total word count | More words overall in deceptive conversations | More words overall in deceptive conversation |
| Words per conversational turn | Fewer words per conversational turn in deceptive conversations | Receivers used *marginally* fewer words per conversational turn than senders in deceptive conversation |
| Personal pronouns | Senders used *marginally* more third-person pronouns when deceptive compared to when telling the truth | Senders used more third-person pronouns when deceptive compared to when telling the truth |
| Questions | Receivers ask more questions during deceptive conversations, senders fewer | Receivers ask more questions during deceptive conversations |
| Negation | None | None |

*Note:* LIWC, Linguistic Inquiry and Word Count (software program).

another pattern, receivers asked marginally more questions in deceptive conversations than the senders who asked fewer questions.

### *Brief summary of Coh-Metrix and LIWC comparison*

Table 2 provides a side by side comparison of the results using the indices that were similar in Coh-Metrix and LIWC. Although these indices are not exact replications because of differences in algorithmic operationalization, they are quantifications of the same linguistic features. Overall, five indices were comparable; of these five, *total word count*, *negation,* and *personal pronouns* had the same result. This convergence confirms that more words are used in deceptive conversations, that there are no differences in the use of negation, and that deceptive senders use more third person pronouns. The multimethod alignment lends greater credibility to the Coh-Metrix and LIWC indices, as well as to the quantity and immediacy constructs in general.

The two remaining indices, *words per conversational turn* and *questions,* did not converge completely; the difference in the indices most likely results from the different definitions of a word used by the two tools. For the words per conversational turn index, the Coh-Metrix analysis revealed that both sender and receiver used fewer words in each utterance during deceptive conversations. With LIWC, only receivers used fewer words in each utterance during deception. For the questions index, the Coh-Metrix analysis revealed that receivers asked more questions while being deceived and senders asked fewer questions while being deceptive. LIWC showed only that the receivers asked fewer questions during deception. In general, for both of the nonconverging indices, the Coh-Metrix analysis found a statistically significant effect that was not found in the LIWC analysis.

## GENERAL DISCUSSION

Both this study and Hancock et al. (2008) demonstrate that at least one type of deception is detectable through NLP tools. For our analysis, we compared the Coh-Metrix and LIWC tools on the original corpus of deceptive conversations used by Hancock and colleagues. Using this approach, we were able to evaluate the effectiveness of each NLP tool in a common context of social interaction. In addition, we were also able to use Coh-Metrix to build a more complete catalogue of the linguistic features that emerge during deception. In this discussion, we first turn to the expanded analysis and the identification of eight Coh-Metrix indices that distinguish deceptive conversations from truthful conversations. Using this winnowed set of indices, we provide new insights into the cognitive and social constraints that are hypothesized to influence deceptive behavior in both the deceiver and their naive conversational partner. Turning next to the comparison with Hancock and colleagues, we discuss complimentary insights provided by the LIWC analysis. In particular, we consider the findings of Coh-Metrix and LIWC within the context of CMC. Throughout this discussion, we address the limitations of our current research and end with suggestions for future work.

There is a well-established conversational maxim of quality that a speaker should avoid saying what the speaker knows to be false (Grice, 1975). When lying to a friend, colleague, or foe, a speaker often violates this maxim; as a consequence, new goals and task demands are introduced into the conversation. The deceiver must now maintain representations of both the truth and a falsified version of that truth. In doing so, the deceiver must also appear convincing while avoiding unintentional "slips" of the truth. The cumulative effect is that deception requires increased cognitive control in the presence of social scrutiny. Previous research suggests that even with the best attempts to maintain control, the inner states brought on by deception are manifested in subtle changes of language use (Pennebaker et al., 2003).

For the Coh-Metrix analysis, we created six theoretically guided categories to represent these changes in language. Each category is composed of two or three Coh-Metrix indices. The categories include (a) the amount of information in the conversation (i.e., *quantity*), (b) the readiness to identify with message content (i.e., *immediacy*), (c) the breadth of detail used to describe a narrative (i.e., *specificity*), (d) the change in semantic memory retrieval (i.e., *accessibility*), (e) the change in grammatical phrasing (i.e., *complexity*), and (f) the repetition of given information (i.e., *redundancy*). We then compared truthful and deceptive conversations for changes in the six categorical dimensions. We found statistically significant results for all categories.

Several of our findings provide novel contributions to the relationships between deception and language. A key discovery is that *quantity* of word use changes for the level of analysis. For example, in deceptive conversations, fewer words were used at the level of conversational turn. Based on this finding alone, we might conclude that deceivers use fewer words to minimize opportunities for incrimination; however, in the same conversations, there are also more conversational turns and more words used overall. This result challenges the original conclusion and suggests that the deceivers are attempting to establish rapport with their conversational partner. Because our results also show that receivers ask more questions of deceptive senders, an alternate interpretation might be that the deceivers do use fewer words per conversation turn to minimize opportunities for incrimination; however, the "paucity" of information in these restricted turns require the receivers to ask for additional information or clarification, which then generates more overall turns and words. Unfortunately, there is not enough information to make a conclusive interpretation either way. However, the results do highlight the rich interplay between the often conflicting goals of cautiously limiting information and the appearance of affability.

Another new discovery is that the words used in deceptive conversations are more meaningful than those used in truthful conversations; for the sender in particular, the words are also more concrete. The *accessibility* of meaningful and concrete words from semantic memory indicates that the deceiver is using an unconscious strategy to decrease burdens on cognitive processing. Because meaningful and concrete words are highly associated to other words in semantic memory, these words are easier to retrieve and in turn allow cognitive resources to be redirected to the more difficult task of maintaining deception in conversation.

Related to increased difficulty, we also found evidence for *redundancy* in deceptive conversations. The redundancy is the repetition of content from contiguous utterances. Previous research that has investigated linguistic features of redundancy has failed to find significant effects because it applied a strict lexical overlap criterion (e.g., Zhou, 2005). Instead, redundancy in deception appears to be more subtle. In our analysis, we used an algorithm that compares the semantic similarity of two words based on their likelihood to appear in similar contexts. This algorithm, called the LSA given-new value, revealed that words similar in meaning are used more often in deceptive conversations than in truthful conversations. This redundancy in meaning suggests the deceiver may find it simpler to focus on consistent themes. Part of the reason for such focus is that the deceiver may have difficulty in using the same interconnected memory representations that are formed with real experiences. Deceivers instead rely more on local conversational cues for what information can or cannot be reasonably fabricated. This orientation toward local context decreases the likelihood of using novel information and increases the chances of repeating what has already been stated.

Deceptive conversations are also characterized by a change in the *complexity* of grammatical constructions. A complex sentence is defined in Coh-Metrix as having more words before the main verb of the main clause. In deceptive conversations, we found that this type of sentence complexity increases. It is important to note that complex grammatical constructions identified by Coh-Metrix are not necessarily more difficult to produce, and may actually be preferred when attempting to generate a spontaneous lie. For instance, consider a lie about what you did yesterday. If you were telling the lie in conversation, it might take some time to think of a false response, such as *I watched TV at my house.* While constructing the response, it would be useful to buy some time with a stalling strategy that provides genuine information. Thus, you begin with *It was really cold outside . . .* and continue with the lie . . . *so, I'd thought I'd stay in and watch TV.* By doing so, there would be a higher occurrence of words before the main verb, and as such, greater evidence for our hypothesized stalling strategy.

Finally, for *specificity* of the deceptive narrative, we found that deceptive conversations were marked by the receiver asking the sender more questions. This result implies that the sender lacked specificity and that the receiver was requesting greater clarification. During these exchanges, the deceptive sender also asked fewer questions compared to when they were telling the truth. Again, these findings can be interpreted in multiple ways. The receiver may ask more questions because of an unconscious suspicion of the sender's deception. Likewise, the deceptive sender may ask fewer questions to defend against the suspicion. It may also be that because deceivers use fewer words in each conversational turn, receivers need to ask for more clarification. The receivers may be responding to a perceived violation of the maxim of quantity rather than the maxim of quality that deals with truthfulness (Grice, 1975). In either interpretation, the important finding is that receiver linguistic behavior systematically varies from that of the sender in terms of specificity.

For most of the analyses, receiver behavior was mostly aligned with the deceptive sender, with the exception of *wh*-adverbs, concreteness, and third person pronouns. The alignment of linguistic features is not uncommon between conversational partners. There is extensive research that shows implicit alignment can occur and cut across lexical, syntactic, and conceptual levels (Garrod & Anderson, 1987; Pickering & Garrod, 2004). The underlying mechanism is because of *priming*, whereby the linguistic features used by one partner elicit a similar representation in the other. In this way, coordination of form and meaning is automatically generated and maintained. In the current study, we find evidence for alignment in the number of words used, the meaningfulness of words, the repetition of similar words and concepts, and the complexity of grammatical constructions. A possible limitation in the alignment is not knowing whether the sender or the receiver is predominantly priming or being primed. The limitation is a particular concern because we are interested in the linguistic features generated by the deceptive sender. As such, we assume that it is the deceptive sender's linguistic behavior that is most influential. We base this assumption on two factors. First, the design of the experiment gives the sender more control by allowing the sender to introduce new topics (total = 4) into the conversation. Second, being deceptive may invoke a greater desire for the sender to be convincing, where an equivalent desire is not present in the receiver. As a result, this unique desire may translate into greater linguistic influence.

In detecting linguistic features of deception, the problem of who influences whom is heightened in conversational interactions. Unlike monologues or scripted interviews, there are cognitive and social constraints that present additional and novel challenges. Moreover, our use of a CMC corpus of deceptive and truthful conversations adds to these challenges. Despite the increasing difficulty, the CMC conversational context is an ecologically important domain that is gaining in popularity and use. However, we must be careful in generalizing our findings from the CMC context to other domains with their own constraints. Face-to-face conversations, for example, are not the same thing as IM conversations, and thus the linguistic features characterizing each conversation may substantially differ. For these reasons, we felt justified in using and extending the Hancock et al. (2008) study. Importantly, their data provided a common context to compare and contrast Coh-Metrix with LIWC.

Our first step in the comparison was simply to assess the degree to which the systems differed in their analysis of deception. Our results suggest that Coh-Metrix was largely able to reproduce LIWC results (e.g., in areas of *quantity* and *immediacy*) and offer many areas of deception detection in addition to LIWC (e.g., *accessibility*, *complexity*, and *redundancy*). For these reproduced results, the replication occurred despite two different computational approaches for operationalization. The alignment gives greater credence to the original findings in the Hancock et al. (2008) study, specifically their findings that more words are used in deceptive conversations and that deceptive senders project the focus of conversation onto others (as evidenced by the greater use of third person pronouns).

Our study also showed where LIWC and Coh-Metrix were not able to reproduce the same results on similar indices: words per conversational turn and questions. For Coh-Metrix, deceptive conversations were marked by fewer words from both the sender and receiver, as well as more questions from the receiver and fewer questions from the sender. In contrast, LIWC did not find a difference of word use for senders, and found only a marginal difference for the receiver. In addition, no difference in question use was found for senders. This inability to reproduce the same results using the identical corpus might suggest that one NLP tool is superior to the other. However, we take an alternative perspective. The algorithmic operationalization for each tool is a matter of preference that should be chosen to best address a research question. In other words, the operationalization does not capture a "truer" representation of reality. More than anything, the operationalization is a manifestation of computational expediency. For example, LIWC uses computationally inexpensive algorithms to process texts. During processing, words are identified by surrounding white space and matched to an internal set of words that are coded for linguistic and psychological features. In contrast, Coh-Metrix goes beyond a predefined set of words and incorporates sophisticated algorithms to maximize the scope of analysis. By including syntactic parsers and psycholinguistic databases, linguistic features can be distinguished at the word, sentence, and discourse levels. To understand this in practice, we consider the operationalization of word count. Words are not separated by white spaces alone (as in LIWC), but are expanded from contraction form (e.g., *don't* > *do not*) and distinguished from a trailing ellipsis to create unique entries for ellipsis occurrence. Furthermore, for the operationalization of questions, instead of counting the number of question marks (as in LIWC), detailed part of speech information, like *wh*-adverbs (e.g., *where*, *what*), can be used as an index of question use. Based on these differences in operationalization and given the current task, the Coh-Metrix analysis may have an advantage over LIWC. Because the data are typed conversations, there are a large number of ellipsis occurrences that might be an important linguistic feature of deception (e.g., pauses, incomplete thoughts). In addition, participants tend to use multiple question marks at the end of a sentence. By just counting question marks, there is a risk for over exaggerating the number of questions. This miscount is not a problem with *wh*-adverbs.

For the current task, LIWC does have an important advantage over Coh-Metrix. LIWC codes words for psychological dimensions, such as sensory information like "see," "touch," and "listen," that may be related to a deceiver's goals of convincing storytelling. In Hancock et al. (2008), the deceptive conversations were reported as having a greater degree of these sensory words. Future work will require adapting Coh-Metrix to include linguistic features that have been successful in detecting deception in multiple contexts. Other candidates include positive and negative word connotations, as well as content word diversity measures (Zhou, 2005). Additional work also needs to be conducted in predicting the likelihood that a narrative is deceptive or truthful. Given our established set of deceptive linguistic features, we can include these features as variables into statistical prediction models (e.g., logistic regression, discriminant function analysis). By doing so, we can also evaluate how well our linguistic features collected in a CMC context explain the variance in other interactions, such as business negotiations or even police interrogations.

Finally, it would be naive for us to argue that a straight and easy road lies between identifying linguistic features of deception and using them in real-world practice. There are many individual differences to account for, as well as the consideration of ethical and legal concerns. Nevertheless, we begin the journey with the current study. We have shown that deception is a feature of language that is identifiable through many variables, established that Coh-Metrix is a computational system that can identify deception, and revealed that there is insight to gain by comparing computational NLP tools.

## ACKNOWLEDGMENTS

This research was supported in part by the Institute for Education Sciences (IES R305G020018-02, IES R305a080589), Department of Defense/CIFA (H9C104-07-C0019), and a National Science Foundation Graduate Research Fellowship awarded (to N.D.D.). The views expressed in this paper do not necessarily reflect the views of the Institute for Education Sciences, Department of Defense, or National Science Foundation. The authors thank Jeff Hancock for providing us with the conversational transcripts used in this study. The authors also acknowledge the contributions to this project made by Arthur Graesser, Zhiqiang Cai, and Joe Weintraub.

## NOTES

1. Coh-Metrix is available from http://cohmetrix.memphis.edu
2. To interpret the $\eta_p^2$ values, Stevens (2002) suggests the following: 0.01 is considered a small effect, 0.06 is considered a medium effect, and 0.14 is considered a large effect. However, such interpretations are merely "guides," and the importance of any effect size is always relative to the task at hand.

3. We report standard errors in this study to be consistent with Hancock et al.'s (2008) results section.

## REFERENCES

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly, 26*, 251–276.

Brown, B. (1977). Face saving and face restoration in negotiation. In D. Druckman (Ed.), *Negotiations: Social–psychological perspectives* (pp. 275–299). Beverly Hills, CA: Sage.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 3*, 203–242.

Burgoon, J., Buller, D., & Floyd, K. (2001). Does participation affect deception success? A test of the interactivity principle. *Human Communications Research, 27*, 503–534.

Burgoon, J. K., Buller, D. B., Floyd, K., & Grandpre, J. (1996). Deceptive realities: Sender, receiver, and observer perspectives in deceptive conversations. *Communication Research, 23*, 724–748.

Carlson, J. R., George, J. F., Burgoon, J. K., Adkins, M., & White, C. H. (2004). Deception in computer-mediated communication. *Group Decision and Negotiation, 13*, 5–28.

Charniak, E. (1997). Statistical techniques for natural language processing. *AI Magazine, 18*, 33–44.

Charniak, E. (2000). *A maximum-entropy-inspired parser*. Paper presented at the 1st Meeting of the North American Chapter of the Association for Computational Linguistics, Seattle, WA.

Clark, H. (1996). *Using language*. New York: Cambridge University Press.

Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language & Cognitive Processes, 2*, 19–41.

Coltheart, M. (1981). The MRC psycholinguistics database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.

Colwell, K., Hiscock, C. K., & Memon, A. (2002). Interviewing techniques and the assessment of statement credibility. *Applied Cognitive Psychology, 16*, 287–300.

Crossley, S., Louwerse, M., McCarthy, P. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *Modern Language Journal, 91*, 15–30.

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

Duran, N. D., Bellissens, C., Taylor, R., & McNamara, D. S. (2007). Quantifying text difficulty with automated indices of cohesion and semantics. In D. S. McNamara & G. Trafton (Eds.), *Proceedings of the 29th Annual Meeting of the Cognitive Science Society* (pp. 233–238). Austin, TX: Cognitive Science Society.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27*, 181–218.

Greene, J. O., O'Hair, H. D., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. *Human Communication Research, 11*, 335–364.

Graesser, A. C., McNamara, D. S., & Louwerse, M. M. (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82–98). New York: Guilford Press.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, and Computers, 36*, 193–202.

Graesser, A. C., Zhiqiang, C., Louwerse, M. M., & Daniel, F. (2006). Question Understanding Aid (QUAID): A web facility that helps survey methodologists improve the comprehensibility of questions. *Public Opinion Quarterly, 70*, 3–22.

Grice, P. H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic Press.

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

Hempelmann, C., Rus, V., Graesser, A., & McNamara, D. (2006). Evaluating state-of-the-art treebank-style parsers for Coh-Metrix and other learning technology environments. *Natural Language Engineering, 12*, 131–144.

Hempelmann, C. F., Dufty, D., McCarthy, P. M., Graesser, A. C., Cai, Z, & McNamara, D. S. (2005). Using LSA to automatically identify givenness and newness of noun phrases in written discourse. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 941–946). Mahwah, NJ: Erlbaum.

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

Johnson, R., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia, 42*, 878–901.

Juul Andersen, T. (2005). The performance effect of computer-mediated communication and decentralized strategic decision making. *Journal of Business Research, 58*, 1059–1067.

Kahn, J., Tobin, R., Massey, A., & Anderson, J. (2007). Measuring emotional expression with the Linguistic Inquiry and Word Count. *American Journal of Psychology, 120*, 263–286.

Landauer, T., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *LSA: A road to meaning*. Mahwah, NJ: Erlbaum.

McCarthy, P. M., Dufty, D., Hempelman, C., Cai, Z., Graesser, A. C., & McNamara, D. S. (in press). Evaluating givenness/newness. *Discourse Processes*.

McCarthy, P. M., Lewis, G. A., Dufty, D. F., & McNamara, D. S. (2006). Analyzing writing styles with Coh-Metrix. In G. C. J. Sutcliffe & R. G. Goebel (Eds.), *Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference* (pp. 764–770). Menlo Park, CA: AAAI Press.

McCarthy, P. M., Renner, A. M., Duncan, M. G., Duran, N. D., Lightman, E. J., & McNamara, D. S. (2008). Identifying topic sentencehood. *Behavior Research Methods, 40*, 647–664.

McNamara, D. S., Kintsch, E., Songer, B. N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from texts. *Cognition and Instruction, 14*, 1–43.

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (in press). Coh-Metrix: Capturing linguistic features of cohesion. *Discourse Processes*.

McNamara, D. S., Ozuru, Y., Graesser, A. C., & Louwerse, M. (2006). Validating Coh-Metrix. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 573–578). Austin, TX: Cognitive Science Society.

Miller, G. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography, 3*, 235–312.

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.

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology, 21*, 337.

Ozuru, Y., Best, R., & McNamara, D. S. (2004). Contribution of reading skill to learning from expository texts. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Cognitive Science Society* (pp. 1071–1076). Mahwah, NJ: Erlbaum.

Paivio, A. (1969). Mental Imagery in associative learning and memory. *Psychological Review, 76*, 241–263.

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

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*, 1296–1312.

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. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Reviews in Psychology, 54*, 547–577.

Pickering, M., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169–226.

Quan-Haase, A., Cothrel, J., & Wellman, B. (2005). Instant messaging for collaboration: A case study of a high-tech firm. *Journal of Computer-Mediated Communication, 10,* article 13. Retrieved from http://jcmc.indiana.edu/vol10/issue4/quan-haase.html

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Erlbaum.

Underwood, B. J., & Schulz, R. W. (1960). *Meaningfulness and verbal learning*. Philadelphia, PA: Lippincott.

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

Vrij, A., & Heaven, S. (1999). Vocal and verbal indicators of deception as a function of lie complexity. *Psychology, Crime & Law, 5*, 203–215.

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

Zhou, L. (2005). An empirical investigation of deception behavior in instant messaging. *IEEE Transactions on Professional Communication, 48*, 147–160.

Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004). Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group Decision and Negotiation, 13*, 81–106.

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.