In CMC We Trust: The Role of Similarity

Lauren E. Scissors, Alastair J. Gill, Kathleen Geraghty, Darren Gergle
Center for Technology and Social Behavior
Northwestern University
2240 Campus Drive, Evanston, IL 60208 USA
{l-scissors, alastair, k-geraghty, dgergle}@northwestern.edu

ABSTRACT
This paper examines how different forms of linguistic similarity in a text-chat environment relate to the establishment of interpersonal trust. Sixty-two pairs played an iterative social dilemma investment game and periodically communicated via Instant Messenger (IM). Novel automated and manual analysis techniques identify linguistic similarity at content, structural and stylistic levels. Results reveal that certain types of content (some positive emotion words, task-related words), structural (verb tense, phrasal entrainment), and stylistic (emoticons) similarity characterize high trusting pairs while other types of similarity (e.g., negative emotion words) characterize low trusting pairs. Contrary to previous literature, this suggests that not all similarity is good similarity.

Author Keywords
Trust, Language, Linguistic Similarity, Lexical Entrainment, Instant Messaging (IM), Social Dilemma

ACM Classification Keywords
H.5.3 **[Information Interfaces and Presentation]**: Group and Organization Interfaces – collaborative computing, computer-supported cooperative work

INTRODUCTION

In face-to-face (FTF) communication, speakers often adapt their language and non-verbal behaviors to become more similar to those of their conversational partner. For example, pairs have been shown to converge in word use, pitch, rate of speech, and even facial expressions [9]. By adapting their communication to be more similar to that of their partners, speakers indicate that they are socially receptive and accommodating. These adaptations have been associated with positive outcomes such as more successful negotiations and increased liking between speakers (e.g., [1],[27]). However, in text-based environments such as Instant Messaging (IM) or text-chat, many of the interpersonal cues on which similarity adaptations are based are no longer available (for a recent review see [29]). In less rich environments, do individuals still make linguistic adaptations that are associated with positive social gains?

The objectives of this paper are to determine whether individuals make these linguistic adaptations, what forms they may take, and how they are associated with the degree of trust a pair achieves. This research is important given our increasing reliance upon text-based media to establish both professional and personal relationships. For example, a recent case study demonstrated that, in a high-tech company, text-chat and email were each used more than FTF and telephone communication combined [21]. Similarly, a recent survey of American teenagers found that text-based communication, (e.g., text-chat), was used almost as much as FTF and telephone, respectively [14].

In order to better understand trust establishment, we examine inter-speaker similarity across characteristics present in a text-only environment at the content (semantic), structural (tenses or syntactic components), and stylistic (surface) levels. At a content level, individuals may increase in similarity by discussing concepts with related meanings (e.g., "unhappy" and "sad"). At a structural level, similarity could be indicated by using the same verb tense (e.g., "tried" and "played"), or more precisely by verbatim repetition at a phrase level. At a stylistic level, speakers may imitate specific words or abbreviations (e.g., *lol*), or punctuation (e.g., *!!!*). We explore such forms of linguistic similarity in text-chat through data collected from participants engaged in an iterative social dilemma game—where success requires a willingness to trust one's partner.

Given the widespread use of text-based communication, it is important to understand how trust is established in these mediated environments. To increase our understanding, we examine the use and adaptation of linguistic forms to increase inter-speaker similarity in text-based communication, and present an analytical framework to capture this behavior. We also present a more sophisticated measure of trust and defection than has been traditionally used in interpersonal trust research. At an applied level, we discuss how our findings can be incorporated into automated or agent-based communication technologies. We also note how our findings can inform interfaces to improve interactions between geographically distributed workers, technical support workers and customers, or online daters.

CHI 2009, April 4–9, 2009, Boston, MA, USA.

BACKGROUND LITERATURE

Trust in CMC

Trust is important in social interactions. The current study employs Rousseau and colleagues' definition of trust as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" [25].

Early computer-mediated communication (CMC) research suggests that text-based environments are not rich enough to facilitate the establishment of trusting relationships (see [29] for a critique). This claim was largely supported by observations that CMC lacks many of the rich, nonverbal social cues present in FTF interactions. Research suggests that other types of communication with richer audio and visual cues, such as telephone or video-mediated communication have been more effective than text-chat for trust formation [2]. More recent research, however, has demonstrated that individuals can develop trust in less rich environments, though it might take longer to develop [31]. Additional research suggests that other factors, such as the communication of enthusiasm, can also increase the amount of trust developed in less rich environments [12].

While the majority of research on trust in mediated environments has tended to focus on trust *outcomes* rather than the *processes* by which trust is developed and maintained, scholars have recently begun to examine the relationship between similarity (verbal and nonverbal) and trust. For instance, Maddux and colleagues [15] observed that during negotiations, increasing nonverbal mimicry was associated with more trust reported by a speaking partner. Taylor and colleagues also found that general linguistic similarity, of topics and of content words, was significantly higher for successful than for unsuccessful negotiation pairs [27]. None of these studies, however, examined the association between text-based language and trust.

In our own work [26], we examined the relationship between lexical entrainment (e.g., [4]) and trust in text-based CMC and found that, within IM chat sessions, high trusting pairs exhibited greater repetition of shared terms than did low trusting pairs. This work, however, did not differentiate between various forms of lexical entrainment (e.g., nouns vs. adjectives). Furthermore, this work did not examine linguistic similarity at the semantic level. This paper builds on previous work by examining similarity in text-based CMC across different linguistic levels.

Linguistic Similarity

Modification and variation in communication style can occur at "all levels of language", namely pragmatics, semantics, syntax and phonology [3]. This paper focuses on linguistic adaptations in a text-only environment at three levels, and how these adaptations relate to trust. In the following, we describe what we term *linguistic similarity*, using Communication Accommodation Theory (CAT) [9] as our basis. Where relevant, we describe its relationship to other theories of similarity.

In general, CAT suggests that similarity can be used to signal solidarity or affiliation (cf. [5],[28],[3]). Convergent linguistic behavior, where two speakers become more similar, would indicate greater solidarity and reduced social distance. By contrast, two speakers who diverge and become increasingly different appear to indicate reduced solidarity and increased social distance. Consider two interlocutors who speak different varieties of English: American English and British English. To demonstrate their increasing affiliation, the former may start to call the "elevator" the "lift", whereas the latter may begin to refer to a "flat" as an "apartment". Given that we expect groups who have established trust to also exhibit solidarity and affiliation, we would generally expect to find greater similarity of language use in trusting groups.

Linguistic similarity can occur at three broad levels of language: content, structural, and stylistic. At a content level, previous work by Niederhoffer and Pennebaker [18] has shown that individuals match their language to that of their partner at a semantic level (which they term "Linguistic Style Matching"). For instance, if speaker A says "football" and speaker B says "soccer", speaker B would be engaging in style matching to speaker A in the semantic category of "sports." Though they do not discuss affiliation, Niederhoffer and Pennebaker propose that the more engaged the interlocutors are in conversation, the greater their semantic convergence should be.

Measuring similarity at this broad level can miss finer levels of detail, such as whether the speakers are converging on sports in general (American football and American soccer/British football) or one sport in particular (American soccer/British football). Yet analysis of semantic similarity allows for understanding how individuals may converge on whole categories of words, which may provide unique insights. For our analysis, we focus on the psychologically defined categories of emotion- and task-related content, since emotion detection in text-based CMC has been demonstrated [10], since emotion can influence trust formation (e.g., [8]), and since participants were told to refrain from discussing topics unrelated to the task.

Similarity can also be measured at the structural level. Like the measures adopted by Niederhoffer and Pennebaker [18], structural similarity could be measured in terms of verb tense, and whether speakers similarly refer to things in past, present, or future forms of verbs. Similarity in verb tense usage might suggest that two speakers show affiliation by focusing on the same point in time. A stricter type of structural similarity is phrase repetition between speakers, known as lexical entrainment[1], which has been related to

[1] The phenomena of speakers standardizing on phrases and repeating their use over time is described by several terms including lexical or linguistic entrainment, alignment, and mimicry. While this behavior is widely established, the explanations for its existence are actively debated in the psycholinguistics literature.

pro-social outcomes. For example, one study demonstrated that waitresses who repeated their customers' orders verbatim received bigger tips [1]. Another study demonstrated that verbal mirroring was associated with improved negotiation outcomes [6].

Linguistic similarity may also occur at a stylistic level. As noted by Bradac (e.g., [3]), in FTF this type of similarity is most commonly reflected by phonological variation. For example, individuals may become more similar to their speaking partners in pronunciation or accent (e.g., [13]). In a text-only environment, phonological information cannot be easily communicated. However, stylistic similarity can be communicated using the same jargon as one's communicative partner, and may serve to signal affinity [30]. Going back to our previous example of soccer and football, two interlocutors might be similar in the semantic meaning of the two terms (if they are both referring to American soccer/British football) but are *dissimilar* in their surface level expressions. Lastly, variations in speaker word choice can result in different levels of within-speaker repetition, with such repetition levels shown to converge over time between speakers [3].

In the following experiment, we examine linguistic similarity at the content, structural and stylistic levels, as it relates to trust establishment in text-based CMC.

METHOD

Participants

Participants (N=124, 47% male, 53% female) were students and staff at a mid-sized Midwestern university[2]. Participants were randomly assigned to pairs and did not know one another prior to the study. Of the 62 participant-pairs, 20 were male-male, 24 were female-female, and 18 were male-female. Participants were all native English speakers. The average age of participants was 20.4 years.

Procedure

Participants engaged in a variant of the DayTrader task paradigm originally developed by Bos and colleagues [2] and modified by Nguyen & Canny [17]. The DayTrader paradigm is an iterative social-dilemma game during which participants may invest tokens into a "market" and get returns on their investments. Like other social dilemma games, individuals playing the DayTrader game must cooperate in order to perform well (for a critique see [23]).

Participants played the DayTrader game via IM, on computers in separate experiment rooms. They could not see or hear each other and did not meet. Participants played 28 rounds of the investment game and could invest up to 60 tokens per round in the group market. Withheld tokens guaranteed individuals a two-fold return, while the tokens in the group market were tripled and then split between both participants. Thus, the pair as a whole could make the most by investing all of their tokens in the group market. However, a participant could earn more individually by defecting or withholding tokens from the group market. In other words, an individual who invests 58 tokens when his or her partner invests 60 tokens will receive higher earnings than if both individuals invest 60 tokens.

Participants received a spreadsheet outlining this payoff structure and, prior to playing, described the game to the experimenter as a check for understanding (although we do note that it took some pairs a while to get it right, with some groups settling for a sub-optimal – but "trusting" – investment strategy throughout). After every 5 investment rounds, participants could chat via IM with their partners for up to 5 minutes. Participants were told they could use this time to discuss strategy or other topics related to the game. The chatting and the investment rounds all took place within the same IM window. After each investment round, participants received their personal earnings but were not notified of their partner's earnings. Participants did not know the total number of investment rounds.

In addition, we included two features to facilitate and encourage defection. The first was a random market fluctuation (+/-3 tokens), which encourages defection by allowing participants to hide it within the market noise ([32],[17]). This also makes the DayTrader task more like real world situations [23]. The second was a 200 token bonus for the participant who earned the most after every five rounds (the bonus was split if participants earned the same amount). Again, to encourage defection, bonus earnings were not revealed until after all investment rounds were completed. Participants were paid between $10 and $25 based on their individual earnings.

Analysis

Our corpus contained a total of 24,002 words, with an average of 387 words per pair (ranging from 70 to 1027). Overall, there were 3,821 lines of chat, with an average of 61.6 lines of chat per pair.

Calculating Defection

Previous research examining defection in iterative social dilemma games has used game behavior as a measure of trust level achieved among a pair or group. Many studies interpret the existence of pro-social behavior in dilemma games as cooperation, not trust (e.g., [7]). However, consistent with prior literature (e.g., [2],[23]), we suggest that individuals' cooperation or defection is a proxy for the level of trust achieved since cooperating requires that one trusts that his/her partner will also cooperate. Furthermore, since our paradigm contains several factors that make the amount of partner cooperation ambiguous, we argue that trusting participants are in a vulnerable state but act based on the positive expectations they have of their partners [25].

Studies of trust in CMC typically use group earnings as a measure of trust, reasoning that if groups are cooperating, they are prospering. Cooperation, in theory, stems from the

[2] A subset of this data (N=52) was used in previous work [26].

mutual trust that develops between group members. High group earnings, therefore, are associated with high amounts of trust. This measure, however, does not take into account whether or not each individual defects in any given round. For instance, if Player A invests 60 tokens and Player B invests 58 tokens to earn the bonus, the group's earnings are quite high. Yet, assuming the players agreed to invest 60 tokens each, Player B has defected, suggesting that the level of trust is not necessarily high and most likely not mutual. Alternatively, if two players agree to each invest 30 tokens and follow through on their promise, no defection occurs yet the group earnings measure would identify this pair as lower trusting than the 60/58 pair.

A more sensitive and accurate measure of defection is "invest minus guess", developed in our previous work [26]. Using this approach, we define defection as an instance where a participant invests less than he or she expects his or her partner to invest. To make this calculation, participants were asked before every investment round how much they thought their partner would invest. This "guess" measure allows us to get a closer look at participants' expectations of how their partner will behave. This measure classifies the pairs who follow through on investing mid-range amounts to be counted as high trusting pairs, and classifies high-investing, defecting pairs as low trusting. For each round, pair-level defection is either zero (neither partner defected) or one (one or both partners defected).

There are potential limitations to our measure. It may be that matched investments of low amounts are not a proxy for trust based on previous agreements. For example, participants may expect their partner to defect (invest low amounts) and are merely matching this behavior. It is possible that this scenario occurs in our data set, though it seems more likely that a participant would actually undercut their partner if they suspected defection, rather than just match the defection. Furthermore, thorough examination of the transcripts revealed that many pairs did not recognize that the 60/60 investment strategy was optimal. Rather, many pairs chose to invest the same lower amounts (e.g., 30/30) in efforts to achieve mutual profit.

Despite limitations, we believe this measure is the best automated way to measure defection in investment-type social dilemma games. To evaluate our measure, a research assistant tagged instances of defection by hand. The research assistant read the transcripts of the chat sessions prior to each set of investment rounds and looked for the plan agreed upon by the two participants. If a participant agreed to a plan but did not follow through, that instance was scored as a defection. There were some cases where it was unclear whether or not defection occurred. For instance, some pairs did not talk at all during the early chat sessions. Other times, participants agreed to leave their earnings up to chance (i.e. agreed to no strategy). These instances were omitted from the analysis of inter-rater reliability between the hand coded and automated measures. Forty percent of the data was hand-coded and compared to the automated measure scores. With the omission of ambiguous cases, reliability was acceptable (κ=.77). This indicates the suitability of our automated defection measure for use as the dependent measure in the following analyses.

Linguistic Similarity - LIWC

Output from the Linguistic Inquiry and Word Count (LIWC) program [20] provided features from which we calculated linguistic similarity. LIWC calculates the rate, or percentage, of words in a text belonging to a number of pre-defined categories. LIWC categories include "linguistic processes" (e.g., words per sentence), "psychological processes" (e.g., positive emotion words), and "personal concerns" (e.g., words related to money).

Researchers using LIWC (e.g., [18]) typically compare LIWC category rates to determine whether the rates are higher or lower for a given group. These rates, however, do not assess the similarity in usage patterns between two speakers (i.e. if speaker A uses more positive emotion words, does speaker B use more of these words as well?). To measure this similarity, we conducted correlations of LIWC scores between players A and B within high and low defecting pairs, in each of the five chat sessions. The correlation between A and B on each category could then be used to assess the degree of overlap on particular LIWC categories. Mehl and Pennebaker [16] used a similar approach though their analysis did not examine the relationship between overlap and interpersonal outcomes. Taylor and Thomas [27] examined LIWC score correlations over time and related them to negotiation outcomes. We extend this work by examining inter-speaker correlations of LIWC scores as they relate to trust.

Linguistic Similarity – Entrainment Coding Scheme

We define lexical entrainment as the process by which both partners come to use the same word(s) during a chat session. Thus, if Partner A used a word and Partner B repeated it, then Partner B's use and all subsequent uses of the word(s) by both Partner A and Partner B were coded as lexical entrainment.

Our previous research isolated instances of word or word phrase entrainment across all parts of speech and labeled them as lexical entrainment [26]. Our current coding scheme isolates instances of entrainment at more specific structural and stylistic levels: noun phrases, verb phrases, adjectives (including adjective phrases), and interjections. Following a list of repeated words and phrases identified using the OpenNLP Treebank Tokenizer and Parser, two research assistants coded instances of entrainment. Unless contained as part of a larger phrase, the following were not included: pronouns (e.g., *I, we, you*), possessive pronouns (e.g., *mine, yours, ours*), determiners (e.g., *this, that, the*), being verbs and function words.

Inter-rater reliability was calculated on 10% of our corpus using Cohen's Kappa. Structural entrainment reliability ranged from acceptable to very good across each category

(κ=.71 to κ=.86) and at the overall level (κ=.94). For stylistic entrainment, reliability ranged from acceptable to very good across each category (κ=.78 to κ=.96) and at the overall level (κ=.92). Table 1 outlines the different forms of linguistic similarity and their respective components.

Category	***Examples***
Content	
Positive Emotion†	*happy, pretty, good*
Negative Emotion †	*hate, worthless, enemy*
Occupation†	*work, class, boss*
Leisure†	*house, TV, music*
Money†	*cash, taxes, income*
Structural	
Past tense verb †	*walked, were, had*
Present tense verb†	*walk, is, be*
Future tense verb†	*will, might, shall*
Noun Phrase‡	*that first round, the same amount*
Verb Phrase‡	*would be beneficial, I think so*
Interjections‡	*okay, sure, hi*
Adjectives‡	*higher, pretty trustful*
Stylistic	
Chat Abbreviations‡	*b/c, lol, w/*
Emoticons‡	*:-) :-(;-) :P*
Exclamation Points‡	*yeah!, sounds good!*
Question Marks‡	*start with 10?, move on?*

(Note: *† = from LIWC analysis, ‡ = from hand-coding*)

Table 1. Examples of Similarity and Entrainment at Content, Structural, and Stylistic Levels

Statistical Methods and Analysis

To assess the various degrees of similarity we performed two different analyses. The first analysis was a technique used to assess the degree of similarity between Speaker A and B on their production of words for the various LIWC categories presented as (†) in Table 1. The first stage was to calculate a correlation (*r*) for each pair on the various categories. This measure tells us whether or not the within-pair speakers' rates were significantly correlated (i.e., greater or less than zero) for a particular linguistic dimension. The second stage allowed us to test for differences between the degree of correlation for the high and low defecting groups. To do so, the correlation scores (*r*) were converted to *z'*-scores using Fisher's *r* to *z'* transformation. Then a confidence interval and standard error were generated on the difference of *z'*-scores, telling us whether or not one group exhibited a greater degree of correlation on a particular dimension.

The second approach was used to model the relationship between the hand-coded entrainment categories and overall counts of defection. The number of defections has a distribution form typical of count data, (i.e. high frequency of zeros, long tail), best modeled using a Poisson technique. We employed a repeated measures Poisson regression that adjusts for the within-group correlation present in the data [24]. The dependent variable for this model was number of defections. The independent variables were those marked with (‡) in Table 1. Overall word count was also included as a covariate to control for the overall amount of language produced. This allowed us to interpret the findings as rates of production.

RESULTS

Similar to previous work using this paradigm [26], defection rates across pairs were highly variable. Figure 1 illustrates the defection rate for "high defecting" (low trust) pairs, the rate for "low defecting" (high trust) pairs, and the average defection rate across all pairs.

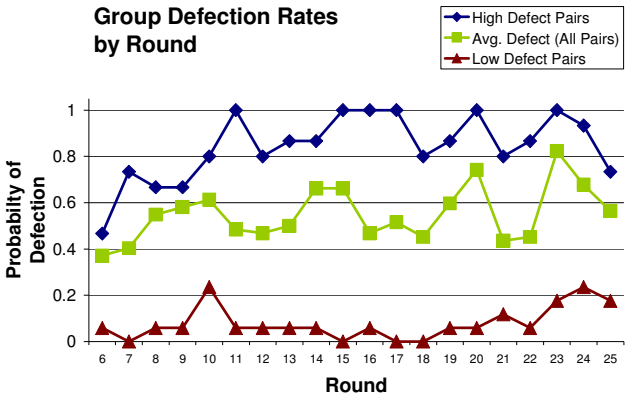

Figure 1. Defection by Round.

Rounds 1 through 5 are omitted from Figure 1 because they were "practice" rounds. Pairs were categorized as high defecting or low defecting based on whether they fell into the upper or lower quartiles of defection rates, respectively. Based on this classification there were 17 high defecting pairs and 15 low defecting pairs.

Analysis of LIWC Category Rates

As a preliminary analysis, we examined whether or not group *rates of usage* in various LIWC categories were significantly different between high and low defecting pairs. For the majority of the LIWC categories, there was no significant difference in the amount of words used in each category (as a proportion of total words used) between high and low defecting pairs. Because these analyses measured group-level rates, they could not provide insights into the amount of linguistic similarity between speakers. In order to assess similarity, we used the technique previously described to determine whether or not an increase in word use by speaker A was associated with an increase or decrease in word use by speaker B across categories.

General Corpus Similarity

To begin, we examined the amount of similarity in general word count patterns of high and low defecting pairs (total

word count and words per sentence). As previously described, amount of similarity was calculated by correlating speaker A's LIWC scores with speaker B's LIWC scores across all five chat sessions. For both high defecting and low defecting pairs, word count and words per sentence were significantly and positively correlated between speaker A and speaker B (Table 2). In other words, the more words and words per sentence speaker A used, the more words and words per sentence speaker B used.

We then examined the difference in speaker correlations between low and high defecting pairs to determine whether high defecting pairs differed from low defecting pairs in the amount of similarity exhibited. For both categories however, there was no significant difference.

	Low Defecting Pairs (r)	*High Defecting Pairs (r)*	*Difference Between Low & High?*
Word Count	0.74***	0.58***	No
Words/Sentence	0.35**	0.32**	No

* = $p<.05$, ** = $p<.01$, *** = $p<.001$

Table 2. Correlations of Speakers A and B within Pairs, and Correlations across Pairs.

Content Similarity – LIWC Categories

Scores from several LIWC categories were analyzed to determine how similar pairs were on the content or meaning of their utterances. We examined positive and negative emotion words as well as task-related categories that contained words relevant to the DayTrader paradigm.

Positive Emotion Words: There was no evidence of similarity in the use of positive emotion words between speaker A and B for either the high or low defecting pairs, nor was there a difference between their patterns of similarity (Table 3). However, we note greater similarity in use of the "optimism" subcategory of positive emotion words by low defecting pairs. Low defecting pairs exhibited a significant amount of similarity in their use of optimism words such that the more optimism words speaker A used, the more speaker B used (r=43, $p<.001$). Furthermore, low defecting pairs exhibited significantly more similarity in their use of optimism words than did high defecting pairs ($p<.05$); there was no correlation for high defecting pairs.

	Low Defecting Pairs (r)	*High Defecting Pairs (r)*	*Difference Between Low & High?*
Positive Emotion	-0.24	-0.06	No
Negative Emotion	-0.06	0.59***	Yes*

* = $p<.05$, ** = $p<.01$, *** = $p<.001$

Table 3. Content Similarity – Emotion Words by Defection.

Negative Emotion Words: As shown in Table 3, results revealed a positive correlation between speakers' use of negative emotion words in high defecting pairs (r = 0.59, $p<.01$), such that the more negative emotion words speaker A used, the more negative emotion words speaker B used. Furthermore, there was significantly more similarity between speakers in high defecting pairs than in low defecting pairs ($p<.05$); there was no correlation between speakers in the low defecting pairs. In other words, when there was an increase in the use of negative emotion words by Speaker A in the low defecting pairs, speaker B did not reciprocate with negative emotion words.

Task-related Words: We also analyzed the amount of similarity among speakers in the use following task-related LIWC categories: Occupation, Leisure Activity (Leisure), and Money and Financial Issues (Money). We selected these categories since they contain words that are relevant to the game and are present in our corpus (see Table 4).

	Examples of Relevant Words in LIWC Dictionary
Occupation	award, benefits, best, bonus, market, pay, team*, test, try, win, work
Leisure	game*, play, played, playing, team
Money	bet, betting, cost, greed*, invest, pay, paying, profit, rich*, wage*, worth

** denotes wildcard matching e.g., game** with *game, games, etc.*

Table 4. Words in LIWC Categories Relevant to Corpus.

When examining similarity in the use of these words between speakers, results revealed that while individuals in low defecting pairs showed significant similarity in their use of occupation and leisure words, individuals in high defecting pairs did not exhibit similarity in these categories (see Table 5). For low defecting pairs, there was a positive correlation between speakers on the use of words in these categories such that the more occupation or leisure words speaker A used, the more occupation or leisure words speaker B used, respectively. Additionally, low defecting pairs exhibited significantly more similarity in these categories than did high defecting pairs ($p<.05$).

	Low Defecting Pairs (r)	*High Defecting Pairs (r)*	*Difference Between Low & High?*
Occupation	0.31**	-0.04	Yes*
Leisure	0.53***	0.00	Yes*
Money	0.19	0.37**	No

* = $p<.05$, ** = $p<.01$, *** = $p<.001$

Table 5. Content Similarity – Task-related Words by Defection.

Conversely, results revealed that high defecting pairs exhibited a significant amount of similarity in their use of money-related words; the more money-related words used

by speaker A, the more money-related words used by speaker B. There was no significant correlation among low defecting pairs nor was there a significant difference in the amount of similarity between high and low defecting pairs.

Structural Similarity – LIWC Categories

In contrast to the focus of content similarity on the intended meaning of the words used, structural similarity focuses on the structure of the words used. Accordingly, we examined the similarity of verb tenses (past, present, and future) used between speakers in both high and low defecting pairs. For speakers in the low defecting pairs, use of both past and future tenses by speaker A were positively and significantly correlated with uses of past and future tenses by speaker B, respectively (see Table 6). For the high defecting pairs, however, there were no significant correlations between use of past and future tense by speakers A and B. There was also a significant difference between the amount of similarity in the use of future tense ($p<.05$) between high defecting and low defecting pairs, as well as a marginally significant difference between the amount of similarity in the use of past tense ($p<.10$).

	Low Defecting Pairs (r)	*High Defecting Pairs (r)*	*Difference Between Low & High?*
Past Tense	0.29*	-0.04	No
Present Tense	-0.08	0.17	Yes^
Future Tense	0.33**	-0.03	Yes*

^ = $p<.10$, * = $p<.05$, ** = $p<.01$

Table 6. Syntactic Similarity – Verb Tense by Defection.

Structural Similarity - Entrainment

In order to examine the relationship between entrainment and defection, we examined entrainment on the following categories: verb phrases, noun phrases, adjectives, and interjections (structural entrainment), and emoticons, chat abbreviations, question marks, and exclamation points (stylistic entrainment).

Results from our model revealed that, within individual chat sessions, low defecting pairs had higher rates of noun phrase entrainment than did high defecting pairs ($p=.029$). The following is an example of noun phrase entrainment:

Speaker A: we could just both do 50/50 **the whole time** to build up our money
…
Speaker A: okay, so 50/50
Speaker B: yes **the whole time**

This entrainment does not result from pairs simply using frequently occurring words. The log-likelihood statistic [22] showed that low defecting pairs used specific nouns, often in entrainment, significantly more often than would be expected by chance (e.g., "time": low-defect rate=3.98, baseline rate=0.62, $G^2=17.32$, $p<.005$). Furthermore, entrainment of noun phrases such as "the whole time," would be even less likely to occur by chance. Results were similar for other entrained nouns and phrases. Furthermore, we did not include high frequency function words and focused on the much larger, and lower frequency, open class of content words.

Contrary to expectations, other forms of structural entrainment (verb phrases, adjectives, interjections) did not predict significant portions of the variance in defection (see Table 7).

Variable	***Coeff.***	***z-score***	***Sig.***	***% change defection***
Noun phrases	**-0.249**	**-2.19**	**0.029***	**-15.3**
Verb phrases	-0.100	-1.12	0.27	
Adjectives	0.105	1.47	0.14	
Interjections	-0.053	-0.59	0.55	
Emoticons	**-0.803**	**-2.16**	**0.031***	**-14.6**
Chat Abbreviations	0.058	1.76	0.08	9.7
Exclamation Points	-0.146	-0.56	0.58	
Question Marks	0.028	0.73	0.46	

* = $p<.05$, $N=310$

Table 7. Syntactic and Stylistic Entrainment by Defection.

Stylistic Similarity - Entrainment

For stylistic entrainment, results revealed that higher amounts of emoticon entrainment were associated with a lower defection rate (see Table 7). The following is an example of emoticon entrainment from our corpus:

Speaker B: shall we continue?
Speaker A: **:-)** yes
Speaker B: **;-)**

Despite the significant finding, the amount of emoticons and emoticon entrainment present in our corpus was quite low and these results should be interpreted with caution. Furthermore, there were no significant relationships uncovered between defection rate and the other types of stylistic entrainment for which we tested.

DISCUSSION

Our findings provide us with a more nuanced picture of how different levels of linguistic similarity relate to the establishment of trust in a text-chat environment (summarized in Table 8). Building on our previous work which identified the importance of entrainment to trust [26], we examined linguistic similarity at the content, structure, and stylistic levels. Previous approaches to similarity acknowledged that convergence can occur across a range of levels [9],[3], but studies rarely examine this convergence at more than one level. By examining different levels of linguistic similarity, we demonstrated that not *all* convergence is *"good"* when viewed in the context of trust

establishment in a synchronous text environment. For example, at a general text level (word count and words per sentence) both high and low defecting pairs are similar in interaction. This reflects a fundamental property of language and conversation, as described in previous empirical and theoretical work [18],[11]. In the following, we discuss our findings, methodological and theoretic contributions, and potential applications of our work.

Category	*Findings*
Content Similarity	High trust pairs exhibited similarity in the use of optimism (positive emotion) words, occupation words, and leisure words; low trust pairs did not. Low trust pairs exhibited similarity in the use of negative emotion words and money words; high trust pairs did not.
Structural Similarity	High trust pairs exhibited similarity in the use of past and future tense verbs; low trust pairs did not. High trust pairs engaged in noun phrase entrainment; low trust pairs did not.
Stylistic Similarity	High trust pairs engaged in emoticon entrainment; low trust pairs did not.

Table 8. Summary of Findings.

At a content level, we note divergent findings for expressions of emotion: references indicating optimism were used more similarly by low defecting pairs, whereas high defecting pairs were more similar in their use of negative emotion words. For task-related language, there is also a discrepancy between low and high defecting pairs, with the former exhibiting more similarity in their use of words belonging to the occupation and leisure categories, and the latter showing similarity in their use of words relating to money.

In the case of emotion, low defecting pairs may similarly use optimism words as a way of jointly maintaining and reinforcing their positive, productive relationship. Moreover, low defecting pairs are unlikely to match their partner's use of negative emotion words, perhaps as a way of repairing the relationship and avoiding an escalation of a conflict. This makes sense in light of evidence [8] that links certain positive emotions (happiness and gratitude) to increased trust.

In the case of task-related words, examination of the data reveals that the low defecting pairs show greater similarity on categories containing words such as "playing" and "game", "testing" and "trying", and "awards" and "benefits". This appears to show that participants in low defecting pairs engage with each other on the friendlier, positive aspects of the game. By contrast, high defecting pairs exhibit greater similarity in their use of money words, (e.g., "betting", "cost", "pay"), indicating more of a focus on (their own personal) outcome. For both the results on positive and negative emotions and positive and negative aspects of the game, accommodation of language occurs (e.g., [5],[28],[3]). Although we only observe examples of convergence rather than divergence, the type of similarity upon which individuals converge is significant for predicting trust establishment. In terms of mechanisms for establishing trust, it may be that accommodation of positive semantic categories contributes to increased liking and affinity, which in turn builds trust. In contrast, accommodation of more negative semantic categories contributes to a lack of affiliation and friendliness, which in turn prevents trust from developing.

Turning now to the syntactic level, we find that low defecting pairs exhibit similarity in their use of past tense, future tense, and noun phrases while high defecting pairs do not. The verb tenses indicate that the trusting pairs are converging in their discussions of past and future activities. By accommodating to each other's verb tense usage, individuals may be indicating that they are similarly coordinated in the process of the task. The trusting pairs' similarity in use of noun phrases shows that entrainment occurs on their use of longer, specific, construction, such as determiner-modifier-nouns (e.g., "the whole time"). This indicates not only similarity, or convergence, at a general or content level, but also greater coordination in the specific references and ways of referring that are used [4].

Finally, at a stylistic level, our results revealed that there was significantly more emoticon entrainment for low defecting pairs than for high defecting pairs. This type of stylistic entrainment may be particularly effective in creating a bond between partners or embodying convergence for at least two reasons. First, emoticons are easily produced and can provide a simple and effective way of indicating similarity to one's interlocutor, perhaps in the same way as modifying pronunciation, or choosing a particular jargon word in FTF contexts [3]. Secondly, it may be that since emoticons graphically represent facial expressions that are present in richer environments, they may trigger some of the emotional or interpersonal responses achieved in richer environments such as FTF or video-mediated communication, (e.g., [8],[9]). We leave this to future investigation.

To summarize, at a content level, similarity in words relating to positive emotion and the task processes relate to trust, whereas matching a partner in expression of negative emotion words or talk of money are associated with defection. Structurally, trusting pairs show similarity, perhaps indicating greater coordination, in their use of past and future verbs, and also noun phrase references. Finally, at a stylistic level, trusting pairs mimic each other's use of emoticons. From these results we can see that "good" similarity exists in areas which strengthen the relationship: reciprocating positives, focusing on positive aspects of the task, and using expressions which can minimize ambiguity and aid understanding. By contrast, "bad" similarity

reiterates and reinforces negative expressions, with individuals focusing on (personal) gain, rather than collaboration. Although we note more findings for content and structural levels, similarity of stylistic level features apparently provides a simple way of indicating or observing trust in groups.

Theoretically, this work provides greater insight into theories of language accommodation. While the vast majority of these theories appear to assume that convergence is good and divergence is bad, our work demonstrates that this is not the whole story. While our measures of convergence at the structural and stylistic levels support the notion that greater convergence is associated with positive social outcomes, convergence at the content level is more complex. Our findings suggest that convergence at the content level can relate to high and low levels of trust, depending upon the topic. By examining different forms of similarity, we note that not all similarity is good, and that it can operate independently across different communicative levels.

Methodologically, we have developed an analytical framework to examine linguistic similarity across three broad levels. We adapt automated content analysis methods in order to produce measurement of similarity at content and structural levels, and use a manual coding scheme for entrainment at the structural and stylistic levels. In addition, we also provide further validation for a more accurate and sensitive automated measure of defection.

Potential applications of this work include improving the design of natural language generation systems, so that similarity can be appropriately matched or not. For example, matching an angry user's language containing negative emotion expression would likely undermine the relationship, whereas reciprocating positive expressions, noun phrases or emoticons might help encourage trust establishment. Additionally, for distributed work teams or online daters, an interface which can remind interlocutors to stay on the topic or task may help to establish trust in the early stages of text-based relationships.

LIMITATIONS AND FUTURE WORK

While we feel that linguistic similarity is important to understanding trust development, certain strategies may also contribute to the amount of trust or cooperation achieved between individuals. For instance, Orbell and colleagues [19] argue that communicated promises increase cooperation in social dilemma games. While we recognize the effects of matched promises on trust establishment, we suggest that they could occur with or without linguistic similarity. For instance, in our first entrainment example, Speaker B agrees with Speaker A that they will both invest 50 tokens "the whole time." In this example, Speaker B could have said "sounds good", or "yeah, half-and-half" instead of repeating Speaker A's phrase. We think that linguistic similarity and matched promises are distinct constructs, although there is certainly some relation between the two.

While our findings provide insights into the relationship between language and trust, they are correlational. The findings do not explain whether trust influences linguistic similarity or whether linguistic similarity influences trust. However, as we learn more about the characteristics of the language used by high and low trusting pairs, we can begin to incorporate these characteristics into automated text generators in order to test specific hypotheses about causation. We might, for instance, experimentally manipulate the level of entrainment or repeated phrases an individual receives by intercepting and modifying their partner's IM responses. It would also be interesting to integrate behavioral responses during such interactions with physiological measures. As previously mentioned, we aim to increase our understanding of entrainment of particular noun phrases and how this relates to interlocutors' shared referential framework. This increased understanding would allow us to better integrate our findings within accommodation and entrainment theory frameworks.

CONCLUSION

Using an iterative social dilemma investment game, we examined how different levels of linguistic similarity relate to the establishment of trust in a text chat environment. We find that high trusting pairs exhibit "good" similarity, which is characterized by reciprocating positivity, focusing on the task and minimizing ambiguity. Low trusting pairs exhibit "bad" similarity which reinforces negativity and focuses on personal gain rather than collaboration. Results suggest that in CMC we *sometimes* trust, and that linguistic similarity plays a role in determining those times.

ACKNOWLEDGMENTS

This research was funded by National Science Foundation grant #0705901. We thank David Nguyen and Nathan Bos for software and comments on modifications. We also thank Alex Markov, Jason Sandler, Josh Waytz, Isaac Wilson, and our reviewers for their valuable insights.

REFERENCES

[1] van Baaren, R.B., Holland, R.W., Steenaert, B. & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology, 39,* 393-398.

[2] Bos, N., Olson, J.S., Gergle, D., Olson, G.M., & Wright, Z. (2002). Effects of four computer-mediated communications channels on trust development. In *Proceedings of CHI 2002*, 135-140.

[3] Bradac, J. (1990). Language attitudes and impression formation. In Giles, H. and Robinson, W. (Eds.), *Handbook of Language and Social Psychology*, 387–412. Wiley, Chichester.

[4] Brennan, S. E. & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *22*, 1482-1493.

[5] Chartrand, T.L., & Bargh, J.A. (1999). The chameleon effect: The perception behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893-910.

[6] Curhan, J.R. & Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes. *Journal of Applied Psychology, 92*(3), 802-811.

[7] Dawes, R.M. (1980). Social dilemmas. *Annual Reviews in Psychology, 31*, 169-93.

[8] Dunn, J.R. & Schweitzer, M.E. (2005). Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology, 88*(5), 736-748.

[9] Giles, H., Coupland, J., & Coupland, N. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland & N. Coupland (Eds.), *Contexts of Accommodation. Developments in Applied Sociolinguistics*. Cambridge University Press.

[10] Gill, A.J., Gergle, D., French, R.M., & Oberlander, J. (2008). Emotion rating from short blog texts. In *Proceedings of CHI 2008,* 1121–1124.

[11] Grice, H.P. (1975). Logic and conversation. In Cole, P. & Morgan, J. (Eds.) *Syntax and semantics*, *Vol. 3*. New York: Academic Press.

[12] Jarvenpaa, S., Knoll, K., & Leidner, D. (1998). Is anybody out there? Antecedents of trust in global virtual teams. *Journal of Management Information Systems, 14*(4), 29-64.

[13] Labov, W. (1972). *Sociolinguistic Patterns*. University of Pennsylvania Press, Philadelphia.

[14] Lenhart, A., Arafeh, S., Smith, A. & Rankin-Macgill, A. (2008). Writing, technology and teens, Pew Internet & American Life Project, April 24, 2008, http://www.pewinternet.org.

[15] Maddux, W.W., Mullen, E., & Galinsky, A.D. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology, 44*, 461-468.

[16] Mehl, M.R., & Pennebaker, J.W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology, 84*(4), 857-870.

[17] Nguyen, D.T., & Canny, J. (2007). Multiview: Improving trust in group video conferencing through spatial faithfulness. In *Proc. of CHI 2007*, 1465-1474.

[18] Niederhoffer, K., & Pennebaker, J.W. (2002). Linguistic style matching in social interaction. *Journal of Language & Social Psychology, 21*(4), 337-360.

[19] Orbell, J.M., van de Kragt, A.J., & Dawes, R.M. (1991). Covenants without the sword: The role of promises in social dilemma circumstances. In Koford, K.J. & Miller, J.B. Social Norms and Economic Institutions, University of Michigan Press.

[20] Pennebaker, J.W. and Francis, M. (1999). *Linguistic Inquiry and Word Count (LIWC).* Lawrence Erlbaum Associates, Mahwah, NJ.

[21] Quan-Haase, A., Cothrel, J., and Wellman, B. (2005). Instant messaging for collaboration: A case study of a high-tech firm. *Journal of Computer-Mediated Communication, 10*(4), article 13.

[22] Rayson, P. (2003). *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. Unpublished doctoral dissertation, Lancaster University, UK.

[23] Riegelsberger, J., Sasse, A.M., & McCarthy, J.D. (2003). The researcher's dilemma: Evaluating trust in computer-mediated communication. *International Journal of Human-Computer Studies, 58*(6), 759-781.

[24] Rogers, W.H. (1993). Regression standard errors in clustered samples. *Stata Technical Bulletin, 13*, 19-23.

[25] Rousseau, D.M., Sitkin, S.B., Burt, R.S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Mgmt. Review, 23*, 393-404.

[26] Scissors, L.E., Gill, A.J, & Gergle, D. (2008). Linguistic mimicry and trust in text-based CMC. In *Proceedings of CSCW 2008*, 277-280.

[27] Taylor, P. J., & Thomas, S. (2008). Linguistic style matching and negotiation outcome. *Negotiation and Conflict Management Research, 1*, 263-281.

[28] Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry, 1*, 285-293.

[29] Walther, J.B. & Parks, M.R. (2002). Cues filtered out, cues filtered in: Computer-mediated communication and relationships. In M.L. Knapp & J.A. Daly (Eds.), *Handbook of Interpersonal Communication* (3rd edition), 529-563. Thousand Oaks, CA: Sage.

[30] Willemyns, M., Gallois, C., & Callan, V.J. (2003). Trust me, I'm your boss: Trust and power in supervisor supervisee communication. *International Journal of Human Resource Management, 14*(1), 117-127.

[31] Wilson, J., Straus, S., & McEvily, B. (2006). All in due time: The development of trust in computer-mediated and face-to-face teams. *Organization Behavior and Human Decision Processes, 99*(1), 16-33.

[32] Zheng, J., Veinott, E., Bos, N., Olson, J.S., Olson, G.M., 2002. Trust without touch: Jumpstarting long-distance trust with initial social activities. In *Proceedings of CHI 2002*, 141–146.