Journal of Applied Communication Research
Vol. 39, No. 1, February 2011, pp. 19–37

Routledge
Taylor & Francis Group

Peer to Peer Lending: The Relationship Between Language Features, Trustworthiness, and Persuasion Success

Laura Larrimore, Li Jiang, Jeff Larrimore, David Markowitz & Scott Gorski

This study examined the relationship between language use and persuasion success in the Peer-to-Peer (P2P) lending environment where unaffiliated individuals borrow money directly from each other using a textual description to justify the loan. Over 200,000 loan requests were analyzed with Linguistic Inquiry and Word Count (LIWC) software. The use of extended narratives, concrete descriptions and quantitative words that are likely related to one's financial situation had positive associations with funding success which was considered to be an indicator of trust. Humanizing personal details or justifications for one's current financial situation were negatively associated with funding success. These results offer insights into how individuals can optimize their persuasiveness by monitoring their language use in online environments.

Keywords: Persuasion; Peer-to-Peer Lending; Online; Language; LIWC

The Internet has provided individuals with a wide audience of potential interaction partners. Online communication allows people to interact with others who are not geographically copresent (Walther & Bazarova, 2007). Some individuals enter the online world highly suspicious of strangers and interact primarily with those they know in the face-to-face world (Hancock, 2007). Others are too trusting and misjudge the risks of interacting with strangers, becoming victims of identity fraud or

Laura Larrimore is in the Department of Communications at Ithaca College. Li Jiang, David Markowitz, and Scott Gorski are in the Department of Communications at Cornell University. Jeff Larrimore is in the Department of Economics at Cornell University. Many thanks to Jeff Hancock and Amy Gonzales for their helpful comments and suggestions on earlier drafts of this paper. Correspondence to: Laura Larrimore, Ithaca College, Department of Communications, 311 Park Hall, Ithaca, NY 14850USA. E-mail: laura.larrimore@gmail.com

ISSN 0090-9882 (print)/ISSN 1479-5752 (online)
DOI: 10.1080/00909882.2010.536844

malicious computer viruses (Freidman, Kahn, & Howe, 2000). As trust is crucial for high-stakes interactions, difficulties in establishing trust online can limit an individual's ability to utilize this medium to find others willing to work with them in partnership. Therefore, it is important to determine how individuals can best facilitate trust online so that these online interactions can be more beneficial. In the current study, peer to peer (P2P) lending interactions are examined. The inclusion of non-financial information in P2P loan requests and the reliance on individual lay people to choose whom they will fund provides a natural setting for testing the impact of language use in creating a trusting relationship as granting a loan request can be considered an indicator of trust.

Humans have developed many persuasive strategies for use during face-to-face interactions such as direct selling or charitable fundraising. For example, individuals can increase the likelihood that their request will be complied with by employing foot-in-the-door or door-in-the-face techniques, encouraging liking by presenting a physically attractive image, or stressing similarities between the requester and the listener (Cialdini, 1993). However, the mechanisms via which individuals persuade one another to comply in computer mediated settings may not be the same. Some strategies known to build trust and increase compliance in face-to-face interactions are either inapplicable or do not work as expected in online spaces (Bos, Olson, Gergle, Olson, & Wright, 2002). Some theories primarily developed for face-to-face interactions, though not necessarily in regard to persuasion *per se*, such as the uncertainty reduction theory (Berger & Calabrese, 1975) and the elaboration likelihood model (Petty & Cacioppo, 1986), can be adapted to online interpersonal interactions (Tidwell & Walther, 2002) and virtual teams (Bos et al., 2002).

The current study examines trust building and persuasion tactics in online environments. It does so by considering how the uncertainty reduction theory and the elaboration likelihood model can be applied to the high-stakes online environment of peer-to-peer (P2P) lending. Although previous research has considered the applicability of face-to-face communication theories to online spaces, P2P lending provides two characteristics that make it uniquely beneficial for this analysis. Like other e-commerce applications, P2P lending is a high-stakes environment where parties are highly involved in the interaction and do not previously know their online partners. Although P2P lending is similar to other e-commerce applications in this regard, it differs in that all transactions are functionally equivalent—a promise to repay a loan with interest over a specified period of time. Thus, P2P lending allows exploration of the impact of language on trustworthiness in an online environment without concerns regarding differences in the products exchanged in the transactions.

Peer-to-Peer Lending

P2P lending is a relatively new online phenomenon, starting in 2005 in the United Kingdom with the creation of Zopa (Kupp & Anderson, 2007), and quickly spreading to the United States in 2006 with the founding of San Francisco-based Prosper

Marketplace, Inc. ("Prosper"). P2P lending companies have flourished amidst the global credit crunch, in part because they are an attractive alternative option for borrowers who are unlikely to qualify for a loan from a traditional bank. By 2009, Prosper alone had funded more than $178 million dollars (US) in loans (Prosper.com, 2008) and it is estimated that by 2013 the P2P loan market as a whole will have facilitated over 5 billion dollars (US) in loans (Gartner Inc., 2010).

In P2P online banking, borrowers appeal directly to a pool of individual lenders using a request which can include textual narratives to justify the loan. Normally in a P2P transaction, individuals who would like to borrow money (hereafter referred to as "borrowers") and those who would like to lend money (hereafter referred to as "lenders") have no previous relationship. Furthermore, almost all interactions between the lender and borrower occur through the website interface where a borrower submits a loan request and a lender chooses to fund that request or not. Beyond the issuance of the loan and the subsequent repayment, the borrower and lender are unlikely to have any future interactions. P2P websites act as the conduit, facilitating the requesting and bidding process and coordinating the payment process if a loan is made. However, unlike traditional banks, P2P sites are not funding the loans; it is the individual lenders who provide the capital and carry the default risk. Thus, P2P lending interactions are online, one-time interactions where the stakes are high for both parties and where the variables in the decision making process are almost exclusively defined by the borrower's profile and request. In such a high-risk context, lenders are highly involved and are motivated to deeply process the loan descriptions and all other information available about the borrower when making funding decisions.

Borrowers seeking loans through Prosper create a loan request which includes traditional financial information such as a credit grade and debt-to-income (DTI) ratio. This financial information is pulled directly from the borrower's credit report and is therefore verified for accuracy. The request can also include non-financial information such as a picture, loan title, and loan description. Borrowers can write as much or as little as they want in the loan description to persuade lenders of the important need for the loan or how they are going to repay it. Lenders read the requests and choose which borrowers to fund and how much funding to offer each borrower.

One major advantage of P2P lending for studying trust in the online environment is that the perceived trustworthiness of the borrower is easily quantifiable based on whether lenders are willing to loan money to a borrower. As described by Flanagin (2007), trust is "the perception of the degree to which an exchange partner will fulfill the transactional obligations in situations characterized by risk or uncertainty" (Flanagin, 2007, p. 406). As the transactional promise between two loans is nearly identical to a lender (a promise of monthly repayment of principal with interest at a defined rate), the reason a lender will choose to fund one loan over another is presumably the perceived trustworthiness of the borrower and the belief that the borrower will fulfill their repayment obligations. As such, funding success should directly reflect the perceived trustworthiness of the loan request.

The popular press has likened these loan profiles to those found on online dating websites (Frier, 2009; Mogul, 2007; Quinn, 2008) because both types of profiles are carefully edited for desirable self-presentation in order to achieve either dating or financial goals. Several previous studies have explored the impacts of personal characteristics that can be observed through the picture included in the profile, such as the physical attractiveness or race of the borrower (see e.g., Herzenstein, Andrews, Dholakia, & Lyandres, 2008; Pope & Sydnor, in press; Ravina, 2008). However, verbal elements such as the language use in loan requests have been under-researched. Previous studies have included basic information about the readability of the loan description (i.e., average word and sentence length) as control variables (Pope & Sydnor, in press), but research has not looked in depth at the language use in P2P lending and how certain linguistic features affect funding success.

The language a borrower uses to communicate the request could be used by lenders to make funding decisions. Word use is generally considered to be a meaningful marker of cognitive and social processes (Pennebaker, Mehl, & Niederhoffer, 2003), and the way people express themselves can often be more persuasive or informative than the content of what they say (Pennebaker & King, 1999; Tan, Swee, Lim, Detenber, & Alsagoff, 2007). Therefore, this investigation focuses on how language use in loan requests contributes to P2P lending success and trust-building. It explores what kinds of linguistics strategies can effectively present a trustworthy image and hence increase the likelihood of getting funded.

Uncertainty Reduction Theory

According to uncertainty reduction theory, exchanging and collecting information on each other reduces uncertainty and allows one to predict others' attitudes and behaviors (Berger & Calabrese, 1975). In initial encounters, strangers go through specific verbal and nonverbal steps to create positive impressions in others, and to help them to make judgments about people and situations. However, many nonverbal actions described by uncertainty reduction theory, such as gauging warmth based on facial expressions, body posture and eye contact, are limited in online settings. Tidwell and Walther (2002) argue that passive and interactive strategies frequently used in face-to-face uncertainty reduction become unavailable or less effective in online interactions. Lacking many of the common cues to reduce uncertainty, individuals in online environments tend to "pay keen attention to the few communication cues available . . . successfully attending to the subtle (but important) information that does exist" (Flanagin, 2007, p. 418).

Research on e-tailing and consumer-to-consumer selling (on sites such as eBay) indicate that reducing uncertainty is crucial for online sellers who hope to convince consumers of their legitimacy and trustworthiness and thus persuade consumers to buy from them during these one-shot, high-risk online interactions. For example, Flanagin (2007) claims that longer product descriptions are integral to uncertainty reduction as longer descriptions were correlated with increased bids and higher selling prices on eBay. Given that one of the primary methods of uncertainty reduction in other online contexts is to provide a lengthier description, this should also be true for online lending. High involvement in lending decisions should make lenders particularly attentive to the additional information provided by longer descriptions. Taken together, lengthier loan descriptions should lead to more effective uncertainty reduction and trust establishment, which are manifested through increased loan success.

> H_1: The use of longer loan descriptions will be positively associated with funding success.

Another linguistic feature closely related to uncertainty reduction is the use of concrete words. Concrete words are associated with more contextualized and detailed representations of objects (Doest, Semin, & Sherman, 2002; Schwanenflugel & Stowe, 1989; Seifert, 1997) and allow faster processing through both verbal and nonverbal semantic systems (Paivio, 1986, 1991). We reason that a concrete persuasive argument provides more specific information for representations of financial situations, and hence more effectively reduces uncertainty and builds lender confidence regarding whether the borrower has the means and desire to repay the loan. Consequently, lenders are more confident in their understanding of the borrower's situation when the profile provides additional concrete information and are more likely to fund loan requests that use more concrete words. In the present study, we pay attention to three language dimensions that signal the concreteness of persuasion messages, namely the use of article (e.g., a, an, and the), prepositions (e.g., in, at, of, on, etc) and quantifiers (e.g., many, lots of, etc.), and predict that these three dimensions contribute to funding success.

> H_2: The use of language specifying concreteness will be positively correlated with funding success.

Persuasion and Elaboration Likelihood Model

Although descriptions that reduce uncertainty are important for building the trust of lenders, it is also important to consider the approach through which borrowers are trying to persuade lenders of their creditworthiness. According to the elaboration likelihood model (Petty & Cacioppo, 1986), there are two main routes of persuasion: the central route and the peripheral route. The central route focuses on the message quality to persuade, while the peripheral route uses heuristics to help influence individual decisions regarding a topic. The engagement of either central or peripheral route depends on processing capability and message involvement. Individuals are more likely persuaded via the central route if they have the ability to process the information and if they are highly involved in the decision. They are more likely persuaded via the peripheral route if involvement is low and information processing capability is diminished.

P2P lenders could potentially be persuaded either by a centrally processed argument regarding ability to repay or by a peripherally processed argument that humanizes the loan. For example, a lender may be persuaded by a centrally processed argument which includes additional concrete financial details in the loan description, or the lender may be persuaded by a peripherally processed argument such as how the borrower is a parent and how the loan would benefit their young children. Both strategies provide increased information about the borrower and present a trustworthy image.

However, due to the importance, seriousness, and financial risk associated with P2P lending transactions, lenders should be more likely persuaded by central cues as opposed to peripheral cues. Hence we expect loan requests that provide strong, centrally processed arguments will be more likely to be successful. Within the context of a loan request, a strong, centrally processed argument usually elaborates on the borrower's financial situation and discusses repayment of a loan. This type of elaboration manifests itself through quantitative descriptions and increased uses of financial terms. Thus we predict that lenders will be swayed by loan descriptions that include number words, such as "second," "two," or "thousand" and money words such as "cash" or "owe."

> H_3: Providing quantitative information increases funding success.

In contrast to the straightforward financial quantitative descriptions, many borrowers also often provide humanizing details about themselves, such as information about their spouses or children, occupation, or membership in religious or social groups. Though the rationale for why borrowers choose to include this information is currently unknown, the use of humanizing details may be an attempt by borrowers to establish likability and trustworthiness by seeking similarities with potential lenders. This strategy would be successful primarily if lenders process the loan requests peripherally, since these details are likely unrelated to the borrower's capacity to repay the loan. However, to the extent that such details include meaningful information about one's financial situation—such as having a large family that may limit one's financial resources available for loan repayment—they could activate central processing. In such cases, the humanizing details provided may contribute to a bias towards individuals in certain demographic or social groups, activating central rather than peripheral processing. Of course, these biases may or may not represent true information about the trustworthiness of the borrowers.

As described above, due to high involvement in the decision-making process, lenders are more like to follow a central processing route and attend to the key issue of whether the borrower is capable of repaying the loan. Thus, if humanizing details are processed centrally and arouse lender biases, they may be an effective addition to the loan appeal. However, the impact of this humanizing information could be positively or negatively correlated with funding, depending on whether the details reveal information which lenders believe makes the borrower more or less able to repay the loan.

However, if humanizing details are processed peripherally, rather than centrally, then it is likely that humanizing details will decrease funding success. Too many peripherally processed arguments could act as a red-flag for lenders who value information about the ability to repay but only observe humanizing details. It could also serve as an indication that the borrower may be attempting to avoid revealing their poor financial situation by focusing on these peripheral arguments instead. As a result, providing humanizing details (i.e., words related to family, achievement, leisure activities, home religion, work, friends or death) may reduce loan success.

> H4a: Providing humanizing details will be associated with a decreased likelihood of funding success.

In addition to providing humanizing details in an attempt to establish a common background with potential lenders, some borrowers choose to justify their requests by explaining what caused the financial predicament. This strategy might work in some contexts, for example, Langer, Blank, and Chanowitz (1978) found a greater rate of compliance when a request was coupled with a justification. However, Langer's study examined a low-involvement decision that requires less careful thinking. We argue that providing justifications such as rationalizing financial predicaments cannot satisfy the need for centrally processed persuasion and are likely to be rejected by lenders. Making justifications could even deter persuasion in that it does not center the argument on the ability to repay and implicitly suggests that the borrower may continue to make excuses for defaulting in the future. Therefore, we expect that providing justifications (e.g., causal words) reduces funding success.

> H4b: Providing justifications will be associated with decreased funding success.

Method

To examine our hypotheses, we analyzed the loan descriptions seen by the Prosper lenders via text analysis, and then linked the categorized linguistics features to loan success.

Text Analysis Procedure

Data files containing loan requests and loan outcome information are publicly available from Prosper. Each loan request contains financial variables such as the borrower's credit grade, loan-specific variables such as the amount of money requested, and language variables such as the request's textual loan description. Loan request data encompassing 220,257 completed loan requests from 2005–2008 were downloaded from Prosper. Loan requests with less than 20 words in the loan description were removed, resulting in a total of 213,510 requests for the text analysis.

The linguistics features were first extracted using Linguistic Inquiry and Word Count (LIWC; Pennebaker, Booth, & Francis, 2007), an automated language analysis program that analyzes written or spoken samples on a word-by-word basis. LIWC analyzed the loan descriptions by checking the linguistic content against an internal dictionary (LIWC 2007 English dictionary) and grouping words into 45 predefined linguistics categories and subcategories (see Table 1). The LIWC output provided the percentage of total words belonging to each category. Of these LIWC categories, 15 were directly related to the theoretical hypotheses described above and were included in our primary analysis. The mean percentages of words in each of these categories in funded and unfunded loan requests are reported in Table 2. Additionally, Table 2 illustrates the uncontrolled correlation coefficient between each theoretically relevant variable and funding success.

Measures

Funding success. As noted earlier, funding success directly reflects perceived trustworthiness of the borrowers in P2P lending. On Prosper a loan request is only funded when it receives offers equal to 100% funding. If the funding percent is less than 100%, the borrower does not receive any money and the loan request expires without a loan transaction occurring. Since the outcome is the same for a loan request that receives 99% funding as one that receives no funding, the dependent

Table 1 Examples of words in each LIWC category analyzed

Category	Examples	Category	Examples
H_1 variables: length of description		*Additional Language Variables*	
Word count	N/A	Function words	I, me, they, those, the
		1st person singular	I, me, mine
H_2 variables: concreteness		1st person plural	We, us, our
Articles	A, an, the	2nd person	You, your, thou
Prepositions	To, with, above	3rd person singular	She, her, him
Quantifiers	Few, many, much	3rd person plural	They, their, they'd
		impersonal pronoun	It, it's, those
H_3 variables: quantitative content		Common Verbs	Walk, went, see
Numbers	Second, thousand	Auxiliary verbs	Am, will, have
Money	Audit, cash, owe	Past tense	Went, ran, had
		Present tense	Is, does, hear
H_{4a} variables: humanizing details		Future tense	Will, gonna
Family	Son, husband, aunt	Adverbs	Very, really, quickly
Friends	Buddy, friend, neighbor	Conjunctions	And, but, whereas
Work	Job, majors, xerox	Negations	No, not, never
Achievement	Earn, hero, win	Positive emotion	Love, nice, sweet
Leisure	Cook, chat, movie	Negative emotions	Hurt, ugly, nasty
Home	House, kitchen, family	Insight	think, know, consider
Religion	Altar, church, mosque	Tentative	maybe, perhaps
Death	Bury, coffin, kill	Certainty	always, never
		Inhibition	block, constrain, stop
H_{4b} variables: justifications		Inclusive	And, with, include
Causation	because, effect, hence	Exclusive	But, without, exclude
Discrepancy	should, would, could	Perceptual process	See, touch, listen
		Biological processes	Eat, blood, pain
		Relativity	Area, bend, stop

Table 2 Summary of statistics of theoretically relevant language variables analyzed and correlations of language variables with funding success

Variable	Mean for unfunded loans	Mean for funded loans	Correlation with funding success
H_1 variables: length of description			
Word count	183	228	.116
H_2 variables: concreteness of description			
Articles	4.51	5.15	.099
Prepositions	11.65	12.58	.094
Quantifiers	2.12	2.47	.077
H_3 variables: quantitative content/centrally processed arguments			
Numbers	.57	.67	.042
Money	12.83	11.62	−.073
H_{4a} variables: humanizing details/peripherally processed arguments			
Family	.62	.55	−.025
Friends	.06	.08	.028
Work	6.95	7.07	.017
Achievement	2.32	2.28	−.011
Leisure	1.71	1.34	−.102
Home	1.84	1.51	−.083
Religion	.10	.08	−.019
Death	.03	.02	−.012
H_{4b} variables: justifications/peripherally processed arguments			
Causation	2.44	2.17	−.060
Discrepancy	1.30	1.24	−.015

Note: All means except word count are percent of total words in loan requests in the language category.

variable is a binary outcome of whether the loan request received 100% of the requested funding. Slightly over 13% of loans during the period of analysis received full funding. Prosper data is valuable in its ability to measure trustworthiness online, as differences in funding success more clearly indicates perceived differences in the trustworthiness of the borrower than is possible in many other online situations where the products exchanged are less homogeneous.

Financial variables. Every loan request contains two types of financial information: profile variables and request variables. Profile variables included items that describe the borrower, such as *credit grade* (a categorical representation of an individual's credit score), *debt-to-income ratio* (*DTI ratio*, the percentage of a consumer's monthly gross income that goes toward paying debts), and dummy variables for *owning a home* and *presenting a profile image*. Prosper creates many of these variables based on the borrower's credit report and they could not easily be changed by borrowers, with the profile picture being an exception. Request variables described the loan, including items such as *amount requested* (the amount of money requested), *borrower maximum rate* (the highest interest rate that the borrower is willing to pay) and *duration* (the duration of the loan requested). Profile variables and request variables are similar to the information typically used by a bank in assessing loan requests and thus are extremely important to the probability of a loan receiving funding. Thus, they were included as controls in all analyses.

Results

Because of the binary nature of funding success, a Probit regression was performed using categorized linguistic features as predictors and financial variables as controls. To test the sensitivity of results to the specification of the dependent variable as a binary outcome, we also performed a Tobit regression on the actual percent funded, but the results showed similar patterns compared to those of the Probit regression.[1] This confirmed that using the binary dependent variable instead of using the actual percent funding did not greatly impact the results. In order to control for all of the financial and linguistic information seen by lenders when making their decisions to fund loans, a single Probit regression model was computed and the results from that regression are used to test each of the hypotheses.

The results of the Probit regression which included the financial variables and linguistic features relevant to our theoretical predictions are shown in Table 3. For ease of interpretation, all results are reported as marginal effects (dF/dx), representing the change in the probability of funding (where a probability of 1 is a loan certain to be funded) from an infinitesimally small change in the explanatory variable at its mean value. Thus, for example, increasing the length of the average request by one word is expected to improve funding success by .017 percentage points. A more substantial change in description length, such as an increase in the description by 60 words, can therefore be quite important. This 60 word increase in loan description would increase funding success by about 1 percentage point for the average loan, which is equivalent to the improvement from reducing the borrower's maximum acceptable interest rate for the loan by 1.7 percentage points.

As expected, all financial variables such as borrower's credit grade and the amount requested were strongly predictive of funding success. Additionally, much of the fit of the model came from the financial variables—when estimating the model with just financial variables the McFadden's Pseudo R^2 is .3588 compared to .3764 when using the financial variables and the theoretically relevant language variables together (if you include just the theoretically relevant language variables without the financial variables, this results in a McFadden's Pseudo R^2 of .0386).[2] Given that traditional banks almost exclusively use these financial variables in funding decisions, neither the significance of financial variables in the regression nor the importance of financial variables to the model fit should be surprising. However, since many of these variables were pulled directly from the borrower's credit report and thus could not be easily adjusted by the borrower in a short period of time, the knowledge that financial variables are important is not particularly valuable to borrowers attempting to improve their funding success.

In contrast, the language in the loan description could be easily changed by the borrower. While the language information is secondary to the borrower's financial situation, given that loan descriptions can be easily adjusted, it is notable that the fit of the model improves from the inclusion of this additional information (from a Pseudo R^2 of .3588 to .3764) and that many of the language variables have a

Table 3 Probit regression results for the impact of language on funding success including only theoretically relevant language variables and financial variables

	Marginal effect[a]	Standard error
H_1 variables: length of description		
Word count	.00017*	(.00001)
Word count squared	−.0000001*	(.0000000)
H_2 variables: concreteness of description		
Articles	.00085*	(.00019)
Prepositions	−.000003	(.000135)
Quantifiers	.00139*	(.00025)
H_3 variables: quantitative content/centrally processed arguments		
Numbers	.00170*	(.00042)
Money	.00079*	(.00012)
H_{4a} variables: humanizing details/peripherally processed arguments		
Family	−.00297*	(.00043)
Friends	.00072	(.00136)
Work	.00139*	(.00016)
Achievement	−.00156*	(.00028)
Leisure	−.00367*	(.00037)
Home	−.00114*	(.00033)
Religion	−.00469*	(.00098)
Death	−.00160	(.00214)
H_{4b} variables: justifications/peripherally processed arguments		
Causation	−.00205*	(.00029)
Discrepancy	−.00355*	(.00033)
Financial and profile variables[b]	Yes	
Additional language variables	No	
Observations	213,510	
McFadden's Pseudo R^2	.3764	

[a]Marginal effect is the increased probability of funding success from an infinitesimal increase in the value of the explanatory variable.
[b]Included financial and profile variables are credit grade, debt-to-income ratio, total credit lines, delinquent accounts, delinquencies in the past seven years, credit inquiries, public records in the past 10 years, homeowner status, length of employment, amount requested, maximum acceptable interest rate, loan purpose, number of days the listing is posted for, whether a quick close on the loan is requested, date of the listing, current prime rate, and whether a picture is included.
*$p < .01$.

significant impact on funding success. The specific impacts of language in the loan descriptions as it relates to our hypotheses are discussed in more detail below.

Hypotheses Testing

Recall that H_1 predicted that longer loan descriptions would increase funding success. Among the language variables, word count, the most direct indictor of description length, was a positive predictor of funding success ($dF/dx = .00017$, $p < .001$[3]). However, the effect of word count squared was significantly negative, indicating that there were decreasing returns to longer descriptions ($dF/dx = -.0000001$, $p < .001$). Thus adding words to lengthen a short description was more beneficial than adding additional words to an already lengthy description. Despite these decreasing returns, it is important to note that within the range of word-count observations (20 to 817 words) the total effect of additional words was always positive. This significant positive impact of word count on funding success supported H_1.

H_2 predicted that the language dimensions that reflect concreteness would increase loan success. This study considered three linguistics features that reflect concreteness: articles, quantifiers, and prepositions. While prepositions were not significant at the 1% level, the results indicated that both articles ($dF/dx = .00085$, $p < .001$), and quantifiers ($dF/dx = .00139$, $p < .001$) positively predicted loan success. Thus, H_2 was partially supported.

H_3 predicted that quantitative information related to the ability to repay the loan would increase funding success by appealing to the lenders' central processing. The analysis considered the use of number words (e.g., "second," "two," or "thousand") and money words (e.g., "cash" or "owe") as the measures of quantitative information. Number words were positively associated with funding success ($dF/dx = .00170$, $p < .001$), and the magnitude of the effect was the largest positive effect among all the linguistics features. Money words similarly demonstrated a positive, significant effect ($dF/dx = .00079$, $p < .001$). Therefore our results supported H_3.

H_4 hypothesized that information that is irrelevant to the ability to repay and appeals to peripheral processing would decrease funding success since this information may distract rational reasoning. Two commonly observed strategies in P2P lending, namely providing humanizing details (H_{4a}) and providing justifications (H_{4b}), were expected to decrease the likelihood of funding success.

As predicted, increased discussion of about one's family ($dF/dx = -.00297$, $p < .001$), achievements ($dF/dx = -.00156$, $p < .001$), leisure activities ($dF/dx = -.00367$, $p < .001$), home ($dF/dx = -.00114$, $p < .001$), or religion ($dF/dx = -.0469$, $p < .001$) all had significant negative impacts on funding success. Two additional humanizing details—discussions of friends or death—were both not statistically significant at the 1% level. The only humanizing detail that had a significant positive impact on funding success was discussion of work activities ($dF/dx = .00139$, $p < .001$). However, this may be due to borrowers offering work details when discussing their financial status and repayment potential rather than due to the humanizing details that the work discussion provides. Thus, in support of H_{4a}, the data showed that in general providing humanizing details actually harmed funding success.

For making justifications, the results indicated a negative relationship between causal words and funding success ($dF/dx = -.00205$, $p < .001$), suggesting that providing explanations about why the borrowers needed the loan actually reduced the chance of being funded. Similarly, discrepancy words such as "should" or "could" were negatively associated with funding success ($dF/dx = -.00355$, $p < .001$). This was also consistent with justifications being negatively related to funding success as individuals are likely to use these words when attempting to justify why things went wrong or how they would prevent similar mistakes in the future. Hence, H_{4b} was also supported.

Additional Language Indicators

In addition to the theoretically supported language variables from our first regression, an additional 25 LIWC categories (presented in the second column of Table 1) were added into the initial regression model to test the robustness of our results. The additional language categories included personal pronouns as well as those that encompassed an average of at least .5% of the language in loan requests. The results of this extended regression are provided in Table 4.

While the addition of 25 more word categories in the second regression did not greatly change the results for the 15 main linguistic features from the primary regression, the second regression did demonstrate that there were other influential categories beyond our hypotheses. As these variables were not included in our hypotheses, they will not be discussed further in this paper. However, given that numerous additional word categories were found to be significant at the 1% level it is clear that the language used in the loan descriptions is important to lenders. Thus these additional observed effects warrant further investigation to determine how their results coincide with other communications theories beyond those explored in this paper.

Discussion

Building trust is critical to many online interactions. For various activities, including lending money, the online medium provides unique advantages such as an increased access to interested participants, increased ability to monitor cues in order to successfully self-present, and ample time to compose messages. Each of these elements can make online interactions more attractive than traditional interactions. However, users have struggled to determine which strategies work best in establishing the trust necessary to facilitate online interactions. This study investigated how people can be more successful in their attempts to establish trust verbally in P2P lending interactions and persuade others to interact with them. The results provide some insights into the linguistic features of successful trust-building and persuasion and illustrate that strategies based on uncertainty reduction theory and elaboration likelihood model can be applied to online persuasion contexts.

Specifically, in accord with the uncertainty reduction theory, the results show that providing more information by either increasing word count or using more concrete expressions (e.g., articles, quantifiers, and prepositions) can increase trust in the P2P lending context, presumably by reducing uncertainty regarding the borrower and the transaction. The findings also demonstrate that linguistic features related to a borrower's ability to pay back a loan (e.g., number words and money words) contribute to loan success, suggesting that the lenders are following a central processing route. In contrast, elements of loan descriptions that would appeal to the lenders' peripheral processing, including providing humanizing details (e.g., friends, family, religions, leisure activities, etc.) and justifications (e.g., rationalizing financial predicament) actually decreases the likelihood of funding success.

Table 4 Probit Regression Results including all language variables and financial variables

	Marginal effect[a]	Standard error		Marginal effect[a]	Standard Error
H_1 variables: length of description			Insight	.00040	(.00042)
Word count	.00017*	(.00001)	Tentative	−.00056	(.00036)
Word count	−.0000001*	(.00000)	Certainty	.00172*	(.00040)
H_2 variables: concreteness			Inclusive	.00057	(.00026)
Articles	.00188*	(.00035)	Exclusive	−.00018	(.00043)
Prepositions	.00076	(.00030)	Perceptual process	−.00039	(.00048)
Quantifiers	.00199*	(.00036)	Biological processes	−.00213*	(.00038)
H_3 variables: quantitative content			Relativity	.00059*	(.00012)
Numbers	.00257*	(.00051)	*Financial and profile variables*		
Money	.00040*	(.00013)	AA Credit grade	.05883*	(.00386)
H_{4a} variables: humanizing details			B credit grade	−.02481*	(.00071)
Family	−.00221*	(.00045)	C credit grade	−.04182*	(.00066)
Friends	.00060	(.00134)	D credit grade	−.05996*	(.00086)
Work	.00066*	(.00017)	E credit grade	−.08503*	(.00109)
Achievement	−.00149*	(.00028)	HR credit grade	−.32243*	(.00353)
Leisure	−.00360*	(.00038)	NR credit grade	−.03327*	(.00055)
Home	−.00150*	(.00033)	Borrower maximum	.58300*	(.00672)
Religion	−.00448*	(.00097)	Amount requested	−.00001*	(.00000)
Death	−.00091	(.00210)	Current prime rate	−.20815*	(.07198)
H_{4b} variables: justifications			Current delinquencies	−.00351*	(.00012)
Causation	−.00021	(.00031)	Delinquencies last 7 yrs	−.00049*	(.00003)
Discrepancy	−.00217*	(.00037)	Inquiries last 6 months	−.00140*	(.00009)
Additional language variables			Homeowner (1 = yes)	.00419*	(.00079)
Function words	−.00118*	(.00030)	Duration of listing	.00279*	(.00018)
1st person singular	.00137*	(.00033)	Length of employment	.00002*	(.00001)
1st person plural	.00088	(.00038)	Public records last 10 yrs	−.00422*	(.00035)
2nd person	.00158*	(.00055)	Debt-to-income ratio	−.00950*	(.00038)
3rd person singular	−.00020	(.00060)	Debt-to-income missing	−.03361*	(.00059)
3rd person plural	−.00122	(.00081)	Quick close-on-funding	.02709*	(.00107)
Impersonal pronoun	−.00154*	(.00037)	Total credit lines	−.00030*	(.00003)
Common Verbs	.00070*	(.00015)	Picture provided	.01801*	(.00073)
Auxiliary verbs	.00145*	(.00038)	Purpose—consolidate	−.00575*	(.00168)
Past tense	−.00102*	(.00032)	Purpose—home	.00297	(.00311)
Present tense	−.00150*	(.00025)	Purpose—business	−.00928*	(.00167)
Future tense	−.00157*	(.00051)	Purpose—personal	.00047	(.00196)
Adverbs	.00052	(.00032)	Purpose—student loan	−.00300	(.00283)
Conjunctions	−.00080	(.00031)	Purpose—auto loan	.00592	(.00401)
Negations	.00180*	(.00053)	Purpose—other	.00395	(.00264)
Positive emotion	.00108*	(.00019)	Credit data missing	−.02136*	(.00098)
Negative emotions	.00009	(.00047)	Month of listing	−.00420*	(.00011)

[a]Marginal effect is the increased probability of funding success from an infinitesimal increase in the value of the explanatory variable.
*$p < .01$.

Practical Application

Our research indicates that there are several strategies borrowers' should use to increase trust and thus the likelihood of getting funded. First, individuals should work to reduce uncertainty and present more information to lenders. Simply extending the length of their loan description can help the lenders understand how the money is to be spent and repaid. However, the benefit of extending the narrative is most effective for short descriptions and provides a smaller benefit as the description increases in length. An additional effective strategy to attract lenders is to describe the loan more concretely by including details. For example, "I can pay back the loan pretty soon using my first two months' salary" sounds more persuasive than simply saying that "I can pay back the loan using my salary." Providing concrete details help lenders better understand what they can expect after funding the money.

Secondly, borrowers should appeal to lenders by using the central route of persuasion. These results suggest borrowers should use more specific, rational arguments to demonstrate their credit worthiness, and support their arguments with more factual details. The results also suggest minimizing the space spent discussing non-financial personal details or rationalizing one's financial predicament. A rational, detailed and informative request seems to be consistently preferred over a pathetic sob story. Although tales of woe are common in P2P lending appeals, the results show that they do not elicit the sympathy or trust of lenders or increase funding likelihood. It is much better to describe a loan by saying, "my job pays $2,500 a month, and I can save $500 per month to pay back the loan," than saying, "I love my children more than anything and go to church regularly."

Linguistic Features and Persuasion across Contexts

Another question of interest is what the present study tells us, beyond the P2P lending context, about persuasion strategie and more generally about computer-mediated communication. The linguistic features identified in this study, although serving to achieve context-specific goals (e.g., reduce risk and uncertainty involved in P2P lending), may also be predictive of persuasion success in other contexts. Providing more information and details to reduce uncertainty has previously been found to be successful in other online spaces. For example, longer and more detailed product descriptions are associated with better reviews and greater purchasing intention in e-commerce (Flanagin, 2007; Yang, Hung, Sung, & Farn, 2006) and online dating profiles that have longer and more concrete self-descriptions are perceived as more trustworthy (Toma & Hancock, in press). This study similarly finds that longer loan descriptions are correlated with increased funding success in peer-to-peer lending. Consistent with the results in this investigation, argument quality appears to matter greatly in high-involvement online commerce (Park, Lee, & Han, 2007). High-involvement consumers are more motivated to think critically about issue-related arguments and scrutinize the relative merits and relevance of those arguments, prior to making decisions or judgments.

Additionally, recent research has attempted to apply communication theories that originally pertained to face-to-face settings in computer-mediated contexts. Both uncertainty reduction theory and elaboration likelihood model have shown a wide range of applications over various media contexts, with slight modifications to accommodate context differences (Flanagin, 2007; Tidwell & Walther, 2002). These applications have focused on exchanges of product information (Flanagin, 2007; Yang et al., 2006), socio-emotional contents (Tidwell & Walther, 2002), and presentation of nonverbal cues (Pope & Sydnor, in press; Ravina, 2008). The present investigation contributes to this research line by applying the uncertainty reduction theory and the elaboration likelihood model to an online context. Our results find support for these theories' applicability to the online environment, strengthening the view that many of the face-to-face strategies such as uncertainty reduction strategies and dual processing strategies can also be used in an online environment.

Avenues for Future Research

Because the phenomenon of P2P lending is only five years old, there are numerous avenues for future research. Current research is establishing a base understanding of how the P2P world operates, and hopefully future research will be able to build upon this and employ increasingly sophisticated linguistic theories to understand the many nuances of the borrower-lender interaction. Theories such as the linguistic category model (Semin & Fiedler, 1988; Semin, 2008) hold promise for investigating how psychological processes of the borrower may impact how they construct messages and help explain why borrowers tend to devote a large amount of space to non-persuasive information. Once the socio-cultural norms of P2P lending are better understood, language expectancy theory may also be a fruitful avenue for further research, in helping to provide a more nuanced understanding of what language lenders expect from borrower as influenced by individual differences, social norms, culture, and behavioral patterns (Burgoon, 1995; Burgoon & Miller, 1985; Burgoon & Stewart, 1975; for a recent overview of the LET literature, see Burgoon, Denning, & Roberts, 2002).

A second avenue for future research is exploring how learning occurs from others' past interactions in an online environment. As the P2P lending market expands, borrowers can observe the credit profiles and loan descriptions of successful and unsuccessful loans from earlier generations of borrowers. One might expect that borrower techniques may evolve as they learn from others' mistakes. Similarly, lenders may also learn from earlier interactions and update the types of profiles that they perceive as trustworthy, leading to changes in the types of language they find persuasive. As a result, future research may consider the impact of both borrower and lender learning on persuasion techniques.

Finally, an important area of research in P2P lending is to determine whether language is predictive of loan repayment for fully funded loans. This paper has focused on the interaction between borrowers and lenders to determine whether borrowers are able to improve their perceived trustworthiness and funding success through loan descriptions. Having observed that these descriptions do impact funding success, it is now worth considering whether lenders are using this information appropriately and whether a borrower's perceived trustworthiness accurately reflects their true trustworthiness.

Notes

[1] As exceptional loans can never receive more than 100% of requested funding and particularly poor loans can never receive less than 0% funding, there is a floor and ceiling on the possible range of values for the dependent variable. Therefore, if the data are treated as purely linear as is the case in a standard OLS regression, the large number of results bunched at 0 and 100% funding will bias the results. Thus, for the robustness check, a Tobit regression was used instead which accounts for this censoring of possible observations of the dependent variable. Results using the alternate Tobit specification are available from the author upon request.

[2] The Probit regression does not have an equivalent R^2 to that found in an OLS regression so the McFadden's Pseudo R^2 value is reported instead. This is default Pseudo R^2 value provided by Stata for a Probit regression but like most Pseudo R^2 measures cannot be directly compared to standard R^2 values. Results for alternate goodness-of-fit measures are available upon request from the authors.

[3] All values reported in the text are from Table 3, which includes just the financial variables and theoretically supported language variables.

References

Berger, C.R., & Calabrese, R.J. (1975). Some exploration in initial interaction and beyond: Toward a developmental theory of communication. *Human Communication Research*, *1*(2), 99–112.

Bos, N., Olson, J.S., Gergle, D., Olson, G.M., & Wright, Z.(2002). Effects of four computer-mediated communications channels on trust development. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 135–140. Minneapolis, Minnesota, USA.

Burgoon, M. (1995). Language expectancy theory: Elaboration, explication, and extension. In C.R. Berger & M. Burgoon (Eds.), *Communication and social influence processes* (pp. 29–52). East Lansing, MI: Michigan State University Press.

Burgoon, M., Denning, P.V., & Roberts, L. (2002). Language expectancy theory. In J.P. Dillard & M. Pfau (Eds.), *The persuasion handbook: Developments in theory and practice* (pp. 117–137). London: Sage Publications.

Burgoon, M., & Miller, G.R. (1985). An expectancy interpretation of language and persuasion. In H. Giles & R.N. St. Clair (Eds.), *Recent advances in language communication and social psychology* (pp. 199–229). London: Lawrence Erlbaum.

Burgoon, M., & Stewart, D. (1975). Empirical investigations of language intensity: The effects of sex of source, receiver, and language intensity on attitude change. *Human Communication Research*, *1*(3), 244–248.

Cialdini, R.B. (1993). *Influence: The psychology of persuasion* (revised ed.). New York: Quill.

Doest, L., Semin, G.R., & Sherman, S.J. (2002). Linguistic context and social perception: Does stimulus abstraction moderate processing style? *Journal of Language and Social Psychology*, *21*, 195–229.

Flanagin, A.J. (2007). Commercial markets as communication markets: Uncertainty reduction through mediated information exchange in online auctions. *New Media & Society*, *9*(3), 401–423.

Freidman, B., Kahn, P., & Howe, D. (2000). Trust online. *Communications of the ACM*, *43*(12), 34–40.

Frier, S.(2009, June 30). Peer-to-peer lending is flourishing amid the credit crisis. *Sacremento Bee*, p. 6b.

Gartner, Inc. (2010). *Gartner says 50 per cent of banks will still lack an innovation programme and budget by 2013*. Retrieved from http://www.gartner.com/it/page.jsp?id=1272313.

Hancock, J.T. (2007). Digital deception: Why, when and how people lie online. In K. McKenna, T. Postmes, U. Reips, & A. Joinson (Eds.), *Oxford Handbook of Internet Psychology*. Oxford: Oxford University Press.

Herzenstein, M., Andrews, R., Dholakia, U.M., & and Lyandres, E. (2008). The democratization of personal consumer loans? Determinants of success in online peer-to-peer lending communities. [Unpublished working paper]. Retrieved from http://ssrn.com/abstract = 1147856

Kupp, M., & Anderson, J. (2007). Zopa: Web 2.0 meets retail banking. *Business Strategy Review*, *18*(3), 11–17.

Langer, E.J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of "placebic" information in interpersonal interaction. *Journal of Personality and Social Psychology*, *36*(6), 635–642.

Mogul, M. (2007, July). Look for a higher return? Consider lending to a peer. *Kiplinger Business Forecasts*, *8*. http://www.kiplinger.com/businessresource/forecast/archive/consider_lending_to_a_peer_070709.html.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.

Park, D., Lee, J., & Han, I. (2007). The effect of online consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, *11*(4), 125–148.

Pennebaker, J.W., Booth, R.J., & Francis, M.E. (2007). *Linguistic inquiry and word count (LIWC2007): A computer-based text analysis program*. Mahwah, NJ: Lawrence Erlbaum.

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

Pennebaker, J.W., Mehl, M.R., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547–577.

Petty, R.E., & Cacioppo, J.T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Pope, D.G., & Sydnor, J.R. (in press). What's in a picture? Evidence of discrimination from prosper.com. *Journal of Human Resources*.

Prosper.com. (2008). *Prosper releases market survey results for September 2008*. Retrieved from http://www.prosper.com/about/media_press_releases.aspx?t = Prosper_Releases_Market_Survey_Results_for_September_2008

Quinn, J. (2008, August 27). Financing is as close as the nearest angel. Retrieved from http://www.bloomberg.com/apps/news?pid = 20601039&sid = a7LI2Jh3XdHM

Ravina, E. (2008). Love and loans: The effect of beauty and personal characteristics in credit markets. [Unpublished working paper]. Retrieved from http://ssrn.com/abstract = 1101647

Schwanenflugel, P., & Stowe, R.W.(1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, 114–126.

Seifert, L.S. (1997). Activating representations in permanent memory: Different benefits for pictures and words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 1106–1121.

Semin, G. (2008). Language puzzles: A prospective retrospective on the Linguistic Category Model. *Journal of Language and Social Psychology*, *27*(2), 197–209.

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

Tan, K.W.P., Swee, D., Lim, C., Detenber, B.H., & Alsagoff, L. (2007). The impact of language variety and expertise on perceptions of online political discussions. *Journal of Computer-Mediated Communication*, *13*(1), article 5. Retrieved from http://jcmc.indiana.edu/vol13/issue1/tan.html

Tidwell, L.C., & Walther, J.B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research*, *28*(3), 317.

Toma, C., & Hancock, J.T. (in press). What lies beneath: the linguistic traces of deception in online dating profiles. *Journal of Communication*.

Walther, J.B., & Bazarova, N.N. (2007). Misattribution in virtual groups: The effects of member distribution on self-serving bias and partner blame. *Human Communication Research*, *33*(1), 1–26.

Yang, S.H., Hung, W.C., Sung, K., & Farn, C.K. (2006). Investigating initial trust towards e-tailers from the elaboration likelihood model perspective. *Psychology & Marketing*, *23*(5), 429–445.