Mind Your Ps and Qs: The Impact of Politeness and Rudeness in Online Communities

Moira Burke and Robert Kraut
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213
{moira, robert.kraut}@cmu.edu

ABSTRACT
Little is known about the impact of politeness in online communities. We use an inductive approach to automatically model linguistic politeness in online discussion groups and determine the impact of politeness on desired outcomes, such as increased reply rates. We describe differences in perceived politeness across a variety of groups and find that, controlling for group norms of responsiveness and message length, politeness increases reply rates in some technical groups, but rudeness is more effective in some political groups. The perceived politeness scores will be used to validate linguistic politeness strategies from theory and to inform the creation of a machine learning model of linguistic politeness that can be applied as a "politeness checker" to educate newcomers to write in ways likely to elicit response from specific communities or as a rudeness detection tool for moderators.

Author Keywords
Linguistic politeness, computer-mediated communication, community responsiveness.

ACM Classification Keywords
H.5.3 [Information Interfaces]: Group and Organization Interfaces - Collaborative computing, Web-based interaction, Computer-supported cooperative work.

INTRODUCTION
Though our mothers advised us to mind our p's and q's, little is known about the effect of politeness in computer-mediated communication. This is especially true for online communities, in which people attempt to start conversations and make requests of strangers. Does polite conflict resolution lead a Wikipedia editor to be promoted to admin status? Do polite responses to newcomers in health support groups cause those newcomers to help others in the future? Does it get you killed in World of Warcraft?

To answer these questions, we need ways to measure politeness in online communities. The current project combines two approaches to train a machine learning algorithm to automatically model polite language: (1) A deductive approach based on Brown and Levinson's linguistic politeness theory [2], and (2) an inductive, survey-based method that identifies text perceived as polite. The present paper describes progress on both fronts: (1) an extension of Brown and Levinson's politeness theory in a coding manual for online discussion groups and (2) a survey of perceived politeness, in which current newsgroup messages are rated, to use as a gold standard for a machine learning model of polite language. Using reply count data for these messages, we determined the impact of politeness on community responsiveness, and found large differences: Politeness tripled reply counts in some technical groups, while rudeness was more effective in eliciting replies in some political issue groups.

CSCW'08, November 8–12, 2008, San Diego, California, USA.

LINGUISTIC POLITENESS THEORY

Face and Linguistic Politeness Strategies
Linguistic politeness theory begins with Goffman's theory of "face" [7]. He claims people present an identity with positive social value and want to have that identity validated by others. However, in the presence of others, they are subject to face-threatening actions, such as impositions and criticisms. So, they engage in face-work when communicating to help maintain each other's identities. Grice's maxims for efficient conversation describe the most direct forms of speech (e.g. "Take out the trash.") [8]. Yet these forms are often lengthened (e.g. "Would you please take out the trash?"), and Brown and Levinson propose that this inefficiency is an attempt to save another's face [2]. By being indirect, people imply some degree of politeness, which the hearer recognizes while still understanding the underlying meaning of the utterance [10].

Based on observations of language in three cultures, Brown and Levinson describe a typology of linguistic politeness strategies. In this project, we focus on two categories of strategies: "negative politeness" in which the speaker attempts to minimize the imposition on the listener (e.g. "If you have the chance, would you close the window?"), and "positive politeness" indicating a social connection between the speaker and listener (e.g. "Let's close the window."). Specific strategies from each category are listed below.

Negative politeness strategies:

N1. Be conventionally indirect
N2. Question, hedge
N3. Be pessimistic
N4. Minimize the imposition
N5. Give deference
N6. Apologize
N7. Impersonalize the speaker and hearer
N8. State the face threatening action as a general rule
N9. Nominalize
N10. Go on record as incurring a debt

Positive politeness strategies:

P1. Notice, attend to the hearer's needs
P2. Exaggerate interest, approval, sympathy
P3. Intensify interest to the hearer
P4. Use in-group identity markers
P5. Seek agreement
P6. Avoid disagreement
P7. Presuppose/raise/assert common ground
P8. Joke
P9. Assert/presuppose knowledge of hearer's concerns
P10. Offer, promise
P11. Be optimistic
P12. Include both speaker and hearer in activity
P13. Give or ask for reason
P14. Assume or assert reciprocity
P15. Give gifts to the hearer

There are two main criticisms of Brown and Levinson's model. First, the strategies are ambiguous, partially overlapping, and fall at many different levels of communication, from syntactic (e.g. question-form) to pragmatic (e.g. joking) [14]. The lowest-level strategies may be relatively easy to detect automatically, while the higher-level ones will be difficult even for human coders. Second, the focus is on the speaker's perception of politeness, rather than the recipient's [14]. Yet speakers and writers often overestimate their ability to convey subtler cues, such as sarcasm [12]. Given this overestimation and myriad cultural norms regarding politeness, it is possible that intended politeness is not always received. Therefore, the current project matches ratings of perceived politeness with intended politeness.

POLITENESS RESEARCH IN CMC

Politeness research in computer-mediated communication generally falls into two camps: One applies small subsets of Brown and Levinson's typology to medium-sized corpora, while the other applies all or most of the typology to very small datasets[1]. The small subsets of Brown and Levinson's typology commonly applied are those terms that are easily countable, such as "please," "thank you," "would," and hedges [3,14]. Brennan and O'Haeri [1] counted hedges and questions in instant messaging, and suggested that the belief that people sound less polite in CMC can be attributed to production costs: It takes more time to type hedges and indirect requests in fast-paced CMC, and so people use balder, shorter forms. Yet adding a question mark takes little extra effort, so question forms were as common in instant messaging as in face-to-face communication. These kinds of studies have successfully applied small, easily countable portions of Brown and Levinson's model to computer-mediated communication, but have not delved into a more comprehensive application of the model, as the current project proposes.

A few studies have applied Brown and Levinson's model in its entirety to very small datasets from specific domains. Carlo and Yoo [5] compared transcripts of 14 face-to-face transactions between reference librarians and students to 15 reference sessions via online chat. They found significantly more negative and fewer positive politeness strategies online than in face-to-face transactions. Simmons coded ten weeks of messages from an online bulletin board on censorship and described face-threatening actions, most of which were threats to negative face [13]. He suggested that over time people show more positive face-saving strategies online, as people adjust to this "faceless" medium, though this is contrary to Carlo and Yoo's findings. Duthler compared the politeness strategies used in email to voicemail when students had to make low- and high-imposition requests of a fictitious professor, and found differences in strategies within email between low- and high-imposition requests, but no differences in voicemail messages, suggesting that email is more tailorable, though extraneous phrases were correlated with decreased perceived politeness [6]. In general, both camps of politeness research in CMC are descriptive, but few connect politeness strategies with desired outcomes, such as getting a reply, or increasing member retention. One goal of the current project is to estimate the impact of specific politeness strategies on these desired outcomes.

METHOD

To build a model of linguistic politeness, we harvested a set of 576 messages posted to 12 discussion groups from 2004 to 2006. The groups cover a wide variety of topics, including diabetes, depression, multiple sclerosis, atheism, economics, life extension, C programming, math, electronics design, piloting, quilting, and general discussion by people over fifty. Each message was the first in its thread, and thus likely an attempt to start conversation, rather than a reply to an ongoing conversation. The number of replies to each message was counted, and usernames and signatures were replaced with same-gender pseudonyms.

For each of these messages we need two pieces of information: how polite readers perceive it to be, and which linguistic politeness strategies it contains. The present paper focuses on the first metric: perceived politeness. To measure perceived politeness, we surveyed 194 readers,

[1] Additional studies of politeness in CMC have focused on gender, culture, or domains such as computer tutoring, but these are outside the scope of this project.

described below. We describe the second piece, the development of a coding manual extending Brown and Levinson's politeness strategies to computer-mediated communication, in the Discussion section.

Measuring Perceived Politeness

To measure perceived politeness, we recruited 225 participants for a thirty-minute web-based survey with a random raffle for one of five $80 gift certificates. The survey was advertised in online classified ads across the U.S., and participants were required to answer 4 of 5 randomly selected grammar questions from the Test of English as a Foreign Language (TOEFL) correctly to proceed. Each participant read 48 randomly ordered messages counterbalanced from the 12 discussion groups and rated each message on a seven-point scale from very rude to very polite. Thirty-one participants were excluded for finishing the survey in less than 15 minutes, including two who selected "4" on the politeness scale for nearly every message, leaving 194 participants. Overall, an average of 14 politeness ratings per message were gathered, with good inter-rater reliability (Cronbach's alpha = 0.93 and mean correlation between any two judges of 0.41). To control for individual biases, each individual's score for a message was standardized by subtracting her mean score across all messages and divided by her standard deviation. Each message's perceived politeness score was then the mean of the standardized scores from each participant.

RESULTS

We found significant differences both in the perceived politeness in the 12 newsgroups and the impact of politeness on community responsiveness. Table 1 shows the results of a linear regression on politeness controlling for the message length (Mean=654.5 characters, SD=603.3). Because length was non-normally distributed, we included it after computing the natural log transformation of length plus 1. We use deviation coding for the 12 groups so that each coefficient represents the difference between the average perceived politeness in that group compared to the grand mean across all groups. We find that politics groups are generally perceived as significantly ruder than other groups, technical groups are more polite, and health support and hobby groups are mixed.

To determine the impact of politeness on community responsiveness, we performed a negative binomial regression on the number of replies the messages in those groups received, controlling for the standardized message length and clustering within groups to control for intragroup correlation. Negative binomial regression is appropriate for non-negative count variables with overdispersion, as is commonly the case in online discussion groups. Table 2 reports the expected number of replies for two kinds of average-length messages in each group: a message that has a standardized politeness score that is positive (and thus more polite than the average across all groups), and one that is negative (and thus ruder). Table 2 shows that there is an interaction between politeness and group topic: In some, such as the life extension and math discussion groups, rude messages receive almost no replies, but polite messages receive an average of 0.34 and 0.52 replies, respectively. Previous research has shown that receiving even a single response has dramatic outcomes on an individual's future behavior, including increased likelihood of posting again [4] and a greater participation answering others' questions in the future [11]. Rudeness, on the other hand, nearly quadruples reply counts in the atheism group and more than doubles them in the economics group. We see mixed results in the health and hobby groups, with politeness helping in diabetes and quilting groups, but hurting in depression and aviation groups.

DISCUSSION AND FUTURE WORK

The overall goal of this work is to determine how politeness affects the experiences people have in online communities. To understand that, we are building a model of linguistic politeness driven both by theory and bottom-up perceptions of politeness. Using a survey to obtain politeness scores for a set of messages on diverse topics, we found large differences in perceived politeness between communities and in the effectiveness of politeness on responsiveness.

Why was rudeness more effective in eliciting response in some groups, and politeness in others? Informal analysis of messages to the atheism group indicate that rude messages often stimulate long threads in which repliers form coalitions on either side of an argument, such as whether the United States should mix church and state issues. Or,

		Perceived politeness	S.E.
	Message length in chars (ln)	0.06 *	.03
Group type	**Group topic**		
politics	Atheism	-0.83 ***	0.03
	Economics	-0.91 ***	0.08
	Life extension	-0.15 +	0.08
health	Diabetes	0.16 *	0.08
	Depression	-0.17 *	0.08
	Multiple Sclerosis	0.37 ***	0.08
technical	C programming	0.22 **	0.08
	Math	0.56 ***	0.08
	Electronics design	0.33 ***	0.08
hobby	Aviation	-0.05	0.08
	Quilting	0.49 ***	0.08
	Over-50 chat	-0.02	0.08

N=576 messages *** p < .001 ** p < .01 * p <.05 + p < 0.1

Table 1. Linear regression on perceived politeness. Group topics use deviation coding, so coefficients represent the number of standard deviations (SDs) from the grand mean across all groups. Thus, messages to the math group were 0.56 SDs more polite than average messages across all groups.

		Expected # replies		
		Rude	Polite	S.E.
Group type	**Group topic**			
Politics	Atheism	2.90	0.75***	0.06
	Economics	1.32	0.54***	0.03
	Life extension	0.01	0.34***	0.03
Health	Depression	1.33	1.12**	0.01
	Diabetes	2.07	2.15*	0.06
	Multiple Sclerosis	1.20	1.25	0.03
Technical	C programming	1.08	1.25**	0.01
	Electronic design	1.80	1.31*	0.00
	Math	0.04	0.52***	0.02
Hobby	Aviation	2.24	1.69**	0.01
	Quilting	1.20	1.49*	0.03
	Over-50 chat	1.98	1.97	0.01

N=576 messages *** p < .001 ** p < .01 * p <.05

Table 2. Negative binomial regression on the number of replies controlling for message length (not shown). For clarity, we report the expected number of replies to a "rude" and a "polite" message of mean length in each group.

single trolls antagonize large numbers of responders by posting content like Christian children's song lyrics. These results are consistent with earlier findings that many political discussion groups act like virtual debating societies, where participants form strong antagonistic relationships by arguing with each other [16]. However, in the math group, posters are typically seeking assistance with statistics packages or formulas, rather than attempting to incite arguments, and thus post requests for help, often including linguistic politeness strategies such as giving deference or offering thanks in advance. The differing effect of politeness across groups is consistent with Harper and colleagues' findings in question and answer sites that "thank yous" were received differently depending on local culture [9]. One of the major limitations of the current study is the use of simple reply counts as a proxy for community responsiveness; future work will examine the emotional and informational content.

We are extending the work reported here in two ways: First, we are training single bag-of-word machine learning models to determine which words and phrases best predict the average perceived politeness scores. Second, we are hand-coding the messages for the presence of Brown and Levinson's 25 linguistic politeness strategies to determine which strategies are perceived as most or least polite, and to compare the performance of a theory-based machine learning model to one built from a simple bag of words.

To identify specific politeness strategies, such as hedging or seeking agreement, we've created a coding manual that extends Brown and Levinson's strategies to online discussion groups, adding keywords and examples from modern groups. Each code will be applied independently, allowing for overlapping codes within a message (e.g. joking, apologizing, and minimizing imposition). Examples for two strategies are included below (bold added to highlight strategy):

N6. Apologize

"So, if you could, please send whatever healing energies you can. . . . This is the scariest thing I've ever been through . . . Thanks in advance and ***apologies for the imposition.****"*

*"****sorry, normally wouldn't ask but*** *a few days ago, I made the announcement that PIF had broken the 3000 pipe sent 'barrier', I thought this was kinda cool, but not one response"*

P2. Exaggerate interest, approval, or sympathy

"I'm still glad you had an ***AWESOME*** *time. That concert was so RARE and you guys got something a lot of fans will never get.* ***CONGRATS!!****"*

*"****Congrats on the Mac****. . . It will help with your quilting. Really. I promise****! ;-P****"*

Previous politeness research has relied upon human codes of small sets of data; this project includes a machine learner that can be applied to much larger corpora, for greater generalizability and the design of automatic interventions, such as a "politeness checker" that suggests linguistic strategies to newcomers before they post their first messages. The machine learner can also be applied to other kinds of messages, such as replies, to determine if people who receive polite replies in their early group interactions go on to contribute more to the group in the future (such as replying to others). The learner can be applied to other domains, such as Wikipedia or SourceForge, to determine if politeness in production communities leads to greater or higher-quality products. Automatically detecting linguistic politeness in online communities will increase our understanding of how strangers make successful requests and become integrated into communities through conversation.

ACKNOWLEDGEMENTS

We would like to thank Sue Fussell and Cristen Torrey for politeness theory guidance and Alla Zats, Neel Shah, and Fahd Arshad for research assistance. This work is supported by NSF grants IIS-0325049 and IIS-0729286 and an NSF Graduate Research Fellowship.

REFERENCES

1. Brennan, S., & Ohaeri, J. (1999). Why do electronic conversations seem less polite? The costs and benefits of hedging. *Proceedings of International Joint Conference on Work Activities, Coordination, and Collaboration*, p. 227-235.
2. Brown, P., & Levinson, S. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
3. Bunz, U., & Campbell, S. (2004). Politeness accommodation in electronic mail. *Communication Research Reports*, 21(1), 11-25.
4. Burke, M., Joyce, E., Kim, T., Anand, V., and Kraut, R. (2007). Introductions and requests: Rhetorical strategies that elicit response in online communities. *Proc. Communities & Technologies*.
5. Carlo, J., & Yoo, Y. (in press). "How may I help you?" Politeness in computer-mediated and face-to-face reference transactions. *Information and Organization*.
6. Duthler, K. W. (2006). The politeness of requests made via email and voicemail: Support for the hyperpersonal model. *Journal of Computer-Mediated Communication*, 11, Article 6. http://jcmc.indiana.edu/vol11/issue2/duthler.html
7. Goffman, E. (1959) *The Presentation of Self in Everyday Life.* Garden City, NY: Doubleday Anchor.
8. Grice, H. (1957). Meaning. Philosophical Review, 67, 377-388.
9. Harper, F.M., Raban, D., Rafaeli, S., & Konstan, J. (2008). Predictors of answer quality in online Q&A sites. *Proc. CHI 2008*.
10. Holtgraves, T. (2002). *Language as Social Action: Social Psychology and Language Use*. Mahwah, NJ: Erlbaum.
11. Joyce, E. and Kraut, R. Predicting continued participation in newsgroups. *Journal of Computer Mediated Communication* 11,3(2006).
12. Kruger, J., Eppley, N., Parker, J., & Ng, Z-W. (2005). Egocentrism over email: Can we communicate as well as we think? *Journal of Personality and Social Psychology*, 89, 925-936.
13. Simmons, T. (1994). Politeness Theory in Computer-Mediated Communication: Face-Threatening Acts in a "Faceless" Medium. Unpublished masters thesis, Asthon University, Birmingham, England. Retrieved April 23, 2007 from http://eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED381005
14. Spertus, E. (1997). Smokey: Automatic Recognition of Hostile Messages. *Proc. AAI 1997*.
15. Watts, R. (2003). Modeling linguistic politeness (II): Brown and Levinson and their critics. In *Politeness*, (pp. 85-116). Cambridge: Cambridge University Press.
16. Welser, H.T., Gleave, E., Fisher, D. and Smith, M. Visualizing the Signatures of Social Roles in Online Discussion Groups *Journal of Social Structure*, 2007, np.