

Perceptions of author based on text style in computer-mediated communication

1 Introduction

Throughout the history of language attitudes research, researchers have devised experimental techniques to investigate attitudes towards different forms of speech. The linguistic object under examination usually pertains to spoken variation: different dialects, different degrees of nonstandardness of accent or fluency, different acoustic characteristics such as rate or pitch, and so forth. Measures are taken of subjects' social judgments or evaluations of speakers with these different linguistic characteristics: what are the social attributes given to an accent or speech style, and what stereotypes or sociocultural facts contribute to those attributions?

Although a broad array of languages, dialects, and acoustic properties have been investigated, there are components of linguistic practice underrepresented in the language attitudes literature. Namely, very little work has taken *written* language as a site of variation and an area in which to explore perceptions of and attitudes toward difference. In recent years, the proliferation of widespread text-based computer-mediated communication (CMC) has piqued the interest of scholars in the fields of communication, social psychology, and sociolinguistics. However, the issue of how different *textual* features serve as social cues in online interpersonal communication has gone largely unexplored.

The present project will explore perceptions of speakers based on written variables, by replicating CMC messages and asking subjects about the person who has produced them. With both quantitative and qualitative response components, the project is designed to enable an understanding of these attitudes as well as an indicator of possible topics for future investigation.

Because of the paucity of prior experimental research on this topic, the study is necessarily guided by a set of broad research questions, rather than specific hypotheses.

2 Background

Although there has been much research on CMC, experimental work is lacking on its sociolinguistic aspects. The majority of CMC work has tried to understand how CMC compares to face-to-face communication, or how it compares to either spoken or written language. Both of these streams of research are relevant to my goals, although the present project's aim differ from prior research in crucial ways.

Centered in communication-related fields, theorists have compared CMC to face-to-face communication, often arguing that CMC represents a medium bare of the nonverbal cues that provide social and contextual information in face-to-face interaction (for a good review and critique of this literature, see Walther 1993). Sometimes it was argued that CMC's lack of cues would lead to inferior or socially unsatisfying communication, or that its text-based nature allowed CMC to act as a more anonymous and socially neutral form of communication. Interestingly, CMC was said to lack precisely the features studied by language attitudes researchers (e.g., tone of voice, speech rate, prosody). However, it has since been recognized that text itself does have nonverbal cues as well, but that they are of course different from those found in speech (Lea and Spears 1992). Communicators will interpret some paralinguistic features as social cues, regardless of what modality they are communicating in (see Sierpe 2005). This project seeks to understand how such cuing works in a text-based environment, and what components of text are likely to serve as cues to social identities.

Centered in more language-oriented fields, other investigations into CMC have focused on the relationship between speech and writing and how the two seemed to blur into one another

online, often evinced by nonstandard features that are presumably the result of a more casual or conversational approach to online writing. A number of features have been identified as being either present in or even stereotypical of online text: lack of standard punctuation and capitalization; in-text iconic graphical symbols, such as emoticons; shortened forms of words such as acronyms or abbreviations; phonetic and stylized spellings (see Collot and Belmore 1996; Crystal 2001). As linguists explored the use of these features, they commonly characterized the features as internet-specific, although many of them had been in use for some time in other media (in particular, "eye dialect" and related practices, where nonstandard spelling is used to represent spoken dialect; Balhorn 1998). This was reflected in public discussion about online discourse as well as academic research (see Thurlow 2006).

More recently, sociolinguists have become interested in the question of linguistic variation internal to CMC, rather than seeking to identify features that make CMC as a whole the site of a distinctive dialect or register. This overlaps with the notion that language in CMC does carry nonverbal cues, albeit in a different form from spoken cues. In terms of stylized spellings, Paolillo (2001) found that social network was predictive of linguistic variation in an IRC chat channel. In terms of punctuation, in a study of college students' instant messaging, Squires (2007) found that females were significantly more likely to use apostrophes than were males. And, using a community of practice model to analyze variation in an IRC chat channel, Raclaw (2006) found that core community members were more likely to use an innovative style of ellipses than were peripheral community members. At the same time, public discourse often presents the same types of nonstandard uses as age- and gender-specific, focusing most intensely on teenage girls as innovators in online talk (see Thurlow 2006).

Aside from descriptive studies, many experimental studies have examined language, personality, and/or gender in CMC. However, none has looked at the use of text-based variables (henceforth "text style") as social cues in the systematic way that the present study aims to. Gill, Oberlander, and Austin (2006) had raters rate the personality of authors of email messages, where authors had taken personality tests prior to composing the messages in an experimental setting. Judges were good at rating authors' level of Extraversion, and not as good at detecting Neuroticism or Psychoticism. Markey and Wells (2002) looked at personality perception in online chats, finding high agreement among judges. However, neither of these studies isolated particular variables to examine their functions as either verbal or nonverbal cues.

More specifically in terms of gender, Savicki, Kelley, and Oesterreich (1999) found accurate judgment of whether email authors were male and female. Herring and Martinson (2004) also looked at gender and language use in an online gaming environment, but they did not examine text style; rather, they were looking at discourse structure. Colley and Todd (2002) used a number of "style" categories to categorize experimentally elicited emails, and among those found to be significant were multiple exclamation marks (e.g. !!!!), but no other gender differences were found. Somewhat more "typical" forms of punctuation were not included in their analysis.

The findings described above indicate that there is rich work to be done surrounding attitudes toward and stereotypes about language use in CMC. We can begin looking at these by adapting extant methodologies to the written form. As already noted, while some experimental work has been performed on CMC, it has for the most part not focused on graphical aspects of text as possible social cues or components of written style. And, while some language attitudes experiments have used written materials, they have predominantly been as merely a means to

studying attitudes toward speech, with written transcripts taken to represent spoken interaction. For instance, Leets (2001) presented subjects with written descriptions of interactions involving racist speech, where subjects were supposed to interpret the scene they had read as if it had been *spoken* between two people. Nigro *et al.* (1989) had subjects pretending to be jurors read trial transcripts of children's and adult's eyewitness testimonies, where the text was altered to represent more "powerful" or "powerless" speech, for instance through hedges or ellipses indicating hesitant pauses.

While such written stimuli may utilize some of the same written means of representing different styles, it's important to recognize that their ultimate goal is measuring attitudes toward differences in *spoken* language. Using written materials for the sake of understanding attitudes towards writing is a largely unexplored task. Examining how people react to standardness of text as expressed in text variables, such as punctuation, represents an important and lacking component of CMC perception research. I view such text variables - ultimately, marks of orthography - as being akin to nonverbal characteristics such as speech rate or to phonological variables such as vowel alternations, and as such as a rich sociolinguistic resource.

A few experimental researchers have taken written text features as perceptually salient, though they have also related such variables to either speech or verbal information. Preston (1985) used written transcripts to investigate subjects' evaluations of eye dialect, juxtaposing standard "speakers" with nonstandard "speakers," with written representations of their nonstandard or standardness achieved through spelling. He indeed found results that indicated that eye dialect effectively rendered speakers as substantively socially different; however, note that although eye dialect is a written modality, it is still directly related to spoken phonological patterns. That is, Preston was testing the efficacy of written forms to represent spoken variation.

Also exploring the relation of text variable to spoken language, Iorio (2006) looked at emoticons but focused on their prosodic content as performed by a speaking author. By contrast, Walther and D'Addario (2001) studied emoticons and their impact on recipients' interpretation of CMC messages; although looking at a written variable as it directly affected the message, this study looked at a particular CMC feature and its meaning as a *semantic* cue, rather than as a *social* cue. In other words, the meaning and interpretation of the emoticon were being studied, not the social properties of emoticon users.

On the other hand, Lea and Spears (1992) examined perceptions of speakers based on sample internet bulletin board texts by introducing orthographic variation into the stimuli. They found that both novice CMC users and experienced CMC users rated speakers more negatively when their messages included misspellings, typos, and extra punctuation (exclamation marks or ellipses). Though they were examining only personality characteristics, and not demographic ones, their results are revealing as to the potential of graphical cues to signal social information, taking "the form of typographical marks and other features of the text that, although they have no lexical meaning, nevertheless signify socially shared meanings" (324). Unfortunately, their viewpoint has not been much extended in the past 15 years.

All of these studies, both descriptive and experimental, suggest that just as people have attitudes toward spoken language, or attitudes toward social groups that are triggered by spoken language, people can form attitudes toward written styles that are ideologically and socio-psychologically motivated. Methodologically, CMC's rising popularity provides a modality that is relatively easy to manipulate and relatively realistic for subjects to evaluate, and thus is a good place to begin to understand attitudes toward textual variation.

3 Research Goals

The goal of this study is to explore how the use of specific text variables may contribute to different perceptions of a speaker in English-language CMC. Text variables are grouped into three categories: punctuation, capitalization, and letter repetition. Each of these somehow relates to standard usage in English writing, and their nonstandard usage is often cited as a prototypical component of online communication. The many dependent variables chosen for study are based on previous findings about language use in CMC and public discourse about users of CMC, as well as a general curiosity about social perceptions of internet use and linguistic practice—about which we currently have little empirical evidence.

The discussion in section 2 reveals that some research has found gender differences in terms of written features (e.g. Squires 2007; see also Baron 2004, Price & Graves 1980). Additionally, gender is often found to be a social parameter along which language use varies, and also along which social impressions given off by language use vary (see Labov 2001 for an overview of gender in variationist findings). Although becoming much more widespread, internet use is still stratified by age and younger people are more likely to use many interactive communicative applications online (see Fox & Madden 2005). In addition, younger users, and particularly females, are often publicly represented as heavy adopters of nonstandard written usage in online communication, and denigrated as such as linguistic destroyers (Thurlow 2006; see also Squires 2006). Thus, it seems as though at least age and gender are potentially salient characteristics that are associated with text cues. Based on what we know from the history of language attitudes research, personality measures are often inferable—or stereotypically inferred—from linguistic or paralinguistic cues as well. I also want to explore the possibility that

text cues can contribute to perceptions of other personal characteristics such as race, internet use, and education level.

Thus, this study is designed to address the following basic research questions:

- How do subjects perceive speakers' age, gender, and other characteristics according to use of standard or nonstandard orthographic features?
- Are some orthographic features more salient than others?
- Does speaker presumed age or gender make a difference as to subjects' perceptions of textual variation?

4 Method

4.1 Procedure

The experiment will consist of stimuli sets of sample instant messaging and email conversations, each set of which contains a combination of text variables. After reading through the set of messages, subjects will answer a brief questionnaire asking for responses about the messages' author, as well as the subject's demographic characteristics. The design is between-subjects, so each subject will be exposed to only one stimuli set and will answer questions about that set of messages only.

Subjects will run the experiment through the internet, using a personal computer or in some cases a researcher-provided laptop computer (see section 5). This method will enable the experiment to be administered outside of the confines of a laboratory setting and independent of the presence of the researcher, thus opening up participation to a wider variety of subjects and populations. Additionally, because the experiment is designed to elicit reactions to simulated *internet* messages, it is hoped that administration online will lend a greater level of realism to the setting, as well as ensure that subjects have some awareness of the forms of communication being presented. After giving consent, subjects will read an introduction to the study, at which point brief explanations of *email* and *instant messaging* will also be given (because of

administration via the internet, it is presumed that most if not all subjects will already be familiar with instant messaging and email; however, this explanation will be done as a precaution).

Once the experiment begins, subjects will first see an instructions screen, where the context condition will be manipulated (see Appendix A). The instructions also provide the frame for the experiment, which is intentionally minimal. Subjects will then read the messages and complete a series of questions pertaining to their impression of the author. There will be eight messages per stimuli set: four messages each of email and instant messaging (IM). The text of the messages will be the same across stimuli sets; sets will differ only in the use of the independent variables and the context condition (see Appendix B). Messages will occur on the screen in screen-shot image format, and subjects will read each message for as long as they wish before moving on to the next. Message order will be randomized.

After reading through the messages, subjects will be given a multi-component questionnaire to complete (see Appendix C). First, subjects will enter their perceptions of such characteristics as speaker's age, gender, race, internet use, education level, and personality. Second, subjects will be offered open-ended text boxes to fill in about any further perceptions of the speaker. Finally, subjects will complete demographic information about themselves. The entire experiment should take only 10-15 minutes per subject.

4.2 Stimuli

We had two options to choose from in terms of stimuli creation: either retrieve naturalistic data or create artificial data. The primary barrier to using naturalistic data is how to acquire the data, since email and instant messaging are private forms of interaction (they are not publicly posted online as, for instance, message board postings are). Additionally, while there

are some publicly available corpora for email (such as the Enron Corpus), and while I have obtained instant messaging corpora for other projects, the data from these sources are strongly identifiable as coming from specific social settings (for instance, my IM corpora clearly come from college students; the Enron corpus clearly comes from a business environment).

Although it is probably impossible to get message content that evokes no social reaction whatsoever, because of the compounding problem of data acquisition, the present experiment will use constructed messages, attempting to keep the content vague and neutral so as to allow attention focused on the text variables. The text of the messages will be designed to carry one or two tokens of each text variable per message, discussed further below. Messages will be altered by only the independent variables, creating stimuli sets with different text styles (approximating a written version of the matched guise technique; see Lambert *et al.* 1960).

4.3 Independent Variables

The study will comprise a 2 (punctuation) x 2 (capitalization) x 2 (letter repetition) x 4 (author) design, with four categories of independent variables. The first three variables below will be manipulated in the messages themselves, while the last variable (author) will be manipulated in the instructions to subjects.

PUNCTUATION

Condition 1: Standard punctuation

Condition 2: Nonstandard punctuation

CAPITALIZATION

Condition 1: Capitalization

Condition 2: No capitalization

LETTER REPETITION

Condition 1: Repetition

Condition 2: No repetition

AUTHOR

Condition 1: No information

Condition 2: Author gender - Male

Condition 3: Author gender - Female

Condition 4: Author gender - Teenager

There will thus be 32 different stimuli sets, each with a distinct combination of the independent variables. For instance, the most extremely "standard" stimuli set with no context would consist of standard punctuation, capitalization, no letter repetition, and no author information. The most extremely "nonstandard" stimuli set with an age context condition would include nonstandard punctuation, no capitalization, letter repetition, and a description of the author as a teenager.

Examples of stimuli are in Appendix B.

4.4 Operationalization of Independent Variables

Because of the different formats of email and instant messaging, it is necessary to operationalize the text variables (punctuation, capitalization, letter repetition) somewhat differently across formats.

EMAIL:**PUNCTUATION**

Standard: commas used in opening and closing greetings, quotation marks around direct quotes, apostrophes used in contractions and possessive nouns.

Nonstandard: commas not used in opening and closing greetings, quotation marks omitted around direct quotes, apostrophes omitted in contractions and possessive nouns.

CAPITALIZATION

Capitalization: proper nouns capitalized, sentence-initial words capitalized, first-person pronoun "I" capitalized.

No capitalization: capitalization never used.

LETTER REPETITION

Repetition: "so" is written as "soooo," "hello" is written as "helloooo," "really" written as "reallllly."

No repetition: "so" is written as "so," "hello" is written as "hello," "really" written as "really."

INSTANT MESSAGING:

PUNCTUATION

Standard: apostrophes used in contractions and possessive nouns, sentence-final periods or question marks.

Nonstandard: apostrophes not used in contractions and possessive nouns, no sentence-final punctuation.

CAPITALIZATION

Capitalization: proper nouns capitalized, sentence-initial words capitalized, first-person pronoun "I" capitalized.

No capitalization: capitalization never used.

LETTER REPETITION

Repetition: "so" is written as "soooo," "hello" is written as "helloooo," "really" written as "reallllly."

No repetition: "so" is written as "so," "hello" is written as "hello," "really" written as "really."

Note that the inclusion of a letter repetition variable is extremely exploratory and has not been studied in-depth, though it has been noted as a strategy for emphasis in CMC (see, among others, Hård af Segerstad 2002). It is not known whether vowels or consonants are more likely to be repeated, whether repetition is more likely in some words than others, or whether repetition's pragmatic functions differ across word, word-type, or letter. Hence the choice of these words is based on my personal experience as a CMC user as well as my research experience with various CMC corpora.

The fourth independent variable, manipulated in the instructions to subjects rather than the message text, is the author condition, designed to manipulate subjects' perceptions of one characteristic of the author. This will consist of four possible conditions: one in which no information is given about the speaker, one in which the subject is told that the speaker is male, one in which the subject is told that the speaker is female, and one in which the subject is told that the speaker is a teenager. This will allow us to see whether subjects' knowledge of one

author trait causes a change in the perception of any other, and how variation might be perceived differently given different contextual information.

4.5 Dependent Variables

Subjects will complete a short questionnaire after reading the messages, including the following response categories: quantitative impressions, qualitative impressions, and demographic information. See Appendix C for the questionnaire.

The first part of the questionnaire will directly address the research goals of this study, by asking subjects for their impressions of the author. Questions are designed to address speakers' perception of author's gender, age, race, level of internet use, education level, and personality. The personality measures were chosen in an attempt to reflect some of the stereotypes of CMC prevalent in public discourse (such as "nerdy" or "flaky"), attitudes toward writing ("good at writing"), and some more traditional personality traits culled from language attitudes research (such as "likeable" and "intelligent"). To prevent subjects from changing prior answers based on questions asked of them later, questions will be split into different screens, and subjects will not be able to revert back to previous screens once completing them.

In addition to forced choice and Likert-style ratings, open-ended questions are designed to elicit from subjects more information about what caused their impressions of the author. It is hoped that this can help identify relevant properties of the messages, serving as an internal manipulation check both for the independent variables and the message content. Finally, subjects will enter information about themselves, including similar demographic and internet-use items as those they were asked to provide about the author.

5 Subjects

Subjects will be recruited predominantly online, and will comprise subjects from multiple age groups, from teenagers to the elderly. Because the experiment will be designed so as to be easy to take from any internet connection, there will not be a controlled number of subjects recruited; rather, the link for the experiment will be targeted to a number of online email listservs, message boards, personal email distribution, and online communities. It is hoped that this method will yield subjects from a wide variety of ages and other social characteristics, enabling analysis about as broad a population as possible. Despite having no numerical cut-off for number of subjects, the target number of subjects will be 20 per stimuli set, or 640 subjects in all; because of the between-subjects design, a large number of subjects is preferable for raising the level of statistical power.

Recruiting subjects of especially younger and older age ranges (e.g., teenagers and the elderly) may require making personal contact with those groups, perhaps via visiting high school or teen settings and visiting nursing homes or senior centers in person. For these groups it will also be possible to have subjects perform the experiment on a portable laptop computer provided by the researcher, as an alternative to performing the experiment online. For the most part, however, it is expected that subjects will be able to be identified online and perform the experiment online as well. Generalizations will thus be limited to internet users, which seems appropriate for an exploratory experiment of this kind, though it may be nice for future researchers to examine more closely the impact of internet experience on perceptions of text style (though note that Lea and Spears [1992] did not find a difference in judgments between novice and experienced CMC users).

6 Analysis

The goal of this project is to measure whether perceptions of message author will vary depending on the independent variables, including message conditions (text style) and context conditions (author information). The main quantitative result, then, will be any relationships between the dependent and independent variables; that is, between the responses of subjects and the stimuli set characteristics. Multivariate analysis of variance should illuminate such correlations, and several different sets of relationships will need to be investigated.

First, the clearest relationship should be between each stimuli set and subjects' ratings. Second, and more complex, will be the relationship between subjects' ratings and separate message conditions, regardless of which stimuli set is used. This should yield the most important results, showing which variables are related to different responses (e.g., all sets with nonstandard punctuation, or all sets with letter repetition). Third, the relationship between context conditions and subjects' responses will be analyzed to see if knowing either the presumed age or gender of the author makes a difference in perception on other variables. Because of the large number of independent variables, it is expected that there will be interactions between variables, which can also be identified through analysis of variance. Finally, analysis of variance between subjects' demographic characteristics and their rating patterns will be performed, to examine whether any social characteristics of subjects show a relationship to certain attitudes; subject demographics may also provide interaction effects.

In addition to analysis of variance, factor analysis will be utilized for two main purposes: to identify groupings of text variables that elicit perceptions of a certain demographic category (e.g. gender, age, race), and to interpret the personality ratings. The former will indicate whether certain realizations of variables present together might better predict responses than any of the

variables alone. The latter might prove useful for understanding what profile of the author subjects form, though because there are not very many (15) personality items, this will be a relatively small component of the analysis.

The experiment will also elicit qualitative responses in the open-ended question portion of the questionnaire, and these responses will be content analyzed, coding for mentions of the different independent variables and the attitudes toward those variables (similar to Giles *et al.* 1990). Coding will be done so as to enable quantitative treatment of the qualitative data, producing the additional possibility of including some qualitative responses in their relation to the independent variables and the quantitative dependent variables. Such an analysis should yield information about what text features the subjects consciously noticed the most or least, whether the manipulation of text style was effective, whether content had an impact on subjects' perceptions, and potential directions for future research.

7 Conclusion

Although much research has investigated language attitudes and social perception, few studies have considered written forms of linguistic practice as worthy of such attention in their own right. I believe that the present project will provide an adequate starting point for understanding what parts of text can serve as social cues, and especially how standard and nonstandard written features may be perceived as indicators of demographic or personality traits. In doing so, it will contribute to the broader goal of understanding ideologies and stereotypes about CMC in particular, and about written English in general.

8 References

- Balhorn, M. (1998). Paper representation of the non-standard voice. *Visible Language* 32: 1, 56-74
- Baron, N.S. (2004). See you online: Gender issues in college student use of instant messaging. *Journal of Language and Social Psychology*, 23(4): 397-423.
- Colley, A. and Todd, Z. (2002). Gender-linked differences in the style and content of e-mails to friends. *Journal of Language and Social Psychology*, 21, 380-392.
- Collot, M. and Belmore, N. (1996). Electronic language: A new variety of English. In S.C. Herring (ed.). *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives* (pp. 13-28). Amsterdam: John Benjamins.
- Crystal, D. (2001). *Language and the Internet*. Cambridge, UK; New York, Cambridge University Press.
- Fox, S. and Madden, M. (2005). Generations Online. Pew Internet and American Life Project. < http://www.pewinternet.org/pdfs/PIP_Generations_Memo.pdf>
- Giles, H., Coupland, N., Henwood, K., Harriman, J., and Coupland, J. (1990). The social meaning of RP: An intergenerational perspective. In *Studies in the Pronunciation of English: A Commemorative Volume in Honor of A. C. Gimson*, ed. by S. Ramgaran, chapter 17, 191-211. New York: Routledge.
- Gill, A.J., Oberlander, J., and Austin, E. (2006). Rating e-mail personality at zero acquaintance. *Personality and Individual Differences* 40(3): 497-507.
- Hård af Segerstad, Y. (2002). Use and adaptation of written language to the conditions of computer-mediated communication. (Doctoral dissertation. Göteborg University, 2002). Retrieved from http://www.ling.gu.se/%7eylvah/dokument/ylva_diss.pdf.
- Herring, S.C. and Martinson, A. (2004). Assessing gender authenticity in computer-mediated language use: Evidence from an identity game. *Journal of Language and Social Psychology*, 23 (4): 424-446.
- Iorio, J. (2006). The Interaction of Pitch Contour, Pause Structure, and Emoticons in Performed Online Narratives. University of Texas at Austin (ms).
- Lambert, W.E., Hodgson, R.C., Gardner, R.C., and Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60(1): 44-51.
- Lea, M. and Spears, R. (1992). Paralanguage and Social Perception in Computer-mediated Communication. *Journal of Organizational Computing* 2: 321-341.

Leets, L. (2001). Explaining perceptions of racist speech. *Communication Research* 28: 676–706.

Markey, P. and Wells, S. (2002). Interpersonal perception in internet chat rooms. *Journal of Research in Personality*, 36, 134–146.

Nigro, G.N., Buckley, M.A., Hill, D.E., and Nelson, J. (1989). When juries “hear” children testify: The effects of eyewitness age and speech style on jurors’ perceptions of testimony. In *Perspectives on Children’s Testimony*, ed. by S. J. Ceci, D. F. Ross, and M. P. Toglia, chapter 4, 57–70. New York: Springer-Verlag.

Paolillo, J. (2001). Language variation on Internet Relay Chat: A social network approach. *Journal of Sociolinguistics* 5 (2):180-213.

Preston, D. (1985). The Li'l Abner Syndrome: Written Representations of Speech. *American Speech* 60(4): 328-336.

Price, G.B. and Graves, R.L. (1980). Sex differences in syntax and usage in oral and written language. *Research in the Teaching of English*, 14 (2), 147-153.

Raclaw, J. (2006). Punctuation as social action: The ellipsis as a discourse marker in computer-mediated communication. 32nd Annual meeting of the Berkeley Linguistics Society, UC Berkeley.

Savicki, V., Kelley, M., and Oesterreich, E. (1999). Judgments of gender in computer-mediated communication. *Computers in Human Behavior*, 15, 185–199.

Sierpe, E. (2005). Gender distinctiveness, communicative competence, and the problem of gender judgments in computer-mediated communication. *Computers in Human Behavior* 21: 127–145.

Squires, L. (2006). Publicity, folk linguistic artifacts, and noninstitutional institutions: Locating language ideologies in online metadiscourse. University of Michigan, Ann Arbor (ms).

Squires, L. (2007). What's the use of apostrophes? Gender difference and linguistic variation in instant messaging. *American University TESOL Working Papers* 4.

Thurlow, C. (2006). From statistical panic to moral panic: The metadiscursive construction and popular exaggeration of new media language in the print media. *Journal of Computer-Mediated Communication*, 11(3), article 1. <<http://jcmc.indiana.edu/vol11/issue3/thurlow.html>>

Walther, J.B. (1993). Impression Development in Computer-Mediated Interaction. *Western Journal of Communication* 57(4): 381-398.

Walther, J.B. & D’Addario, K.P. (2001). The Impacts of Emoticons on Message Interpretation in Computer-Mediated Communication. *Social Science Computer Review* 19, 324-347.