

Predicting gender from electronic discourse

Rob Thomson and Tamar Murachver*

University of Otago, New Zealand

There is substantial evidence of gender differences in face-to-face communication, and we suspect that similar differences are present in electronic communication. We designed three studies to examine gender-preferential language style in electronic discourse. In Expt 1, participants sent electronic messages to a designated 'netpal'. A discriminant analysis showed that it was possible to successfully classify the participants by gender with 91.4% accuracy. In Expts 2 and 3, we wanted to determine whether readers of e-mails could accurately identify author gender. We gave participants a selection of messages from Expt 1 and asked them to predict the author's gender. It was found that for 14 of the 16 messages used, the gender of author was correctly predicted. In the third experiment, six messages about gender-neutral topics were composed. Using a subset of the variables identified in Expt 1, female and male versions of each message were created. When participants were asked to rate whether a female or a male wrote these messages, their ratings differed as a function of the message version. These findings establish that people use gender-preferential language in informal electronic discourse. Furthermore, readers of these messages can use these gender-linked language differences to identify the author's gender.

Gender is a salient and pervasive dimension that people, old and young, use to classify others. Even without physical cues, speakers in face-to-face communication provide information about their gender classification by the way they talk (Mulac & Lundell, 1986). On the net, when a person posts a message to a newsgroup or contacts a colleague about a research query, how salient is his or her gender? We suspect that even without a gender-informative signature, people convey their gender in electronic discourse. They do so, in part, by generalizing established gender-preferential language patterns from spoken communication to the electronic medium.

Gender-preferential communication is evident in many contexts of face-to-face interaction (see Coates, 1986; Holmes, 1984, 1995; Lakoff, 1973, 1990; and Tannen, 1990, 1994, for reviews). Mulac and Lundell (1986), for example, asked 40 university students to describe landscape photographs to researchers. A discriminant analysis found that the frequencies of a set of 17 linguistic variables could be used to classify the participants' gender with 87.5% accuracy. In general, women are more likely than men to refer to emotions (Mulac, Studley, & Blau, 1990), use

* Requests for reprints should be addressed to Dr Tamar Murachver, Department of Psychology, University of Otago, Box 56, Dunedin, New Zealand (e-mail: tamar@psy.otago.ac.nz).

intensive adverbs (McMillan, Clifton, McGrath, & Gale, 1977; Mulac & Lundell, 1986; Mulac, Wiemann, Widenmann, & Gibson, 1988), make compliments (Holmes, 1988), and use minimal responses (Carli, 1990), modals and tag questions (McMillan *et al.*, 1977), personal pronouns, subordinating conjunctions and oppositions (Mulac & Lundell, 1986). Women are also more likely to ask questions (Tannen, 1994) and use more linguistic forms associated with politeness (Holmes, 1995). Men are more likely to make references to quantity, use non-standard forms or make grammatical errors, and provide more opinions and justifiers (Mulac & Lundell, 1986; Mulac *et al.*, 1990). Some research has also shown that men use more words, and talk more often in conversation (Dovidio, Brown, Heltman, Ellyson, & Keating, 1988; Swacker, 1975; Wood, 1987; but see James & Drakich, 1993).

Few would argue that all women use one set of language features and all men use a different set of features. Language features are gender-preferential, rather than exclusive to one sex or the other (Fitzpatrick, Mulac, & Dindia, 1995). For example, both women and men provide minimal responses while listening, but the overall frequency, timing and context of these might vary with gender (Fishman, 1977; Hannah & Murachver, 1999; Zimmerman & West, 1975). In addition, some members of a gender category might use a different set of gender-preferential features than other members of the same gender. Some men, for example, might use a relatively high frequency of teases and insults, whereas other men might speak in longer utterance lengths. Another way of saying this is that gender-preferential language is not definitional. Gendered linguistic behaviour seems to vary in much the same way as other gendered behaviours. This way of thinking about category membership is neither new nor exclusive to gender categories. The idea that members of a category do not share an exact set of necessary features has been argued in a variety of forms, from fuzzy set theory (Zadeh, 1982) and prototype theory (Rosch, 1975; Rosch & Mervis, 1975) to more recent connectionist theories of representation (Rumelhart, Smolensky, McClelland, & Hinton, 1986). Given our understanding of categories, and gender categories in particular, it is not surprising that there is wide variation in adherence to gender-preferential styles of talk within each gender group.

The context in which communication takes place markedly influences language style. For example, most people would adopt different communication styles when talking to friends at a bar than to an elderly grandmother. People modify their language style to what is appropriate for the situation. This is true for language on the informal-formal continuum, as well as for language on the female-male continuum. Gender-preferential language features are more common in same-sex dyads than in mixed-sex dyads (Bilous & Krauss, 1988; Fitzpatrick *et al.*, 1995; Mulac *et al.*, 1988). When conversing with a member of the other sex, both women and men change some aspects of their language style towards the gender-preferential style of their partner. This phenomenon is explained by communication accommodation theory (CAT; Coupland, Coupland, Giles, & Henwood, 1988). According to CAT, individuals in friendly exchanges who are motivated to accentuate similarities will converge towards the communication style of their partner. In contrast, it follows that if individuals are motivated to accentuate differences, then their language style will diverge from that of their partner.

There is a general consensus that women and men differ in the way they communicate in face-to-face interactions. Although contextual factors shape the nature of the difference, it is rare for the difference to disappear completely. Not surprisingly, the limited research examining gender differences in written language has found similar effects. For example, Mulac *et al.* (1990) examined differences in primary and secondary students' essays. Results indicated that by 9 years of age boys and girls show gender-linked language differences in their writing. Boys used greater frequencies of action verbs, contradictions and judgmental phrases, whereas girls used more hedges and references to emotion.

A relatively recent development in how people communicate is the rise in the popularity of electronic discourse or computer-mediated communication (CMC). It involves typing text on computers and sending the text to others via computer networks. Electronic discourse shows elements of both written language and speech (Foertsch, 1993; Lea, 1991). Although electronic discourse is written, it is sent and received so rapidly that it nears the interactivity of speech, and conversations can occur. Electronic messages are typically composed on the spot, and are less censored and more informal than writing intended for hard copy. Yates and Orlikowski (1993) document some of the similarities between spoken conversation and e-mail correspondence. For example, both make use of informal vocabulary and incomplete sentences. Users of electronic communication mimic the paralinguistic features of spoken conversation with written forms such as 'sigh', 'gasp', and 'yawn'. However, Yates and Orlikowski cite examples where electronic communication is similar to written correspondence. Both show evidence of editing, careful formulation and use of formatting devices such as headers and lists. Electronic messages also contain features unique to CMC. These include emoticons or 'smileys' (e.g. ;-), :-o) and trying to signify physical actions (e.g. jon hugs jane).

Given that gender differences in language emerge in oral and written communication, it is likely that similar differences are present in electronic communication. Recent reports suggest that there are different styles used by females and males in electronic communication (Herring, 1993, 1994; Savicki, Lingenfelter, & Kelley, 1996; Witmer & Katzman, 1997). Herring (1993, 1994), for example, studied on-line discussion groups and found that there were differences in both topics and features used by women and men. The most common topics for men were concerned with issues and information, whereas personal topics and queries were more common with women. Furthermore, Herring reported that men's postings to newsgroups tended to show self-promotion, sarcasm, insults and strong assertions, whereas women's postings reflected supportiveness and attenuation. The women were more likely to hedge, express doubt, apologise, ask questions and suggest ideas rather than assert them. Kaplan and Farrell (1994) studied a small group of female authors who regularly spend time communicating using electronic discourse. Analysis of their messages led Kaplan and Farrell to suggest that the females adopted a rapport style, as opposed to a report style, similar to that identified by Tannen (1990) in face-to-face interaction. Herring (1994) suggests that for publicly posted messages, there are gender differences in electronic message style as well as different assumptions about netiquette. Herring notes that men are likely to 'flame' and engage in harsh personal insults on the net; these are behaviours that they

might not engage in as readily face-to-face (but see Witmer & Katzman (1997) for contrary findings, and Lea, O'Shea, Fung, & Spears (1992) for an alternative view of flaming).

There is a growing body of evidence for gender differences in language style for CMC. We suspect that people simply take the gender-preferential styles used in other forms of communication and adapt them to electronic discourse. To what extent these styles are adapted is an area of debate. Some theorists have argued that CMC has fewer cues to identity and a lack of regulatory feedback, which may reduce normative influence (Kiesler, Siegel, & McGuire, 1984; Sproull & Kiesler, 1986). This in turn would reduce social barriers and differences in status, and have an 'equalization effect' on participants (Kiesler & Sproull, 1992; Sproull & Kiesler, 1986). Because of the equalization effects one might predict a decrease in the use of gender stereotypical language in CMC.

Others have argued, however, that CMC is not as equalizing for gender as might be hoped for (Herring, 1993, 1994; Kramarae & Taylor, 1993; Savicki *et al.*, 1996; Spears & Lea, 1994). Because of the veil of anonymity there is some suggestion that differences might actually be amplified in electronic discourse (Herring, 1994; Kramarae & Taylor, 1993; Spender, 1995). These authors report that men monopolize on-line discussion groups, both in number of participants and length of messages. Participation rates may, however, be more complex. Graddol and Swann (1989), for example, found equal contribution by men and women, and Jaffe, Lee, Huang, and Oshagan (1995) argue that women are more likely than men to use a pseudonym that masks gender. This would distort any observations of male and female participation.

A recent approach offers an alternative explanation for social psychological factors operating in CMC. Proponents of the Social Identity and Deindividuation (SIDE) model. (Postmes, Spears, & Lea, 1998; Spears & Lea, 1992, 1994; Spears, Lea, & Lee, 1990) claim that although CMC limits the number of interpersonal cues, it provides cues to social identity. People generally give cues to gender, ethnicity, location and occupation through usernames, server and context of communication. Because of the reduced number of personal identity cues, social identity becomes more salient in CMC compared with other communication media. Consequently, one would predict an increase in conformity to group normative behaviour such as gender-preferential language in electronic discourse.

The aims of the present experiments were to examine gender-preferential language styles used in informal electronic discourse, and to examine if author gender could be predicted from language style. The first experiment was designed to examine whether specific language features that are characteristic of males and females are found with electronic discourse. Participants were taught how to use an e-mail program, and were then instructed to send messages to a designated netpal. Participants were assigned to same-sex dyads to maximize the amount of gender-preferential language (Bilous & Krauss, 1988; Fitzpatrick *et al.*, 1995). At the conclusion of the experiment these messages were extracted and coded for 14 linguistic variables that have been identified in previous experiments as characteristic of male or female speakers or writers. It was anticipated that author's gender

could be discriminated on the basis of these language features. The hypotheses were thus as follows:

H₁: Females will be identified by a constellation of features that include higher frequencies of references to emotion, personal information, questions, compliments, self-derogatory statements, apologies, intensive adverbs, modals and subordinating conjunctions.

H₂: Males will be identified by a constellation of features that include higher frequencies of opinions, adjectives and insults. Males will also tend to produce longer sentence units.

EXPERIMENT 1

Method

Participants

Forty-two participants were recruited for this study, of which 35 completed the experiment. Participants were excluded if they did not return to write e-mail messages after the initial session. There were 19 females of mean age 18.8 years ($SD = 1.3$) and 16 males of mean age 20.8 ($SD = 4.8$). Participants for the present experiment were undergraduates enrolled in introductory psychology courses at a New Zealand university. Participants volunteered in order to fulfil part of their course requirements.¹

Procedure

Participants were informed that the experiment was designed to determine how people communicate using computers, and were given a network usercode and password. This allowed them to access the psychology VAX network using IBM compatible computers. They were taught how to log on to the network and start the mail program. Participants were also shown how to read, reply and send messages. They then opened up a message from the experimenter and sent a reply. When they were comfortable with using the computer, they were shown how to exit the mail program and log off the network. Participants were then asked to log on to the network in a few days, at which time there would be a message from the experimenter giving the user-name of their netpal. The experimenter randomly assigned the netpals with the constraint that all dyads were same-sex pairs. Participants were instructed to send at least six messages to their netpal over the next two weeks. All participants were informed that the experimenters would have access to their mail folders. At the conclusion of the experiment, participants' mail folders were accessed and all the messages that they had received were extracted.

Coding

The number of messages sent by each participant and the total word count were calculated. Messages were coded for 13 additional language features, and the frequencies of these features were counted for each participant. The coded variables included: (1) references to emotion (e.g. 'I was *annoyed*'); (2) requests for information (e.g. 'what kinds of music do you like?'); (3) personal information (e.g. 'I work at Farmers'); (4) opinions (e.g. 'I think the Registry protest was worthwhile'); (5) self-derogatory comments (e.g. 'I'm useless with computers'); (6) insults (e.g. 'you were stupid to take that paper');

¹Undergraduate students are encouraged to participate in research projects in the Psychology Department for course credit. There are many experiments offered and they can choose which ones to participate in.

Table 1. Mean frequencies (and standard errors) per 100 words of language variables elicited from female and male participants

	Female participants		Male participants	
	<i>M</i>	SE	<i>M</i>	SE
Emotion*	.89	.15	.35	.17
Questions	1.12	.20	.89	.22
Personal information*	2.52	.33	1.42	.36
Opinions	.54	.20	.93	.22
Self-derog. comments	.16	.06	.00	.06
Insults	.00	.02	.05	.02
Compliments	.03	.01	.00	.01
Apologies	.49	.02	.00	.02
Subord. conjunctions	1.70	.28	1.66	.30
Modals/hedges**	.63	.10	.20	.11
Intensive adverbs**	1.52	.20	.66	.22
Adjectives	2.90	.36	2.95	.40
Oppositions	.00	.00	.00	.00

* $p < .05$; ** $p < .01$.

(7) compliments to their netpal (e.g. 'your poem was good'); (8) apologies (e.g. 'sorry I haven't written'); (9) subordinating conjunctions (e.g. 'the play went on for ages *although* I had a good time'); (10) modals or hedges (e.g. 'it was *sort of* interesting'); (11) intensive adverbs (e.g. 'the game was *really* good'); (12) adjectives (e.g. 'this *stupid* computer'); and (13) oppositions (e.g. 'the lecture was boring, but OK'). Two raters independently coded 25% of the messages. Inter-rater reliabilities (alphas) for each language characteristic were all at or above .85.

Results and discussion

The mean number of messages, the total word count and the mean words per message were calculated for both male and female participants. Both male and female participants sent between four and five messages on average to their netpal (females: $M = 4.8$, $SD = 2.5$; males: $M = 5.6$, $SD = 2.8$). The total word count (words per message) was found to be slightly higher for male participants ($M = 95.2$ words, $SD = 42.2$) than for female participants ($M = 83.6$ words, $SD = 46.8$). A t test on mean words per message showed that this difference was not statistically significant. The mean unit lengths or mean sentence length were then calculated. This was done by dividing the total number of words produced by each participant by the total number of sentences. Although the male group had slightly longer mean unit lengths ($M = 12.60$ words, $SD = 4.6$) than did females ($M = 10.54$ words, $SD = 2.7$), a t test performed on MLU showed no effect of gender.

The frequencies of the remaining coded variables were converted into proportions by dividing the total frequency by the total word count and multiplying by 100 to yield a frequency per 100 words. These values are presented in Table 1. This shows that only small differences in the frequency of language variables were found

Table 2. Structure matrix of discriminant analyses to be predictive of author gender

Language feature	Correlation with discriminant function ^a	Language feature	Correlation with discriminant function ^a
Intensive adverbs	-.38	Self-derog. comments	-.22
Modals	-.36	Opinions	.18
Emotion	-.31	Compliments	-.12
Personal information	-.30	Questions	-.10
Insults	.24	Adjectives	-.01
Apologies	-.23	Subord. conjunctions	-.01

^aCanonical discriminant functions at group means: female = -1.49, male = 1.41.

between female and male participants. No oppositions were produced by any of the participants, so oppositions were omitted from further analyses. A multivariate analysis of variance (MANOVA) with gender as the between-participant factor was computed for the language variables.² This analysis indicated a main effect of gender across the 12 language variables ($F(12,22) = 3.25$, $p < .01$). Univariate analyses indicated that females made significantly more references to emotion ($F(1,33) = 5.56$, $p < .05$), provided more personal information ($F(1,33) = 4.96$, $p < .05$) and used more modals ($F(1,33) = 7.69$, $p < .01$) and more intensive adverbs ($F(1,33) = 8.58$, $p < .01$) than did male participants.

The data were then further analysed using a discriminant analysis. This procedure identifies weighted combinations of language features that distinguish group membership (Mulac *et al.*, 1990). Results from the analysis are presented in Table 2, and indicated that the 12 linguistic variables provided substantial differentiation between female and male authors (Wilk's $\lambda = .25$, $\chi^2 = 27.53$, $p < .01$). Based on a weighted combination of these 12 variables, 91.4% of the authors (female = 94.7%, male = 87.5%) could be correctly reclassified into their gender group. Only one female and two males were classified incorrectly using the weights derived from the discriminant analysis. Examples of messages written by female and male netpals are shown in Appendix 1.

When the coded language characteristics were examined independently, there were few significant differences between male and female participants. Those differences are consistent with previous research which has shown that females make more references to emotion (Mulac *et al.*, 1990), provide more personal information (Herring, 1993, 1994; Tannen, 1990), use more modals (McMillan *et al.*, 1977; Mulac *et al.*, 1990) and use more intensive adverbs (McMillan *et al.*, 1977; Mulac & Lundell, 1986; Mulac *et al.*, 1988). Differences in other language variable frequencies between male and female participants, although non-significant, were

²A MANOVA with gender as the between-participants factor was also computed on the raw frequencies for the 12 features. This analysis showed a main effect of gender across the 12 features ($F(12,22) = 2.55$, $p < .05$). Univariate analyses showed that female participants made significantly more self-derogatory references and marginally more apologies than did male participants. There was also a marginal effect of male participants using more insults than female participants.

also in the predicted direction. Females showed a higher mean frequency per 100 words of questions, self-derogatory comments, compliments, apologies and subjective conjunctions. Males were more likely to convey opinions, make insults and to write more.

The results from this experiment show that although there are only small gender differences in individual features, it is possible to classify accurately e-mail messages by author's gender using a combination of those features. This is consistent with several recent studies where although only small effects for individual features were observed, a strong effect of style was shown (Fitzpatrick *et al.*, 1995; Mulac, 1998; Mulac *et al.*, 1990).

EXPERIMENT 2

In the first experiment, although there were only small differences in the frequency of individual features, gender could be predicted by a discriminant analysis model based on the combination of these features. As such, it is the speech pattern that is predictive of gender rather than any individual feature. This raises the question of whether readers of e-mail messages are sensitive to these gender-preferential speech patterns. Many people mask or impersonate gender in CMC (Bruckman, 1993; Jaffe *et al.*, 1995; Truong, 1993; Turkle, 1990; Witner & Katzman, 1997). If gender can be determined by linguistic style alone, then gender impersonation may be less straightforward than simply using a gender-informative label.

Herring (1994) suggests that readers use their knowledge of gendered language styles to classify messages in CMC. Thus, readers should be able to judge accurately the author gender of electronic messages, especially those written within same-sex dyads. Yet evidence from Savicki, Kelley, and Oesterreich (1999) leads us to question whether author gender is readily inferred from linguistic style. Savicki *et al.* asked undergraduate students to judge the author of messages written in a high or low communication style, where the high communication style (HCS) was defined as more typical of women and the low communication style (LCS) was more typical of men. As expected, male authorship was more correctly identified in LCS messages (more typical of men) than in HCS messages (more typical of women). However, female authors of LCS messages (male typical) were more often correctly identified than were female authors of HCS messages (female typical). The reason for this pattern of results is unclear. It is possible that HCS messages were not particularly characteristic of women authors, and that other gendered linguistic features were present in the LCS messages written by women that made gender more identifiable. One way to clarify whether participants can predict author gender from electronic messages is to have participants judge messages already shown to be linguistically female-preferential or male-preferential. This was the basis for Expt 2.

Method

Participants

Participants were undergraduate students enrolled in a psychology course at a New Zealand university. There was a total of 78 participants in this experiment. None of the participants in Expt 2 served

as participants in Expt 1. Participants were recruited during a lecture, and participation was voluntary.

Materials

Sixteen e-mail messages that had been written by the participants in Expt 1 were used. These messages were selected randomly from all of the transcripts that satisfied inclusion criteria. Only messages that were correctly classified by the discriminant analysis and had more than 50 words were used. Furthermore, messages were not included if it was impossible to mask gender or identifying characteristics without altering the structure of the message. Eight messages written by males and eight written by females were selected and these were ordered randomly. The messages were presented in either forward or reverse order to counterbalance for order effects.

Procedure

Participants were recruited at the end of a lecture and were informed that participation in the experiment was voluntary. Each participant was given 16 brief e-mail messages to read and asked to rate the likelihood that the message was written by a female or a male using a 6-point scale where 1 = definitely written by a female and 6 = definitely written by a male. When they had finished, participants were thanked for their participation, and were informed that the results would be made available in a later lecture.

Results and discussion

The mean ratings for the messages written by females and the messages written by males were calculated for each rater. A mixed-factor nested ANOVA with gender of participant as the between-participants factor and message (1 through 8) nested within author gender as the repeated measures was calculated. This showed a significant effect of author gender ($F(1,76) = 71.89, p < .001$) and no effect of gender of rater. The female versions ($M = 3.04, SE = .08$) were rated as more likely to be written by a female than were the male versions ($M = 3.99, SE = .07$). For the individual messages, author gender was correctly predicted by more than 60% of raters for 14 of the 16 messages (percentage correct for these ranged from 62% to 95%). Two messages written by females were incorrectly classified as male, and one message written by a male was correctly classified by only 50% of raters.

These results indicate that people are sensitive to gender differences in language style in e-mail messages and can make accurate judgments about author's gender. It could be argued, however, that the participants based their judgments about gender on other factors (such as topic) in the message. As such, they may not only be using gender-preferential speech patterns to determine gender. To separate out gender-preferential speech style from other cues to gender, in Expt 3 message templates were composed and then rewritten to have either female-preferential language or male-preferential language features. Other than the gender-preferential features, the female and male versions of each message were identical. As in Expt 2, we predicted that participants would be able to discriminate between e-mail messages more typical of female authors and those more typical of male authors.

EXPERIMENT 3

Method

Participants

Participants were undergraduate students enrolled in a psychology course at a New Zealand university. None of the participants in Expt 3 served as participants in either of the two earlier experiments. There was a total of 70 participants, both male and female. Participants were recruited during a lecture, and participation was voluntary.

Materials

Six messages were composed based on the style used by participants in Expt 1. Each message was approximately 50 words long, and the topics were designed to be gender neutral. Using a subset of the variables from Expt 1, a male and a female version of each message was created. Only small changes were made to each message to create male and female versions. There were two–five minor changes to each version, and hence content and structure were preserved. Examples of the female and male versions of a message are presented in Appendix 2. In this example, the female version contains an extra reference to emotion ('annoyed'), an apology ('Sorry I couldn't make it last night') and an intensive adverb ('really'). The male version had an insult ('lazy sod') and the sentences are longer.

Participants rated one version of each of the six messages. Participants always rated three female versions and three male versions, and all passages were presented in the same order. However, the ordering of the gender (i.e. version) for each message was counterbalanced to produce four combinations.

Procedure

Participants were recruited at the end of a lecture and were informed that participation in the experiment was voluntary. Each participant was given six brief e-mail messages to read. Participants were asked to rate the likelihood that the message was written by a female or a male using a 6-point scale where 1 = definitely written by a female and 6 = definitely written by a male. When they had finished, participants were thanked for their participation, and were informed that the results would be made available in a later lecture.

Results and discussion

The mean ratings for the female message versions and the male message versions were calculated for each participant and submitted to a 2 (female or male version) \times 4 (ordering of passage versions) ANOVA. This revealed a significant effect of passage version on participant's ratings ($F(1,51) = 50.25, p < .001$). The female versions ($M = 2.79, SD = .64$) were rated as more likely to be written by a female than were the male versions ($M = 3.65, SD = .78$). No other significant effects were present. These results indicate that participants can correctly discriminate gender on the basis of linguistic-style in the absence of other cues to gender.

GENERAL DISCUSSION

The results from Expt 1 showed that gender-preferential language is expressed in electronic communication, and it is similar to that found in face-to-face

communication. The results from Expt 2 indicate that participants are able to discriminate the gender of the author based on a subset of gender-preferential language features. This is especially interesting because the changes made to create the male and female versions were minor. This suggests that people are quite sensitive to the gender cues in language. These results have implications for people who adopt gender-neutral identities on the Internet. If people are sensitive to gender-linked language effects, readers would be able to predict the gender of users even when they used gender-neutral identities. One participant in the Kaplan and Farrell (1994) study described how the absence of social cues (e.g. appearance) and salient power differentials (e.g. gender, age and power) created a more comfortable social space for her. The results of the present study imply that these social cues are not absent in electronic discourse. In fact, when inferring gender, people might place more weight on these cues than on the name given in the signature. Herring (1994) reports instances where the gender of a message writer is questioned because there is a mismatch between the gender of the pseudonym and the style of writing.

If anything, this study might underestimate the extent of gender-preferential language in electronic communication because of the sample used. University students contributed data, and they might not be representative of either the general population or the cybercommunity. Mulac and Lundell (1980), for example, found that university students demonstrated fewer gender-linked language differences than did other participants. Mulac (1989) also suggests that university students are less likely to display sex role stereotypical behaviour than the general population.

The results from this set of studies show that there are cues to gender in CMC. This supports claims that social identity cues are provided in electronic discourse (Spears & Lea, 1992; Spears *et al.*, 1990). The extent to which social identity cues are used will depend on the social context of the communication. When the task is primarily informal and social in nature, gender should be a salient social category. However, gender might be a less salient category when the goal of the communication is to exchange information or complete a set task. According to the SIDE model, when gender is a salient category for social identity, gender-typed behaviour and discrimination are more likely to occur (Postmes *et al.*, 1998) and power relations are reinforced (Spears & Lea, 1994). Although computer-mediated communication has been hailed as the great equalizer, in many instances it is likely to be just the opposite. Claims that the anonymity provided through CMC will free people of their social roles appears to have ignored the incredible propensity and ability of humans to construct social realities.

These findings raise some interesting questions about how readers will reply to electronic messages that are characteristically female or male in style. One theoretical perspective on communication influence is conversation accommodation theory (CAT; Coupland *et al.*, 1988). We did not examine accommodation in the present study, but are doing so in our current research. We expect that in computer-mediated communication, writers will adapt their style of communication to be more consistent with that of their target audience. Evidence of accommodation to gender-preferential styles has been shown in face-to-face communication with adults (Fitzpatrick *et al.*, 1995; Hannah & Murachver, 1999) and more recently

with children (Farry, 1999; Stephenson & Murachver, 1999). Coupland *et al.* report that individuals sometimes converge towards what they expect of their partner rather than their partner's actual behaviour. For example, young people might overaccommodate when speaking to a much older person because of beliefs they hold about social stereotypes of the aged. Although we do not know of similar findings of overaccommodation to gender stereotypes, this is an interesting avenue to explore. It is particularly relevant given that message writers often post to larger groups and it might be the emergent characteristics of the group that influence style of language rather than characteristics of individual group members. These points highlight the need to move beyond studying individuals who communicate and towards how individuals mutually influence, react and change with one another in the process of communicating.

In conclusion, the results of this study indicate that, as in written essays (Mulac *et al.*, 1990) and face-to-face communication (Coates, 1986; Fitzpatrick *et al.*, 1995; Holmes, 1984, 1995; Tannen, 1990, 1994), people show gender-linked language differences in electronic discourse. Moreover, people are sensitive to these differences, and even in the absence of gender-specific topics and physical indicators of gender they are able to identify the gender of the author of e-mail messages.

Acknowledgements

We would like to thank Meagan Stephenson, Victoria Johnstone, Sasha Farry, Ruth Walker, Briar Mackenzie, James Green, Russell Spears and three anonymous reviewers for their many helpful comments. Finally, thanks to the many students enrolled in psychology courses who volunteered their time and contributed data to this study.

References

- Bilous, F. R., & Krauss, R. M. (1988). Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads. *Language and Communication*, 8, 183–194.
- Bruckman, A. S. (1993, August). *Gender swapping on the Internet*. Paper presented at The Internet Society, San Francisco, CA. Available at (ftp.media.mit.edu/pub/asb/papers).
- Carli, L. L. (1990). Gender, language, and influence. *Journal of Personality and Social Psychology*, 5, 941–951.
- Coates, J. (1986). *Women, men, and language: A sociolinguistic account of sex differences in language*. London: Longman.
- Coupland, N., Coupland, J., Giles, H., & Henwood, K. (1988). Accommodating the elderly: Invoking and extending a theory. *Language in Society*, 17, 1–41.
- Dovidio, J. F., Brown, C. E., Heltman, K., Ellyson, S. L., & Keating, C. F. (1988). Power displays between women and men in discussions of gender-linked tasks: A multichannel study. *Journal of Personality and Social Psychology*, 55, 580–587.
- Farry, S. J. (1999). *An investigation of adolescent speech behaviour in different contexts*. Unpublished master's thesis, University of Otago, Dunedin, New Zealand.
- Fishman, P. M. (1977). Interactional shitwork. *Heresies: A Feminist Publication on Arts and Politics*, 2, 99–101.
- Fitzpatrick, M., Mulac, A., & Dindia, K. (1995). Gender-preferential language use in spouse and stranger interaction. *Journal of Language and Social Psychology*, 14, 18–39.

- Foertsch, J. (1993). The impact of electronic networks on scholarly communication: Avenues for research. *Discourse Processes*, 19, 301–328.
- Graddol, D., & Swann, J. (1989). *Gender voices*. Oxford: Basil Blackwell Ltd.
- Hannah, A., & Murachver, T. (1999). Gender and conversational style as predictors of conversational behaviour. *Journal of Language and Social Psychology*, 18, 153–174.
- Herring, S. (1993). Gender and democracy in computer-mediated communication. *Electronic Journal of Communication* [Online], 3(2). Available via e-mail message addressed to {Comserv@cios.org}. Message: Send Herring V3N293.
- Herring, S. (1994). *Gender differences in computer-mediated communication: Bringing familiar baggage to the new frontier*. Keynote talk at panel entitled 'Making the net*work*: Is there a Z39.50 in gender communication?' American Library Association annual convention, Miami, FL.
- Holmes, J. (1984). 'Women's language': A functional approach. *General Linguistics*, 24, 149–178.
- Holmes, J. (1988). Paying compliments: A sex-preferential politeness strategy. *Journal of Pragmatics*, 23, 445–465.
- Holmes, J. (1995). *Women, men, and politeness*. Harlow: Longman.
- Jaffe, J. M., Lee, Y., Huang, L., & Oshagan, H. (1995). *Gender, pseudonyms, and CMC: Masking identities and baring souls*. Paper presented at the annual conference of the International Communication Association, Albuquerque, New Mexico. Available at (<http://research.haifa.ac.il/~jmjaffe/genderpseudocmc.html>).
- James, D., & Clarke, S. (1993). Women, men and interruptions: A critical review. In D. Tannen (Ed.), *Gender and conversational interaction* (pp. 231–280). Oxford/New York: Oxford University Press.
- James, D., & Drakich, J. (1993). Understanding gender differences in amount of talk: A critical review of research. In D. Tannen (Ed.), *Gender and conversational interaction* (pp. 281–312). Oxford/New York: Oxford University Press.
- Kaplan, N., & Farrell, E. (1994). Weavers of webs: A portrait of young women on the net. *Arachnet Electronic Journal on Virtual Culture*, 2, 26 June.
- Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist*, 39, 1123–1134.
- Kiesler, S., & Sproull, L. (1992). Group decision making and communication technology. *Organizational Behaviour and Human Decision Processes*, 52, 96–123.
- Kramarae, C., & Taylor, H. J. (1993). Women and men on electronic networks: A conversation or a monologue? In C. Kramarae, H. J. Taylor, & M. Ebben (Eds.), *Women, information technology and scholarship* (pp. 52–61). Urbana-Champaign, IL: Centre for Advanced Studies, University of Illinois.
- Lakoff, R. (1973). Language and women's place. *Language in Society*, 2, 45–80.
- Lakoff, R. (1990). *Talking power: The politics of language*. New York: Basic Books.
- Lea, M. (1991). Rationalist assumptions in cross-media comparisons of computer-mediated communication. *Behaviour, & Information Technology*, 10, 153–172.
- Lea, M., O'Shea, T., Fung, P., & Spears, R. (1992). 'Flaming' in computer-mediated communication. In M. Lea (Ed.), *Computer-mediated communication* (pp. 89–109). London: Harvester Wheatsheaf.
- McMillan, J. R., Clifton, A. K., McGrath, D., & Gale, W. S. (1977). Women's language: Uncertainty or interpersonal sensitivity and emotionality. *Sex Roles*, 3, 545–559.
- Mulac, A. (1989). Men's and women's talk in same-gender and mixed-gender dyads: Power or polemic? *Journal of Language and Social Psychology*, 8, 249–270.
- Mulac, A. (1998). The gender-linked language effect: Do language differences really make a difference? In D. J. Canary & K. Dindia (Eds.), *Sex differences, & similarities in communication* (pp. 127–153). Mahwah, NJ: Erlbaum.
- Mulac, A., & Lundell, T. L. (1980). Differences in perceptions created by syntactic-semantic productions of male and female speakers. *Communication Monographs*, 47, 111–118.
- Mulac, A., & Lundell, T. L. (1986). Linguistic contributors to the gender-linked language effect. *Journal of Language and Social Psychology*, 5, 81–101.
- Mulac, A., Studley, L. B., & Blau, S. (1990). The gender-linked language effect in primary and secondary students' impromptu essays. *Sex Roles*, 23, 439–469.

- Mulac, A., Wiemann, J. M., Widenmann, S. J., & Gibson, T. W. (1988). Male/female language differences and effects in same-sex and mixed-sex dyads: The gender-linked language effect. *Communication Monographs*, 55, 315–335.
- Postmes, T., Spears, R., & Lea, M. (1998). Breaching or building social boundaries? SIDE-effects of computer-mediated communications. *Communication Research*, 25, 689–715.
- Rosch, E. H. (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–233.
- Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rumelhart, and the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models* (pp. 7–57). Cambridge, MA: MIT Press.
- Savicki, V., Kelley, M., & Oesterreich, E. (1999). Judgments of gender in computer-mediated communication. *Computers in Human Behaviour*, 15, 185–194.
- Savicki, V., Lingenfelter, D., & Kelley, M. (1996). Gender language style and group composition in Internet discussion groups. *Journal of Computer-Mediated Communication* [Online], 2(3). Available at (<http://www.ascusc.org/jcmc/vol2/issue3/savicki.html>).
- Spears, R., & Lea, M. (1992). Social influence and the influence of the 'social' in computer-mediated communication. In M. Lea (Ed.), *Computer-mediated communication* (pp. 30–64). London: Harvester Wheatsheaf.
- Spears, R., & Lea, M. (1994). Panacea or panopticon: The hidden power in computer-mediated communication. *Communication Research*, 21, 427–459.
- Spears, R., Lea, M., & Lee, S. (1990). De-individuation and group polarisation in computer-mediated communication. *British Journal of Social Psychology*, 29, 121–134.
- Spender, D. (1995). *Nattering on the net: Women, power and cyberspace*. North Melbourne, Victoria: Spinifex Press.
- Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organisational communication. *Management Science*, 32, 1492–1512.
- Stephenson, M., & Murachver, T. (1999, September). *Gender-preferential speech and accommodation in children*. Paper presented at the New Zealand Psychological Society Conference, Dunedin.
- Swacker, M. (1975). The sex of the speaker as a sociolinguistic variable. In B. Thorne & N. Henley (Eds.), *Language and sex: Difference and dominance* (pp. 76–83). Rowley, MA: Newbury House.
- Tannen, D. (1990). *You just don't understand: Women and men in conversation*. New York: William Morrow.
- Tannen, D. (1994). *Talking from 9 to 5*. London: Virago Press.
- Truong, H. (1993). Gender issues in online communications. *Bay Area Women in Telecommunications*. Available at (<ftp://ftp.cpsr.org/cpsr/gender/bawit.cfp93>).
- Turkle, S. (1990) Parallel lives: Working identity in virtual space. In D. Grodin & T. R. Lindlof (Eds.), *Constructing the self in a mediated world: Inquiries in social construction* (pp. 156–175). Thousand Oaks, CA: Sage.
- Witmer, D. F., & Katzman, S. L. (1997). On-line smiles: Does gender make a difference in the use of graphic accents? *Journal of Computer-Mediated Communication* [Online], 2(4). Available at (<http://www.ascusc.org/jcmc/vol2/issue4/witmer.html>).
- Wood, W. (1987). Meta-analytic review of sex differences in group performance. *Psychological Bulletin*, 102, 53–71.
- Yates, J., & Orlikowski, W. J. (1993). *Knee-jerk anti-LOOPism and other e-mail phenomena: Oral, written, and electronic patterns in computer-mediated communication*. Paper presented at the 53rd Annual Meeting of the Academy of Management, Atlanta, GA.
- Zadeh, L. (1982). A note on prototype theory and fuzzy sets. *Cognition*, 12, 291–297.
- Zimmerman, D. H., & West, C. (1975). Sex roles, interruptions, and silences in conversation. In B. Thorne & N. Henley (Eds.), *Language and sex: Difference and dominance* (pp. 105–129). Rowley, MA: Newbury House.

Received 2 December 1998; revised version received 7 December 1999

Appendix 1

Messages between female–female netpal pairs and male–male netpal pairs^a*Female–female messages*

From: SYCO322

To: SYCO330

Subj: Sorry I have'nt written

Dear syco330

thank you very much for all your lovely letters. I'm sorry, I'm not as good a letter sender as you. I have had a very busy morning. Three lectures in a row (in the same lecture theatre) and then quite a fun tutorial. The tutorial was for English 124, the most horrible paper you can take. I'm doing health science, if you hadn't guessed. Anyway I'm getting hungry so I think I'll finish this message and go and have some lunch. I apologise for any spelling mistakes, of which I'm sure there are plenty. I can't spell for peanuts. Have a good day, I'll write again soon—SYCO322

From: SYCO330

To: SYCO322

Subj: THANKYOU

DEAR SYCO322

How are you? Thankyou for your letter, I think I finally have got this thing sussed./ I notice that everyone who is taking english 124 doesn't seem to be thrilled with it. I read all the benches in the lecture theatres about how much people hate it.

So, are you trying for medicine, pharmacy or what? I eventually want to do Marine biology after I have done my degree. I am doing psych this year, because I want to do some research into animal behaviour in the long term . . . Well, that's my exciting career aspirations; for now I have to put up with hostel life, and boring lectures. But I like Dunedin so far. There seems to be a good student lifestyle down here.

Can't think of much else to say . . . so I'll talk to you again soon,

Have a great weekend, don't party too hard!

From your friend, SYCO 330

Male–male messages

From: SYCO329

To: SYCO307

Subj: netpal

Hello Syco307, what's up? I've just totally stuffed up the message that I was trying to send to you, so don't be suprised if you get something crazy before this. How's things with you anyway? Should I keep calling you syco, or should I call you —? (How's that for interactive?) My name is syco329 but my friends call me Sam. Take your pick. Well, I'm off now as my pinkies are sore from all this writing. Catch you later,

Sam

From: SYCO307

To: syco329

Subj: crazy message senders

Well, Sam, my friends call me Peter, or socks (to long a story to go into at present). I think Peter will do just nicely; no-one will get offended, and I won't have to send you hate mail. As it happens, I did get a rather crazy message just before I got the sane one. Actually, it was a triple-message, but hey,

^aThese messages are presented as written by participants (including grammatical and spelling errors). Personal names, when mentioned, have been changed to protect anonymity.

who's counting, right? Well, I've gotta fly like Orville, I've got a Russian (my major) translation due at 5. Love to the folks.

Peter

Appendix 2

Sample female and male messages

Female version

Hi, this is just a brief note, because I am really stressed at the moment. I had to drop some boxes of stuff off at a friend's before school today. Unfortunately the car wouldn't start, so I had to ring the AA. After ages they turned up and got it going. By that time I was late for my lecture, which was annoying. Sorry I couldn't make it last night. I have a big assignment that is due tomorrow, and I haven't started it yet. See ya later.

Male version

Hi, Write to me you lazy sod! This is just a brief note, because I am stressed at the moment. I had to drop some boxes of stuff off at a friend's before school today but the car wouldn't start, so I had to ring the AA. After ages they turned up and got it going, but I was late for my lecture. I have a big assignment that is due tomorrow, and I haven't started it yet. See ya later.