Gender, language and computer-mediated communication

J. Miller & A. Durndell

Department of Psychology, Glasgow Caledonian University, Scotland

Abstract

This research aimed to investigate gender and language in the context of educational computer-mediated interactions. As the use of online discussion groups in campus-based education increases, so does the need to investigate the impact of gender on language and communication styles in this context. It has been claimed that the lack of social context cues in computer-mediated communication (CMC) equalises participation, resulting in a more democratic environment than face-to-face communication. However, gender differences in power and language use have been found online. In total, 197 students (148 females, 49 males) participated in open-ended online discussion as an optional part of their introductory psychology module. Informed consent was sought from the online participants to download the electronic discourse for analysis using Atlas.ti 4.2. Participation and language use were analysed using a combination of quantitative and qualitative methods. A detailed coding scheme was developed during the pilot study that incorporated various linguistic, stylistic, paralinguistic and task variables. This coding scheme was used in subsequent studies and 699 messages from four studies were coded in total. Males and females were similar overall with regard to participation as shown by the total number of posts and mean number of words per post. However, gender-related patterns in language use and interaction styles were found. The results are discussed in relation to previous literature on gender and language in both face-to-face and computer-mediated contexts and in relation to the democratising theories of CMC. The implications of the results for the use of CMC in Higher Education will also be considered.

Keywords: computer-mediated communication, electronic discourse, online discussion groups, gender, language use, qualitative content analysis.
1 Introduction

Asynchronous computer-mediated communication (CMC) refers to text-based communication that takes place over computer networks in delayed time. This includes electronic mail and online discussion forums such as listservs, newsgroups, bulletin boards or asynchronous computer conferences. Online discussion groups allow the exchange of text-based messages surrounding a particular topic. Asynchronous discussion forums are increasingly being incorporated into campus-based Higher Education (HE) courses in order to support peer interaction and cooperative learning. The term computer-supported cooperative learning (CSCL) can be used to cover any form of cooperative learning communication that occurs over a network of computers (McConnell [1]).

It has been claimed that the loss of social context cues, such as gender, in text-based CMC equalises participation, making it a more democratic form of communication, in comparison to traditional face-to-face methods (Sproull and Kiesler [2]). It is proposed that the loss of nonverbal indicators of status and power enable participants, who would otherwise defer speaking turns to higher-status participants in face-to-face interaction, become uninhibited and participate more. For example, it has often been found that males dominate mixed-sex interaction, participating more frequently than females and for longer in public and formal contexts (Coates [3], Tannen [4]).

Furthermore, it has been suggested that under conditions of anonymity males and females may be less likely to feel that they have to project the socially expected qualities corresponding to their gender (Matheson and Zanna [5]). Indeed, early reviews concerning online communication implied that cyberspace was ‘a utopian place, where ideas, not social factors, were the key features of these new discourse settings’ (Gurak and Eberltof-Kraske [6]). This makes CMC an attractive option for use in educational contexts, in addition to the benefits of increased time for reflection and the articulation of ideas into words that are shared and built upon through the responses of others (Harasim et al. [7]).

The filtering out of cues in CMC is also claimed to result in communication that is less personal, compared to face-to-face interaction (e.g. Sproull and Kiesler [2]) and can lead to increased disinhibited behaviour such as flaming (e.g. swearing and personal insults). However, Joinson [8] notes that disinhibited behaviour can also be in the form of self-disclosure. The lack of social presence that is assumed to characterise CMC is thought to lead to a lack of socioemotional discourse and responsiveness to one another’s ideas, relative to face-to-face communication (Sproull and Kiesler [2]). Therefore this suggestion has implications for the use of CMC to support peer interaction in HE and responding to another’s ideas is deemed to be an important cognitive skill. Socioemotional content includes use of emoticons, expressions of supporting references, self-references, references to others and self-disclosure (Jaffe et al. [9]). However, Walther’s [10] social information processing theory assumes that CMC participants are affected by the same internal drive of ‘affiliation’ as
participants in face-to-face communication and argues that text-based CMC can support socioemotional communication. Gender may influence the expression of socioemotional behaviour online. Males are believed to value status more through the process of gender role socialisation, whereas females are thought to value connection or affiliation, leading to gender-preferential communication styles, differentiated as ‘competitive’ and ‘cooperative’ (Coates [3], Tannen [4]). Therefore, females may be more likely than males to express socioemotional responses in CMC such as those described above. However, negative forms that would also qualify as responses to another’s ideas, such as expressions of disagreement, are not always included in definitions of socioemotional content.

The Social Identity Explanation of Deindividuated Effects (SIDE) model (Reicher et al. [11]) predicts that visual anonymity of the self to others leads to heightened self-awareness and greater adherence to group norms when a social identity is salient. For example, if gender is made a salient social identity in CMC then this could invoke behavioural norms and stereotypes regarding gender appropriate behaviour, influencing expectations and perceptions of CMC users. Similarly, Matheson and Zanna [5] argue that differences in status may actually be accentuated in CMC if cues to gender are available. Low public-awareness levels in CMC are associated with lower social pressures that make the expression of internalised gender biases unacceptable. Therefore, it is possible that CMC could exacerbate existing asymmetrical power differences, or even create them.

Furthermore, cues to gender may not be restricted to usernames and signatures in CMC. Experimental work by Thomson and Murachver [12] found evidence for gender-preferential language in informal CMC and showed that readers of email messages used gender-linked language differences within messages to correctly identify the author’s gender. Therefore, if linguistic cues to gender are found in other CMC contexts, this will have implications for the assumption of gender anonymity online and CMC as a gender-free environment.

Gender-related patterns in language style have been reported in Internet discussion groups such as email listservs (e.g. Herring [13], 14)). Similar to face-to-face research, these results suggest that males dominate interaction, which may deter women from participating in CMC or force them to seek women-only groups online. Herring found large differences in language style along the same task-orientated versus socioemotional dimensions as Tannen [4] described. Female postings tended to display features of attenuation, such as hedging, apologising, asking questions and a personal orientation, revealing thoughts and feelings and interacting with and supporting others. On the other hand, male postings were lengthy and/or frequent, adversarial and featured strong assertions, self-promotion, sarcasm and flaming. It is suggested that gender-based communication styles and the power dynamics associated with these styles carry over to electronic environments, despite the loss of overt face-to-face cues to gender [14].

Savicki et al. [15] provided some evidence for the gender-related communication styles identified by Herring [14] in their study of Internet discussion groups. Soukup [16] observed traditional masculine and feminine
forms of discourse in Internet chatrooms. However, Yates [17] notes that many studies have failed to replicate these findings. For example, Michaelson and Pohl [18] did not find differences along the supportive/emotional versus adversarial/task-oriented dimensions in their study of an email problem-solving task. Mixed results in terms of gender variations in language use in CMC make in unclear whether CMC moderates or magnifies the gender differences reported in face-to-face research. Perhaps these differences could be attributed to the varying online contexts that have been studied. There is also the extent to which gender can reliably be inferred by researchers from usernames or email addresses. For example, Jaffe et al. [9] found that females are more likely to choose a pseudonym to mask their gender when communicating online, which could distort the findings of studies using samples taken directly from the Internet.

The present research is concerned with formal use of CMC on an undergraduate module. This is a context in which gender equality is of major importance. If differences in participation and communication style are found to exist in this context then this will have implications for the increasing use of CMC in HE. Therefore this research aims to investigate the existence of gender-related patterns in participation and language use in educational, mixed-gender, online discussion groups.

2 Method

2.1 Design

The computer-mediated discourse of students using CMC for open-ended discussion of psychological topics was downloaded with informed consent for a qualitative content analysis procedure (Mayring [19]). Qualitative content analysis seeks to conserve some methodological advantages of quantitative content analysis and broaden them to a concept of qualitative procedure. It involves quantitative methods such as defining the unit of analysis and the coding categories and carrying out a check of inter-rater reliability [19]. In the present research, the unit of analysis was at the level of the message, following Rourke et al. [20] recommendation. A coding scheme was developed during the initial study that was used to code 699 messages, obtained through four consecutive studies. The coding scheme was extensive and covered 12 linguistic, 8 paralinguistic, 19 stylistic and 19 task codes. A second independent rater coded 20 per cent of the sample and inter-rater reliability was good (Cohen’s kappa = 0.9). The results were analysed for overall effects of gender on choice of user identification (real name or pseudonym), quantitative measures of participation (mean number of posts and words per post) and frequencies of coded categories relating to linguistic and stylistic usage (at varying levels of abstraction using Atlas.ti 4.2).
2.2 Participants

The participants were 197 campus-based introductory psychology students (149 females, 48 males) at a Scottish university who had all chosen to take part in the online discussion groups as part of the module. The age range was 17-46 years and the mean age of the sample was 22 years (SD = 6.52).

2.3 Apparatus

The coding list was developed using Atlas.ti 4.2 and mainly consisted of variables coded as potential discriminators of gender in previous studies (e.g. Herring [14]). The coding scheme is listed in Miller [21], along with a full analysis of individual variables. Codes were grouped into eight major code families. These were female language (e.g. self-disclosure, intensifiers), male language (e.g. humour, rhetorical question), task-orientated (e.g. answers question), socioemotional (e.g. references by name), attenuated (e.g. personal opinion, qualifiers), authoritative (e.g. strong assertion, presuppositions), negative socioemotional (e.g. disagreement) and positive socioemotional (e.g. agreement). A full description of these code families and examples is in Miller [21]. Analyses were then conducted on ‘supercodes’, which are stored queries constructed from combinations of code families using Boolean operators in Atlas.ti 4.2 that are used to search the data and explore patterns of language use.

2.4 Procedure

Students on an introductory psychology module were invited to participate in online discussion groups for extra coursework credit. Students were given instructions on how to access the online forums. They used their real name or matriculation number as a user identification. Ethical clearance permitted access to background details of the students, such as gender. Coding was carried out in Atlas.ti 4.2, as described above. Statistical analyses were carried out using SPSS.

3 Results

A total of 699 postings (538 female, 161 male) were analysed. Table 1 shows the mean number of posts and words per post by gender.

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>48</td>
<td>3.35</td>
<td>3.64</td>
<td></td>
<td>.441</td>
</tr>
<tr>
<td>Female</td>
<td>149</td>
<td>3.61</td>
<td>3.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>3.55</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Words per post</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>48</td>
<td>128.48</td>
<td>89.09</td>
<td></td>
<td>.245</td>
</tr>
<tr>
<td>Female</td>
<td>149</td>
<td>125.23</td>
<td>76.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>197</td>
<td>126.03</td>
<td>79.58</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 shows the online contributions of the 197 participants, in terms of the mean number of posts and words per post. The mean number of posts was 3.55 (SD = 3.50) and words per post was 126.03 (SD = 79.58). Gender was not found to significantly influence the frequency of posts ($t = -.441$, d.f. = 196, n.s.) or the length of posts ($t = .245$, d.f. = 196, n.s.).

Table 2: Gender and user id.

<table>
<thead>
<tr>
<th></th>
<th>Gives name</th>
<th>Pseudonym</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>41</td>
<td>7</td>
<td>48</td>
</tr>
<tr>
<td>Female</td>
<td>88</td>
<td>61</td>
<td>149</td>
</tr>
<tr>
<td>Total</td>
<td>129</td>
<td>68</td>
<td>197</td>
</tr>
</tbody>
</table>

Table 3: Supercode analysis.

<table>
<thead>
<tr>
<th>Supercode</th>
<th>Male (%) (N=48)</th>
<th>Female (%) (N=149)</th>
<th>Male (%) (N=161)</th>
<th>Female (%) (N=538)</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attenuated NOT Authoritative</td>
<td>31% (N=15)</td>
<td>64% (N=96)</td>
<td>16% (N=26)</td>
<td>41% (N=218)</td>
<td>16.26**</td>
</tr>
<tr>
<td>Authoritative NOT Attenuated</td>
<td>69% (N=33)</td>
<td>37% (N=55)</td>
<td>40% (N=64)</td>
<td>16% (N=85)</td>
<td>14.88**</td>
</tr>
<tr>
<td>Male NOT Female Language</td>
<td>33% (N=16)</td>
<td>12% (N=18)</td>
<td>12% (N=20)</td>
<td>4% (N=19)</td>
<td>11.50**</td>
</tr>
<tr>
<td>Female NOT Male Language</td>
<td>33% (N=16)</td>
<td>69% (N=103)</td>
<td>20% (N=32)</td>
<td>44% (N=235)</td>
<td>19.43**</td>
</tr>
<tr>
<td>Male AND Female Language</td>
<td>73% (N=35)</td>
<td>65% (N=97)</td>
<td>61% (N=98)</td>
<td>46% (N=246)</td>
<td>1.00</td>
</tr>
<tr>
<td>Positive NOT Negative</td>
<td>35% (N=17)</td>
<td>58% (N=86)</td>
<td>19% (N=30)</td>
<td>35% (N=190)</td>
<td>7.24*</td>
</tr>
<tr>
<td>Negative NOT Positive</td>
<td>56% (N=27)</td>
<td>19% (N=28)</td>
<td>27% (N=44)</td>
<td>7% (N=38)</td>
<td>25.32**</td>
</tr>
<tr>
<td>Task-Oriented NOT Socioemotional</td>
<td>56% (N=27)</td>
<td>63% (N=94)</td>
<td>36% (N=58)</td>
<td>30% (N=161)</td>
<td>0.72</td>
</tr>
<tr>
<td>Socioemotional NOT Task-Oriented</td>
<td>33% (N=16)</td>
<td>37% (N=55)</td>
<td>16% (N=25)</td>
<td>13% (N=71)</td>
<td>0.21</td>
</tr>
<tr>
<td>Task Oriented AND Socioemotional</td>
<td>65% (N=31)</td>
<td>73% (N=109)</td>
<td>44% (N=71)</td>
<td>52% (N=279)</td>
<td>1.29</td>
</tr>
</tbody>
</table>

*significant at $p < .01$, **significant at $p < .001$

Table 2 shows the frequencies of male and female participants who chose to give their real name and those who opted for a pseudonym in the form of their matriculation number. Chi-square analysis showed a significant gender difference in the choice of user identification ($\chi^2 = 11.16$, d.f. = 1, $p < .01$). Males were more likely to give their real name online in comparison to females.

The electronic discourse was characterised by extensive first person pronoun usage as 87 per cent of all postings analysed contained some form of first person
or plural pronouns (e.g. ‘I’, ‘we’). Also, 38 per cent of student contributions directly responded to the ideas of other students by expressing agreement and/or disagreement. Just over a fifth of all postings contained references to own emotions or self-disclosure.

The results of the supercode analysis are shown in Table 3. The supercode queries are listed, as are the corresponding proportions of males and females making each type of contribution, on which chi-square analyses were conducted. The percentages of each type of posting, expressed as a percentage of the total male and female postings are also given in the table. Significantly more females made contributions that were attenuated but not authoritative ($x^2 = 16.26$, d.f. = 1, $p < .001$) and consisted of female language features only ($x^2 = 19.43$, d.f. = 1, $p < .001$) than males. Significantly more males made contributions that were authoritative but not attenuated ($x^2 = 14.88$, d.f. = 1, $p < .001$) and consisted of male language features only ($x^2 = 11.50$, d.f. = 1, $p < .001$) than females. There were no significant gender differences found along the task-orientated versus socioemotional dimensions. However, the socioemotional code family was further divided into positive socioemotional and negative socioemotional families. This revealed that significantly more females engaged in positive socioemotional behaviour than males ($x^2 = 7.24$, d.f. = 1, $p < .01$), whereas significantly more males engaged in negative socioemotional behaviour than females ($x^2 = 25.32$, d.f. = 1, $p < .001$).

4 Discussion

This research investigated participation and language use in educational CMC. As stated in the introduction, males have been found to dominate mixed-sex interaction in both face-to-face (Coates [3]) and computer-mediated contexts (Herring [13, 14]). However, in this study no significant gender differences were found in the frequency and length of posts. Males and females were similar regarding these quantitative measures of participation, supporting democratising theories of CMC (e.g. Sproull and Kiesler [2]). Perhaps equal participation is more likely to be a feature of this context, as opposed to a public online forum, as females may be more motivated to participate in the context of education. Females may also have felt more comfortable participating as they were the majority group and thus may have influenced the overall linguistic norms.

It is also possible that the pseudonymity offered could have been conducive to female participation online. It was found that females were significantly more likely to choose a pseudonym over their real name than males. This suggests that females had a greater wish to appear ‘anonymous’ than males, which could reflect, as Jaffe et al. [9] suggested, an effort to maintain equality of status. Females may experience a lack of confidence or implicit social pressure when participating in mixed-sex interaction that the loss of face-to-face cues and pseudonymity may have reduced, which could in turn explain the equal participation observed online.

The computer-mediated discourse also showed extensive first person pronoun use and over a third of postings expressed agreement or disagreement with other...
students. It should be noted that tutors also contributed online and many student postings were in response to questions set by tutors. Just over a fifth of student postings contained references to own emotions or self-disclosure. As these factors are all said to contribute to social presence, this suggests that the computer-mediated context is perhaps not as impersonal as previously claimed (e.g. Sproull and Kiesler [2]) and provides support for social information processing theory (Walther [10]), in that CMC can support socioemotional communication. It is possible that participants, especially females, developed ways of overcoming what could be perceived as an impersonal environment through the development of a norm of explicitly personal writing.

Although no gender differences were found in terms of the amount of participation, the qualitative content analysis revealed gender-related patterns in language use. Females were more likely to make attenuated contributions and use only traditional female language features in their postings, whereas males were more likely to make authoritative postings and use only male language features. These findings support the communication styles identified by Herring [14] to some extent. Females were found to employ personal and emotional forms of language more than males, who in turn used more authoritative language.

However, the results do not support the task-orientated versus socioemotional distinction in relation to gender. Males and females were just as likely to send messages of each type and, more commonly, messages containing both task-orientated and socioemotional content. However, the distinction between these categories was problematic as participants were frequently engaging in task behaviours due to the specific formal purpose, some of which could also be classed as socioemotional behaviour (e.g. requests opinions). Furthermore, the definition stated in the introduction appears to include only positive elements of socioemotional discourse such as expressions of support. However, disagreement and challenging utterances were also categorised as socioemotional discourse in this research, as they also involve references to others and reacting to others. Thus a distinction was made between positive and negative socioemotional behaviour. Gender-related patterns were found as males had a tendency to post negative responses, whereas females were more likely to respond positively to other participants in the online discussion.

The tendency to agree and support, could be interpreted as taking a low-power role in the discussion. Tannen [4] states that women typically use more supportive language patterns, which have the effect of diminishing the power of their own contributions. The issue of the perceived credibility of a contribution that uses personal and attenuated language is worthy of further investigation, as these linguistic forms are often associated with negative stereotyping. Furthermore, the finding that males tend to express disagreement more, whereas females tend to express agreement supports Coates [3] and Tannen [4] that the male style is based on competitiveness and the female style is based on cooperativeness. It could also reflect gender-related learning preferences, as females may prefer to learn through connectedness and cooperativeness, whereas males may prefer a more independent and argumentative learning environment.
It is possible that the gender-related patterns reported here, which are similar to traditional gender role stereotypes, were found as the salience of gender as a social category may have invoked norms and stereotypical expectations regarding gender appropriate behaviour (Matheson and Zanna [5], Reicher et al. [11]). Gender was made a salient social category to some extent, as some participants used their real name online and the tendency for females to opt for a numerical identification more than males could perhaps have acted as a gender marker in itself. It was also found that some participants revealed their gender in their contributions (e.g. ‘as a gay guy’, ‘as a 19-year old girl’).

Finally, the unequal male and female sample sizes were unavoidable in this context as female students often outnumber male students on psychology modules. It is argued that studying gender and CMC in a meaningful context outweighs the potential drawbacks. Future research should investigate the extent to which the gender-related patterns found here can be extrapolated to other educational computer-mediated contexts, for example male-majority courses such as engineering. The relationship between gender identity and expression of that identity in online text-based contexts is also deemed worthy of attention. In conclusion, the results were positive for use of CMC in education in terms of equal participation, however they also suggest that CMC does not guarantee a gender-free environment. Cues to gender were found to exist in the language used by CMC participants and as the work by Thomson and Murachver [12] suggests, CMC users may be able to identify gender based on these cues alone.

References


