

## **Gender, Communication, and Self-Presentation in Teen Chatrooms Revisited: Have Patterns Changed?**

Sanja Kapidzic

Susan C. Herring

*Indiana University, Bloomington*

### **Abstract**

This study evaluates empirically the proposition implicit in much recent gender and CMC research that expressions of gender distinctness among teens in online environments are becoming less frequent and less traditional. Gender preferences were analyzed in linguistic features and communication styles in synchronous text messages, along with self-presentation in user profile pictures, drawing on data from popular English-language teen chat sites collected in 2010. Significant differences were found in speech acts, message tone, and in physical stance, dress, and social distance in profile pictures that generally conform to traditional gender stereotypes. These findings are interpreted in light of previous gender and teen CMC research, adolescent development and socialization patterns, mass media representations, and trends towards media convergence in chat platforms.

### **Introduction**

During the teenage years, identity and sexuality start to play major roles in the lives of young people, as they seek to define and explore who they are (Erickson, 1968). Subrahmanyam, Greenfield, and Tynes (2004) identified teenage chatrooms as a rich source of data for the study of adolescent development, in that they provide insight into the kind of unmonitored interaction to which researchers normally lack access. A study conducted in 2007 by the Pew Internet and American Life Project found that of the estimated 93% of teens who used the Internet in America, nearly 20% visited chat sites, despite growing competition from instant messaging and social network sites. Multiparticipant text chat is also common in virtual worlds and online gaming environments (Paolillo & Zelenkauskaitė, in press), some of which, such as World of Warcraft, are popular with teens. In this study, we examine teenage chat sites to identify possible gender preferences in the online communication and self-presentation strategies of young people.

The study of gender differences in computer-mediated communication (CMC) has a long history, relative to the study of CMC as a whole. As early as 1991, Selfe and Meyer reported gendered power dynamics in an asynchronous academic discussion list, with men and high-profile members of the community dominating communication, even under conditions of pseudonymity. In the early 1990s, Herring (1992, 1993, 1994) identified gender preferences in discourse style among adults posting messages to academic discussion lists: Women tended to use more hedges and politeness markers and manifest a more supportive attitude towards their addressees in comparison with men, who tended to make more strong assertions, violate conventional politeness norms, and adopt a more adversarial stance towards their interlocutors. These observations were supported in later studies of asynchronous CMC (e.g., Guiller & Durdell, 2006; Thompson & Murachver, 2001). They also resemble previously-reported findings of gender differences in spoken communication (cf. Coates, 1993; Lakoff, 1975;

Tannen, 1994), suggesting that Internet users transfer their offline communication patterns into asynchronous computer-mediated messages.

The picture for chat environments has always been less clear. Some scholars writing in the mid-1990s contended that traditional gender binaries were blurring and breaking down in synchronous chat environments such as MUDs and MOOs<sup>1</sup> (Danet, 1998) and Internet Relay Chat (Rodino, 1997), due to the greater anonymity afforded by these text-only technological environments, which renders them conducive to playful experimentation with identity. For these scholars, pseudonymous chat environments bring out the inherently fluid, performative nature of gender identity (Butler, 1990), in keeping with—and possibly even accelerating—a trend toward the deconstruction of gender categories in postmodern society. Some intriguing anecdotal evidence of creative play with gender identity was indeed reported for recreational chat environments by, e.g., McRae (1996); however, the findings of empirical chat research mostly resemble those of previous studies. For example, Cherny (1994) found that stereotypically gendered patterns of behavior were reproduced in playful textual actions performed in a social MOO, and Herring (2003) found gendered discourse patterns in the pseudonymous Internet Relay Chat channels she studied, including teen channels, during the mid-1990s: Females smiled and laughed more, while males were more aggressive, including sexually.

More recently, empirical evidence has surfaced of some non-traditional patterns involving teenagers online. Subrahmanyam et al. (2004) found that girls were quite sexually assertive in the two teen chatrooms they studied (although Subrahmanyam, Smahel, and Greenfield [2006] observed that similar to offline romantic pursuits, those who identified as female online were more likely to use sexually implicit communication, whereas those identifying as male were more likely to use sexually explicit communication). In a corpus-based study of adolescent blogs, Huffaker and Calvert (2005) found no significant gender differences in frequencies of words expressing cooperation and passivity, which they had predicted females would use more of based on past gender and language research (although males used more resolute and active language, as predicted). Moreover, in a study of teenagers' Internet use, Gross (2004) found that male and female teenagers did not differ greatly in their online habits or behavior.

These last observations raise an important issue, which is that times—and technologies—have changed since the gender and CMC research conducted in the 1990s. Young people today, heirs to the benefits of the second wave of the feminist movement in the late 20th century, have been claimed to be more gender egalitarian in their interactions and androgynous in their self-presentation than previous generations (Twenge, 1997). At the same time, the development of Web 2.0 technologies, with their convergence of multimodal, communicative, and collaborative features, has given rise to mega-sites popular with teens, such as the social network site MySpace (boyd, 2008). Web-based teen chat sites, too, have taken on a new complexion, incorporating such features as asynchronous discussion forums, polls, and testimonials; 'avatars,' or visual representations of each user attached to their posted messages; and user profiles, in which personal information about users may be presented along with real-life photographs.

One consequence of these technological changes is that anonymity is less a feature of social media sites now than it was in the past. This raises the question of how and to what extent gender identity is expressed on such sites. Indeed, although all of the chat sites analyzed for the

present study offer users the possibility of hiding their gender, most teenage users choose to identify themselves as male or female (e.g., through the choice of distinctly gendered nicknames and/or photographs). This gives researchers the possibility to examine and compare male and female communication and self-presentation strategies systematically.<sup>2</sup>

The purpose of this study is to evaluate empirically the claim implicit in much recent gender and CMC research that expressions of gender distinctness among teens in online environments are becoming less frequent and less traditional. Discourse and content analysis methods are employed to examine gender preferences in linguistic features and communication styles in synchronous text chat messages, along with self-presentation in user profile images, drawing on data from five popular (default heterosexual) teen chat sites collected in early 2010. The findings suggest a distinction according to the nature of the linguistic and communicative features analyzed: Word choice, as well as some speech acts, appear to be determined more by the topic of conversation at hand than by the message producer's gender, but significant differences are found for the use of other speech acts (especially directive and reactive acts) and overall message tone, as well as for physical stance, dress, and social distance in profile images, which appear to function as broader, more stable signals of gender identity across topics. Where differences are found, they generally conform to traditional gender stereotypes.

These findings indicate that despite changes in technology and purported feminist advances in society over the past 20 years, traditional gender patterns in communication style and self-presentation persist in CMC, at least in heterosexual teen chat sites. We suggest that this is because the patterns are perceived by the teens who employ them to serve useful purposes. Symbolic gender differentiation via language and images aims to heighten mutual (sex-differentiated) attractiveness in teen chatrooms, in which direct physical actions are unavailable, and thus it can be read as socially facilitative (cf. Eckert, 1996). At the same time, expectations for what constitutes female and male attractiveness are not random; rather, they are ingrained in western society and reinforced by mass media representations (cf. Durham, 2008).

## **Background**

Much research has been conducted on the topic of whether males and females communicate differently and if so, at what level of communication differences are evident. Unlike in languages such as Japanese, where men and women use different forms for the personal pronouns 'I' and 'you', or in the Romance languages, where agreement with the gender of the speaker is obligatorily marked on adjectives, gender differences in English tend to be a matter of preference rather than grammatical requirement (Coates, 1993). Thus Argamon, Koppel, and Shimoni (2003) were able to train a machine-learning algorithm to identify the gender of writers of various genres of English texts with an 80% degree of accuracy, based solely on the frequency of use of grammatical function words: Female writers used more personal pronouns, while male writers used more noun determiners (articles, demonstratives, and quantifiers). The researchers explained this finding in terms of females having a more interactive, interpersonal style as compared to males, who tend to communicate more about things than about people. Relatedly, Pennebaker, Mehl, and Niederhoffer (2003), using an automated analysis program they developed based on word frequencies (LIWC), identified significant gender differences,

including in personal pronoun use, in writing samples of various print genres and speech transcripts of conversations on a diverse range of topics.

Interestingly, however, little evidence has been found of gender differences on the grammatical or word level in CMC. Guiller and Durndell (2006) studied students' language use in computer-mediated discussion groups and found few gender differences in linguistic features, although stylistic gender differences were evident. Herring and Paolillo (2006) analyzed the frequency of the grammatical features identified by Argomon et al. (2003) in adult blogs, and found that when blog genre (personal diary vs. 'filter' blogs focusing on external events) was taken into account, gender differences disappeared. Herring and Paolillo suggest that it is the genre of writing, rather than the gender of the writer, that determines the extent to which lower-level grammatical features are used. They further suggest that Huffaker and Calvert's (2005) finding of few gender differences in lexical choice in teen blogs may be due to the fact that the latter's data were all of the same genre, diary blogs, in which teens of both genders wrote about the same topics: the writers' thoughts, likes and dislikes, love lives, and daily routines. Koch et al. (2005) also found few gender differences in their experimental study of gender construction in chat groups, where all the undergraduate subjects were engaged in discussing the same topic.

On the face of it, these findings might be taken to support Rodino's (1997) and Danet's (1998) claims that traditional gender binaries are breaking down in CMC, were it not for the fact that research has repeatedly found evidence of gender differences in CMC at the discourse and stylistic levels. Herring (1992, 1993, 1994, 2003) found that in academic discussion groups women tend to apologize, appreciate, and thank—as well as to perceive and be upset by violations of politeness—more than men, whereas men are less likely to be concerned with politeness and more readily violate online etiquette. In Internet Relay Chat rooms, Herring (2003) observed that females tend to type representations of laughter and smiling more, while males tend to use more profanity, more sexual language, and to be more aggressive. Cherny (1994) found similar patterns in a social MOO: Females performed more affectionate textual acts such as 'hugs,' while males performed more violent acts such as 'kills.' Analogously, in asynchronous discussion groups, Guiller and Durndell (2006) found that although male and female students were similar in their use of individual linguistic variables (with the exception of intensifiers, which more females used than males), significant gender differences were evident in their use of many stylistic variables: Males were more likely to use authoritative language and to respond negatively in interactions, while females were more likely to agree explicitly, support others, and make more personal and emotional contributions. Similarly, in an analysis of positive and negative message tone on MySpace profiles, Thelwall, Wilkinson, and Uppal (2010) found that female messages had a positive tone significantly more often than did male messages.

Thelwall et al.'s study is one of a relatively small number of studies to examine gender differences in social network sites. Sites such as Facebook are rapidly gaining in popularity, including among teens, and many have incorporated synchronous chat to enhance interaction. A common characteristic of such sites is that users can create profiles and upload pictures of themselves (Ellison & boyd, 2007). A few studies have analyzed users' visual self-representations in multimodal CMC environments. Siibak (2009) studied Estonian teenagers' motivation for profile picture choice in social network sites and found that female users had a higher tendency to base their choice on looking good, whereas the motivations of males were

more varied. Their results are consistent with the findings of Ellison, Heino, and Gibbs (2006) for online dating sites, which found that in written self-presentations women tended to heighten desirability by describing an idealized version of themselves. In her analysis of self-presentation on online dating sites, Whitty (2008) also found a tendency for users to present an idealized version of themselves through “glamour shots;” women more than men presented themselves with images that emphasized their attractiveness. Relatedly, Toma, Hancock, and Ellison (2008) found that users of online dating sites reported the photograph as being the least accurate feature of the user profile, implying that the pictures were edited to appear more attractive.

Except for Siibak (2009), most studies of visual self-presentation have not focused on young people. Another exception is Scheidt (2004), who analyzed visual self-presentation in a moderated graphical teen chat environment. She found that female avatars tended overwhelmingly to be seductively posed and partially clad, while male avatars tended to be fully covered, including hiding their eyes—which tended to look down or away from the viewer—under hair or caps. In their study of photographic self-representations of college students on MySpace, Manago, Graham, Greenfield, and Salimkhan (2008) also found a pervasiveness of sexualized female self-presentations. However, a new finding was that males in Manago et al.’s MySpace corpus also tended to present themselves in attractive, sexualized ways.

### **Research Question and Hypotheses**

The overarching question that guides this study is: To what extent, and in what ways, are gender differences manifested in contemporary teen chat sites? We address this question by analyzing four levels of communication: micro-linguistic, discourse-pragmatic, stylistic, and visual. Based on the literature surveyed in the previous section, we posit the following hypotheses:

- H1: Few, if any, gender differences will be found on the level of individual word choice (micro-linguistic level) in teen chatrooms.
- H2: Gender differences will be found on the discourse-pragmatic level. Specifically:
  - H2a. Boys will use language that is more assertive, resolute, and active.
  - H2b. Girls will use language that is more passive, cooperative, and accommodating.
- H3: Stylistic gender differences will be found. Specifically:
  - H3a. Boys’ communication will tend to adopt a more flirtatious and overtly sexual tone.
  - H3b. Girls’ communication will tend to adopt a more positive, friendly tone.
- H4: Gender differences will be found in self-representation in profile photographs. Specifically:
  - H4a. Girls more often than boys will be shown in suggestive clothing or undress.
  - H4b. Girls more often than boys will present seductive behavior.

The first hypothesis is based on the findings of Huffaker and Calvert (2005), Guiller and Durdell (2006), and Herring and Paolillo (2006); the second on findings by Herring (1992, 1993, 2003) and others; the third on findings reported by Cherny (1994), Herring (2003), and Thelwall et al. (2010); and the hypotheses about visual self-representation are based on the findings of Scheidt (2004) and Siibak (2009). These hypotheses do not posit that a change has occurred over time (e.g., towards fewer or less traditional gender distinctions). Rather, as a

heuristic, we adopt the conservative assumption that what has been found in previous studies of gender and CMC is likely still to be true for teen chat. However, the hypotheses are subject to disconfirmation, in which case, that assumption would be called into question.

### Data Sample

The sites for analysis were chosen by first attempting to compile an exhaustive sample of English-language teen chat sites from multiple sources: a Google search for the key phrase “teenage chat,” two chat directories, and by following links from sites identified by the preceding searches as teenage chat sites. The 27 sites identified in this manner were subsequently narrowed down to a judgment sample based on popularity. Site popularity was defined as user frequency (individual visits per month) and identified using the website rank page [www.quantcast.com](http://www.quantcast.com). The five highest-ranked sites in the sample were selected for analysis. None of these sites specified the sexual orientation of its target population of users (e.g., they were not ‘gay’ or ‘LGBT’ forums); in the absence of such specification, the sites were assumed to be heterosexual by default. This assumption was supported by the active presence of heterosexual flirtatious interactions, and the absence of same-sex flirtatious interaction, that we observed on the sites. Three of the sites seemed to invite flirtation through the names of the chat rooms they made available (e.g., ‘singles’), whereas the other two sites seemed to be devoted to general chatting.

A sample of one hour of chat was collected from all five sites simultaneously in January 2010, and the first 200 messages from each were selected for analysis, excluding overt advertisements (e.g., “Are You Single? Meet cute singles now!”), for a total of 1,000 messages. All messages were coded for the demographic variable *gender* (male/female/NA). Gender was identified by analyzing userIDs, as well as the thematic content of messages. Messages from the userID “MizSweetGirl,” for example, were coded as female, those from “RandomBoy15” as male, and a message with the content “Bored 18/m/PA here with PICS, any one wanna hit me up?????????” from the user “bonanza2142” was also coded as male. Messages from users with non-gender-distinct IDs such as “dsds” were coded as NA (gender information ‘not available’) if further examination of their content did not reveal the gender of the author. Overall, 614 messages in the corpus were from male users, 339 were from female users, and the sender’s gender could not be identified for 47 messages.

Three out of the five originally selected sites contained either no images apart from the site logo or had only graphical images (ranging in number from 1 to 16), while two of the chat sites supported profiles where users could upload profile pictures. For the purpose of this study, profile images from one of these sites were analyzed. Images from the other site containing profile images were considered for analysis but discarded, because the site crops images uploaded by users, such that a full analysis of those images was not possible.

The selected site has a search feature that randomly retrieves user profiles according to specified search criteria. For this study, two searches were conducted, one for female and one for male users between the ages of 16 and 19. The first 100 profile images from each search were taken as the sample for analysis. The term ‘image’ in this article refers to the box in the profile interface where photographs or other graphic representations may be uploaded by the user. Images that did not contain a photograph were not excluded, since we were also interested in

knowing what proportion of users of each gender chose to represent themselves with photographs and what proportion did not. The final sample consisted of 200 profile images (100 male, 100 female).<sup>3</sup>

## **Methodology**

The data were analyzed using discourse analysis and content analysis methods on four communicative levels: micro-linguistic, discourse-pragmatic, stylistic, and visual. The methodological procedures followed for each level of analysis are described below.

### ***Micro-Linguistic Features***

The first analysis focused on word frequencies. All of the chat messages were analyzed using the free online version of the Linguistic Inquiry and Word Count (LIWC) tool developed by Pennebaker, Booth, and Francis (2001). This text analysis application, which contains a default dictionary that defines which words should be counted in the target text file, had previously been used by Pennebaker et al. (2003) to identify gender differences in language use. The online version provides automated counts of seven key linguistic features: self-references (I, me, my), social words, positive emotions, negative emotions, overall cognitive words, articles (a, an, the), and big words (> 6 letters). The application was run for all the female messages and all the male messages separately for each of the five chat sites.

### ***Discourse-Pragmatic Features***

The approach adopted to analyze discourse-pragmatic features in this study was speech act analysis, which is concerned with the intended meaning or illocutionary force of utterances (Levinson, 1983). All of the chat messages were manually coded by the authors according to a *CMC act taxonomy* developed by Herring, Das and Penumarthy (2005). The CMC act taxonomy is derived from Bach and Harnish's (1979) classic classification of speech acts combined with Francis and Hunston's (1992) classification scheme for acts used in spoken conversation, adapted to fit the medium of online communication and simplified for ease and reliability of coding. After coding all messages according to the 16 categories of the original taxonomy, the categories 'thank' and 'manage' were excluded because no messages in the data occurred in those categories, and an additional category, 'summon,' was identified as being relevant to the data and added. The resulting adapted taxonomy thus consists of the following 15 CMC acts: inquire, request, direct, invite, inform, claim, desire, elaborate, accept, reject, react, repair, apologize, greet, and summon. Each message was coded for only one act; in cases where more than one act label could have applied (e.g., reject and react), the most specific applicable label was assigned (in this example, reject).

### ***Stylistic Features***

All messages were also coded for *message tone*, using categories adapted from Herring (2007)—aggressive, friendly, and neutral. Herring's original categories were an adaptation of Hymes' (1974) notion of "key," which describes the tone or manner of speech. Using a grounded theory approach, we added three other categories that emerged from our data: mild negative, flirtatious,

and sexual, for a total of six coding categories. Each message was coded for only one tone. In cases where more than one tone could apply, the most marked tone was assigned according to the following hierarchy: sexual > aggressive > flirtatious > friendly/mild negative > neutral.

### ***Image Features***

The profile images were coded for whether or not they contained a photograph, and those with photographs were further coded for three variables, following established methods of visual content analysis. The photographs were first coded for the variable *social distance*, originally identified by Hall (1966) as part of his ideas on “proxemics,” the study of how people use and perceive space. Kress and van Leeuwen (1996) applied Hall’s categories of individuals’ perception of the space around them as intimate, social, or public to the perceived social distance of figures in an image from the perspective of the viewer of that image. They defined six values: intimate (head only), close personal (head and shoulders), far personal (from the chest up), close social (from the knees up), far social (entire figure visible), and public (multiple figures visible).

Next, the photographs were coded for the variable *behavior*. Building on the observations of Goffman (1979) about the differing representations, in terms of gaze and posture, of men and women in magazine advertisements, together with Kress and van Leeuwen’s (1996) observation that the gaze of a subject in an image can affect the viewer’s perception of that subject, Bell (2001) formulated the variable *behavior*, defining its values as: offer/ideal, demand/affiliation, demand/submission, and demand/seduction. To these, we added the value: other.

Last, all photographs were coded for the variable *dress*. This variable was first put forward by Soley and Reid (1988) to analyze the degree to which models in advertisements were dressed, and to argue that models were wearing less in the 1980s than they had been 20 years earlier. The variable was later adopted by Lambiase (2003) in a study of the portrayal of male and female celebrities on their official Web homepages and fan sites. The values used are: demure, suggestive, partially clad, and nude. The category NA was added for images at intimate distance (head only) in which the subject’s clothing (or lack thereof) was not visible.

The data were coded by two coders. To assess inter-rater reliability, 400 messages (all messages from two sites in the sample) were coded independently by both authors; an inter-rater agreement level of 80% was achieved for CMC acts and 81% for message tone. After the coded samples were discussed and all disagreements resolved, the remaining data were coded by the first author. Both authors coded 100 images (50 male and 50 female). For the 89.5% of images that were photographs, inter-rater agreement was 83% for the variable *social distance*, 94% for the variable *dress*, and 84% for the variable *behavior*. After the coded images were discussed and all disagreements were resolved, the remaining image data were coded by the first author.

## **Findings**

### ***Word Counts***

The Linguistic Inquiry and Word Count (LIWC) analysis of the linguistic features of 1,000 messages reveals that the chat messages are overall high in social expression, while being low in



cognitive expression, use of definite articles, and big words, compared with the measures for both formal and personal texts provided by the analysis program (Table 1). No strong differences are evident between the word usage of male and female teenagers, and where differences appear in the overall values for each gender, there is considerable variability across the chat samples (see Appendix, Table A), suggesting that the frequencies of the keywords measured by LIWC are sensitive to local contextual factors such as topic of discussion. Nonetheless, boys tend to use more self-reference words, social words, articles, and big words in most of the samples. Girls, in contrast, tend to express more negative emotions and positive emotions, although the values for emotion words across the chat samples are quite variable. These trends are italicized in Table 1.

Statistical analysis was not conducted, because the LIWC program does not favor the testing of individual messages, and messages as short as those in the sample (6.0 words on average for males, 4.8 words on average for females, and 5.5 words on average for the total chat corpus, including individuals whose gender could not be determined) do not provide enough data to allow for reliable analysis. Nor was it possible to group together all messages from each participant to provide a larger data set for LIWC, because more than 50% of participants in the sample contributed only one or two messages.

Table 1. Key linguistic features (values normalized per 1,000 words)

	Self- references (I, me, my)	Social words	Positive emotions	Negative emotions	Overall cognitive words	Articles (a, an, the)	Big words (>6 letters)
<b>Formal texts*</b>	<b>4.2</b>	<b>8.0</b>	<b>2.6</b>	<b>1.6</b>	<b>5.4</b>	<b>7.2</b>	<b>19.6</b>
<b>Personal texts*</b>	<b>11.4</b>	<b>9.5</b>	<b>2.7</b>	<b>2.6</b>	<b>7.8</b>	<b>5</b>	<b>13.1</b>
Male (N=614)	7.68	<i>11.86</i>	2.26	1.87	3.43	<i>3.41</i>	<i>10.73</i>
Female (N=339)	5.63	10.08	<i>3.1</i>	<i>3.0</i>	4.19	3.26	9.56

\* reference provided by LIWC

### *CMC Acts*

The CMC act analysis reveals that ‘inviting’ was the most frequent act in the chat messages overall, followed by ‘claiming’ and ‘reacting’ (see Table 2). All messages were normalized by individual user, to avoid users posting a large number of messages being over-represented in the sample. Non-parametric Mann Whitney tests were conducted for each individual act. The acts ‘react’ and ‘request’ (both favored by females) achieved statistical significance, with values of  $p=.033$  and  $p=.042$ , respectively.

Chatroom communication takes place in near-real time, and when a large number of users are present, messages scroll by quickly. Individual identity tends not to be salient in this context, especially when users are represented by pseudonymous nicknames, as was the case for the chatrooms analyzed in this study. Therefore all messages were further analyzed distinguishing between ‘messages posted by males’ and ‘messages posted by females.’ Chi squares were calculated for each act individually, and a high significance was achieved for ‘invite,’ an act favored by males:  $\chi^2(1, N=952) = 14.43, p < .001$ . Females, in contrast, tended to favor ‘react’:  $\chi^2(1, N=952) = 3.71, p=.054$ .

Because the numbers for some individual acts are small, overall gender patterns emerge more clearly when functionally-related acts are grouped together into act categories. These patterns are represented graphically in Figure 1.

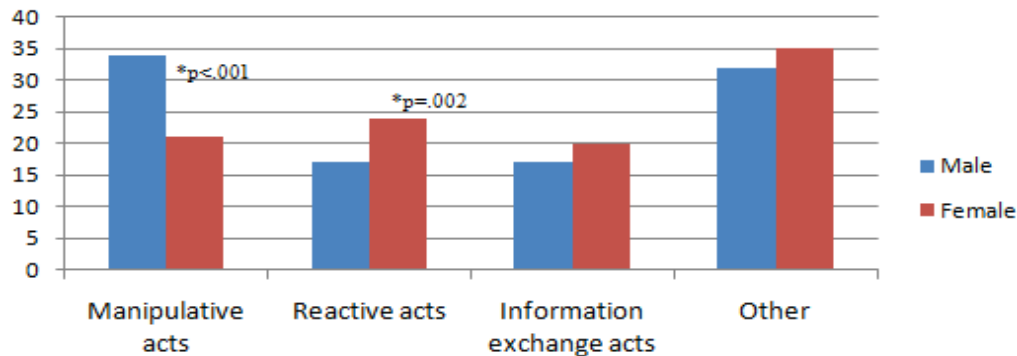


Figure 1. CMC act categories (results normalized as percentages per gender)

These groupings show that males use more manipulative acts, while females use more reactive acts and tend to use more acts that contribute to information exchange. Chi squares calculated for each of the categories showed a significant gender difference in use of manipulative acts,  $\chi^2(1, N=952) = 16.54, p < .001$ , and in reactive acts,  $\chi^2(1, N=952) = 9.61, p = .002$ .

Examples of act categories used more by one gender than the other are given below:

Invite (manipulative): “Any hot chicks want to get dirty a hot 15/m?? msg mee hun=)” (M)

React (reactive): “wow” ; “ugh” ; “wat!” ; “Imao” ; “omgggg”<sup>4</sup> (F)

Inquire (information exchange): “marionette u not a mod, r u ?” (F)

The frequency of individual CMC acts varied considerably across the five chat sites (see Appendix, Table B), suggesting that like word choice, choice of CMC act is locally conditioned to some extent by the context and topic of discussion.

### ***Message Tone***

The content analysis results for message tone also show significant gender differences. Again, all messages were normalized by individual user, to avoid users posting a large number of messages being over-represented in the sample, and non-parametric Mann Whitney tests were conducted for each tone. Statistically significant differences were found for the use of flirtatious tone ( $p=.04$ ), which was used more frequently by males, and for friendly tone ( $p=.002$ ), which was adopted much more often by female users.

All messages were further analyzed using a general ‘message posted by male/message posted by female’ distinction. Similar to the previous results, the teenage boys used more aggressive and flirtatious tones, whereas the girls much more often adopted a friendly tone in their messages. Sexual message tone was used slightly more in male than in female messages (see Figure 2). Chi squares were calculated for each tone individually, and a high significance

was achieved for flirtatious tone,  $\chi^2(1, N=952) = 19.46$ ,  $p < .001$ , friendly tone,  $\chi^2(1, N=952) = 27.36$ ,  $p < .001$ , and aggressive tone,  $\chi^2(1, N=952) = 7.2$ ,  $p = .007$ .

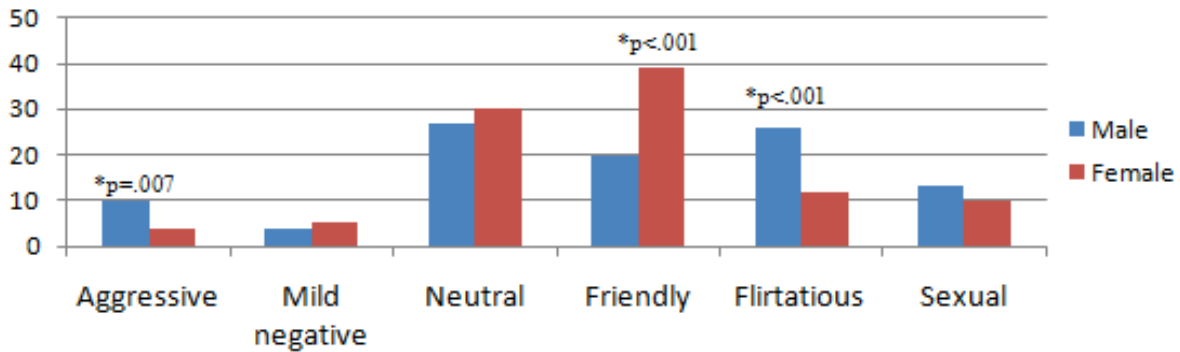


Figure 2. Message tone (results normalized as percentages per gender)

Examples of the tone categories that showed gender differences or tendencies<sup>5</sup> are given below.

Flirtatious: “any hotties want to chat??” (M)

Friendly: “how do u feel?” (F)

Aggressive: “Wha the fuccc?” (M)

Sexual: “any girls like blackk and 9 inches?” (M)

### ***Image Characteristics***

For the last part of the analysis, 200 images from a single site were analyzed. Only 21 (10.5%) of users did not have a profile photograph: One male had a graphical (cartoon) avatar, another had a graphical representation of an alien, and 12 had only the default image provided by the site, whereas two females had cartoon avatars, one had an image that was unidentifiable, and four had the default profile image. (The default image is a featureless gray shadow of the head and shoulders of a humanoid shape, not differentiated by gender.) Overall, 14% of male users chose not to represent themselves with a photograph, as compared to 7% of female users.

The relative reticence of boys to show themselves is reflected in the social distances of the pictures that each gender chose to display. Analysis of the profiles that had photographs revealed that female users were more likely to choose images of themselves at intimate (male 1%, female 11%) and close personal (male 30%, female 52%) distances. In contrast, male users preferred far personal (male 40%, female 20%) distance by a large margin. All values were scaled from 1 to 6 and analyzed using a non-parametric Mann Whitney test, which revealed a high statistical significance for females to choose representations at a closer distance to the camera and males to choose representations that are further away ( $p=001$ ). Chi squares were calculated and showed significant gender differences in representation,  $\chi^2(6, N=200) = 18.12$ ,  $p < .006$ . These results are represented graphically in Figure 3.

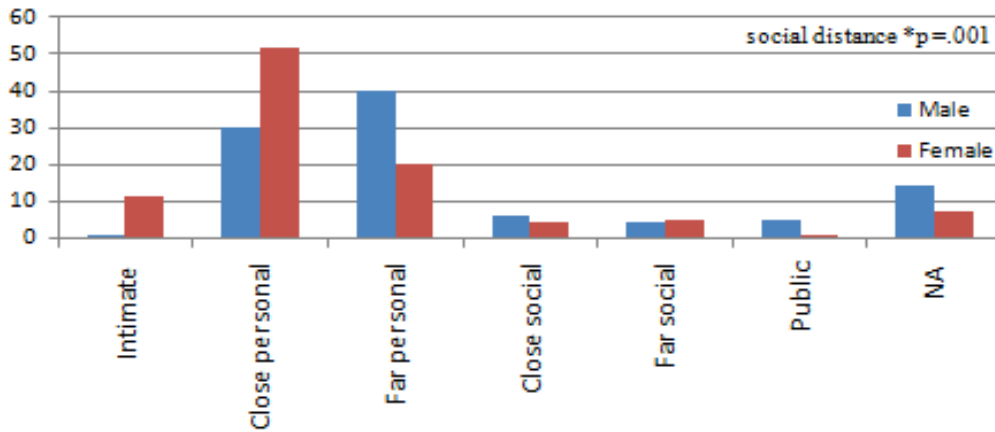


Figure 3. Social distance (results normalized as percentages per gender)

The behavior analysis also showed strongly gender-skewed results. The overwhelming majority of teenage girls (71%) chose to present themselves in photographs with seductive behavior—head tilted, body angled, eyes looking up or sideways at the viewer—, in comparison with 28% of male users who chose to present themselves that way. Males were more likely to choose photographs of themselves depicting behavior classified as offer (looking away in the distance), demand/submission (looking down at the viewer), and demand/affiliation (looking straight at the viewer). See Figure 4.

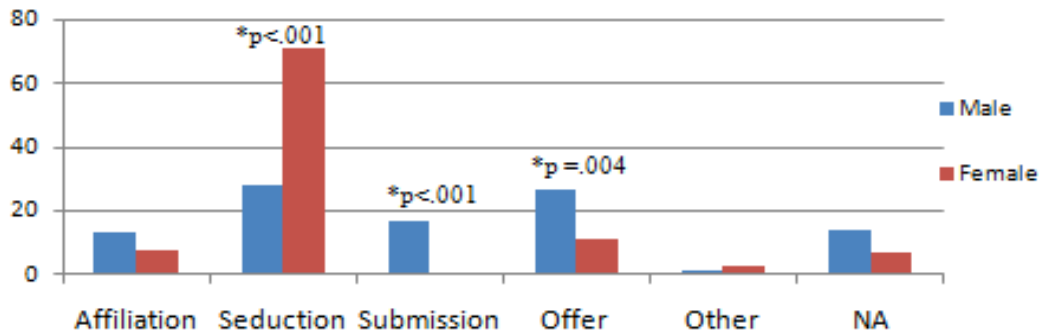


Figure 4. Behavior (results normalized as percentages per gender)

Strongly significant gender differences were found for overall behavior,  $\chi^2(5, N=200) = 46.94$ ,  $p < .001$ , as well as for specific behaviors such as seduction,  $\chi^2(1, N=200) = 36.98$ ,  $p < .001$ ; submission,  $\chi^2(1, N=200) = 18.58$ ,  $p < .001$ ; and offer,  $\chi^2(1, N=200) = 8.32$ ,  $p = .004$ .

The last set of image analysis results concerns dress. Most of the teens tended to present themselves in demure dress, especially the boys (male 66%, female 45%). However, one-third of all girls (32%) chose suggestive dress, and 8% chose photographs in which they were partially clad, as compared to 15% of male users whose pictures showed their nude upper body. Since a male showing his upper body is not socially equivalent to a female showing her breasts, the suggestive and partially clad categories may be combined to compare more accurately across genders. Doing so still shows that more girls than boys chose photos of themselves revealingly (un)dressed, although the results do not achieve significance,  $\chi^2(1, N=200) = 1.7$ ,  $p = .19$ .

However, strongly significant gender differences were found for overall dress,  $\chi^2(4, N=200)=27.59, p < .001$ , as well as for demure dress,  $\chi^2(1, N=200)=8.93, p=.003$ , and suggestive dress,  $\chi^2(1, N=200)=26.56, p < .001$ . These results are displayed in Figure 5.

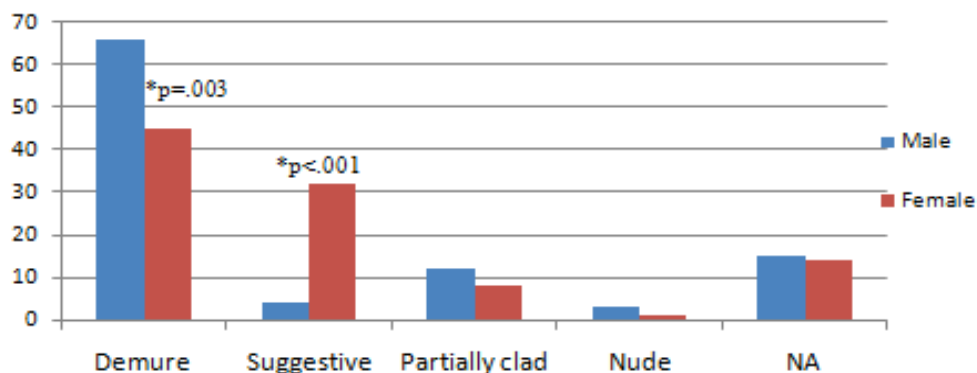
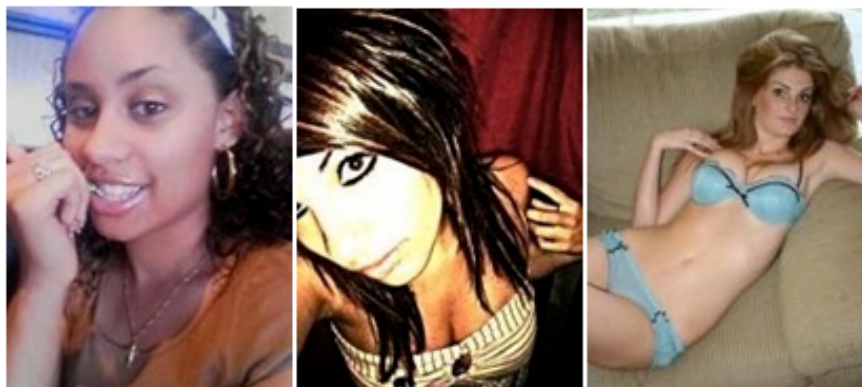


Figure 5. Dress (results normalized as percentages per gender)

The most common gender patterns found in the profile photographs are illustrated below:<sup>6</sup>

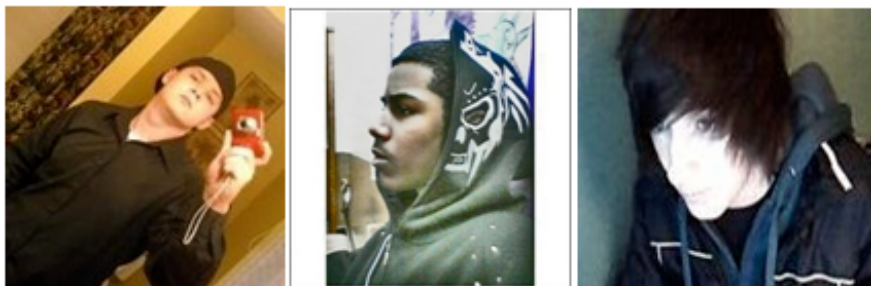


a)

b)

c)

Girls: seductive behavior with a) demure, b) suggestive, and c) revealing dress



a)

b)

c)

Boys: Demure dress with a) demand/submission, b) offer, and c) seductive behavior

## Discussion

Analysis of the teen chat sites reveals significant gender differences at three out of the four levels of communication analyzed. These findings are discussed below in relation to the research hypotheses.

We first predicted that few, if any, gender differences would be found on the level of individual word choice in the teen chat sites. In fact, some weak tendencies related to gender were observed, some of which are consistent with previous research (e.g., males using more articles; females expressing more emotion) and others of which are not (e.g., males using more 1<sup>st</sup>-person pronouns). However, the results were variable across chat samples, and the chat data as a whole differ more from the formal and personal text data provided by LIWC as points of comparison than males and females in the chat differ from each other for most word categories. Overall, then, we consider H1 to be supported.

Our second hypothesis posited that differences would be found on the discourse-pragmatic level; specifically, that boys would use more assertive and active language, while girls would use language that is more cooperative and accommodating. The CMC act analysis provided support for this prediction, in that boys were found to use more manipulative acts and girls more reactive acts. Specifically, boys used more ‘invite’ acts and girls used more ‘react’ acts in the corpus as a whole. Both genders made claims (subjective assertions) about equally, and other acts were used infrequently and showed little gender difference. The act results were also quite variable across samples, suggesting that what participants in a particular chat session happened to be talking about at the time we sampled our data conditioned the choice of verbal acts, in addition to gender. Overall, however, we consider H2 to be supported.

Our third hypothesis predicted that stylistic gender differences would be found in the chat messages; specifically, that boys would tend to adopt a more flirtatious and overtly sexual tone, while girls’ communication would be less sexual and more friendly. This hypothesis was strongly supported, with the exception of sexual messages, which were contributed by boys only slightly more often than by girls. However, boys were decidedly more flirtatious—a finding also supported by their higher use of ‘invite’ CMC acts—consistent with the gender norm of males as the initiators of heterosexual relations. Girls were decidedly less flirtatious and more friendly, although girls sent flirtatious messages, too, in keeping with the overall flirtatious tone of several of the chat rooms.

All of the hypotheses related to images were strongly supported. Significant gender differences were found in self-representation in profile photographs in both dress and behavior, with girls presenting themselves seductively in posture, gaze, and clothing. In contrast, boys varied little in their dress, but adopted a greater range of behaviors in their profile photos, include presenting themselves as remote (offer) and dominant (demand/submission). In addition, gender differences were found for the social distance of the subject from the viewer of the photographs, although we did not advance a hypothesis about social distance, due to the lack of previous research examining this variable in relation to gender. However, the fact that males more often showed themselves at a distance from the viewer, combined with the greater number of male profiles that contained no photo, can be interpreted in light of past research by Scheidt (2004),

which found that male avatars tended to look withdrawn—to be hiding, even, in their hair and clothes. Overall, the image findings in the present study conform well with Scheidt’s earlier observations about avatars in graphical teen chat rooms.

At the same time, a not insignificant number (15%) of all male profiles depicted the user with a nude upper body. This appears to support a trend identified by Manago et al. (2008), who in their study of MySpace profiles observed an “increasing pressure for men to display their physical attractiveness” (p. 455). Whether the number of male profile pictures that display partial nudity constitutes an “increase,” however, cannot be determined through comparison, because Manago et al. did not provide frequencies of different self-representations in their study, which was based on focus groups.

Finally, it is important to acknowledge the variation in results across the sample sites. Although our initial assumption in sampling teen chat sites was that the sites would show more similarities than differences, and chat site was not a dimension of variation in our hypotheses, the frequencies of words and CMC acts, especially, were observed to vary across the five chat sites in the sample (see Appendix, Tables A and B). In part, this is because of the different purposes the sites serve. Some sites seem to be devoted primarily to social chat, as evidenced by the larger variety and more even distribution of CMC acts in their chat rooms. Others tend to function as a starting point for private text or video chat and attract more flirtatious interactions. This is evidenced by the larger number of invitations issued and sexual and/or flirtatious tone used (by both genders) in two out of three such sites. However, the chat sample from one site that presents as a “hook-up” environment lacks these characteristics, and the interaction on one social chat site is quite flirtatious. Moreover, the “hook-up” sites do not exhibit stronger gender differences, contrary to what one might expect. For the three levels of analysis that we conducted on all five sites, the degree of gender contrast varies across sites, but not in a way that causes social chat sites to pattern together in contrast to flirtatious chat sites. Differences in the local topic of discourse, along with site purpose, need to be taken into account systematically in future research, particularly given the evidence from past research (e.g., Herring & Paolillo, 2006) that topic can confound gender differences in online interaction.

## Conclusions

We began by asking to what extent male and female teenagers communicate differently on contemporary, multimodal teen chat sites, and whether the extent and nature of gender differences in their communication have changed since the early findings on gender and CMC reported in the 1990s. Despite some evidence of non-traditional gendered behavior in the recent literature (e.g., females less passive: Huffaker & Calvert, 2005; females more sexually assertive: Subrahmanyam et al., 2004) or no gender differences (e.g., Gross, 2004), the findings of the present study are overall more in line with traditional findings than supportive of these non-traditional results. That is, young females in 2010 still tended to present themselves as emotional, friendly, good listeners (reactive), sexually available, and eager to please males (cf. Magnuson & Dundes, 2008), while young males appear more assertive, manipulative, initiating, and visually dominant, while at the same time more distant. Our empirical results provide some limited support for the self-reported trend in the Manago et al. (2008) study for males to present themselves in photos that emphasize their physical attractiveness, including through sexualized,

seductive images. However, the female images in our data are significantly more sexualized and seductive than the male images, overall.

While these findings are not new, *per se*, they deserve to be reported, both to update the scholarly record and to counter the ongoing tendency for people to imagine that gender differences are continuously receding with each subsequent generation. Rather, the results of this study suggest that the phenomenon of gender differentiation occurs at multiple levels of communication and is relatively stable over time and across media. For adolescents, in particular, it serves a useful social purpose: The search for partners is an important activity in the teen years (Smahel & Subrahmanyam, 2007), and symbolic gender differentiation via language and images can heighten mutual attractiveness, especially in virtual environments such as chatrooms where physical sex characteristics are not visible. More generally, teenagers awaken out of childhood to the realization that they are social and sexual creatures within a gendered society, who must learn to manifest aspects of their identity appropriately in relation to other social and sexual creatures. Eckert (1996) argues that gender identities are complementary and co-constructed within a default heterosexual marketplace, and Herring and Zelenkauskaite (2009) propose that public computer-mediated environments can function as such marketplaces. Evidence that certain gender patterns are continuously and persistently reproduced further supports the view that such patterns are socially facilitative.

At the same time, power dimensions underlie the patterns, which can be read as evidence of socialization (or performances, if one prefers) of males as dominant and in control and females as accommodating and pleasing to males (e.g., Magnuson & Dundes, 2008). These roles are continuously reinforced through the mass media, to which young people are exposed from an early age in the form of television, movies, advertising, and increasingly, the Internet itself (including online pornography). As regards visual presentation, girls, in particular, learn at a young age that looks matter and that they will be judged based on them (Frederickson et al., 1998). They learn that they should be sexually appealing. The analysis of images from the chat site studied here suggests that female users have internalized that message and apply the mass media's simplified formula—"if you've got it, flaunt it" (Durham, 2008)—to appear attractive.

We also raised questions at the outset of the article about the communicative levels at which gender differentiation takes place in teen chat rooms. While past research has reported (not) finding evidence of gender differences for different linguistic phenomena, no previous study to our knowledge systematically examined phenomena at more than two levels of language structure and/or function. By comparing communicative levels ranging from micro-linguistic (word level) to discourse-pragmatic (utterance level) to stylistic (discourse level) to visual (chat room/site level), we found support for the findings of previous research that gender differences are least in evidence at the level of individual word choice, but significantly present at the utterance and discourse levels, as well as in self-presentation style in images. Moreover, the variation across chat samples in our data suggests a reason for this discrepancy: Lower-level language features, including word choice and some speech acts, appear to be sensitive to the local topic of conversation and to vary accordingly, making them less available as signals of (global) gender identity. This is not to imply that gender does not shape word and speech act choice at all; Huffaker and Calvert (2005) found gender differences in word frequencies in teen blogs related to 'resoluteness' and 'activity,' both favoring males, and our findings include



gender differences in the acts of inviting and reacting, related to flirtatious (male initiating, female responding) heterosexual interactions. A tentative generalization might be that gender signaling on lower linguistic levels depends on the extent to which the interactions involve gender-relevant activities. (Thus, for example, we would not expect gender differences in word choice in online academic discussions, consistent with Guiller & Durdell, 2006).

Finally, we asked what effects changing technological affordances are having on gender expression in multimedia, convergent media sites. Specifically, since anonymity was claimed in previous research to encourage the breakdown of traditional gendered forms of expression (e.g., Danet, 1998), we wondered whether the availability of photographs of users would make gender more salient. Profile pictures are a relatively new feature on chat sites, but they have been embraced by users; 89.5% of all randomly selected profiles in the present study contained a photograph of the user. While comparative historical data are lacking to determine relative degrees of salience, it is certainly the case that gender is visually salient—even emphasized—in the present study through choice of profile pictures. We found that gender identity is signaled not only through a photo-realistic depiction of a male or female person, but through stance, eye gaze, dress, and proximity to the camera, all aspects of visual presentation that users control, if only by selecting some photos over others to use as profile images. These aspects of photographic representation have as yet barely been studied for online communication; our study suggests that they are becoming increasingly important.

A limitation of this study is that the data were collected by means of cross-sectional, rather than longitudinal, sampling. A sample obtained longitudinally from a single site would provide direct evidence, rather than inferential evidence as in the present study, that change over time is (or is not) taking place. Moreover, this study was not designed to test the effect of individual chat site on any of the communicative features. Future research should be structured to control for chat site as a potentially predictive variable, for example by collecting larger samples of data and conducting multivariate statistical analyses.

The word frequency analysis was limited by the shortness of the chat messages and the fact that many participants posted only one or two messages; LIWC works best on longer textual units. An analysis of a larger corpus in which all messages by active individuals were grouped together and analyzed statistically, taking topic of discourse into account, would provide a more precise indication of the presence or absence of gender differences in word usage. Further, the profile images were obtained from a single chat site, due to limitations of our sample, which contained only one chat site with uncropped images. A sample of images from different sites would produce more generalizable results and might show cross-site variation that could yield insight. Finally, although the findings of the present study resonate with a large body of previous gender and CMC research, which lends them *prima facie* validity, other sites where young people communicate spontaneously online (including sites where the communication is not flirtatious, although these may be difficult to find) should be analyzed for gender, to determine the extent to which the findings of the present study are characteristic of teens' online communication patterns in general.

These limitations notwithstanding, this study's findings of gender differences in young users similar to those reported in CMC research 20 years ago, and indeed in spoken language and

gender research before that, should give researchers pause. They demonstrate that CMC technologies do not make gender identity irrelevant; if anything, multimodal environments tend to make gender more salient, especially when they include graphical representations of users, as in the case of virtual worlds, social network sites, and other Web 2.0 platforms. Nor does gender appear to be receding in importance in young people's communication. Clearly it is up front and center in flirtatious interactions, which are common among teens, but it also manifests in terms of assertiveness, expression of emotion, politeness, etc. in other types of engagement, including online academic discussions (Guiller & Durndell, 2006). The claim that gender identity is deeply embedded in the human psyche and expressed largely unconsciously in human communication is not new. We recommend that future scholarship examine how the gender identities internalized by social actors work for or limit those actors in different communicative settings.

## References

- Argamon, S., Koppel, M., Fine, J., & Shimoni, A. R. (2003). Gender, genre, and writing style in formal written texts. *Text*, 23(3), 321-347.
- Bach, K., & Hamish, R. M. (1979). *Linguistic communication and speech acts*. Cambridge, MA: MIT Press.
- Bell, P. (2001). Content analysis of visual images. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 10-34). London: Sage.
- boyd, d., (2008). Why youth heart social network sites: The role of networked publics in teenage social life. In D. Buckingham (Ed.), *Youth, identity and digital media* (pp. 119-142). Cambridge, MA: MIT Press.
- Butler, J. (1990). Performative acts and gender constitution: An essay in phenomenology and feminist theory. In S. Case (Ed.), *Performing feminisms: Feminist critical theory and theatre* (pp. 270-82). Baltimore, MD: The Johns Hopkins University Press.
- Cherny, L. (1994). Gender differences in text-based virtual reality. In M. Bucholtz, A. Liang, & L. Sutton (Eds.), *Cultural Performances: Proceedings of the Third Berkeley Women and Language Conference*. Berkeley, CA: Berkeley Women and Language Group.
- Coates, J. (1993). *Women, men and language*. London: Longman.
- Danet, B. (1998). Text as mask: Gender and identity on the Internet. In S. G. Jones (Ed.), *Cybersociety 2.0* (pp. 129-158). Thousand Oaks, CA: Sage.
- Durham, G. M. (2008). *The Lolita effect: The media sexualization of young girls and what we can do about it*. Woodstock & New York, NY: The Overlook Press.
- Goffman, E. (1979). *Gender advertisements*. London: Macmillan.
- Gross, E. F. (2004). Adolescent Internet use: What we expect, what teens report. *Applied Developmental Psychology*, 25, 633-649.
- Guiller, J., & Durndell, A. (2007). Students' linguistic behavior in online discussion groups: Does gender matter? *Computers in Human Behavior*, 23(5), 2240-2255.
- Eckert, P. (1996). Vowels and nail polish: The emergence of linguistic style in the preadolescent heterosexual marketplace. In N. Warner, J. Ahlers, L. Bilmes, M. Oliver, S. Wertheim, & M. Chen (Eds.), *Gender and Belief Systems: Proceedings of the Fourth Berkeley Women & Language Conference* (pp. 183-190). Berkeley, CA: Berkeley Women & Language Group.

- Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication*, 11(2), article 2. Retrieved February 17, 2011 from <http://jcmc.indiana.edu/vol11/issue2/ellison.html>
- Erikson, E. H. (1968). *Identity: Youth and crisis*. Oxford, UK: Norton.
- Francis, G., & Hunston, S. (1992). Analysing everyday conversation. In M. Coulthard (Ed.), *Advances in spoken discourse analysis* (pp. 1-34). London: Routledge.
- Frederickson, B. L., Roberts, T. A., Noll, S. M., Quinn, D. M. & Twenge, J. M. (1998). That swimsuit becomes you: Sex differences in self-objectification, restrained eating, and math performance. *Journal of Personality and Social Psychology*, 75(1), 269-284.
- Hall, E. (1966). *The hidden dimension*. New York: Doubleday.
- Herring, S. C. (1992). *Gender and participation in computer-mediated linguistic discourse*. Washington, D.C.: ERIC Clearinghouse on Languages and Linguistics, doc ED345552. Retrieved March 17, 2010 from <http://ella.slis.indiana.edu/~herring/participation.1992.pdf>
- Herring, S. C. (1993). Gender and democracy in computer-mediated communication. *Electronic Journal of Communication*, 3(2). Retrieved March 17, 2010 from <http://ella.slis.indiana.edu/~herring/ejc.txt>
- Herring, S. C. (1994). Politeness in computer culture: Why women thank and men flame. In: *Cultural Performances: Proceedings of the Third Berkeley Women and Language Conference* (pp. 278-294). Berkeley, CA: Berkeley Women and Language Group.
- Herring, S. C. (2003). Gender and power in online communication. In J. Holmes & M. Meyerhoff (Eds.), *The handbook of language and gender* (pp. 202-228). Oxford: Blackwell.
- Herring, S. C. (2007). A faceted classification scheme for computer-mediated discourse. *Language@Internet*, 4, article 1. Retrieved March 17, 2010 from <http://www.languageatinternet.de/articles/2007/761>
- Herring, S. C., & Paolillo, J. C., (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10(4), 439-459.
- Herring, S. C., & Zelenkauskaite, A. (2009). Symbolic capital in a virtual heterosexual market: Abbreviation and insertion in Italian iTV SMS. *Written Communication*, 26(1), 5-31.
- Herring, S. C., Das, A., & Penumarthy, S. (2005). *CMC act taxonomy*. Retrieved March 17, 2010 from <http://www.slis.indiana.edu/faculty/herring/cmc.acts.html>
- Huffaker, D. A., & Calvert, S. L. (2005). Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication*, 10(2). Retrieved March 17, 2010 from <http://jcmc.indiana.edu/vol10/issue2/huffaker.html>
- Hymes, D. (1974). *Foundations of sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press.
- Koch, S. C., Mueller, B., Kruse, L., & Zumbach, J. (2005). Constructing gender in chat groups. *Sex Roles*, 53(1/2), 29-41.
- Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. New York: Routledge.
- Lakoff, R. T. (1975). *Language and woman's place*. NY: Harper & Row.

- Lambiase, J. (2003). Codes of online sexuality: Celebrity, gender and marketing on the Web. *Sexuality & Culture*, 7(3), 57-78.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge, UK: Cambridge University Press.
- Magnuson, M. J., & Dundes, L. (2008). Gender differences in “social portraits” reflected in MySpace Profiles. *CyberPsychology & Behavior*, 11(2), 239-241.
- Manago, A. M., Graham, M. B., Greenfield, P. M., & Salimkhan, G. (2008). Self-presentation and gender on MySpace. *Journal of Applied Developmental Psychology*, 29(6), 446-458.
- McRae, S. (1996). Coming apart at the seams: Sex, text and the virtual body. In L. Cherny & E. Weise (Eds.), *Wired\_Women* (pp.242-263). Seattle: Seal Press.
- Paolillo, J. C., & Zelenkauskaitė, A. (In press). Real-time chat. In S. C. Herring, D. Stein, & T. Virtanen (Eds.), *The handbook of pragmatics of computer-mediated communication*. Berlin: Mouton.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count (LIWC): LIWC 2001*. Mahwah, NJ: Erlbaum.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54(1), 547-577.
- Pew Internet and American Life Project (2007). Teens and social media. *Pew Internet and American Life Project*. Retrieved March 16, 2010 from [http://www.pewinternet.org/~media/Files/Reports/2007/PIP\\_Teens\\_Social\\_Media\\_Final.pdf](http://www.pewinternet.org/~media/Files/Reports/2007/PIP_Teens_Social_Media_Final.pdf)
- Rodino, M. (1997). Breaking out of binaries: Reconceptualizing gender and its relationship to language in computer-mediated communication. *Journal of Computer-Mediated Communication*, 3(3). Retrieved March 17, 2010 from <http://jcmc.indiana.edu/vol3/issue3/rodino.html>
- Scheidt, L. A. (2004, May). *Buxom girls and boys in baseball hats: Adolescent avatars in graphical chat spaces*. Paper presented at the 54th International Communication Association Conference, New Orleans, LA. Retrieved March 17, 2010 from [http://loisscheidt.com/working\\_papers\\_archive/Buxom\\_girls\\_and\\_boys\\_in\\_baseball\\_hats.pdf](http://loisscheidt.com/working_papers_archive/Buxom_girls_and_boys_in_baseball_hats.pdf)
- Selfe, C. L., & Meyer, P. R. (1991). Testing claims for on-line conferences. *Written Communication*, 8(2), 163-192.
- Siibak, A. (2009). Constructing the self through the photo selection - Visual impression management on social networking websites. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 3(1), article 1. Retrieved October 15, 2010 from <http://cyberpsychology.eu/view.php?cisloclanku=2009061501&article=1>
- Smahel, D., & Subrahmanyam, K. (2007). “Any girls want to chat press 911:” Partner selection in monitored and unmonitored teen chat rooms, *CyberPsychology & Behavior*, 10(3), 346-353.
- Soley, L. C., & Reid, L. N. (1988). Taking it off: Are models in magazine ads wearing less? *Journalism and Mass Communication Quarterly*, 65(4), 960-966.
- Subrahmanyam, K., Greenfield, P. M., & Tynes, B. (2004). Constructing sexuality and identity in an online teen chat room. *Journal of Applied Developmental Psychology*, 25(6), 651-666.

- Subrahmanyam, K., Smahel D., & Greenfield, P. M. (2006). Connecting developmental constructions to the Internet: Identity presentation and sexual exploration in online teen chat rooms. *Developmental Psychology, 42*(3), 395-406.
- Tannen, D. (1994). *Gender and discourse*. New York: Oxford University Press.
- Thelwall, M., Wilkinson, D. & Uppal, S. (2010). Data mining emotion in social network communication: Gender differences in MySpace. *Journal of the American Society for Information Science and Technology, 61*(1), 190-199.
- Thomson, R., & Murachver, T. (2001). Predicting gender from electronic discourse. *British Journal of Social Psychology, 40*, 193–208.
- Toma, C. L., Hancock, J. T., & Ellison, N. B. (2008). Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles. *Personality and Social Psychology Bulletin, 34*(8), 1023-1036.
- Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles, 36*, 305–325.
- Whitty, M. T. (2008). Revealing the 'real' me, searching for the 'actual' you: Presentations of self on an internet dating site. *Computers in Human Behavior, 24*, 1707-1723.

## Appendix

Table A. Key linguistic features (LIWC)

	Self - references (I, me, my)	Social words	Positive emotions	Negative emotions	Overall cognitive words	Articles (a, an, the)	Big words (>6 letters)
<b>Formal texts*</b>	<b>4.2</b>	<b>8.0</b>	<b>2.6</b>	<b>1.6</b>	<b>5.4</b>	<b>7.2</b>	<b>19.6</b>
<b>Personal texts*</b>	<b>11.4</b>	<b>9.5</b>	<b>2.7</b>	<b>2.6</b>	<b>7.8</b>	<b>5</b>	<b>13.1</b>
<b>chat site A</b>							
male (89)	5.09	12.33	2.35	2.15	4.89	5.09	10.76
female (77)	2.58	10.08	1.55	5.17	4.39	2.58	7.75
<b>chat site B</b>							
male (147)	8.51	12.08	1.5	1.41	2.77	3.2	11.75
female (48)	7.62	11.71	3.95	1.97	4.8	3.81	11.28
<b>chat site C</b>							
male (133)	6.91	10.36	2.51	2.83	2.83	2.51	10.2
female (61)	3.56	10.32	3.91	2.14	3.2	2.49	13.17
<b>chat site D</b>							
male(115)	5.11	12.88	3.27	2.45	4.09	3.89	8.38
female (83)	6.44	8.28	0.92	4.29	2.76	3.68	7.06
<b>chat site E</b>							
male (129)	11	11.25	4.5	1.75	5.25	3.25	9
female (70)	6.03	7.33	5.17	1.72	5.17	3.02	6.47
<b>total</b>							
male (614)	7.68	11.86	2.26	1.87	3.43	3.41	10.73
female (339)	5.63	10.08	3.1	3	4.19	3.26	9.56

\* reference provided by LIWC

Table B. CMC acts

	<b>inquire</b>	<b>request</b>	<b>invite</b>	<b>claim</b>	<b>elaborate</b>	<b>reject</b>	<b>react</b>	<b>direct</b>	<b>repair</b>	<b>apologize</b>	<b>summon</b>	<b>greet</b>	<b>inform</b>	<b>desire</b>	<b>accept</b>
<b>chat site A</b>															
Male (89)	8 (9%)	0 (0%)	14 (16%)	18 (20%)	3 (4%)	1 (1%)	11 (12%)	2 (2%)	2 (2%)	0 (0%)	0 (0%)	17 (19%)	12 (14%)	1 (1%)	0 (0%)
Female (77)	9 (12%)	0 (0%)	12 (15%)	12 (15%)	3 (4%)	3 (4%)	14 (18%)	6 (8%)	0 (0%)	2 (3%)	1 (1%)	6 (8%)	9 (12%)	0 (0%)	0 (0%)
NA (34)	6 (18%)	0 (0%)	0 (0%)	6 (18%)	2 (6%)	3 (9%)	7 (21%)	3 (9%)	0 (0%)	0 (0%)	0 (0%)	2 (6%)	3 (9%)	0 (0%)	1 (3%)
<b>chat site B</b>															
Male (147)	18 (12%)	0 (0%)	101 (70%)	4 (3%)	0 (0%)	0 (0%)	3 (2%)	14 (10%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (1%)	1 (1%)	1 (1%)
Female (48)	2 (4%)	1 (2%)	27 (56%)	6 (13%)	0 (0%)	2 (4%)	3 (6%)	2 (4%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (9%)	0 (0%)	1 (2%)
NA (5)	0 (0%)	0 (0%)	0 (0%)	3 (60%)	1 (20%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (20%)	0 (0%)
<b>chat site C</b>															
Male (133)	6 (4%)	0 (0%)	39 (29%)	25 (19%)	13 (10%)	5 (4%)	17 (13%)	8 (6%)	0 (0%)	1 (1%)	2 (1%)	5 (4%)	11 (8%)	1 (1%)	0 (0%)
Female (61)	10 (17%)	0 (0%)	13 (22%)	7 (12%)	0 (0%)	3 (5%)	9 (15%)	4 (7%)	0 (0%)	0 (0%)	1 (2%)	5 (8%)	6 (10%)	0 (0%)	1 (2%)
NA (6)	2 (33%)	0 (0%)	0 (0%)	1 (17%)	0 (0%)	0 (0%)	0 (0%)	2 (33%)	0 (0%)	0 (0%)	0 (0%)	1 (17%)	0 (0%)	0 (0%)	0 (0%)
<b>chat site D</b>															
Male (115)	21 (18%)	0 (0%)	9 (8%)	32 (28%)	7 (6%)	1 (1%)	16 (14%)	7 (6%)	1 (1%)	0 (0%)	4 (3%)	3 (3%)	9 (8%)	0 (0%)	5 (4%)
Female (83)	4 (5%)	0 (0%)	3 (4%)	18 (21%)	18 (21%)	5 (6%)	17 (20%)	2 (2%)	0 (0%)	0 (0%)	3 (4%)	2 (2%)	7 (8%)	1 (1%)	3 (4%)
NA (2)	0 (0%)	0 (0%)	1 (50%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (50%)	0 (0%)	0 (0%)	0 (0%)
<b>chat site E</b>															
Male (129)	7 (6%)	0 (0%)	3 (2%)	25 (19%)	12 (9%)	8 (6%)	30 (23%)	11 (9%)	0 (0%)	0 (0%)	5 (4%)	10 (8%)	10 (8%)	5 (4%)	3 (2%)
Female (70)	16 (24%)	2 (3%)	0 (0%)	14 (21%)	4 (6%)	1 (1%)	15 (22%)	3 (4%)	1 (1%)	0 (0%)	2 (3%)	6 (9%)	0 (0%)	0 (0%)	4 (6%)
NA (1)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	0 (0%)	0 (0%)
<b>Total</b>															
Male (614)	60 (10%)	0 (0%)	166 (27%)	104 (17%)	35 (6%)	15 (2%)	77 (13%)	42 (7%)	3 (1%)	1 (1%)	11 (2%)	35 (6%)	44 (7%)	8 (1%)	9 (1%)
Female (339)	41 (12%)	3 (1%)	55 (16%)	57 (17%)	25 (8%)	14 (4%)	58 (17%)	17 (5%)	1 (5%)	2 (1%)	7 (2%)	19 (6%)	26 (8%)	1 (5%)	9 (3%)
NA (47)	8 (17%)	0 (0%)	1 (2%)	10 (21%)	3 (6%)	3 (6%)	7 (15%)	5 (11%)	0 (0%)	0 (0%)	0 (0%)	5 (11%)	3 (6%)	1 (2%)	1 (2%)

## Notes

- <sup>1</sup> A MUD (Multi-User Dungeon or Multi-User Dimension) is a text-based virtual reality environment in which communication is via synchronous chat.
- <sup>2</sup> This assumes that the graphical representations and nicknames accurately represent the users' real-life genders. For discussion of this methodological point, see, e.g., Herring (2003).
- <sup>3</sup> Permission was obtained from the authors' institutional review board to collect and analyze these data, which are publicly available on the Web. Nonetheless, to protect the privacy of the subjects, no names of chat sites or users are mentioned in this article.
- <sup>4</sup> "lmao" = laughing my ass off; "omggg" = oh, my god ('g's repeated for emphasis).

- <sup>5</sup> The other two tone categories—mild negative (e.g., “damn i have only one...” [F]) and neutral (“so I am rich” [M])—did not differ in use appreciably by gender in the corpus.
- <sup>6</sup> Only images from subjects whose profiles state that they are 18 or 19 years old are used in these examples. The photos are presented here without any link to the users’ profiles. Although we do not believe they could be traced, as a check, we searched for the profiles of these subjects one year after data collection and found none of the same photos; all the subjects had either left the chat site and removed their profiles, or changed their profile pictures.

### **About the Authors**

Sanja Kapidzic is a Master’s student in the Department of Telecommunications at Indiana University Bloomington. Her research interests are computer-mediated communication, mediated social interaction, and identity construction and self-presentation in online environments.

Address: Department of Telecommunications, 1229 East 7th St., Bloomington, IN 47405. Email: [sankapid@indiana.edu](mailto:sankapid@indiana.edu).

Susan C. Herring is a Professor of Information Science and Adjunct Professor of Linguistics at Indiana University Bloomington. Her primary research interests are computer-mediated communication and computer-mediated discourse, especially issues of gender, genre, interaction management, methodology, multimodality, and change over time. A past editor of the *Journal of Computer-Mediated Communication*, she currently edits the online journal *Language@Internet*.

Address: SLIS, Wells Library LI 011, 1320 E. 10<sup>th</sup> Street, Bloomington, IN 47405. Email: [herring@indiana.edu](mailto:herring@indiana.edu).