

Learning the Lingo? Gender, Prestige and Linguistic Adaptation in Review Communities

Libby Hemphill and Jahna Otterbacher*

Illinois Institute of Technology

3301 S Dearborn St, Chicago, IL 60616

{libby.hemphill, jahna.otterbacher}@iit.edu

*The authors contributed equally to this work.

ABSTRACT

Women and men communicate differently in both face-to-face and computer-mediated environments. We study linguistic patterns considered gendered in reviews contributed to the Internet Movie Database. IMDb has been described as a male-majority community, in which females contribute fewer reviews and enjoy less prestige than males. Analyzing reviews posted by prolific males and females, we hypothesize that females adjust their communication styles to be in sync with their male counterparts. We find evidence that while certain characteristics of “female language” persevere over time (e.g., frequent use of pronouns) others (e.g., hedging) decrease with time. Surprisingly, we also find that males often increase their use of “female” features. Our results indicate, that even when they resemble men’s reviews linguistically, women’s reviews still enjoy less prestige and smaller audiences.

Author Keywords

Gender; linguistic entrainment; online community; social voting.

ACM Classification Keywords

H5.3. Information interfaces and presentation (e.g., HCI): Asynchronous interaction, Web-based interaction.

General Terms

Human Factors.

INTRODUCTION

Much public discussion surrounds issues of gender and participation in online communities (OCs). For example, a March 2010 survey found that less than 15% of Wikipedia participants are women [11]. More recently, a *Room for Debate* session at NYTimes.com [34] questioned why, given that Wikipedia is meant to be the world’s “free encyclopedia that anyone can edit,” there is such a stark difference in participation between men and women. It is often claimed that men and women communicate and

behave differently in their online interactions (e.g., [19,27,30]). Possible explanations for the Wikipedia findings included that women are too busy to contribute or that they are intimidated by the discourse [34]. Technological and social factors likely both contribute to the differences between genders in participation. We contribute to the literature on the social factors that influence participation by examining the differences between men and women’s contributions to a popular review community, the Internet Movie Database (IMDb).

IMDb has been described as a male-majority community, in terms of participation and participant prestige. In a study of IMDb reviews of 250 all-time favorite movies, Otterbacher [27] reported a median utility score of 0.67 for male-authored reviews (i.e., two thirds of people who voted found these reviews useful), and a median of zero for females. Likewise, the contributions of males received more total votes. In short, females receive less attention than males and enjoy less prestige. Given that the default filter sorts on utility, females’ reviews are effectively hidden, presenting users with a predominately male view.

Nonetheless, some women manage to be successful members of IMDb, participating actively and receiving high utility scores. Men and women communicate differently both in face-to-face and in computer-mediated settings, and our data concerning prolific participants reveals a number of differences in language use that the genders exhibit in their reviews. However, even when women resemble males linguistically, the utility of their reviews remains lower.

IMDb provides the opportunity to examine the interplay between technological features such as sorting, social aspects such as language use and gender, in the development of collaborative resources. Like Wikipedia, the reviews on IMDb are a shared resource created by many authors. While authors do not collaboratively write reviews, their behaviors can provide insights about OC engagement and participation. Earlier research demonstrated that it might not be necessary for people to engage in one-to-one communication, in order for them to feel affiliation toward the OC [5]. Our study contributes to the understanding of OCs and participation by identifying relationships between language use, gender, and the perceived value of authors’ contributions.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW’12, February 11–15, 2012, Seattle, Washington, USA.
Copyright 2012 ACM 978-1-4503-1086-4/12/02...\$10.00.

PREVIOUS WORK

In mixed-gender communities, males typically enjoy more power and prestige than females [30]. At the IMDb, participants use plain text to present their views, making the language they use crucial – it is the vehicle through which others form an impression of them [22]. Here, we review findings related to gender, language and OCs.

Gender, power and online communities

In the early days of the Web, there was much anticipation for the creation of democratic, shared information spaces, in which all would get a chance to have their views heard. Years later, it has become apparent that this was an overly optimistic outlook. In the case of gender, Herring [19] claims that there is little evidence that online, power asymmetries differ from those in the offline world.

For instance, just as in face-to-face interactions, men tend to dominate conversations in online forums and are more likely to post aggressive messages. In contrast, female participants post fewer and shorter messages, are more polite, and write in a manner that aligns themselves with others [19]. Gefen and Ridings [16] explain that such differences have to do with how men and women use OCs. While women seek and offer social support (e.g., for problem solving), men seek to achieve and maintain status by showcasing their knowledge to others. Similar findings have been echoed in a variety of domains including e-commerce forums (e.g., [3]) and fan OCs (e.g., [7]).

At IMDb, participants do not comment on one another's postings, but rather, they vote as to whether or not reviews are useful. Therefore, there is less opportunity for harassment or direct criticism of others. However, males dominate the discourse by writing more, and longer, reviews. In addition, overall, males enjoy a more prestigious status, receiving higher utility scores and having their reviews prominently displayed by default.

Gender and language use

Men and women exhibit a number of differences in their language use behavior. Overall, women use more social styles while men use more broadcast styles. For instance, women tend to adapt a style that more directly engages their audiences ("rapport talk") while men tend to report information ("report talk") [33]. In written discourse, women use a more "outer-focused" style, in which they tend to acknowledge their perceived audience [4] and use far more pronouns [2]. Similarly, men are more likely to discuss objects, while women often discuss people [9]. In face-to-face conversation, women also exhibit more use of hedging phrases (e.g., "kind of," "sort of") [25].

Earlier work found similar differences between the style and content of IMDb reviews written by men and women [27]. Male reviewers often discussed the success of a movie and its special effects, while females wrote about characters and people. In terms of style, male-authored reviews were characteristically written in third person, while females more often wrote in the first person voice. However, the

data set analyzed contained reviews written by a wide variety of reviewers, many of whom were not regular contributors. To contrast, we will analyze a set of reviews written by prolific participants at IMDb.

Language adaptation

Communication Accommodation Theory (CAT) provides a possible explanation as to how female reviewers might become more successful at IMDb. CAT holds that interlocutors adjust their linguistic patterns in order to be perceived as being more or less like their conversation partners [18]. This might result in social gains, such as fitting in with a desirable group (e.g., [12]).

While CAT was developed with face-to-face interaction in mind, such "linguistic entrainment" has been noted in computer-mediated contexts. Scissors and colleagues [31] studied pairs of participants playing a game, in which cooperation was required to win. During sessions, participants communicated only through text chat. Pairs with a high level of trust (i.e., who cooperated) exhibited more entrainment than low-trust pairs. Cassell and Tversky [8] studied an international community of young people and found evidence of both convergence and divergence in language patterns. One interesting finding was the increasing use of "we" and decreasing use of "I" over time, suggesting that participants increasingly identified with the community. With respect to gender, while initially, girls expressed emotion more often than boys, over time, girls decreased and boys increased their use of emotion words.

At IMDb, it might be that female reviewers, who want to their voices to be heard, recognize the need to adjust their communication style. In fact, Herring [20] described a "list effect." In her observations of two mixed-gender listserv discussions, she found that all participants posted in a male voice, implying that the voice of the majority had become normative. In such situations, it could be that female participants understand that the less personal, more objective "male" style is better received by others [32].

It is clear from existing research that men and male-marker language dominate online spaces [19,20]. What existing literature does not tell us is whether changes in language use correlate to an increase in the perceived value and visibility of women's contributions.

STUDY CONTEXT

At IMDb, gender is disclosed through a "male/female" review filter.¹ Participants communicate in an asynchronous and indirect manner, exchanging views by posting textual reviews and by rating movies on a scale of one to ten stars. Profile pages are simple, displaying the participant's contributions, an optional self-description and contact

¹ Though the terms "male" and "female" usually refer to one's sex, we use them interchangeably with the gender designations "man" and "woman," as done at IMDb.

	Measure(s)	Explanation
Complexity	-Word complexity: Character-to-word ratio -Sentence complexity: Word-to-sentence ratio	Some claim that women are often stronger writers than men (e.g., [14]). We use two basic measures of complexity used in automatic essay evaluation [15].
Hedging	Number of hedges, normalized by review length (words)	We use a list of 55 hedging phrases, updated from [24]. Examples include: “more or less,” “sort of,” “rather,” “in a way,” “basically.”
Inner vs. outer-focused discourse	-Pronoun rate (#PNs / words) -Proportion of PNs that are first person	We compute reviewers’ overall rate of pronoun use, as well as the proportion of pronouns used that are first person.
Vocabulary richness	Number of unique words / total words	Quantifies the diversity of words reviewers use. [27] reported that using a rich vocabulary is characteristic of female reviewers.
Table 1. Gendered language attributes studied.		

information. Although IMDb collects gender information when the user registers, it is not explicitly noted on one’s profile. When users display reviews for a movie, they may customize the display by various attributes (e.g., reviewer gender, prolific authors). The default filter is “Best,” which orders reviews by utility (i.e., binary social votes collected on the question, “Was the above review useful to you?”)

Data Set and Hypotheses

We began with a list of 250 top-rated movies. For each, we used the “male/female” filter. This returned all reviews written by participants who had disclosed their gender and allowed us to separate male-authored from female-authored reviews. This process resulted in a list of 21,012 unique reviewers. We found each reviewer’s total contributions as listed on her profile, and then identified the most prolific 100 males and females. For each of these 200 participants, we obtained all reviews written (199,166 reviews in total).

Even among these prolific participants, males are much more active than are females, contributing a median of 1,187 total reviews to IMDb, versus only 183.5 for

females². In addition, the males write longer reviews than females, (a median of 249 words versus 223, respectively).

Our first hypothesis is that prolific female reviewers adapt their writing style over time. To this end, we studied attributes previously described as being gendered, detailed in Table 1. We focused on writing style, rather than on the content of reviews, to evaluate whether women change *how* they write, rather than what they write. We study changes in the six attributes, in the reviews that participants contribute over time. We identify trends for our prolific female reviewers, and compare them to those of prolific males.

H1: Prolific females write more like males over time.

We also compare the utility scores received by female reviewers, over the course of their tenure as IMDb participants. We expect that since they are learning how to write in a manner that appeals to the IMDb audience, that they will achieve more prestige with time.

H2: Prolific females improve their utility scores over time.

Defining Time

We define time in terms of the number of reviews written by a participant up to a given point (e.g., 10 reviews, 100 reviews), rather than in terms of absolute time (i.e., days since registration). We do this for two reasons. First, our hypothesis is that with repeated exposure to IMDb, a female will change her writing style. For us, writing a review represents an interaction with the community. The second is that participants in OCs might go through periods where they are inactive. Using the number of reviews written gives us a more consistent measure of the amount of active exposure to the community that participants have had.

Preece and Schneiderman [28] propose stages of participation in OCs. Their “reader-to-leader” framework explains how participants’ initial activities might be limited, but that over time, they become more involved. In our data, this might be indicated by change in language use, signaling that a reviewer is assimilating and accepting the norms of the OC. To observe change over time, for each reviewer, we order her reviews chronologically, and then assign “dummy dates” (i.e., review orderings). It should be noted that since reviewers often write multiple reviews on a given day, dummy dates are not necessarily unique.

RESULTS: ALL REVIEWS WRITTEN

We used regression analysis to determine the impact of gender and the number of reviews written (i.e., dummy date) on reviews’ utility and various features of the language within reviews. The results of our regressions are summarized in Table 2 and discussed below.

² A Wilcoxon rank-sum test for the difference between groups results in a p-value of approximately zero.

Variable	B	η^2	F	R^2	Variable	B	η^2	F	R^2
Hedging			373.08	.01	Vocabulary richness			1499.20	.02
# of Reviews	.00***	.01 ⁺			# of Reviews	-.00***	.02 ⁺		
Gender	.00***	.00			Gender	-.00*	.00		
Gender*Reviews	-.00***	.00			Gender*Reviews	-.00	.00		
First person pronoun use			1233.35	.02	Word complexity			2750.79	.04
# of Reviews	.00***	.00			# of Reviews	-.00***	.03 ⁺		
Gender	.08***	.02 ⁺			Gender	-.13***	.01 ⁺		
Gender*Reviews	-.00***	.00			Gender*Reviews	-.00***	.00		
Aggregate pronoun use			2737.18	.04	Sentence complexity			588.79	.01
# of Reviews	.00***	.00			# of Reviews	-.00***	.00		
Gender	.01***	.02 ⁺			Gender	-6.13***	.01 ⁺		
Gender*Reviews	.00*	.00			Gender*Reviews	.00	.00		
Utility			253.10	.00					
# of Reviews	-.00***	.00							
Gender	-.06***	.00							
Gender*Reviews	-.00	.00							

Note. $N = 199,166$ * $p < .05$. ** $p < .01$. *** $p < .001$. ⁺small effect size.

Table 2. Predictors of review utility and language differences (all reviews by all reviewers).

Using the CMOGRAM Stata module [29], we also plotted medians for each variable over time and display both scatterplots and lines-of-best-fit for males and females (see graphs in Table 3³). The graphs indicate positive trends for both genders on *aggregate pronoun use*, and negative trends for both genders on *vocabulary richness* and *word complexity*. For *hedging*, *first person pronoun use*, and *sentence complexity*, the direction of the usage trend differed between the genders. Finally, Figure 1 indicates a positive trend for both genders over time on review *utility*.

Even among prolific authors, some are much more active than others. The number of total reviews contributed by reviewers in our dataset ranged from only 106 (for the 100th most prolific female) to 8,167 (for the most prolific male). In addition, the most prolific female wrote a total of 2,061 reviews. Therefore, in the graphs of language features and review utility for all reviewers (i.e., Table 3 and Figure 1), the trend lines for females do not span the entire width of the graphs.

³ In each plot, the x-axis represents numbers of reviews written, and the corresponding y-axis values represent the average medians over groups of 150 reviews. In order to make the distributions more visible, we removed axis markers and re-scaled each y-axis to make the axis intercept approximately the minimum for each value rather than 0.

Language Differences

Combined Model

All of our models showed both statistically significant and meaningful effects for reviews written, gender, and/or an interaction term. All those effect sizes are small, according to conventions established by Cohen [10] where an η^2 value of 0.01 constitutes a small effect, 0.06 a medium effect and 0.14 a large effect.

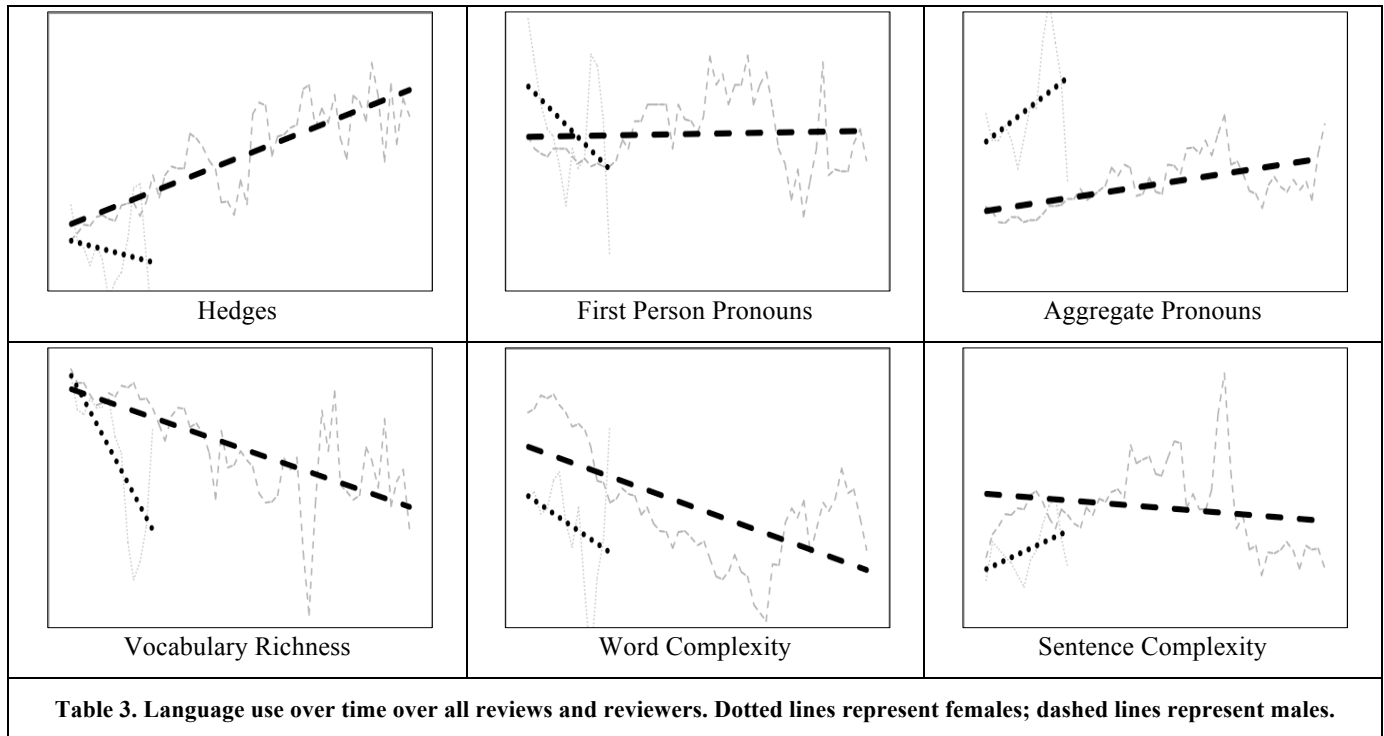
Reviews written

The number of reviews an individual had written (i.e., time) showed main effects on all our measures of language use. Of those six language differences, only *pronoun use*⁴ and *hedging* increased significantly over time; the others decreased. We calculated effect size for each variable in each model as well, and these results indicate that while statistically significant, only *hedging*, *vocabulary richness*, and *word complexity* showed meaningful effects.

Gender

All language features showed a significant main effect for gender. In our data gender is coded as either 1 (female) or 0 (male), meaning that a negative coefficient indicates females used less of the given dependent variable than

⁴ We use the term *pronoun use* to refer to both *aggregate pronoun use* and *first person pronoun use* at once.



males. Our results show that, in general, females exhibited more *hedging* and *pronoun use*, and less *vocabulary richness*, *word complexity*, and *sentence complexity*. Again, we calculated effect sizes for each variable in each model, and found that *pronoun use*, *word complexity*, and *sentence complexity* showed small meaningful differences. Therefore, it is clear that female reviewers, in general, exhibit a relatively outer-focused discourse, which is consistent with previous findings [2,4]. However, we did not find that females wrote with more complex word and sentence structures, nor that they used a richer vocabulary.

Rate of change

We included interaction terms in our regressions to compare the rates of change over time between the genders. Our results indicate females' changing rates of *hedging*, *pronoun use*, and *word complexity* use differed significantly from males' but that differences were not meaningful.

Utility

The results of our regression confirm Otterbacher's [27] findings and indicate that both the number of reviews written and gender had significant effects on *utility*. However, the effect sizes do not indicate meaningful effects for those statistics. Figure 1 shows the trend lines for males and females with respect to review utility. As mentioned, there is a positive trend for both. While the females appear to be increasing the utility of their reviews at a greater rate over time, they do not manage to catch up to the males (i.e., the female trend line never intersects that of the males).

RESULTS: REVIEWS WRITTEN DURING EARLY EXPOSURE

The regression results indicate significant effects for the number of reviews written on a variety of language use measures and utility, but the lack of measureable meaningful effects makes those results difficult to interpret. It might be that males and females do differ in their writing and utility over time but that those differences wash out by the time they've written thousands of reviews. Most reviewers don't write thousands of reviews though, so we examined the first 100 reviews posted by all reviewers to explore whether changes in language use or review utility appear early in a participant's exposure to the community.

For this limited dataset (the first 100 reviews written by each of the 200 reviewers), we used repeated measures ANOVA to determine the impacts of gender and the number of reviews written on reviews' utility and various features of the language within reviews early in a reviewer's tenure (see Table 4). We grouped the number of reviews into sets of 10 (e.g., reviews 1-10, 11-20) following Monk's [26] conventions for handling this kind of timed data. We labeled our effect sizes small, medium, and large according to Cohen's [10] recommendations where *f* values of .10, .25, and .40 indicate small, medium, and large effects, respectively. To illustrate these results, we use graphs of data about just two reviewers (see Table 5) – the most prolific male and the most prolific female.

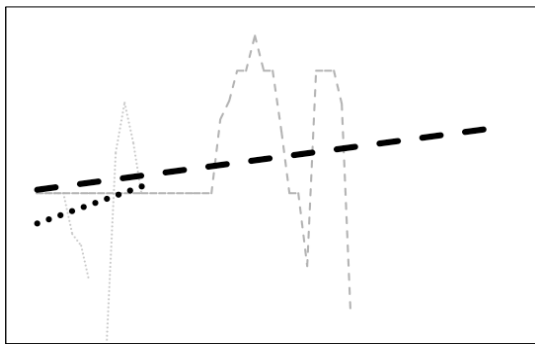


Figure 1. Review utility over time. Dotted line represents females; dashed line represents males.

Language Differences

The ANOVA indicated statistically significant differences in *first person pronoun use* and *vocabulary richness* over time; both effects were also meaningful. Our ANOVA results indicate main effects for gender in all language use variables except *vocabulary richness*, and our effect size analysis indicates that of those statistically significant results, *hedging*, *pronoun use*, and *word complexity* show meaningful differences. The ANOVA results indicate that the interaction between time and gender is not significant for any of the language use variables.

Utility

The ANOVA results indicate that the interaction of time and gender is significant for measures of *utility* and that the mean *utility* for female reviewers is lower than males' ($M_f = .43$ vs $M_m = .49$). In fact, at each interval, the males' mean *utility* is higher than the females', and the gap between them is larger after 100 reviews than after just 10 ($M_m = .48$, $M_f = .47$). Both gender and the interaction of gender and time showed meaningful effects on utility.

RESULTS: EXAMPLES FROM MOST PROLIFIC REVIEWERS

We now illustrate some of our analysis using data from the most prolific male and female. Table 5⁵ shows aspects of language use (using means for each group of 10 reviews) over the first 100 reviews the most prolific author of each gender contributed. For each feature, except *vocabulary richness*, the intercept between the male's trend and female's trend occurs by 100 reviews, indicating that the major changes in writing style happen early on. Although the mean usage between the two reviewers often does not differ significantly, their patterns of use, especially in *hedging*, *pronoun use*, and *word complexity* differ remarkably.

⁵ Each x-axis represents numbers of reviews written, and the corresponding y-axis values represent the means over groups of 10 reviews. To ease visibility, we removed axis markers and re-scaled each y-axis to make the axis intercept approximately the minimum for each value rather than 0.

Variable	F(df)	Cohen's <i>f</i>
Hedging		
Model	4.99(19)***	
Timebin	.18(9)	.009
Gender	76.19(1)***	.061 ⁺
Timebin#gender	0.89(9)	-
First person pronoun use		
Model	40.87(19)***	
Timebin	4.17(9)***	.037 ⁺
Gender	732.95(1)***	.191 ⁺⁺
Timebin#gender	0.67(9)	-
Aggregate pronoun use		
Model	53.47(19)***	
Timebin	0.67(9)	-
Gender	996.63(1)***	.223 ⁺⁺
Timebin#gender	1.46(9)	.014
Vocabulary richness		
Model	3.27(19)***	
Timebin	5.37(9)***	.044 ⁺
Gender	3.28(1)	.011
Timebin#gender	1.16(9)	.009
Word complexity		
Model	29.54(19)***	
Timebin	0.68(9)	-
Gender	544.78(1)***	.165 ⁺⁺
Timebin#gender	1.16(9)	.008
Sentence complexity		
Model	13.07(19)***	
Timebin	0.75(9)	-
Gender	234.56(1)***	.108
Timebin#gender	0.77(9)	-
Utility		
Model	6.81 (19)***	
Timebin	1.42(9)	.014
Gender	92.39(1)***	.068 ⁺
Timebin#gender	2.70(9)**	.028 ⁺

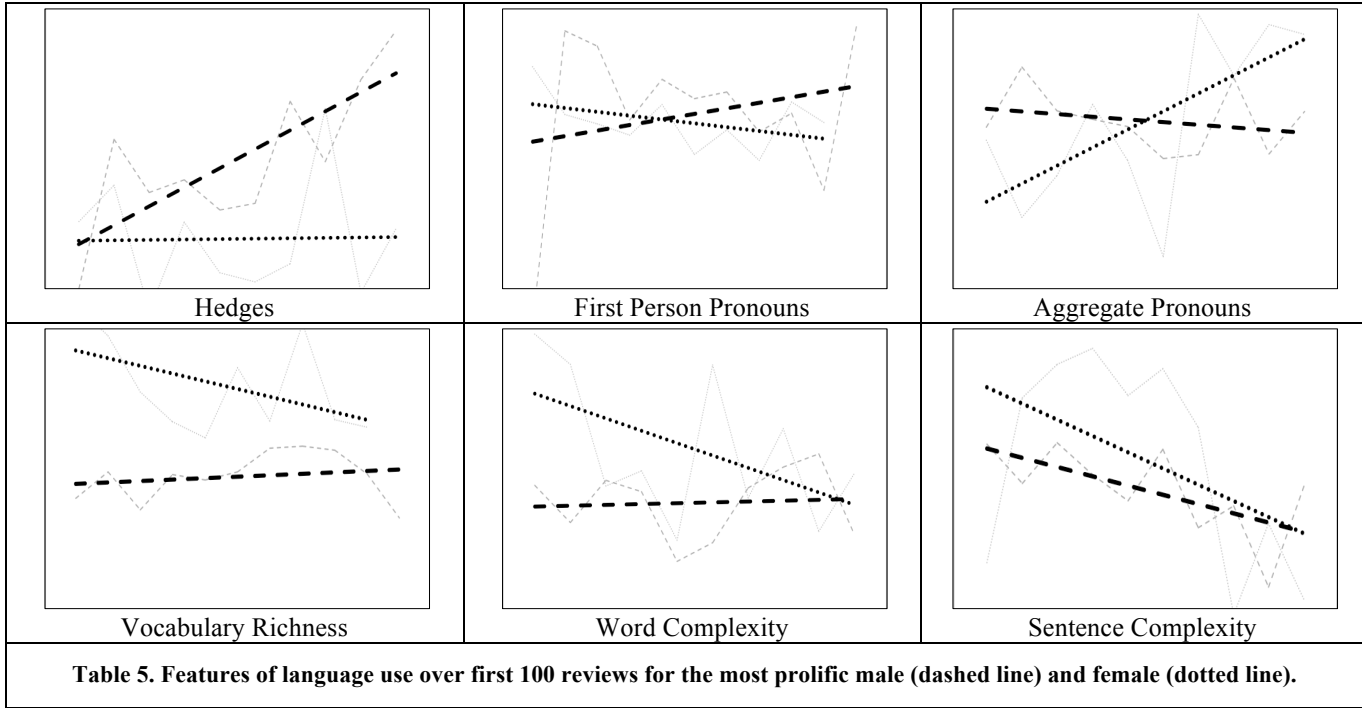
Note. $N = 20,000$.

* $p < .05$. ** $p < .01$. *** $p < .001$

+small effect size. ++medium effect size.

Table 4. Results of repeated-measures ANOVAs for each language use variable and review utility

These distributions for prolific reviewers indicate that the male exhibited a marked increase in *first person pronouns* before settling into a more consistent pattern of use. Also, he dramatically increased his use of *hedges*. Meanwhile, the female showed less variation in her *first person pronoun*



and *hedges* use, but also showed a decrease over time. The two appear to converge on measures of language complexity – both *word* and *sentence complexity*. *Word complexity* differed meaningfully over time, indicating that the trends shown in their first 100 reviews likely continued – the female’s likely decreased while the male’s increased.

Next, we present a qualitative analysis of a set of reviews written by these prolific reviewers at four distinct points in time. In Tables 6 and 7, we show the opening sentences of the first review written, as well as the 25th, 50th and 200th. We also provide utility scores, in the second column.

Reviews by the male are shown in Table 6. Initially, he writes strictly in the third person, as evidenced by the first two reviews. In his first review, he simply provides a critique of the film. To contrast, in the second, while he still writes in a “reporting” voice, he describes the film’s characters. By his 50th contribution, he has begun using first person pronouns to a limited extent. We also find one instance of hedging (“often”). Finally, by his 200th review, he is writing using an outward-focused discourse. In addition to expressing his feelings about the film, he interacts, using second person pronouns to address readers.

Table 7 shows the opening lines of four reviews by the female. We observe that she writes in third person in her first contribution. In contrast, her 25th review shows two markers of feminine language, namely one instance of hedging (i.e., “almost”), and one second person pronoun. Nonetheless, the voice used in the 25th review is not nearly as personal as the voice she uses in her 50th review. Here, she even apologizes to her readers for loving the film. We

also observe extensive use of first person pronouns. Finally, in her 200th review, we find both singular and plural first person pronouns, as the writer interacts with her audience. In fact, she also references Hurricane Katrina, which touched all Americans, and discusses religion.

In summary, both the most prolific male and female move toward a more personal voice in their writing over time. The female explicitly uses more markers of feminine language overall. However, a change in writing style from a more reporting stance to a personal perspective is evident in both reviewers’ styles.

DISCUSSION

Our data indicate support of our first hypothesis – that women write more like men over time. However, our second hypothesis – that women improve their utility scores – was not fully supported. In the first 100 reviews, females increased their scores. However, they never caught up to the males. Our findings are interesting for several reasons.

First, our results suggest that even without direct interaction, language convergence can occur. We saw evidence of linguistic convergence at IMDb that other researchers [8,20] found in more interactive OCs. Convergence has been interpreted as being positive (e.g., an indication that participants view themselves as part of the OC [8]), as well as negative (e.g., evidence that minority groups conform to majority norms [20]). That we find convergence indicates that even without direct interaction, collaborators can develop a common language.

That women reduced their use of some markers of female language is not surprising. Following Eckert [12] and Giles

Opening sentences of review		
1	15/18 (0.83)	<u>The Apple (1980)</u> : Kudos to this German production by Golan/Globus. Who else could have conceived of a Scifi/Disco/Musical/Religious epic? Unlike Ed Wood's flicks (such as Plan 9, Glen or Glenda and Bride of the Monster), this movie proves that an embarrassingly silly and awful picture does NOT need to be made on a shoestring budget!
25	8/12 (0.67)	<u>Tsuma Shimatsu (1967)</u> : This wonderful movie stars the great Toshiro Mifune as the leader of this family in feudal Japan. And, despite starring Mr. Mifune, it is NOT an Akira Kurasawa film! The patriarch, Mifune, is in a tough spot. His son has been ordered to divorce the wife he loves and marry another by his warlord.
50	2/2 (1)	<u>Conte d'automne (1998)</u> : This was a little film with a simple plot and likable characters. In fact, Hollywood would learn a lot from films like this. It's not the dynamic plot, special effects or big name stars that often make a film exceptional, it's the writing and the acting! And this movie is written so lovingly and acted so honestly that I couldn't help but like it.
200	3/4 (0.75)	<u>Laissez-passer (2002)</u> : This is a film directed by Bertrand Tavernier. I loved his film IT ALL STARTS TODAY, and I was quite impressed by this one as well. However, be forewarned that this film will not be for all tastes. If you are French or have a good knowledge of French cinema, then you'll no doubt enjoy this film.
Table 6. Examples of reviews by the most prolific male.		

et. al [18], we expected women to adjust their patterns. What we did not expect was to see a trend among males to adopt *more* feminine markers. Cassell and Tversky [8] reported males adjusting toward female style, but we expected a “list effect” here [19], since we knew that IMDb has a male majority. “List effect” was identified in listserv conversations in which all authors adopted the male voice. In contrast, we saw that men adjusted their language, especially by increasing hedging and pronoun use.

One possible explanation is that men are becoming more comfortable with their place in the OC over time. Rather than continuing to write with a detached voice, they begin to write in a socially engaging manner. This would be consistent with the observation that with increased involvement, one’s personal identity (e.g., male, 28 years old) might influence behavior to a lesser extent than one’s

Opening sentences of review		
1	14/21 (0.67)	<u>Leaving Las Vegas (1995)</u> : A big disappointment for what was touted as an incredible film. Incredibly bad. Very pretentious. It would be nice if just once someone would create a high profile role for a young woman that was not a prostitute.
25	4/5 (0.8)	<u>Elvis (1979/I) (TV)</u> : This magnificently produced bio of Elvis Presley contains an eerie, almost frightening portrayal of Presley by Kurt Russell, who literally seems to be inhabited by Elvis' spirit. Physically, the movie is perfect in casting and location - you could see a freeze of any frame of this film and know it's about Elvis. All that being said, die-hard Elvis fans will be left frustrated by the movies' gaping holes and unnecessary inaccuracies, the biggest of which is that the film stops in 1970 when Elvis lived until 1977.
50	2/5 (0.4)	<u>Phenomenon (1996)</u> : Sorry but I just loved it. I know for some it will be sentimental and trivial. It was on the slow side, a little too easygoing, but utterly charming with a terrific cast. The Robert Duvall role certainly did not need anyone of his caliber playing it, but it's always good to see him just the same. For the moment, it was just what I needed. If you're a man, don't bother.
200	3/4 (0.75)	<u>Brigham Young (1940)</u> : This is a wonderful movie about the struggle of the Mormons and their final settlement in Salt Lake, Utah. The beginning and the ending are especially powerful, and the message is one we all have to be reminded of - God doesn't talk, but he communicates, if we would only listen. As I am writing this in the midst of the horrors going on in New Orleans and the surrounding area due to Katrina, I was especially moved by the Mormons having to leave everything behind and move on after Joseph Smith was assassinated.
Table 7. Examples of reviews by the most prolific female		

social identity (e.g., prolific reviewer of foreign films) [23]. We believe this is possible, even for participants who do not engage in repeated interaction with others, since participants often develop a personal connection to the OC as a whole (i.e., the site itself acts as a primary actor in the participant’s social network [5]).

Likewise, female participants might not feel as attached to the OC, particularly if they do not feel that others value their contributions. Following [27], we observed higher utility scores for men. The gap between men and women was not as pronounced as in earlier work, but still matters. The implications of men having higher average utility, even if the gap isn't meaningful according to effect size calculations, is that men's reviews are more likely to be read than are women's. Because IMDb sorts its reviews according to utility by default, and users rarely read past the first page [1,17], women's reviews remain unseen, buried on later pages. As a result, female participants might adapt their writing patterns toward those of the majority.

Another surprising result is that in all language differences, the convergence between men and women's behavior occurs within the first 100 reviews. However, even though women's utility increases, it never catches up with men's, even after 2,000 reviews. The question remains – why are women's reviews consistently seen as being less “useful,” even once they shift to the majority writing style? The apparent undervaluing of women's contributions may have broader impacts. For instance, women's contributions in other electronic media such as workplace email may be undervalued as well. The increased use of asynchronous communication in which contributions are addressed to a community rather than to an individual (e.g., Wikipedia, corporate knowledge management systems) implies the importance of the implications of the disparity in value and visibility awarded women's contributions.

LIMITATIONS

Our findings are subject to a few limitations with regard to the content and object of the reviews, and the authors' self-report of their genders. Differences in the content of reviews were not investigated in our study. However, Otterbacher [27] previously found that even when she controlled for aspects of reviews' content and stylistic features, reviewer gender and review utility were still highly correlated. We believe that several processes are at work beyond the differences in the way that men and women write, including the use of the social voting mechanism to order reviews, as well as the gender composition of the community and particularly of those who vote on review utility. Unfortunately, we cannot observe who votes for which reviews, but we suspect that just as male participants contribute more reviews than females, that they provide more social feedback as well.

We also do not investigate possible differences between various features of the movies being reviewed (e.g., genre, critics' rating, release date) that may impact reviewers' behavior. Aspects of the movie may influence how reviewers write about it – for instance, reviewers may adopt a different linguistic style for mainstream movies than for lesser-known films, or for romantic versus action films.

It is challenging to observe gender dynamics in OCs because participants often do not disclose their

demographics. Even in communities in which it is customary to reveal one's identity, participants do not necessarily use their offline identities. They may have faceted identities, choosing to present certain personal characteristics in different situations [13], or they may adopt new identities. In fact, Bruckman [6] notes that many women use neutral or male identities (e.g., pseudonyms, avatars) to avoid unwanted attention from men, as well as outright harassment [21]. Accordingly, at IMDb, revealing one's gender is not mandatory. Our results rely on the assumption that users who provide their identity do so accurately.

CONCLUSION AND FUTURE WORK

Our results showed that even if women adopt language more like men's, their contributions might still go unnoticed. We identified a change in behavior over time, but only by asking authors could we really tell whether changes were made intentionally or even consciously. For instance, whether changes occur because a reviewer is trying to communicate more effectively, or wants to appeal to the voting masses is not clear.

Another possible reason for a change in language behavior might be that with time, reviewers' perception of their place in the OC might change. We would especially like to ask prolific reviewers if their *feelings* toward IMDb changed over time, as they contributed more reviews. Also, we would like to know to what extent, if any, the feedback on their reviews influenced these feelings.

IMDb reviews provide an accessible and rich data set for examining a number of questions about user-contributed content (UCC) and its authors. Our study reveals differences between the genders that impact what reviews other users see, and future work should explore other implications of differences in the style of UCC and the characteristics of its authors.

ACKNOWLEDGMENTS

We are grateful for the helpful, constructive feedback the AC and our anonymous reviewers provided.

REFERENCES

1. Adamic, L.A. and Huberman, B.A. The Web's hidden order. *Communications of the ACM* 44, 9 (2001), 55-59.
2. Argamon, S., Koppel, M., Fine, J. and Shimoni, A. R. Gender, genre, and writing style in formal written texts. *Text*, 23, August 2003.
3. Awad, N.F. and Ragowsky, A. Establishing trust in electronic commerce through online word of mouth: an examination across genders. *Journal of MIS* 24, 4 (2008), 101-121.
4. Biber, D. *Dimensions of Register Variation: a Cross-Linguistic Comparison*. Cambridge & New York: Cambridge University Press, 1995.
5. Brown, J., Broderick, A.J. and Lee, N. Word of mouth communication within online communities:

- Conceptualizing the online social network. *Journal of Interactive Marketing* 21, 3 (2007), 2-20.
6. Bruckman, A. Gender swapping on the Internet. In Ludlow, P. (Ed.), *High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace*. Cambridge, MA USA: MIT Press, 1996.
 7. Bury, R. *Cyberspaces of their Own: Female Fandoms Online*. Peter Lang Publishing, New York, NY USA, 2005.
 8. Cassell, J. and Tversky, D. The language of online intercultural community formation in Junior Summit '98. *Journal of CMC* 10, 2 (2005).
 9. Coates, J. *Women, Men and Language*. London: Longman, 1993.
 10. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 1988.
 11. Collaborative Creativity Group of the United Nations University. Wikipedia Survey – Overview of Results. http://www.wikipediaurvey.org/docs/Wikipedia_Overview_15March2010-FINAL.pdf (accessed May 13, 2011).
 12. Eckert, P. *Jocks and Burnouts: Social Categories and Identity in the High School*. Teachers College Press, New York, NY, USA, 1989.
 13. Farnam, S.D. and Churchill, E.F. Faceted identity, faceted lives: social and technical issues with being yourself online. *Proc. CSCW*, ACM Press (2011), 359-368.
 14. Flynn, E.A. Composing as a woman. *College Composition and Communication* 39, 4 (1988).
 15. Foltz, P.W., Laham, D., Landauer, T.K. Automated essay scoring: applications to educational technology. In *Proceedings of EdMedia* (1999).
 16. Gefen, D. and Ridings, C.M. If you spoke as she does, sir, instead of the way you do: a sociolinguistics perspective of gender differences in virtual communities. *ACM SIGMIS Database* 36, 2 (2005), 78-92.
 17. Gilbert, E. and Karahalios, K. Understanding déjà reviewers. *Proc. CSCW*, ACM Press (2010), 225-228.
 18. Giles, H., Coupland, J., and Coupland, N. Accommodation theory: communication, context, and consequence. In Giles, H., Coupland, J. and Coupland, N. (Eds.), *Contexts of Accommodation. Developments in Applied Sociolinguistics*. Cambridge & New York: Cambridge University Press, 1991.
 19. Herring, S. C. Gender and power in online communities. In Holmes, J. and Meyeroff, M. (eds.) *The Handbook of Language and Gender*, pp. 202-228. Oxford: Blackwell, 2003.
 20. Herring, S.C. Posting in a different voice: gender and ethics in computer-mediated communication. In Ess, C. (Ed.) *Philosophical Perspectives on Computer-Mediated Communication*. Albany, NY: SUNY Press, 1996.
 21. Herring, S.C. The rhetorical dynamics of gender harassment on-line. *The Information Society* 15, (1999), 151-167.
 22. Jacobson, D. Impression formation in cyberspace: Online expectations and offline experiences in text-based virtual communities. *Journal of CMC* 5, 1 (1999).
 23. Kim, H-W., Zheng, J.R. and Gupta, S. Examining knowledge contribution from the perspective of an online identity in blogging communities. *Computers in Human Behavior* 27 (2011), 1760-1770.
 24. Lakoff, G. Hedges: a studying meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic* 2, 4 (1973), 458-508.
 25. Lakoff, R. Language and woman's place. *Language in Society* 2, (1973), 45-79.
 26. Monk, T.H., Buisse, D.J., Reynolds III, C.F., Berga, S.L., Jarrett, D.B., Begley, A.E. and Kupfer, D.J. Circadian rhythms in human performance and mood under constant conditions. *Journal of Sleep Research* 6, (1997), 9-18.
 27. Otterbacher, J. Inferring gender of movie reviewers: exploiting writing style, content and metadata. In *Proc. CIKM*. ACM Press (2010), 369-378.
 28. Preece, J. and Schneiderman, B. The Reader-to leader framework: motivation technology-mediated social participation. *AIS Transactions on Human-Computer Interaction* 1, 1 (2009), 13-32.
 29. Robert, C. CMOGRAM: Stata module to plot histogram-style conditional mean or median graphs, *Statistical Software Components*, Boston College Department of Economics (2010). <http://econpapers.repec.org/software/bocbocode/s457162.htm> (accessed November 19, 2011).
 30. Savicki, V., Lingenfelter, D. and Kelley, M. Gender language style and group composition in Internet discussion groups. *Journal of CMC* 2, 3 (1996).
 31. Scissors, L.E., Gill, A.J., and Gergle, D. Linguistic mimicry and trust in text-based CMC. *Proc. CSCW*, ACM Press (2008), 277-280.
 32. Spender, D. *The Writing or the Sex, or, Why You Don't Have to Read Women's Writing to Know it's no Good*. New York, NY: Pergamon Press, 1989.
 33. Tannen, D. *You Just Don't Understand*. New York, NY: Harper Collins Publishers, Inc, 1990.
 34. Where Are the Women in Wikipedia? Room for Debate. <http://www.nytimes.com/roomfordebate/2011/02/02/where-are-the-women-in-wikipedia> (accessed November 19, 2011)