

COMPARATIVE ANALYSIS OF GENDER-BASED CONVERSATIONAL FEATURES IN WHATSAPP ACADEMIC GROUPS

Sana Urooj

Institute of Humanities and Arts,

*Khwaja Fareed University of Engineering and Information Technology, Rahim
Yar Khan. Email: sana1arooj@gmail.com*

Abstract

This study examines gender-based conversational features in WhatsApp academic groups, comparing how male and female university students communicate within task-oriented, peer-to-peer digital exchanges. Adopting a mixed-methods comparative design, the study combines quantitative content analysis with qualitative discourse analysis, drawing on twenty WhatsApp conversation threads (eight male-dominated, eight female-dominated, and four mixed-gender) comprising 20 messages from six male and six female participants. Messages were coded for linguistic features (directness, hedging, politeness, and message length) and interactional/paralinguistic features (topic initiation, question-asking, emoji use, and code-switching). Findings indicate that female participants produced longer, more elaborated messages and used politeness markers considerably more often than male participants, who favored brief, direct, imperative-style requests. Emoji use occurred at similar rates across genders but differed sharply in type and function, with male use concentrated in a single humor-oriented emoji and female use spanning a broader affective range. Male participants asked more frequent, but shorter, questions, while female participants asked fewer, more elaborated ones. Notably, topic-initiation rights were evenly distributed between genders in mixed-gender threads, suggesting that stylistic differences did not correspond to unequal conversational power. These findings extend existing sociolinguistic and computer-mediated communication literature into the underexamined genre of academic WhatsApp groups and carry implications for how educators design and interpret digital group communication in higher education.

Keyword: *gender differences; computer-mediated communication; WhatsApp; academic discourse; sociolinguistics; discourse analysis; digital communication; code-switching*

Introduction

The emergence of instant messaging via smart phones has made WhatsApp one of the most popular apps for informal academic communication amongst students at university, continuing conversations beyond the classroom into an out-of-sync, written, digital space in which familiar gendered norms of talk emerge. Prior sociolinguistic research suggested that women's speech is more hedging, politeness-oriented, and more conversational, whereas men's speech is more direct and status-oriented (Lakoff, 1975; Tannen, 1990), and previous studies of computer-mediated communication (CMC) discovered that many of these distinctions are not lost in electronic discourse (Herring, 2000). Recent research on digital genres, however, indicates a more complex picture in which communicative context has a greater influence on gendered stylistic features than the speaker's identity (Herring & Paolillo, 2006; Huffaker & Calvert, 2005). WhatsApp academic groups are situated in a unique position between an informal, emoji-filled, conversational style and a purposeful academic style, sparking questions about the negotiation of turn-taking, politeness and hedging, as well as of message style between male and female users in a shared space. The latter is the gap addressed in the present study, where the conversational elements based on Gender is compared in academic groups created in WhatsApp.

Significance of the Study

This study is significant on a number of inter-related fronts. It extends the current discussion in sociolinguistics and computer-mediated communication (CMC) in theoretically testing

existing gender-and-language frameworks (Lakoff, 1975; Tannen, 1990) with a particular, relatively under-researched digital genre, the academic WhatsApp group, rather than the overall social or family contexts that have been the focus of WhatsApp research. It does so in response to Herring and Paolillo's 2006 observation that the gendered nature of communication is more easily explained in terms of the role of genre and context than in terms of gender of the speaker per se, and provides new empirical data from a task-oriented purposive setting.

In practice, the results can help teachers and institutions to have a better understanding of student interactions outside the classroom. Understanding how male and female students communicate in WhatsApp groups can assist teachers in developing more gender-inclusive approaches to group-based online learning and anticipate potential gender communication gaps or misunderstandings in the group.

The study also has pedagogic implications for teacher training and curriculum development in ESL and communication studies as it brings attention to students' awareness of authentic and real-life gendered digital discourses and its potential to be used to draw student awareness to politeness strategies, directness, and register-appropriate communication in digital space.

On the social side, the research contributes to a growing awareness of the way digital communication technologies have (and perhaps will continue to) negotiate and alter the traditional gendered communication norms, and thereby provides information of interest to digital literacy and equitable participation in online academic environments for communication scholars, sociolinguists, and digital literacy and equity policymakers.

Lastly, this study provides context-specific results that could provide a basis for further comparative and/or cross-cultural research in the field of gender and digital academic discourse within a specific Pakistani context, where WhatsApp is widely used as a medium for student coordination.

Research Objectives

1. To identify and compare the linguistic features (such as directness, hedging, politeness strategies, and message length) used by male and female students in WhatsApp academic groups.
2. To examine and compare the interactional and paralinguistic patterns (such as turn-taking, topic initiation and response, and emoji use) displayed by male and female students in WhatsApp academic groups.

Research Questions

1. What differences, if any, exist in the linguistic features (directness, hedging, politeness strategies, and message length) used by male and female students in WhatsApp academic groups?
2. How do male and female students differ in their interactional and paralinguistic patterns (turn-taking, topic initiation and response, and emoji use) within WhatsApp academic groups?

Literature Review

There have been many studies of gender and language that have indicated that men and women employ different conversational styles: female speakers tend to use hedges, politeness markers, and connection-oriented language, while male speakers tend to be direct, assertive, and status-oriented (Lakoff, 1975; Tannen, 1990). Given the nature of computer-mediated communication (CMC), scholars wondered if these differences would be eliminated by the less explicit social cues of digital communication, but Herring (2000) found that women's messages were still more supportive and men's more assertive. Later work on digital genres was challenging this perspective: Herring and Paolillo (2006) showed that stylistic aspects

that are commonly associated with male and female authors in weblogs were more likely to be associated with genre than with the actual gender of the author, and Huffaker and Calvert (2005) revealed both gender-typical patterns in the emotional expression of adolescent bloggers and significant overlap between the expressions of male and female writers. Specific studies of WhatsApp have found that female users are likely to communicate in a more expressive and emotionally nuanced manner, whereas male users are likely to be more concise, direct and to the point. This WhatsApp evidence, however, does not come from academic student groups, but rather from general social or family-based groups, thus leaving a gap as to how these gendered conversational features can be observed within the more goal-oriented, task-oriented conversation of academic student groups that the present study attempts to fill.

The findings from further research into digital correspondence have confirmed these gendered trends, not just with blogs and IM, but with email as well: Colley and Todd (2002) noted that men's emails were less detailed than women's, with men writing more about information and women writing more about their feelings and social situations. In a similar vein, Crystal (2001) claimed that the Internet-mediated language is a hybrid that combines characteristics of speech and writing, and can either reinforce or disguise the gendered stylistic features based on the context and nature of the communication. Baron (2010) further developed this argument by focusing on mobile messaging, specifically, the brevity, immediacy, and multimodalities of platforms such as WhatsApp, calling for new empirical investigation of this type of CMC. In summary, this literature suggests that there are inter-indigital differences in communication between men and women, but studies on the relationship between gender and conversational style in academic WhatsApp groups are less extensive. This is a gap that calls for a targeted and comparative study of the communicative behavior (CB) of male and female students in a semi-formal and task-oriented communicative environment for academic group messaging.

Research Design

The study was using mixed methods comparative research design, which is based on two research methods, first quantitative content analysis, and second qualitative discourse analysis, to analyze the features of the gender dimension in academic groups made by WhatsApp. The quantitative component provides for systematic counts and statistical comparisons of linguistic and interactional features of male and female participants in two KPIs: RO1/RQ1 (Linguistic features: Directness, Hedging, Politeness, and Message length); RO2/RQ2 (Interactional and paralinguistic features: Turn taking, Topic initiation/response, and Emoji use). The findings of the quantitative component are complemented by the qualitative component, which places patterns in the context of real-world conversational exchanges, based on Herring's (2004) Computer-Mediated Discourse Analysis (CMDA) framework that integrates linguistic discourse analysis core approaches with the structural components of digital communication platforms. A mixed-methods approach, as recommended by Creswell (2014), has been used here, which is appropriate for research questions involving measurable comparison and interpretative depth.

Population and Sample

University-level students, active in academic groups created on WhatsApp for communication of academic content, represent the population of the present study. Purposive sampling technique is used to select two to three active academic WhatsApp groups where boys and girls are actively involved in the group, ensuring a good representation of both sexes for meaningful comparison. Group administrators and members are approached for informed consent before data is extracted, and only groups in which academic discussion

occurs today (not dormant, or announcement-only groups) are chosen to represent the data, as this is a more representative sample of conversational interaction.

Data Collection

Data are archived WhatsApp conversation logs obtained from a predetermined time frame (e.g. 8 weeks) of chosen academic groups. Chats are exported in text format directly from WhatsApp's built-in export feature, with the added benefit of timestamps, the sender's identity will be anonymized (later), the message content, including emojis. All personally identifiable information is anonymized prior to analysis, by replacing it with coded identifiers (e.g., M1, M2 for male participants; F1, F2 for female participants) to maintain participant anonymity in accordance with good practice for digital discourse research.

Coding Scheme and Instrumentation

The variables are matched with each research question and a structured coding scheme is created to operationalize them. In the case of RQ1, the linguistic features were coded according to categories of directness (imperative or assertive statement vs. mitigated request), hedging markers (please, I think, thank you, apologetic framing), politeness strategies (please, thank you, apologetic framing), and average message length (word count per message). In RQ2, the interactional and paralinguistic features are coded as turn taking frequency, topic-initiation versus topic-response roles, response lateness (if present), number and type of emojis used. The coding scheme is tried out on a small portion of the data and is clarified and adjusted for mutual exclusivity of categories before coding is undertaken on the entire data set.

Data Analysis Procedure

Data from the coded forms are then analysed quantitatively in terms of frequency counts and percentages for each linguistic and interactional category and by gender. As in most comparative CMC studies, independent samples t-tests (or chi-square tests for frequency/categorical data) are used to examine if differences that have been observed between the male and female participants are statistically significant. In parallel to this quantitative analysis, a qualitative discourse analysis of selected conversational leftover materials is conducted to showcase and interpret the quantitative patterns, as to gain more insight into the functioning of directness, hedging, politeness and paralinguistic cues in actual academic exchanges.

Reliability and Validity

A second coder independently codes a subset (about 15-20%) of the data to ensure coding reliability and intercoder reliability is computed using Cohen's Kappa, where intercoder coding agreement of $\kappa \geq 0.75$ is regarded as acceptable. Content validity of the coding scheme is confirmed by having the coding scheme reviewed by faculty members who are knowledgeable about sociolinguistics and CMC research before the data is fully coded.

Ethical Considerations

Appropriate academic/institutional review board approval is obtained before data collection. Group administrators and, if possible, individual group members are informed of the consent, and the individual can withdraw their messages from the data set. Data is anonymised prior to analysis, kept securely, and only used for the purposes of this research.

Delimitations

This study is focused on the academic group conversations that happen on WhatsApp and not on voice notes, video calls or private one-to-one messages. It is limited to an analysis of gender as the key variable for comparison rather than other potentially influential variables like academic discipline or level of digital literacy.

Data Analysis and Presentation of Findings

The results obtained from the content analysis and discourse analysis of the academic-group data that were gathered by the WhatsApp application for this study are presented in this chapter. The data consists of 20 conversation threads (8 male, 8 female and 4 mixed gender threads) and 6 male participants (Ali, Hamza, Bilal, Ahmed, Usman and Daniyal) and 6 female participants (Ayesha, Sara, Fatima, Hina, Maryam and Nida) contributed 107 individual messages. Each message was coded for message length, directness, hedging, politeness marking, emoji use, code-switching, and topic-initiation behavior, as described in Chapter Three, and the results are broken down by gender below. As this is an initial, small-scale representation of a dataset, rather than a large randomly sampled dataset, the results are reported in a descriptive manner as frequencies and percentages, with illustrative examples provided, and are not reported using inferential statistical tests (for example, independent samples t-tests, chi-square tests). This descriptive approach is used to prevent over-precision of the results obtained with few threads and participants, and the patterns that are detected in the following are preliminary and exploratory, and require confirmation in larger corpora in future studies, following the delimitations.

Message Frequency and Average Message Length

Table 4.1 summarizes the total number of messages and average message length, measured in words per message, contributed by each gender group across the twenty threads.

Table: Message Frequency and Average Message Length by Gender

Gender	Total Messages	Total Word Count	Average Words per Message
Male	57	143	2.51
Female	50	220	4.4

The length of the messages produced by women was, on average, almost 75% longer than those produced by men (4.40 words per message compared to 2.51 words per message). This is similar to Baron's (2010) characterization of mobile messaging as an inherently short and economical medium, and somewhat challenging Baron's characterization of genres as independent of conversational context, given that the present data showed that gender may affect the genre-level stylistic norms, and therefore not exhaust them.

Directness, Hedging, and Politeness Markers

Table reports the frequency of directness markers (bare imperatives and unmitigated requests), hedging markers (e.g., "maybe," "I think"), and politeness markers (e.g., "please," "thank you," "JazakAllah") identified in each gender group.

Table: Frequency of Directness, Hedging, and Politeness Markers by Gender

Gender	Bare Imperative/Directive Messages	Hedging Markers	Politeness Markers
Male	Frequent (e.g., "Notes bhej do," "Share kar dena," "Sir extension de den")	0	3
Female	Infrequent; requests typically mitigated (e.g., "Could someone please share...")	1	10

In line with the models of directness and politeness described by Lakoff (1975) and Tannen (1990), most requests by male participants were in the form of bare imperative constructions while female participants more often used mitigated questions or polite appeals. Politeness markers (please, thank you, JazakAllah, etc. and expressions of appreciation) were used more than three times as frequently in female dominated threads as in male dominated threads. Gendered patterns of hedging, however, were not seen in either group, and this may not be a complete lack of gendered hedging tendencies, but rather a result of the register of peer-to-

peer academic exchanges, as suggested by Herring and Paolillo (2006): genre conventions can moderate overall gender patterns found in previous CMC studies.

Emoji Use

Emoji frequency was similar in raw count between the two groups, but the type and apparent communicative function of emojis diverged sharply, as summarized in Table.

Table: Emoji Frequency and Type by Gender

Gender	Total Emoji Count	Emoji Types Observed	Apparent Function
Male	6	😂 (laughing face) only	Humor / deflection (e.g., around deadlines, attendance, teacher's mood)
Female	5	😊, ❤️, 🌸, 🙏	Affective warmth, gratitude, and social bonding

Male emoji use was restricted exclusively to the one emoji: a laugh, which was mainly used to deflect or soften potentially face-threatening situations, including missed deadlines and low attendance; while female emoji use was more widespread but used in very similar ways. Female emoji users, on the other hand, exhibited a broader spectrum of emotions such as appreciation, caring, and interpersonal connection. This qualitative discrepancy corroborates previous research that suggests that even if the frequency of using WhatsApp communication is comparable between genders, emoji sensitivity and the interpretation of emoji play a role.

Question-Asking, Interactional Engagement, and Topic Initiation

Table: Frequency of Interrogative Messages by Gender

Gender	Messages Containing a Question	Percentage of Total Messages
Male	20	35.1%
Female	5	10.0%

The male participants sent interrogative messages over three times more frequently than the female participants. It was, however, more evident from a qualitative analysis that male questions were usually shorter, elliptical, and asked with a "confirm", "time", "deadline kab hai?". However, female questions were rare and more elaborated, socially oriented and were asked with "does anyone know...?" / "does it seem that...?". In the four mixed gender threads topic initiation was evenly distributed between the two genders, with the male participants (Ali) initiating two threads and the female participants, (Maryam) initiating the other two. It indicates that, even in the face of such stylistic difference noted above, topic-initiation rights and turn-taking did not seem to be monopolized by either gender when both were in the same conversation.

Code-Switching Patterns

In the male messages, there were 24 (42%) messages with code switching, which were mostly embedded in overall English language messages, such as requests or questions (e.g. Kis kis ki team bani?). A smaller proportion of female messages (9 out of 50; 18%) were Urdu-origin messages, of which almost all were formulaic religious or social utterances (Assalam-o-Alaikum, JazakAllah, Ameen), as opposed to grammatical Urdu-English code-mixing within task-oriented utterances. This means that male participants used to code-switched more freely inside the content of task talk, whereas the female participants used Urdu-origin language mainly for greetings and closings, the majority of academic negotiation was carried out in English.

Qualitative Discourse Patterns: Conversational Closings

Reciprocated closing sequences, that is, a sequence in which a request for help or information was followed by an expression of gratitude and a subsequent acknowledgment (for instance: a request for notes is followed by "Thank you very much" and "Really appreciate it"), were

another recurring pattern in the female dominated threads. Male threads, on the other hand, tended to close off after the practical information was conveyed or after a humorous deflection, rather than as a consequence of a two-way exchange of politeness. This situation parallels Tannen's (1990) distinction between rapport-talk and report-talk, which was identified with males where the goal of the conversation was to communicate information and with females where the goal was to communicate social information.

Summary of Findings

The results suggest that, broadly speaking, there are differences in the length, directness, politeness marking, function of emoji, framing of questions, and code switching in this academic group of WhatsApp messages, as outlined by Lakoff (1975) and Tannen (1990), and supported by Herring and Paolillo (2006)'s caution that some gendered patterns might be reduced in certain digital genres. Interestingly, the distribution of topic-initiation rights was not significantly different between men and women in mixed-company threads as the distribution of these rights in other threads might suggest, indicating that topic-initiation style and the power to initiate topics are not the same in this set of data.

Discussion

This study confirms broadly the foundation sociolinguistic arguments which highlight the differences in the conversational styles of men and women in terms of directness, elaboration and politeness (Lakoff, 1975; Tannen, 1990), and extends these arguments to the somewhat under-researched genre of the academic WhatsApp group. Female participants' messages were significantly longer and more elaborate, they used politeness markers more than three times as often as male participants and they invariably ended exchanges with a reciprocal expression of gratitude, consistent with Tannen's (1990) rapport-talk concept. Male participants, on the other hand, used a lot of imperative language and very short, transactional messages, ending their exchanges when the practical purpose was accomplished, and in keeping with Tannen's report-talk framework, and Baron's (2010) description of mobile messaging as a "terse, economical" medium, which described the male participants in this sample better than the female participants.

The results also call attention to one point in Herring and Paolillo (2006) that was raised in the study of "women's language": gendered stylistic markers are not consistent across all types of digital genres: hedging, one of the major features of the original account of "women's language" (Lakoff, 1975), is relatively uncommon in the data for both males and females, implying that hedging is not the norm in this particular group of academic WhatsApp exchanges, or that the function of hedging may be muted in this register for both genders. This aligns with Herring and Paolillo's 2006 general claim that the genre of digital communication, rather than gender identity, may better explain stylistic traits of digital communication. Likewise, the more balanced distribution of topic initiation rights in the mixed-gender threads suggests that power and style in the context of a task-driven academic genre are somewhat distinct; males and females seemed to be equally successful in initiating topics despite using quite different registers to do so.

The emoji findings expand on the literature about emoji use in WhatsApp which suggests that the use of emoji at a basic level of frequency alone may mask the qualitative differences in the functions of the emoji, as male participants used just one emoji for humor/deflection, whereas female users used a wider range of emoji for warmth and social bonding. This is in line with previous research, which found gender differences in emoji interpretation and use even when overall emoji use is the same. The results of code switching, on the other hand, indicate that the Internet language as described by Crystal (2001) is a hybrid of spoken and written language, and that these hybrid forms differ between the two gender groups in this bilingual academic setting: male participants freely interspersed into the core of their task-

oriented utterances, whereas female participants restricted Urdu-origin language to formulaic religious and social utterances, employing the majority of their substantive utterances in English.

Conclusion

The aim of this study was to examine the linguistic and interactional characteristics of male and female university students in academic groups on WhatsApp. The results suggest that, as in previous sociolinguistic research on gendered communication style (Lakoff, 1975; Tannen, 1990), male participants preferred to talk briefly, directly, and to deflect with humor, whereas female participants preferred to elaborate on the topic, be polite in a mitigated way, and reciprocate social closings, confirming a genre-sensitive description of gendered digital discourse (Herring & Paolillo, 2006) and extending it to a genre analysis of academic WhatsApp as a discourse community. Stylistically, importantly, these differences did not correlate with unequal distribution of the topic-initiation rights in the mixed-gender exchanges, so that it is important to view the use of gendered style and the use of gendered interactional power separately within this context. The study provides some context-specific evidence from a Pakistani higher-education environment for the scholarly discussion on the influence of gender and genre on communication in digital and academic environments.

Pedagogical and Practical Implications

The results have several practical implications for teachers and schools that use WhatsApp groups as informal communication tools in their classrooms. First, teachers planning for groups should realize that there may be gender differences in how instructions, deadlines, or requests for assistance are perceived and implemented by students and use adapted language to prevent such misunderstandings. It is important to recognize that male and female students may have different communicative styles, with the male students being more brief and direct in their approach, and the female students being elaborated and relationally oriented in their approach, which can lead to gender differences in how instructions, deadlines, and requests for help are perceived and executed. Second, very little hedging was used by both the male and female groups, indicating that the academic WhatsApp exchanges are less linguistically cautious than exchanges in face-to-face classroom interaction, a factor that may be taken into account when analysing students' writing on WhatsApp. Third, the even distribution of topic-initiation rights in mixed-gender threads is a positive sign that digital group communication can contribute to relatively balanced participation, even across the stylistic asymmetries mentioned elsewhere in this document, as part of broader institutional agendas for inclusive digital learning environments.

Limitations

There are a number of limitations to this study. The number of conversation threads and participants that were sampled was twenty and twelve, respectively, which was adequate for descriptive/exploratory content and discourse analysis but inadequate for inferential statistical testing or meaningful generalizations beyond the sampled groups. In the analysis, the potentially confounding variables of participants' academic discipline, year of study, digital literacy, and the specific social relationships among group members were not controlled for, as these variables could affect conversational style independently of gender, as described in the delimitations outlined in Chapter Three. Third, the data are only text-based exchanges, so the study does not address patterns of gendered voice notes, video calls or private one-on-one messaging, which may be different from the patterns of exchange reported here.

Recommendations for Future Research

Future studies could replicate this analysis on a significantly larger collection of academic group conversations discussed on WhatsApp, ideally from a variety of institutions and disciplines, to allow for inferential analysis by statistical testing (including independent

samples t-test or chi-square test) of the descriptive patterns discussed here. Future research may also consider conducting interviews or questionnaires with the participants to better understand if they are consciously aware of the differences, and if they affect their communication. In addition, researchers could consider whether the stylistic trends found in this study are moderated by the composition of the group (mixed-gender or all-female/female) and size of the group (small or large), and whether the relatively equal topic-initiation patterns found in this study are replicated in larger and more varied samples of mixed-gender and all-female/female groups. Finally, the findings of this study could be further replicated across other cultural and non-Pakistani contexts in higher education, to see how many of the patterns found in academic WhatsApp use are genre specific versus how many are more broadly associated with gendered digital communication.

References

- Baron, N. S. (2010). *Always on: Language in an online and mobile world*. Oxford University Press.
- Colley, A., & Todd, Z. (2002). Gender-linked differences in the style and content of e-mails to friends. *Journal of Language and Social Psychology*, 21(4), 380–392.
<https://doi.org/10.1177/026192702237956>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Crystal, D. (2001). *Language and the Internet*. Cambridge University Press.
- Herring, S. C. (2000). Gender differences in CMC: Findings and implications. *Computer Professionals for Social Responsibility Journal*, 18(1).
- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 338–376). Cambridge University Press.
- Herring, S. C., & Paolillo, J. C. (2006). Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10(4), 439–459. <https://doi.org/10.1111/j.1467-9841.2006.00287.x>
- Huffaker, D. A., & Calvert, S. L. (2005). Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication*, 10(2). <https://doi.org/10.1111/j.1083-6101.2005.tb00238.x>
- Lakoff, R. (1975). *Language and woman's place*. Harper & Row.
- Tannen, D. (1990). *You just don't understand: Women and men in conversation*. Ballantine Books.