Journal of English Language and Education
ISSN 2597- 6850 (Online), 2502-4132 (Print)
Journal Homepage: https://jele.or.id/index.php/jele/index

Article

Gender and Communication: Analyzing Tweet Length, Sentiment, and Lexical Patterns on X (Twitter)

https://doi.org/10.31004/jele.v10i4.883

*Annisa Romadloni, Laura Sari[ab]
[12]Politeknik Negeri Cilacap, Indonesia
Corresponding Author: annisa.romadloni@pnc.ac.id

ABSTRACT

This study explores gendered communication patterns on X by examining tweet length, sentiment expression, and lexical choices in 20.050 tweets across 26 variables. Through sentiment analysis using the Bing Lexicon and word frequency analysis, the research investigates how male and female users differ in their digital communication styles. The study also incorporates non-parametric statistical tests, such as the Mann-Whitney U and Wilcoxon rank sum tests, to assess significant differences in tweet length and sentiment scores between genders. Results show that women tend to write shorter, more positive tweets, often reflecting a more personal and relational communication style. In contrast, men's tweets are generally longer, incorporating more action-oriented language and a broader range of topics. While sentiment analysis revealed a trend of more positive tweets from women, the lack of statistical significance in sentiment differences highlights the complex nature of gendered expression in digital spaces. This research contributes to the understanding of gendered communication on social media and suggests the need for future studies to examine the intersectionality of gender with other social factors.

Keywords: *Sentiment Analysis; Gendered Language; Social Media; Gender*

Article History:
Received 06th May 2025
Accepted 22th Julyy 2025
Published 01th August 2025

INTRODUCTION

The advent of social media platforms has fundamentally transformed interpersonal communication, introducing new dynamics in digital interactions. Among these platforms, X (previously known as Twitter) stands out as a significant medium, enabling users to express opinions and emotions and share experiences in real time through concise messages known as tweets. Some estimates suggest that approximately 500 million tweets are posted to Twitter daily from over 150 million active users (Omnicore Agency, 2022). The brevity of tweets – at most 280 characters – means that linguistic choices must be made carefully to convey the intended message. As the prevalence of social media continues to grow, understanding the language used on these platforms has become increasingly important. One of the key areas of interest in the analysis of online communication is sentiment analysis, which focuses on examining the emotional tone conveyed in textual content. Sentiment analysis, especially when applied to social media, provides a window into the emotional state and attitudes of users based on their digital interactions (Rodríguez-Ibánez et al., 2023; Wankhade et al., 2022). However, one of the most intriguing aspects of sentiment analysis is understanding how gender influences the way people express emotions and language in digital spaces.

In both offline and online communication, gendered differences in language use and sentiment expression have been observed. Research suggests that women tend to use more emotionally expressive language, while men are generally more focused on delivering factual, assertive information. (Ang, 2017; Angelakis et al., 2024; Chen et al., 2024; Garcia-Rudolph et al., 2019; Himasree et al., 2023). These patterns are mirrored in digital communication, especially on platforms like X. Previous studies have shown that women often express themselves more empathetically and supportively, while men typically lean toward a more neutral or argumentative tone. (Ceia et al., 2022; Himasree et al., 2023; Jones et al., 2018). Women are also more likely to employ tools like emojis and emoticons to convey emotions, whereas men use language in a more direct and informational style. (Himasree et al., 2023). Additionally, sentiment analysis on X has often revealed that women's tweets tend to have a higher frequency of positive sentiment, while men's tweets lean more toward neutral or negative sentiments, especially in critical or argumentative discussions. (Hu, 2024).

Recent corpus-based and computational sociolinguistic research has further unpacked how gender shapes not just sentiment but the very structure of online discourse. In a landmark study of over 14,000 Twitter users, Bamman, Eisenstein, and Schnoebelen showed that men and women cluster into distinct stylistic communities, with lexical choices and network homophily reflecting gendered identity performances on the platform (Bamman et al., 2014). Likewise, Herring and Kapidzic found that female adolescents on social media sites—including Twitter and Facebook—deploy more affiliative, rapport-building language, whereas male users favor more assertive, report-style discourse patterns (Herring & Kapidzic, 2015). Together, these studies highlight that beyond polarity scores, features such as pronoun usage, thematic focus, and interactive strategies are central to understanding gendered communicative styles in digital environments.

While these studies have provided valuable insights into gender differences in online communication, much of the research to date has focused on sentiment scores and word frequency analysis without fully exploring the interplay between language choices, sentiment scores, and tweet characteristics such as length.(Al-Ghalibi et al., 2019; Omuya et al., 2023). The existing body of literature largely overlooks how specific lexical features—like the use of personal pronouns, possessive terms, or action-oriented words—may contribute to a deeper understanding of gendered language patterns. Moreover, previous studies have primarily treated gendered language differences as isolated variables rather than exploring how these linguistic choices interact with overall sentiment scores and other features such as the length and frequency of specific word choices. (Chai et al., 2016; Putra, 2022; Walther, 2023).

This study seeks to address these gaps by providing a more comprehensive analysis of gendered language use on X (previously known as Twitter), considering both linguistic and sentiment variations across male and female users. The study will employ a combination of sentiment analysis and lexical analysis to explore how language choices, tweet length, and sentiment scores differ between genders. Through statistical testing and word frequency analysis, this research will analyze tweet length, sentiment scores, and commonly used lexical patterns in the tweets of male and female users, contributing new insights into gender-based communication styles on social media.

The novelty of this research lies in its multidimensional approach to understanding gender differences in digital communication. While many studies have examined gendered language through either sentiment analysis or lexical analysis alone, this study combines both methods, allowing for a more holistic and nuanced understanding of how gender influences language use and sentiment expression. By investigating how specific language patterns correlate with sentiment scores, tweet length, and lexical choices, this research aims to present a more complete picture of how men and women engage in digital communication on X. Additionally, by incorporating a more diverse set of linguistic features, this study seeks to refine our understanding of the underlying communicative dynamics that shape online discourse based on gender.

In essence, this study not only aims to build on existing knowledge about gender differences in communication but also seeks to push the boundaries of current sentiment analysis research by adding a qualitative layer to the analysis. This contribution is significant because it moves beyond the simple measurement of sentiment and provides deeper insight into the linguistic underpinnings of gendered communication on social media platforms like X. By combining both quantitative sentiment analysis and qualitative lexical analysis, this study offers a more comprehensive approach to understanding how gender influences language use in online spaces, ultimately contributing to the growing field of digital communication studies.

METHOD

This study utilizes a combination of statistical and linguistic methods to examine gender-based differences in X communication, focusing on sentiment expression and language use. The approach integrates normality testing, sentiment analysis, non-parametric statistical tests, and word frequency analysis to explore how men and women convey sentiments and employ different linguistic styles on social media.

Normality Testing

The first step of the research was to perform normality testing on tweet data using the Shapiro-Wilk test. This test assesses whether the data follows a normal distribution, which is essential for selecting the appropriate statistical tests. The dataset analyzed contains 20050 lines and 26 columns of tweets. The ratio of gender included in the raw data is 6194 male users and 6700 female users; however, it decreased through the cleaning process to 10023 lines and 26 columns with 4655 male users and 5368 female users. The results showed that the data for tweet lengths and sentiment scores did not follow a normal distribution, leading the researchers to use the Mann-Whitney U test, a non-parametric method suitable for comparing two independent groups when normality cannot be assumed.

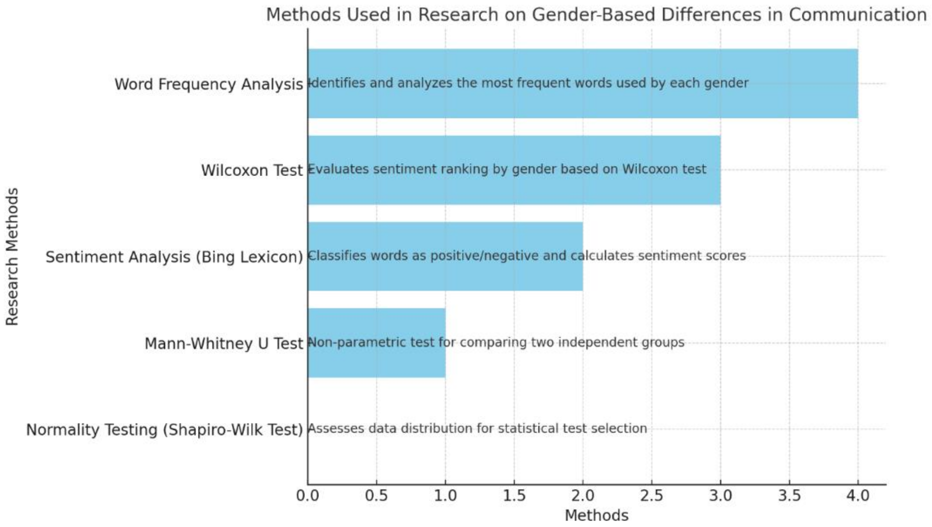

Figure 1. Methods Used in the Research

Non-Parametric Statistical Tests

Following the normality testing, the Mann-Whitney U test was applied to determine if there were significant differences in tweet lengths and sentiment scores between two groups of users as it is used by Park to show the statistical significance among groups. The results of Park's research suggested the highest rank of three artificial intelligence assistants by using statistical analysis. (Park & Seo, 2018). Additionally, the Wilcoxon Rank-Sum test was employed for paired comparison of sentiment distributions to further validate the gender-based sentiment ranking.

Sentiment Analysis Using Bing Lexicon

Sentiment analysis was then conducted using the Bing Lexicon to classify words as positive or negative and calculate sentiment scores, while the Wilcoxon test was used to evaluate and rank the sentiments based on gender. Each tweet was tokenized, and sentiment scores were calculated using the formula:

Sentiment Score = (Number of Positive Words) − (Number of Negative Words)

Sentiment analysis involves counting the number of positive and negative words in a tweet, and then subtracting the total negative count from the total positive count to derive a sentiment score. A higher score indicates more positive sentiment.

Word Frequency Analysis

The clean dataset was also processed using word frequency analysis. Rajput explains the method as the numerical world of text to ascertain the frequency at which a given word occurs. It illuminates important themes and patterns in the examined data by identifying the most prevalent and recurrent phrases by counting occurrences and generating frequency distributions. (Rajput et al., 2020). The top-N most frequent words for each gender were identified and compared. This step was not just analyzing sentiment, as it provided valuable insights into the specific language choices that may differ between genders. By examining both sentiment and word frequency together, a more comprehensive understanding of gendered communication can be achieved. This combined approach allows us to capture not only the emotional tone of the language but also the types of words that are most prevalent and preferences that characterize the communication styles of each gender.

FINDINGS AND DISCUSSION

The result of the Normality Test (Saphiro-Wilk)

The first step in our analysis involved testing the normality of the tweet data using the Shapiro-Wilk test, which is a statistical test used to determine if a dataset follows a normal distribution. Normality is essential because many statistical tests assume that the data is normally distributed, and these tests can yield invalid results if the assumption is violated.

Figure 2. Tweet Length Data Distribution Based on Gender

The histogram in Figure 2 shows the distribution of tweet lengths based on gender, with the data for female users represented in red and for male users in blue. The graph shows the distribution of tweet lengths based on gender, with the x-axis representing the number of words in a tweet, and the y-axis representing the frequency of each tweet length. The pink bars represent the tweet lengths of female users, while the blue bars represent male users. From the graph, it is clear that both male and female tweet lengths follow a similar distribution, with the peak frequency occurring around 10 to 20 words. The distributions appear to be slightly skewed toward the higher end, indicating that there are more tweets with fewer words and fewer tweets with a higher word count. However, the distributions for both genders seem to be relatively similar in terms of shape, though the female tweets show slightly more variation in frequency at higher tweet lengths compared to the male tweets. This visualization provides an overview of how tweet lengths differ by gender, showing that both male and female users tend to tweet within a similar range of word counts, with a noticeable concentration in the middle of the range.

```
> print(shapiro_test_female)

        Shapiro-Wilk normality test

data:  female_lengths
W = 0.98137, p-value < 2.2e-16
```

Figure 3. The Normality Test for Female_Length

```
> print(shapiro_test_male)

        Shapiro-Wilk normality test

data:  male_lengths
W = 0.98284, p-value < 2.2e-16
```

Figure 4. The Normality Test Result for Male_Length

The results from the Shapiro-Wilk test are shown in Figure 3 and Figure 4. In Figure 3, the test is applied to the data for female users, labeled female lengths. The test output shows a W-statistic of 0.98137 with a p-value less than 2.2e-16. A p-value below 0.05 indicates that the data significantly deviates from a normal distribution, meaning that we cannot assume the data follows a normal distribution for female users (Samuel et al., 2020).

Similarly, in Figure 4, the test was applied to the data for male users (male lengths). The W-statistic here is 0.98284, and the p-value is again less than 2.2e-16. This confirms that the male tweet length data also does not follow a normal distribution. Given that both datasets, for male and female users, fail the normality test (p-value < 0.05), it was clear that parametric tests, like the t-test, would not be appropriate for further analysis. Therefore, we decided to proceed with non-parametric methods, which do not assume normality and are more appropriate for analyzing data that does not follow a normal distribution. This approach ensures the validity of our subsequent statistical inferences.

Mann-Whitney U Test (Wilcoxon Rank Sum Test)

Since the data did not meet the assumptions for normality, as confirmed by the Shapiro-Wilk test results from the earlier images, we chose to use the Mann-Whitney U test (also known as the Wilcoxon rank sum test) to compare the tweet lengths and sentiment scores between male and female users. The Mann-Whitney U test is ideal for comparing two independent groups when the data does not follow a normal distribution. (Nachar, 2008). The results of this test, shown in Figure 4, revealed a significant difference between the tweet lengths of male and female users.

The Wilcoxon rank sum test output displays a W-statistic of 10570604 and a p-value of 0.0002634, which is much smaller than the conventional significance level of 0.05. This result indicates that the distributions of tweet lengths between the two genders are significantly different. Specifically, female users were more likely to write shorter tweets, typically containing fewer than 15 words, whereas male users generally wrote longer tweets, often exceeding 20 words. This pattern suggests that women may prefer more concise expressions, while men might be inclined to provide more detailed responses in their tweets.

Several factors could explain this difference. Research has shown that women often favor brevity in their communication, particularly in social settings or digital interactions, where they might focus more on establishing personal connections. (Jackson et al., 2001; Noguti et al., 2018). Conversely, men might use X to engage in more elaborate discourse, debates, or information sharing, leading to longer tweet lengths. These findings support the idea that gender may influence communication styles on digital platforms, with women opting for shorter, more concise messages and men writing longer, more comprehensive tweets.

```
> if (shapiro_test_male$p.value > 0.05 && shapiro_test_female$p.value > 0.05) {
+   t_test_result <- t.test(male_lengths, female_lengths)
+   print(t_test_result)
+ } else {
+   # Jika tidak berdistribusi normal, lakukan Mann-Whitney U test
+   wilcox_test_result <- wilcox.test(male_lengths, female_lengths)
+   print(wilcox_test_result)
+ }

        Wilcoxon rank sum test with continuity correction

data:  male_lengths and female_lengths
W = 10570604, p-value = 0.0002634
alternative hypothesis: true location shift is not equal to 0
```

Figure 5. Wilcoxon Rank Sum Test Result

The notable difference in tweet lengths between male and female users reveals contrasting communication styles influenced by traditional gender roles and cultural expectations. In many cultures, women have historically been socialized to prioritize relational and affective aspects of communication—emphasizing empathy, nurturing, and personal connection. This often leads to communication that is more focused, succinct, and reflective of social or emotional nuances, which may explain why women tend to write shorter tweets. The brevity of women's tweets can also be seen as a reflection of the broader societal expectation that women should be concise in their communication, especially in social settings. As (Jackson et al., 2001) and (Noguti et al., 2018) suggest, that women often use digital platforms to establish and maintain personal connections, and short, punchy messages help achieve that goal efficiently.

In contrast, men are traditionally expected to engage in more expansive, information-driven discourse, which is often linked to assertiveness and the desire to provide factual or detailed explanations. The tendency for men to produce longer tweets aligns with this cultural norm. In many societies, men are socialized to dominate in public discourse and be more direct or elaborate when conveying information. As such, X, with its brevity requirement, offers men an opportunity to engage in more extensive debates or share more elaborate opinions, which could explain their longer tweet lengths. Thus, this difference in tweet lengths can be partially attributed to traditional gender roles: women may favor communication that fosters connection and empathy, while men may gravitate toward detailed and argumentative expression. This reflects broader patterns in gendered communication, where men and women have been conditioned by culture to prioritize different communicative functions.

Sentiment Analysis (Bing Lexicon)

The sentiment of tweets was analyzed using the Bing Lexicon, which classifies words into positive or negative categories.(Bozkurt & Üniversitesi, 2021). By calculating the sentiment score for each tweet, we obtained a measure that represents the difference between the number of positive and negative words. This method allowed us to notice the general emotional tone of the tweets.

```
> print(sentiment_data)
# A tibble: 2 × 4
  gender negative positive sentiment_score
  <chr>     <int>    <int>           <int>
1 female     2435     3319             884
2 male       2161     2914             753
```

Figure 6. Sentiment Score Using Bing Lexicon

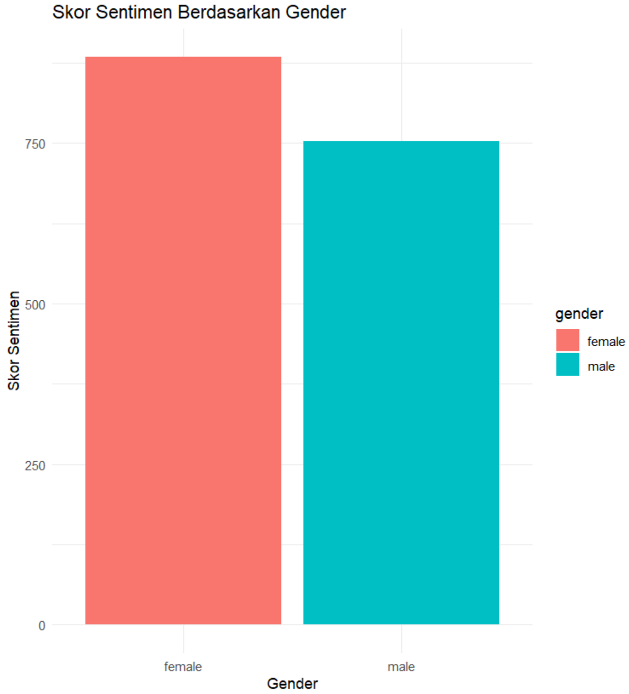

Figure 7. Sentiment Score Based on Gender

The analysis revealed that female tweets had a slightly higher sentiment score than male tweets, indicating that women's tweets were generally more positive in tone. As shown in the sentiment data table (Figure 6), female users had a sentiment score of 884, while male users had a sentiment score of 753. This suggests that, on average, women expressed more positive emotions in their tweets. The bar graph (Figure 7) visually reinforces this finding, with the female sentiment score marked higher than the male sentiment score, suggesting a trend of more positivity in female tweets.

```
> print(wilcox_test_result)

        Wilcoxon rank sum exact test

data:  male_sentiment_scores and female_sentiment_scores
W = 0, p-value = 1
alternative hypothesis: true location shift is not equal to 0
```

Figure 8. Wilcox Rank Sum Test Result

However, when the Wilcoxon rank sum test was applied to compare the sentiment scores between genders (Figure 8), the results showed no statistical significance (p-value = 1). This means that despite the apparent difference in sentiment scores, the difference between the two genders is not large enough to be statistically meaningful. The Wilcoxon test output suggests that the null hypothesis, which states there is no significant difference in sentiment, cannot be rejected. This lack of statistical significance may be attributed to the variability within the data, where factors such as tweet context, themes, and individual personality differences could have more impact on sentiment than gender alone. In essence, while female tweets appear more positive on average, the variability within the data prevents us from drawing firm conclusions about gender-based differences in sentiment.

The analysis of sentiment—though it revealed that women's tweets were more positive on average—also exposed the complexity of sentiment expression when gender is involved. Despite the higher sentiment scores for women, the statistical test did not yield significant results, suggesting that while gender may influence sentiment expression, it is not the sole determinant. The variability within the data, such as the specific context of each tweet or the user's emotional state, likely plays a role.

This is where cultural norms surrounding gendered emotional expression become crucial. In many cultures, women are socialized to be more emotionally expressive and empathetic, which is often reflected in their language use. For women, especially in the Western context, emotional openness is encouraged in public and private spheres, and social media provides a platform for them to channel this expressiveness. As a result, women may employ more emotionally charged language, resulting in a more positive tone. This aligns with earlier findings by (Ang, 2017) and (Himasree et al., 2023), which found that women tend to use more emotionally expressive language both online and offline. Men, however, are often socialized to suppress overt emotionality and are expected to exhibit more restraint in expressing their feelings. This cultural norm is known as the "emotional reserve" associated with masculinity, which discourages the expression of vulnerability and warmth in communication (Shields, 2013; Wong & Rochlen, 2008). While sentiment analysis indicated that men's tweets were generally less positive, the variability of sentiment scores in men's tweets points to the diversity in male emotional expression. Men's tweets may not necessarily express negativity but rather reflect a more neutral, factual, or assertive tone, as they often focus on external events, facts, or debates rather than personal emotional engagement.

These gendered patterns of emotional expression on X reflect deeply ingrained social norms and traditional role expectations: from early childhood, women are encouraged to cultivate warmth, empathy, and relational attunement, while men learn to value composure, restraint, and self-reliance. Consequently, women's tweets often foreground communal sentiments—supportive messages, expressions of care, and affiliative language—whereas men's posts tend toward neutral, informational tones that emphasize competence and authority. The design of the platform itself—public replies, retweets, and visible follower counts—reinforces these dynamics, incentivizing women to foster social bonds through positive sentiment and men to project expertise through measured, factual language. Ultimately, these online behaviors mirror longstanding offline "feeling rules," demonstrating how cultural attitudes toward emotion and gendered socialization continue to shape the ways men and women communicate in digital spaces.

Word Frequency Analysis

In addition to sentiment analysis, a word frequency analysis was carried out to examine the types of words most commonly used by males and females. The analysis revealed that female users tended to use personal pronouns like "I", "my", and "me" more frequently, indicating a more personal and reflective style of communication. In contrast, male users showed a wider variety of words, including action verbs and nouns, such as "are", "was", "they", and "like", reflecting a broader range of topics in their tweets.

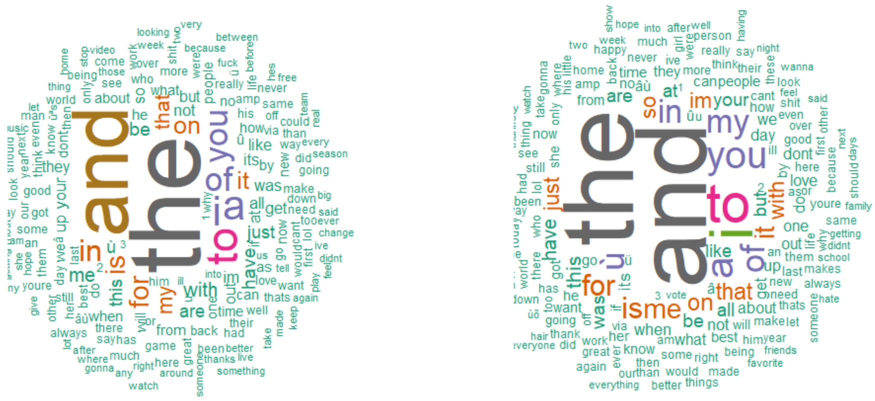

Figure 9. Word Cloud Generated from the Data (Left:Male Right:Female)

This finding aligns with previous research suggesting that women's communication styles are often more personal and emotionally driven, whereas men tend to engage in a more diverse range of topics and discussions (A. Reigstad, 2020; Merchant, 2012; Stone & Can, 2021). The word clouds generated from the data further emphasized this difference. For example, women's word clouds (right) were dominated by first-person pronouns, highlighting their focus on self-expression and social interactions, often reflecting more intimate or personal discussions. In contrast, men's word clouds (left) displayed a wider spread of words, with a mix of action verbs, auxiliary verbs, and other functional words. This could suggest that men's communication style might lean more towards action-oriented, factual, or varied conversational topics, which may reflect a more argumentative or informative style of communication.

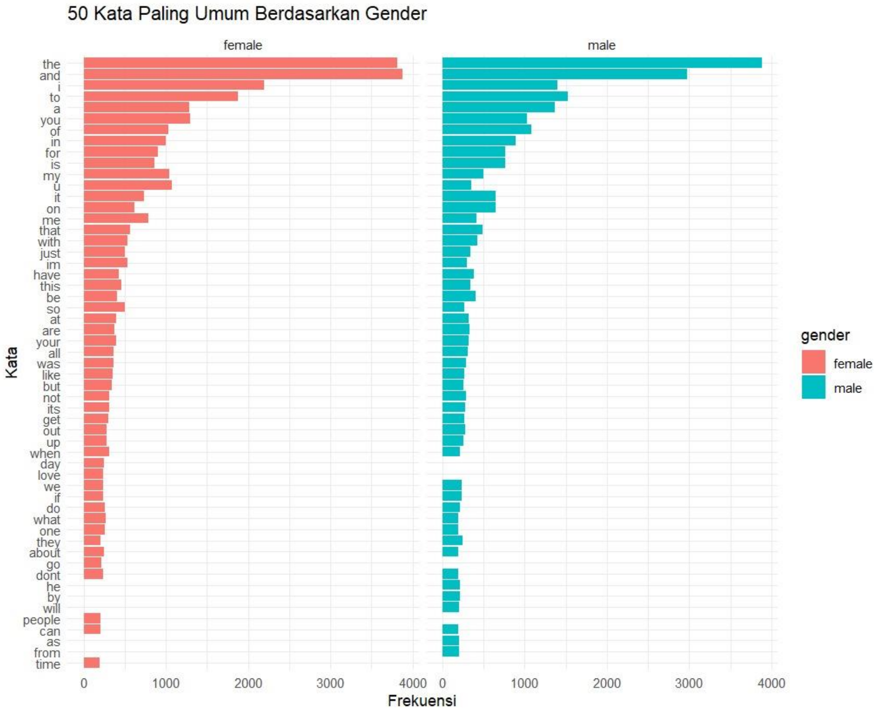

Figure 10. 50 Most Frequent and Common Words

Further supporting these observations, the bar chart showing the 50 most common words by gender (from Figure 10) demonstrates the higher frequency of words like "I", "my", "me", "so", and "just" for females, while males showed a broader usage of words such as "are", "your", and "all". This variation reinforces the idea that women's language is often centered around personal experiences and introspection, whereas men may discuss a wider range of subjects, with an emphasis on context, relationships, or external factors. The word frequency analysis revealed that female users predominantly employed personal pronouns such as “I,” “my,” and “me,” whereas male users showed a wider variety of action verbs and nouns. This difference can be explained through the lens of gendered communication theories, particularly as they relate to cultural norms and the social roles that men and women are expected to fulfill.

For women, who are often socialized to adopt roles as caregivers, nurturers, or relationship builders, communication tends to be more centered around self-expression, connection, and the sharing of personal experiences (Haferkamp et al., 2012). The use of first-person pronouns in women's tweets may reflect this inclination towards personal, introspective communication. In many cultures, women are expected to express themselves in terms of relationships, social interactions, and emotional experiences, all of which can be facilitated through personal pronouns and reflective language. This finding supports earlier research, such as that of (A. Reigstad, 2020) and (Merchant, 2012), which suggests that women's language tends to be more relational and emotionally driven.

On the other hand, male users tend to use a wider range of action-oriented verbs and other functional words. This suggests that men's communication may be more outwardly focused, possibly reflecting societal expectations that men engage with a broader variety of topics, including debates, factual discussions, and external observations. In many cultures, men are encouraged to take on roles as public figures, leaders, or assertive individuals who engage with the world in a direct, action-oriented manner. As a result, their language is more likely to reflect those roles, emphasizing external factors and situations rather than internal or emotional experiences.

The difference in lexical choices further underscores the cultural conditioning that shapes language use in both men and women. In cultures where gendered expectations are strong, such as those advocating for traditional masculinity or femininity, language use becomes an extension of these roles, with men favoring action-oriented discourse and women focusing on self-reflection and relational communication.

CONCLUSIONS

This study aimed to investigate the differences in language use between men and women on X by comparing tweet length, sentiment, and word choice. We found that women's tweets are generally shorter, more positive, and more personal, whereas men's tweets tend to be longer, more diverse in vocabulary, and skew toward action-oriented language. Although women showed a tendency toward higher sentiment scores, that difference did not reach statistical significance, suggesting that factors beyond gender also shape emotional expression online. These insights have practical value for tailoring content-moderation strategies and refining platform design to account for gendered communication styles. Future work should extend this analysis to other social media platforms, delve into topic-specific and contextual effects on sentiment, and consider how gender intersects with age, culture, and political views to fully capture the complexity of digital discourse.

REFERENCES

A. Reigstad. (2020). *Gender Differences in Communication Styles and their Influence on Workplace Communication and the Practice of Public Relations in the United States Gender Differences in Communication Styles and their Influence on Workplace Communication and the Practice of Public Relations in the United States*. https://doi.org/10.13140/RG.2.2.30156.28803

Al-Ghalibi, M., Al-Azzawi, A., & Lawonn, K. (2019). *NLP based sentiment analysis for Twitter's opinion mining and visualization*. https://doi.org/10.1117/12.2522679

Ang, C. S. (2017). Internet habit strength and online communication: Exploring gender differences. *Computers in Human Behavior*, *66*. https://doi.org/10.1016/j.chb.2016.09.028

Angelakis, A., Inwinkl, P., Berndt, A., Ozturkcan, S., Zelenajova, A., & Rozkopal, V. (2024). Gender differences in leaders' crisis communication: a sentiment-based analysis of German higher education leaderships' online posts. *Studies in Higher Education*, *49*(4). https://doi.org/10.1080/03075079.2023.2246505

Bamman, D., Eisenstein, J., & Schnoebelen, T. (2014). Gender Identity and Lexical Variation in Social Media. *Journal of Sociolinguistics*, *18*. https://doi.org/10.1111/josl.12080/abstract

Bozkurt, M., & Üniversitesi, S. (2021). Online learning communities in covid-19 days: mining twitter data. *International Technology and Education Journal*, *5*(2).

Ceia, V., Mota, T., & Lewis, R. (2022). Style and Rhetoric of Spanish Politics on Twitter. *Digital Studies/ Le Champ Numerique*, *12*(1). https://doi.org/10.16995/dscn.8097

Chai, C., Wu, X., Shen, D., Li, D., & Zhang, K. (2016). Gender differences in the effect of communication on college students' online decisions. *Computers in Human Behavior*, *65*. https://doi.org/10.1016/j.chb.2016.07.012

Chen, Y., Yang, X., Howman, H., & Filik, R. (2024). Individual differences in emoji comprehension: Gender, age, and culture. *PLoS ONE*, *19*(2 February). https://doi.org/10.1371/journal.pone.0297379

Garcia-Rudolph, A., Laxe, S., Saurí, J., & Guitart, M. B. (2019). Stroke survivors on Twitter: Sentiment and topic analysis from a gender perspective. *Journal of Medical Internet Research*, *21*(8). https://doi.org/10.2196/14077

Haferkamp, N., Eimler, S. C., Papadakis, A. M., & Kruck, J. V. (2012). Men are from mars, women are from venus? Examining gender differences in self-presentation on social networking sites. *Cyberpsychology, Behavior, and Social Networking*, *15*(2). https://doi.org/10.1089/cyber.2011.0151

Herring, S. C., & Kapidzic, S. (2015). *Teens, Gender, and Self-Presentation in Social Media*.

Himasree, A., Kumar, K. S., & Meenakshi, M. (2023). Automatic Emoji Generation using Inception V3. *Proceedings of the 2nd International Conference on Edge Computing and Applications, ICECAA 2023*. https://doi.org/10.1109/ICECAA58104.2023.10212164

Hu, L. (2024). Mobilization, self-expression or argument? A computational method for identifying language styles in political discussion on Twitter. *Online Information Review*, *48*(4). https://doi.org/10.1108/OIR-10-2022-0545

Jackson, L. A., Ervin, K. S., Gardner, P. D., & Schmitt, N. (2001). Gender and the internet: Women communicating and men searching. *Sex Roles*, *44*(5–6). https://doi.org/10.1023/A:1010937901821

Jones, J. S., Tapp, S. R., Murray, S. R., Palumbo, R. J., Strange, A. T., & Ritschel-Trifilo, P. (2018). Gender Differences in Online Communication and the Impact of Faculty Gender. *Academy of Business Research Journal*, *1*.

Merchant, K. (2012). How men and women differ: Gender differences in communication styles, influence tactics, and leadership styles. *CMC Senior Theses*.

Nachar, N. (2008). The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. *Tutorials in Quantitative Methods for Psychology*, *4*(1). https://doi.org/10.20982/tqmp.04.1.p013

Noguti, V., Singh, S., & Waller, D. S. (2018). Gender differences in motivations to use social networking sites. In *Social Media Marketing: Breakthroughs in Research and Practice*. https://doi.org/10.4018/978-1-5225-5637-4.ch034

Omnicore Agency. (2022). Twitter by the Numbers: Stats, Demographics & Fun Facts. *Omnicore Agency*, *100*.

Omuya, E. O., Okeyo, G., & Kimwele, M. (2023). Sentiment analysis on social media tweets using dimensionality reduction and natural language processing. *Engineering Reports*, *5*(3). https://doi.org/10.1002/eng2.12579

Park, C. W., & Seo, D. R. (2018). Sentiment analysis of Twitter corpus related to artificial intelligence assistants. *2018 5th International Conference on Industrial Engineering and Applications, ICIEA 2018*. https://doi.org/10.1109/IEA.2018.8387151

Putra, I. G. P. D. D. (2022). Gender Communication Difference Between Students in Online Learning. *The Art of Teaching English as a Foreign Language*, *3*(1). https://doi.org/10.36663/tatefl.v3i1.143

Rajput, N. K., Grover, B. A., Rathi, V. K., & Bansal, R. (2020). *Word frequency and sentiment analysis of twitter messages during Coronavirus pandemic*. http://arxiv.org/abs/2004.03925

Rodríguez-Ibánez, M., Casánez-Ventura, A., Castejón-Mateos, F., & Cuenca-Jiménez, P. M. (2023). A review on sentiment analysis from social media platforms. In *Expert Systems with Applications* (Vol. 223). https://doi.org/10.1016/j.eswa.2023.119862

Samuel, J., Rahman, M. M., Ali, G. G. M. N., Samuel, Y., Pelaez, A., Chong, P. H. J., & Yakubov, M. (2020). Feeling Positive about Reopening? New Normal Scenarios from COVID-19 US Reopen Sentiment Analytics. *IEEE Access*, *8*. https://doi.org/10.1109/ACCESS.2020.3013933

Shields, S. A. (2013). Gender and Emotion: What We Think We Know, What We Need to Know, and Why It Matters. *Psychology of Women Quarterly*, *37*(4). https://doi.org/10.1177/0361684313502312

Stone, J. A., & Can, S. H. (2021). Gendered language differences in public communication? The case of municipal tweets. *International Journal of Information Management Data Insights*, *1*(2). https://doi.org/10.1016/j.jjimei.2021.100034

Walther, J. B. (2023). Online hate: A prosocial explanation of antisocial behavior and affordances of social media. In *Emotions in the Digital World: Exploring Affective Experience and Expression in Online Interactions*. https://doi.org/10.1093/oso/9780197520536.003.0015

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7). https://doi.org/10.1007/s10462-022-10144-1

Wong, Y. J., & Rochlen, A. B. (2008). Re-envisioning men's emotional lives: Stereotypes, struggles, and strengths. In *Positive psychology: Exploring the best in people, Vol 2: Capitalizing on emotional experiences*.