



## **Gender Representation in News Headlines: A Computational Corpus Analysis**

**Saira Arshad**

**MPhil Scholar, Department of English, NUML University, Faisalabad Campus**

*Email: [sairaarshad5157@gmail.com](mailto:sairaarshad5157@gmail.com)*

**Dr. Aftab Akram**

**Lecturer, Department of English, NUML University, Faisalabad Campus**

*Email: [aakrum@numl.edu.pk](mailto:aakrum@numl.edu.pk)*

### **Abstract**

*In this research, the linguistic construction of gender representation in newspaper headlines in English is explored using an integrated approach of feminist and computational corpus methodology. Utilising the feminist linguistic theory and the critical discourse analysis, the study focuses on revealing the role of the less apparent lexico-grammatical decisions made in the headlines to either reinforce or challenge the gendered identity. Corpus analysis software was applied to a self-compiled collection of newspaper headlines in order to determine frequency distributions, keyword distributions, and collocational patterns related to gender-related words. The results show that there are systematic asymmetries in terms of representation, as female actors tend to be linked with victimization, domesticity, and relational categories of themes, whereas the male actors are mostly situated in the areas of authority, politics, and economic power. The results of collocational analysis also reveal the presence of repetitive semantic frames, which foreground or background agency based on gender, indicating that headline construction is a strong ideological process. Through quantitative methodologies of computation and qualitative methods of critical interpretation, the research offers empirical information about the ongoing gender biasness in the headline level. The study adds to the increasing literature in corpus-aided studies of discourse by anticipating headlines as a genre with unique and powerful media practices, as well as applying the computational linguistic research to the Pakistani media domain. The results highlight the fact that the study of headline discourse is a very crucial area where language, ideology, and social power can be involved in constructing the way people think.*

**Keywords:** *News headlines, Computational linguistics, collocations, Gender Representation, Feminist Linguistics, Antconc*

### **Introduction**

The language is very strong in influencing our perception of the world. Headlines in the news are highly significant in that they are usually the introduction, and even the only segment of a news item that individuals read. Headlines do not just present the information; they present issues in a way that puts them in perspective, emphasizing some particulars and shaping the opinion of people. Owing to this fact, the language employed in the headlines can influence the presentation of various social groups in the society, such as men and women.

The issue of gender representation in the media has dominated in linguistic and media studies. Scientists have presented the view that not all news is represented equally between men and women. In most instances, leadership, politics, and professional success are more often linked to men, whereas family roles, appearance, or to victimhood are more often associated with



women. These trends could be indicators of more underlying social ideologies and cultural demands regarding gender roles.

In the Feminist Linguistics point of view, there is nothing neutral about language, as gender inequalities can be reflected and reproduced in language. The naming, description or framing used on men and women in headlines might support old stereotypes. In the same manner, Framing Theory implies that the manner and manner in which news is reported determines the perception that the audiences have over events and individuals. Hence, the investigation of gendered news headlines can contribute to the display of hidden ideology and power dynamics.

In the last few years, computational aids and corpus-based techniques have enabled the systematic analysis of large amounts of text. Researchers are now able to examine hundreds or thousands of headlines to determine patterns of frequency of words, collocations, and trends of representation than previously simply by using small examples. This quantitative method will give more valid and objective data regarding the construction of gender in media language.

The current research seeks to explore the gender reflection in online news headlines by way of computational corpus. Based on the frequency of male and female references, collocational patterns, and framing strategies, this study intends to understand the process of gender identities formation and whether there is an imbalance or ideological bias in news coverage. This analysis will aim to add to the current debate on the subject of media language, gender equality, and the social effects of the news discourse.

### **Research Questions**

1. How do the frequencies of male and female references in online news headlines reflect ideological meanings about gender roles?
2. What types of words and collocations are commonly associated with male and female terms in the headlines?
3. Are there noticeable differences in how men and women are framed in news headlines?

### **Research Objectives**

1. To examine the frequency of male and female references in online news headlines to identify ideological representations of gender roles.
2. To analyze the words and collocations associated with male and female terms in the headlines.
3. To investigate differences in the framing of men and women in online news headlines.

### **Statement of the Problem**

The paper focuses on the subject of representation of men and women in online news headlines, which have a strong influence on how the general population perceives the news as they are concise, captivating, and highly read. Media language might not necessarily provide gender balance, and it might often be related to gender identity, i.e. associating power and authority to men, and domesticity, looks, or victimhood to women. Though gender representation has been explored in the past literature, much of these works are based on small samples, and primarily qualitative research, leaving a vacuity of large-scale, systematic, and computational



corpus-based research with feminist and framing attitudes. Hence, this paper will seek to present a detailed, empirical investigation of how frequently, collocating with male and female references appear in the news headlines, and how they are framed, to establish the potential ideological affiliations with the use of media language.

### **Significance of the Study**

This work is important as news headlines are one of the strongest ways to impact the opinion of the people, and as most individuals read only headlines, the manner in which men and women are presented can affect society perceptions regarding gender roles and identities. Through the representation of genders, the study will determine whether the media can showcase men and women equally or can reinforce stereotypes. In linguistics, it leads to corpus linguistics through offering trustworthy, quantitative data of a vast number of data across a computer tool as opposed to subjective views. The Feminist Linguistics, as well as Framing Theory, are also supported in the study since it demonstrates that the use of language and words in headlines could represent gender inequality and influence the way readers perceive it. It can also sensitize journalists and media houses on the unconscious gender bias and promote a more balanced reporting. Lastly, it helps in sociolinguistics, media discourse analysis and gender studies particularly in situations where computational studies of news headlines remain scarce, and as an underpinning of future research on other media platforms.

### **Literature review**

Extensive research in the field of language and gender indicates that news media speech is at the forefront of constructing and reproducing gender identities by making minor linguistic decisions. The argument by the feminist linguists is that language is a mirror of structural power relations, which puts women in subordinate or stereotypical positions and takes male power as natural (Lakoff, 1975). Subsequent critical discourse studies broaden this perspective revealing that media discourses such as headlines encode ingrained gendered ideologies via lexicogrammatical choices and narrative frames of reference that create informational images on the part of the audience (van Dijk, 1998).

Linguistic studies, therefore, have prefigured the place of grammar, agent, and transitivity formations working ideologically to foreground or background actors in media discourses, a point of view underpinning computing corpus research of news coverage. Systematic gender biases in media language are further justified by recent research that is specifically related to the context of political and social media.

Later studies in the field of feminist media studies give empirical evidence of the systematic gender bias in the news discourse of the strongest nature. For instance, Monika Bednarek and Helen Caple (2012) discuss the news discourse and illustrate how the linguistic and visual decisions in the headlines and reports can be used as the way of constructing social meanings, including gendered ones. On the same note, Michelle M. Lazar (2005) asserts that media texts tend to reproduce patriarchal ideologies through the construction of women in constrained and stereotypical ways that uphold systems of greater gender inequality. In addition to this, Sara Mills (2008) presents the argument that discourse structures often place women as passive or relationship figures, and that men tend to be represented as authoritative figures and



agents. Corpus-based discourse research also evidences that gender bias is not only at the lexical level but it is also applied with more comprehensive patterns, like structural and thematic.

As revealed in corpus-assisted discourse studies, gender bias in news coverage is not only restricted to the frequency of lexical expression but its scope also encompasses structural patterns. Investigations into gender in news using topic modelling on a large scale indicate that the voices of women are over-represented in lifestyle or health-related courses, but the voices of men are over-represented in economically or politically important ones, which supports the division of domains by gender in media coverage. There is empirical research on media presentation of women that establishes that women are often sidelined or they are presented in a different manner than men are in news presentation.

Shari L. Collins (2011) proves that women are underrepresented in the hard news like politics and economics, whereas they are overrepresented in lifestyle and human-interest news. In the same way, Elizabeth A. Skewes and Holly E. H. Scott (2003) present the fact that women are more apt to be depicted as the person in relation or as a woman of beauty, men are depicted as people of power and authority. These results are consistent with the more general claim that media talk can help support gendered social orders. Although this body of work is expanding, comparatively little has been done on news headlines as a single genre of discourse.

Headlines, by their concise, salient nature, are primary interpretive frames to their readers, but much of the available corpus research favors the whole article over analysis of headlines. The news texts tend to support the conventional gender roles and expectations (Caldas-Coulthard, 1995). Besides, whereas feminist and framing theories have been extended to newspaper materials on a general basis, limited literature has explicitly combined theoretical frameworks with the methodology of computational corpus tools to examine the role of gendered headline formations in non-Western media settings, creating a knowledge gap as to how such headline formations can be constructed in Pakistani media. . According to research by Paul Baker (2006), corpus linguistics can be used in a systematic way to reveal patterns of representation and bias in large volumes of data, thus allowing researchers to shift beyond anecdotal knowledge.

Similarly, Tony McEnery and Andrew Hardie (2012) emphasize the value of quantitative approaches to determining collocational patterns and semantic associations that will point to underlying ideological structures in the discourse. Media studies are especially one of the fields where these methods are useful as large volumes of news texts can be analyzed to reveal common gendered patterns. In addition to this, the intersectional studies of the recent interdisciplinary work in the field of Computational Linguistics and Critical Discourse Analysis could be highlighted, suggesting the importance of corpus-assisted methodologies in uncovering the ideological trends inherent in the discourse of media.

Through quantitative production coupled with qualitative explanation, researchers will be in a position to transcend claims of anecdotes and prove statistically backed tendencies of presentation. Such a methodological change has been especially notable in the field of gender studies, which now offers computational modelling as a means of scholars to follow recurrent correlations between gendered terms and particular areas of thematic concern: crime, politics,



economics or family life. These methods are particularly applicable to situations in which extensive digital news archives exist but have not hitherto been systematically investigated by means of computational interrogation.

Although such a body has grown, however, there is a relative dearth of research on news headlines as a specific genre of discourse applying specifically a computational framework. The brevity, salience, and framing capability of headlines are very different when compared to full news texts, but much corpus research still gives preference to long texts. Moreover, although the use of computational techniques is generally more widespread within the Western context of media, its combination with the feminist linguistic theory and framing analysis in non-Western contexts has not been developed yet. There are not many studies in the Pakistani context of the media that implement a merge of quantitative corpus tools with fundamental theoretical approaches to explore the construction of gendered meaning through linguistic decisions at the headline level.

Considering this literature, it is possible to combine the theory of feminist linguistics with the computational analysis of corpus, so that the analytical investigation of the formation of gendered subjectivity in news headlines could become strong. This integrative strategy not only develops on conventional critical discourse models, but also utilizes quantitative data of lexical frequency, collocation, and syntactic patterns of language to display systematic predispositions. This kind of holistic inquiry about language is necessary to further the study of media and gender and to assist researchers to trace language, ideology, and social power interactions to form the discourse of the people.

### **Research Methodology**

The study is grounded on Feminist Linguistics and Framing Theory in order to explore the issue of gender representation on online news headlines. Feminist Linguistics considers language as an establishment, which depicts and sustains power relations between men and women. It claims that media discourse practices of visibility, role assignment, and word selection are not neutral but can indicate the voices and authority of more favored voices.

The Framing Theory also describes how the interpretation of reality through the use of headlines influences people by emphasizing some facts and downplaying some of the facts. Gender ideologies may be manifested in how headlines are repeated due to the fact that headlines form the first impression about women and have become associated with victims or family problems and men and leadership or powers. These structures, combined, can be used to describe the unfairness of the representation as well as the meanings that are generated by repetition of linguistic structures.

The research methodologically employs the corpus-based quantitative research approach. The ideological structures that are buried in discourse can be uncovered through quantitative methods of corpus (McEnery and Hardie, 2012). A corpus of online news headlines in the English language was created digitally to showcase the selection of news headlines in the selected newspapers over a certain period. Corpus-based methodologies enable the researcher to discover the patterns of representation in masses of data in a systematic way (Baker, 2006). Words that were related to gender like pronouns, nouns, and job titles were recognized and



analysed with the help of the AntConc software. The analysis will involve three steps, which are frequency analysis to estimate the frequency of occurrence of male and female references, collocation and concordance analysis to determine common word association and context, and comparative analysis to determine variation in thematic patterns and role construction. The quantitative model then gets interpreted using the Feminist Linguistics and Framing Theory to establish whether the patterns portray the traditional gender roles and ideological framing in the media discourse.

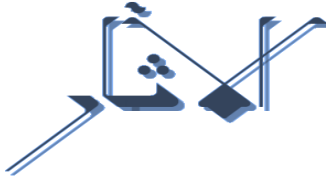
**Theoretical Framework:**

The Feminist Linguistics and Framing Theory are the foundation of this study and make it a critical approach to analyzing the issue of gender in online news headlines. Feminist Linguistics comes into the picture as a social and ideological system and an agent of reinforcing and reflecting the power relationships between men and women. Initial theoretical research by Robin Lakoff (1975) asserted that the subservient status of women in society was usually encoded in the linguistic formations. Lakoff emphasized that language can unfairly place women in less powerful positions by describing things in specific ways, in a polite manner, and expectations of the roles. Further feminist critics developed this line of thought by showing that gender bias is not just present in the vocabulary but wider forms of representation, visibility, and role distribution in discourse. In this perspective, it is not a neutral viewpoint regarding the frequency of the occurrence of men and women in news headlines. Rather, it can signify whose experiences, behaviours and authority are given greater priority within the discourse of the population.

Another focus of feminist linguistic studies is that media discourse is very influential in creating social understanding of gender roles. The newspapers and other media sources do not merely cover events; they also create social realities by creating a word-based construction. The choice of subjects, the actions they are supposed to perform and the settings in which they are represented may either be used to uphold or break down the already existing stereotypes based on gender. In most instances in the media, men tend to be linked with power, leadership, and power as compared to women who are often associated with family themes, victimization or emotional scenes. These trends show how language helps to support wider social hierarchies. Thus, the analysis of gendered linguistic patterns of news headline offers valuable information on the ways in which media discourse can reinforce or oppose patriarchal societal organization.

The Framing Theory adds even more strength to this analytical basis by clarifying the impact of media texts on the way they are perceived by the audience. Framing, according to Robert Entman (1993), is the process of choosing certain things in the reality and highlighting them in the process of communication to enable certain kinds of interpretation to capture a lot of attention. Media framing emphasizes certain aspects of an occurrence at the expense of others and thus it influences the way the audience perceives social issues. Headlines particularly are very strong framing tools since they reduce complex occurrences to brief and effective statements that inform the first impressions of the news story to the reader.

Framing in the news headline context works based on the linguistic decision made like lexical choice, actor framing, and theme focus. The recurrence of some gendered patterns may



have subtle ideological implications of masculinity and femininity. As an example, in case more frequently men are related to leadership, crime, or decision-making processes, whereas women are related to victimhood, home tasks, or emotional contexts, the repetition of such patterns can represent the predominant gender ideologies that are considered inherent in media discourses. Such framing can change the way social events are perceived by the readers and can strengthen the existing gender stereotypes in the long run. This paper will merge Feminist Linguistics with Framing Theory in exploring the structural issue of imbalance in representation as well as the ideological significances that arise out of repetitive linguistic patterns in news headlines. Feminist Linguistics offers the theoretical context of the importance of gender inequalities in language and Framing Theory offers the theoretical explanation of how the media discourse constructs the understanding of the people by choosing some and emphasizing some things and creating a discourse. The combination of these frameworks makes it possible to interpret linguistic patterns in the headlines more deeply.

Moreover, the research incorporates these theoretical viewpoints and combines them with computational corpus analysis in order to analyze gender representation in an empirical manner. The study determines systematic patterns of the representation of male and female actors through the analysis of frequencies and collocational patterns of the gender-related words in a corpus of news headlines. These quantitative results are subsequently understood using the Feminist Linguistics and Framing Theory in order to reveal the ideological context portrayed by the news headline language. This combination enables the study to get past mere observation and offer a more detailed insight into the role that language, media discourse and gender ideology play in the modern news reporting environment.

## **Data Analysis**

### **Introduction:**

In this chapter, the analysis of the news headline corpus will be provided to answer the main research question: How are men and women represented in online news headlines? The four categories used (Crime and Judiciary, Politics and Government, Social issues and Human rights, and Sports, Business and Culture) are used to organize the analysis. Each category under consideration is represented by a combination of a quantitative frequency analysis and a qualitative collocation analysis. The results are then discussed in the twin perspectives of Feminist Linguistics and Framing Theory with the view of identifying the underlying ideological trends. Then finishes by a comparative conclusion to generalize the findings of all categories.

### **4.1 Category 1: Crime & Judiciary**

MALE MENTIONS	18
FEMALE MENTIONS	20



TOTAL MENTIONS	38
%MALE	47.4%
%FEMALE	52.6%

Table:4.1

**Frequency Analysis:**

Women were also slightly mentioned in the Crime & Judiciary category (Female: 20; Male: 18). This demonstrates that men and women were nearly identical in headlines pertaining to crimes. On the face of it, this contradicts the popular notion that men are represented more frequently in the news about crime. However, Feminist Linguistics state that equality is not guaranteed by numerical visibility only, but it is the manner in which people are linguistically depicted that counts more (Robin Lakoff, 1975; Sara Mills, 2008).

Nonetheless, it is not sufficient to simply count the number of times in which men or women are mentioned (Feminist Linguistics). More crucial in this regard is the way they are described and presented. Therefore, the words applied to male and female reference (collocations) must be examined in order to comprehend the implicit meanings and concepts of the headlines.

**Collocation Patterns / Context:**

The collocation analysis reveals that there is a distinct difference in the description and presentation of men and women in the headlines. The words that relate to direct agency and criminal action, which include arrested, confesses, sentenced, suspect, and kills are strongly linked with references to male. These words demonstrate men as the subjects of the criminal justice system and they also follow the trends discovered in discourse analysis (Teun A. van Dijk, 1998). These frames frame men as the main actors or the main perpetrators of the criminal justice system.

On the contrary, references by females are associated with another group of terms that make them victims, according to relations, or objects of certain, gendered crimes. The most important collocates are: scammed, questioning, missing, husband, honour, body and exhumed. The use of such words as husband and honour indicates that women are usually put in relational frame or victims of crimes that are directly related to their sex and social position.

**Interpretation / Ideological Meaning:**

Under the Framing Theory, the existence of such patterns leads to two stories of crime. Men are represented by their action and responsibility through such words as suspect and kills. This gives images of men as dynamic and even violent. Women on the other hand are referred to as scammed, missing and body with such words indicating that they are passive, vulnerable and victims. This disparity resonates with the social concept that men are most often perceived to be



the perpetrators of crimes, and women are mostly depicted as victims. Feminist Linguistics tells us that such representation makes power unequal. Men are associated with action (including criminal action) in the public, and women with sufferings, consequences or the inner world (husband, or honour). Linguistic and narrative ways of presenting media discourse often construct men and women in different ways (Talbot, 2007).

Language in such a manner is a manifestation of and enforcement of patriarchal social organization. These trends cannot be viewed as neutral; headlines construct social cognition and can have an effect on how society interprets gender roles in crime. Such framing being repeated in reports on the news, normalizes the notion that men are violent and women are vulnerable beings. This is in line with the position held by Sara Mills (2008) who considers the discourse as usually portraying women as passive actors. These types of representations support patriarchal systems of power through linking men to action and women to suffering as these arguments underpin the arguments of Robin Lakoff (1975).

#### **4.2 Category 2: Politics & Government**

MALE MENTIONS	9
FEMALE MENTIONS	11
TOTAL MENTIONS	20
%MALE	45.0%
%FEMALE	55.0%

Table:4.2

#### **Frequency Analysis:**

Referring to Politics & Government, the female references were once more a bit more than male ones (Female: 11; Male: 9). Whereas such increased visibility of women could be interpreted as the positive shift to the state of gender equality in the political discourse, the collocational analysis paints a more complicated and, ideologically, more heated picture.

#### **Collocation Patterns / Context:**

Male references in this category are co-located with the words that denote, institutional power and official authority. The use of words like minister, governor, cabinet, lawmaker, and head all position men as members of the systems of power, occupying credible and recognised positions. Conversely, the female allusions are linked to the actions and conditions that represent



them as an outsider or an outcast of the status quo. The portrayal of gender in the media is usually disproportionate with women being underrepresented or stereotyped (Ross & Carter, 2011). Words used alongside it are uproar, threatening, protest, questions, arrested, and activist. This framing does not represent women as those in power but as disruption, objects of state action (arrested) or as activists outside of the official apparatus (activist).

**Interpretation / Ideological Meaning:**

This ideologically is a framing. Women are not presented as the equals in the field of politics, even though they are visible. The Framing Theory exhibits that the repetitive identification of men with labels of authority and women with demonstrations and dissent encourages the understanding of politics as a male domain. Feminist Linguistics would perceive it as a linguistic manifestation of the glass ceiling, where the presence of women is taken into account but the power of women delegitimized. Language in media is very instrumental in the development of social identities and power relation (Conboy, 2007). They are presented as emotional (uproar) and oppositional (protest) instead of being rational and governing. This trend is used to uphold the status quo of male domination whereby male domination is seen as the normal and female political intervention as an exception or a challenge to the norm.

**4.3 Category 3: Social Issues & Human Rights**

MALE MENTIONS	9
FEMALE MENTIONS	17
TOTAL MENTIONS	26
%MALE	34.6%
%FEMALE	65.4%

Table:4.3

**Frequency Analysis:**

The largest and more significant difference in quantitative terms can be seen in the category of Social Issues and Human Rights where female references (17) almost doubled male references (9). This visibility of women is however supposed to be critically looked into instead of accepted at face value. Yet, corpus studies as pointed out by the scholars indicate that frequency should be viewed in context and collocation (Paul Baker, 2006).

**Collocation Patterns / Context:**



The collocational patterns herein the category are very revealing. Males are all but synonymous with the words of perpetration and aggression: torturing, killing, convicted, accused, and attack. Men are positioned as the cause of the social problems that are being reported. References of females are mostly linked with the suffering and victimization and also with gender-related vulnerabilities. The salient collocates are violence, transgender, domestic, divorce, acid, ablaze and struggles. This puts women not as the participants in the realm of social issues, but as the main sufferers of that.

**Interpretation / Ideological Meaning:**

This type gives strong arguments to the main principles of Feminist Linguistics. The fact that the female references are high is not an indicator of empowerment but it shows that females are linguistically associated to the world of social issues and misery. The collocated frame made by these collocates is a frame of vulnerability and trauma. This is in line with feminist criticism that the media tends to relegate women to the so-called domestic realm of experience, displaying their presence in the world not much beyond the situation of crisis and victimization. The ideological implication is that, women issues are the same as women suffering; men are there only to cause this suffering. This framing solidifies a patriarchal ideology which characterizes women and their toleration of hardship and their male counterparts and their ability to cause it. The language of news discourse helps to construct a perception and social meaning among the masses of people (Bednarek and Caple, 2012).

**4.4 Category 4: Sports, Business & Culture**

MALE MENTIONS	14
FEMALE MENTIONS	15
TOTAL MENTIONS	29
%MALE	48.3%
%FEMALE	51.7%

Table:4.4

**Frequency Analysis:**

The most balanced category in terms of frequency is Sports, Business & Culture: the female references (15) are a bit more than the male ones (14). This numerical equality is an indication of a more egalitarian area of representation in these areas.



**Collocation Patterns / Context:**

This relative balance is initially assessed by the collocational analysis. Both the male and the female references are linked to achievement, success, and professional stature. Female collocates such as wins, gold, award, author and artist are reflections of male collocates such as wins, gold, medal, author and director. Both sexes are portrayed as successful in the workplace and overachievers. Moreover, they are both connected with corporate leadership (ceo, chairperson, board). Nevertheless, minor differences remain. Male references bring in such words as tycoon which brings in connotation of aggressive power and wealth. The application of female allusions is distinctly linked to the sphere of fashion, which is traditionally a feminine subject of stereotypes.

**Interpretation / Ideological Meaning:**

This type would be a very important contrast to the rest to prove that the more fair framing can be made. The linguistic predominance of the achievement-oriented language in both genders makes the universal role of the agent / victim dichotomy in other settings questionable. This implies that in areas that are not seen to be directly related to the power of core state or social crisis, traditional gender ideology may be more lax. However, the fact that even in less discriminatory situations, there are some traces of stereotypical signs (tycoon vs. fashion) shows that, even in more equalized situations, there may be some gendered assumptions in the words used by the writers of headlines. This subtext fits in with an advanced feminist interpretation that accepts progress yet still recognizes the ways in which language can still be used to perpetuate old stereotypes.

**4.5 Comparative Summary and Discussion:**

NEWS CATEGORY	TOTAL MENTIONS (Male and female)	%MALE	%FEMALE
1. Crime & Judiciary	38	47.4%	52.6%
2.Politics & Government	20	45.0%	55.0%
3.Social Issues & Human Rights	26	34.6%	65.4%
4.Sports, Business & Culture	29	48.3%	51.7%
TOTALS	113	44.2%	55.8%

Table:4.5



The four category analysis shows that gender representation in news headlines is not a universal problem, but it is very specific to the thematic background. There are a number of major trends that can be identified in making comparisons. To begin with, gender presence, as per category, varies greatly. The most prominent ones are found under the category of Social Issues and Human Rights, whereby women are the important subjects of discussion. Men are not in any single category dominant in numbers, but are uniform in their presence. The visibility, however, does not, according to the arguments by the feminist scholars, amount to equal treatment (Robin Lakoff, 1975). This questions a naive invisibility argument and is directed towards a more intricate ideological pattern concerning the nature of their invisibility.

Second, and most importantly, there is an effective and consistent ideological division in the pattern of framing. In Crime, Politics, and Social Issues, there is an evident dichotomy of agent/victim or an insider/outsider. Men are always cast as criminals (suspect, kills), government officials (minister, governor) and violent men (torturing, killing). Women are always modeled as deceived (swindled, lost), as alien to authority (protest, activist), or as the targets of trauma (violence, acid).

Third, the category Sports, Business and Culture is an exception as it has a more balanced type of representation where both sexes are discussed as successful agents. The difference between the biased frames and the rest indicates that the former is not unavoidable but especially effective in the headings connected with the state power, crime, and social turmoil. Finally, the computational analysis presents strong evidence that news headlines, with the selective representation and framing, promulgate conservative gender ideologies. They also create a methodological reality whereby men are the default participants of the public and political space, and women are characterized by vulnerability, their relation to men, or their victimization in the social context. This is a linguistic pattern that has been uncovered by using corpus analysis and that has been explained by use of Feminist Linguistics and Framing Theory where media language has been seen to perpetuate patriarchal power structures in the society.

### **Conclusion:**

The paper explored the question of gender representation in headlines of English-language newspapers through a computational corpus approach guided by Feminist Linguistics and Framing Theory. The results indicate that the differences in the numerical visibility between men and women differ by categories, but the underlying collocational equations always depict gendered differences. Men are positioned as players in the areas of crime, politics, and governance, and women are often depicted as victims, social characters, and outsiders. There is more even-handedness in the representation in such aspects as sports, business, and culture, but some nuanced stereotypical indicators remain.

The patterns show that language in headlines is not neutral, but rather, they reflect and support the social power systems and conventional gender ideologies. The authors of the study emphasize the necessity to examine the genre of headlines since this is one of the media genres that has a strong impact on the perceptions of people and forms social interpretation of gender roles. The research has got a number of implications. Among the issues that journalists and editors need to consider are the use of words and framing in headlines that can be applied to



reinforce gender stereotypes, and more deliberate editorial decision-making can lead to more leveled reporting. Media literacy programs can also be informed by the findings, whereby students and readers can be in a position to critically analyse the way gender is represented in news.

Moreover, the study is also an effective way of revealing the concealed tendencies in the ideological analysis of big textual sources, as the combination of computational corpus analysis and feminist interpretation can be considered the productive tool. Such evidence can be used by organizations fighting against gender inequality to demand better media practices. The study has limitations even though it has made contributions. The sample is restricted to a set of online newspapers and categories that might not be representative of the overall trends in the media environment. Analysis of only English-language headlines was done, which may have excluded the views that are found in the Urdu or regional-language media. Although the patterns of association are clear using collocation analysis, it might not include all the socio-cultural background or the subtle meaning of each headline.

Lastly, the study focuses on a particular period, and trends can change in times of significant political, social, or cultural occurrences. Further studies can build on the present results and incorporate multilingual corpora that will give a more detailed representation of the gender representation in Pakistan. Longitudinal research would determine how the gender of framing changes over time and cross-cultural comparisons would emphasize how social and institutional contexts of the media discourse shape the media. More sophisticated computational tools, e.g., sentiment analysis, topic modeling, semantic frame analysis might reveal more layers of meaning to the headlines. The combination of the textual analysis and studies of the visual media would also serve purposes of studying the joint efforts of the language and images in creating gendered discourses. To conclude, newspaper headlines, even though being concise are potent locations of ideological meaning. The integration of computational corpus approach with feminist and framing approaches enables this project to provide the empirical evidence that gender in headlines still represents the legacy of the traditional societal norms, and at the same time, it hints at the potential of more balanced and equitable media practices.

### **References:**

- Lakoff, R. (1975). *Language and woman's place*. New York, NY: Harper & Row.
- van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. London, UK: Sage Publications.
- Bednarek, M., & Caple, H. (2012). *News discourse*. London, UK: Continuum.
- Lazar, M. M. (Ed.). (2005). *Feminist critical discourse analysis: Gender, power and ideology in discourse*. London, UK: Routledge.
- Mills, S. (2008). *Language and sexism*. Cambridge, UK: Cambridge University Press.
- Baker, P. (2006). *Using corpora in discourse analysis*. London, UK: Continuum.
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge, UK: Cambridge University Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge, UK: Cambridge University Press.



- Baker, P. (2014). Using corpora to analyze gender. *Bloomsbury Companion to Corpus Linguistics*, 148–163.
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). Discourse analysis and media attitudes: The representation of Islam in the British press. *Cambridge University Press*.
- Ross, K., & Carter, C. (2011). Women and news: A long and winding road. *Media, Culture & Society*, 33(8), 1148–1165.
- Caldas-Coulthard, C. R. (1995). Man in the news: Social representations of men and women. *Discourse & Society*, 6(2), 226–243.
- Conboy, M. (2007). *The language of the news*. London, UK: Routledge.
- Talbot, M. (2007). *Media discourse: Representation and interaction*. Edinburgh, UK: Edinburgh University Press.