

# Journal of Communication and Cultural Trends (JCCT)

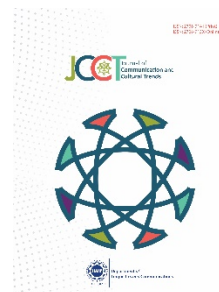
Volume 4 Issue 2, Fall 2022

ISSN(P): 2706-9141 ISSN(E): 2706-915X

Homepage: <https://journals.umt.edu.pk/index.php/jcct>



Article QR



**Title:** Representation of Women: A Corpus-based Analysis of Pakistani English Newspapers

**Author (s):** Salma Asghar<sup>1</sup>, Urooj Akhter<sup>2</sup>

**Affiliation (s):** <sup>1</sup>Women University of Azad Jammu and Kashmir, Bagh  
<sup>2</sup>Department of English University of Poonch, Rawalakot

**DOI:** <https://doi.org/10.32350/jcct.42.04>

**History:** Received: March 3, 2022, Revised: June 27, 2022, Accepted: August 31, 2022

**Citation:** Asghar, S., & Akhtar, U. (2022). Representation of women: A corpus-based analysis of Pakistani English newspapers. *Journal of Communication and Cultural Trends*, 4(2), 61–80.  
<https://doi.org/10.32350/jcct.42.04>

**Copyright:** © The Authors

**Licensing:**  This article is open access and is distributed under the terms of [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

**Conflict of Interest:** Author(s) declared no conflict of interest



**UMT**

A publication of  
Institute of Liberal Arts  
University of Management and Technology, Lahore, Pakistan

# Representation of Women: A Corpus-based Analysis of Pakistani English Newspapers

Salma Asghar<sup>1\*</sup> and Urooj Akhter<sup>2</sup>

<sup>1</sup>Women University of Azad Jammu and Kashmir, Bagh

<sup>2</sup>Department of English, University of Poonch, Rawalakot

## Abstract

This study investigates the use of gendered lexical items and collocates to deconstruct the patriarchal social system of Pakistan with the help of corpus tools. Media frame the mind and thinking of people with its repetitive announcements. Women being a part of the society are also part of media news but they are underreported by reporters and editors as most of the news are male-filtered news. Pakistani English newspapers are one of the popular sources of getting information across Pakistan. The local social system and male dominating mindset have influenced all types of media in Pakistan. It is observed that like all other fields of life, women are marginalized in Pakistani English newspapers too. It is hypothesized that male are over represented as compared to female and women related news are written by female writers and reporters. A corpus of 270 Pakistani English Newspapers has been selected from The News International, The Dawn and Daily Times with 90 articles from each paper. A corpus-based analysis has been carried out to identify gender discrimination and gender stereotypes in Pakistani society. All collocate of women and men are kept in attention in this work. The results reveal that women are represented as a minority and marginalized victims by male dominating society. On the contrary, the modern writers raise their voices to support and highlight the role of women in society. The study is the first attempt of its kind for an amalgam of corpus-based and CDA approaches.

**Keywords:** representation, corpus-driven analysis, newspapers, women, marginalized, minority

## Introduction

Media represent its masses in different ways, where news, stories, and advertisements depict the mindset of a society directly or indirectly. Women are an important component of every society yet they have been given less

---

\* Corresponding Author: [salmakhwaja14@gmail.com](mailto:salmakhwaja14@gmail.com)

attention by news agencies and media groups. Gender-related words such as male, female, men, and women and gendered pronouns like she and he can best tell us how two different genders are represented in Pakistani English newspapers. Although a general attitude toward women is quite obvious in society. According to Jia et al. (2016) reporters and editors have been contending that masculine/ male-centered news is considered to be more newsworthy, which results in highlighting male voice by marginalizing women in English newspapers. In this research the researcher intended to explore status of women's representation in Pakistani English newspapers. Moreover, the gendered vocabulary is taken into consideration to study and uncover a complete picture of Pakistani news media with the help of a corpus and CDA approach.

### **Problem Statement**

Media is the representation of society. Women have a significant number in every society but unfortunately, are under-represented. In under-developing countries like Pakistan, the case is even worse as women are marginalized and under-privileged in almost all domains. Their problems are scarcely focused on both i.e. print media and electronic media. It is observed that Pakistani English newspapers also give least importance to the issues of women and their struggles. Thus, previous literature shows that less research has been conducted in this domain; therefore, this research fills the gap by using critical discourse analysis (CDA) to study the invisible ideologies and social practices which are a part of Pakistani society.

### **Objectives of the Study**

This research aims to investigate certain objectives:

- 1) To explore the most frequently occurring collocates of gendered words in Pakistani English newspapers.
- 2) To deconstruct the patriarchal system of selected editorials of Pakistani English newspapers.

### **Significance of the Study**

This research aims to discuss the representation of women and their issues in Pakistani newspapers. This study will help to understand how Pakistani news reporters and editors represent men and women in the print media. It will also enable novice researcher to analyze data with the help of corpus and CDA approach which is a relatively new method in the field of

research. This research will help editors, and policymakers in building an equal representative relationship for both genders in the print media.

### **Research Gap/ Niche**

Studies showed that a lot of work has been done in both critical discourse analysis and discourse analysis, but little attention has been given to gender-oriented issues in Pakistani news media in general and Pakistani English newspapers in particular. There is no significant work conducted on women's representation in terms of gender discrimination in Pakistani English newspapers thus, the current study incorporates corpus tools to conduct the gendered collocation analysis. The current study exhibits the representation of gender in Pakistani English newspapers.

A corpus-based analysis has been carried out to evaluate the representation of women in Pakistani English newspapers. The results showed that women-related news is quite less in all three selected newspapers. Although women are underrepresented in news media even in the 21<sup>st</sup> century, a good thing observed in the recent practices which reflect change in the trend towards women's representation. Now women are expressing their issues in a quite open manner. The paper is further subdivided into five sections introduction, literature review, research methodology, and findings and discussion, followed by a conclusion.

### **Review of Literature**

Gerbner et al. ([1986](#)) were of the view that media influences the human mind with its repetitive announcements through images and concepts. Gender inequality is a common issue that varies from culture to culture and society to society depending on economic status, social structure, and mindset of the society in which people live.

Corpus linguistics methodology being an emerging field in the study of language helps researcher to analyze data both qualitatively and quantitatively. It saves time, money, and energy of researchers through its advanced computational features by accessing and analyzing linguistic data of real language. British national corpus (BNC) and American national corpus (ANC) are examples of such corpora.

Language constructs identity like gender and social class issues related to society, such as gender bias and discrimination. Numerous types of research have been carried out to expose gender-biased myths and

stereotypes about women. In this regard, multiple types of approaches are used to analyze data, such as CDA (critical discourse analysis), thematic analysis, content analysis, and CL (corpus linguistics) in this regard. Corpus linguistics is an emerging methodology in linguistics that helps a researcher to analyze and describe linguistic data both quantitatively and qualitatively. In addition to this, it enables the researcher to find results with evidence from the data; where the purpose of research is to find out the probabilities, trends, patterns, co-occurrences, and groupings and come up with findings that can be generalized about language. There are numerous studies carried out in the field of the press such as the representation of women in the news media (Caldas-Coulthard, [1993](#)). The study of different corpora to evaluate racial, ethnic, and tribal aspects, an analysis of the British and Polish community in the British press (Galasinski & Marley, [1998](#)) and an analysis of New Labor rhetoric (Fairclough, [2002](#)) are some of them. In another corpus study by Simon-Vandenberg (1997), a set of 65 interviews from political discourse have been taken from a corpus of BBC where the model of certainty was investigated. There was another study conducted by De Beaugrande ([2001](#)) with a corpus-based CDA and frequent words, nouns, verbs, and concordances were measured and described to expose some interpretive possibilities that would not be explored otherwise.

Sadia and Ghani ([2018](#)) in their study highlighted the modality in Pakistani English newspapers by using corpus techniques and stated that modal auxiliaries are very important in Pakistani English newspapers as they show prediction, obligation, prayers, possibility, and ability. Pearce ([2008](#)) conducted a research where he derived word sketches of man and woman by using a sketch engine from British National Corpus and found that woman tends to be an annoying subject with the verbs like cluck, fuss, and nag, whereas, it was the grammatical object of verbs that refer to sexualization, e.g. bed, ravish and shag.

Similarly, another study was conducted in Hong Kong by Flowerdew ([2004](#)) where he collected a corpus of speeches of the special administrator of Hong Kong, appointed by China and he explored that this person was more interested in a knowledge-based economy and free market where he refers to China, not Hong Kong. Flowerdew measured the frequency of certain words by using corpus tools and analyzed the discourse of the administrator thoroughly by describing the concordance lines of his discourse where he was trying to establish a new stance on people of Hong

Kong's identity as Chinese. His lexico-grammatical choices were analyzed by using corpus linguistic tools.

Cheng (2004) criticized the study and summed up that politicians should adopt buzz collocations and patterns which are unique in political discourse and are based on the choice of keywords and collocates measured. Sotillo and Wang-Gempp (2004) focused on both qualitative and quantitative analysis where they examined the hyperbole, lexical choices, word collocations, verb collocations, and personal pronouns by employing corpus tools, and power relations, ideologies, social class, and politics were analyzed with the help of a framework from CDA.

Orpin (2005) analyzed five semantically related words of 'corruption' in a corpus of 800 words in the British corpus of newspapers and found that most of the words with negative connotations were used to refer to Italy, Korea, Pakistan, and India. He claimed that Corpus methodology was not used to ignore qualitative analysis but to join CDA in order to find reliable evidence. In a corpus-based linguistic approach, O'Halloran (2007) studied how metaphors express ideology in news corpus where he examined the collocations in a corpus of 260 million words from the news. Finally, he concluded that the corpus-based explanation of metaphors is very helpful in understanding the deep structure of the meaning-making process. For quantitative analysis, a set of three corpora had been selected and calculations were interpreted with the help of concordances of media coverage, scientific text, and policy documents. Corpus strengthened the findings of CDA with evidence and examples.

Baker et al. (2008) combined CL and CDA where they adopted a corpus-assisted CDA technique to study the representation of refugees, asylum seekers, and migrants from a corpus of 140 million words. Primarily, they focused on the keywords and collocations and measured statistics of frequently occurring keywords and their concordances. Another research has been conducted by Afzaal et al. (2019) on media discourse on CPEC (China-Pakistan Economic Corridor) in Pakistani news by employing corpus-based analysis. Concordance lines and keywords were described and collocates of certain words such as China, Pakistan, economic, and growth were measured with the help of Antconc software. The results showed that there was a positive representation of these keywords in the discourse. Some other studies explored gender by applying

corpus techniques and identified the frequent use of certain adjectives with keywords such as men, man, women, and woman.

A study related to the representation of women in the Olympics identified the use of women and women in a media discourse of sports and concluded that news pre-Olympics and post-Olympics had an underrepresentation of women whereas the news during the Olympics had a significant use of women-related words. For this, a corpus-based study had been conducted by Partington et al. (2013) where at first a corpus of all British newspapers have been collected and analyzed by employing the technique of keywords, concordances, and collocates, and later frequency lists were prepared.

Shore et al. (2014) argued that this gender gap in media and news is not surprising or new as the predominance of men in this world is an old phenomenon. Norberg (2017) was of the view that corpus linguistics can be used as a tool to examine the representation of women in the mining industry and other male-oriented industries. They adopted different tools such as word lists and n-grams to measure the specific use of the word ‘women’ in the context of management-related jobs in mining.

Such stereotypes are discussed in other researches and almost all of the findings indicated that it is not only observed in underdeveloped countries but is a common practice all over the world. This section deals with previous literature conducted in the field of gender on media studies and also scrutinized the techniques used by researchers to investigate the issues related to media representation of males and females in different societies. The prime focus of the current study is to uncover the gender bias and patriarchal social system of Pakistan which leads writers, reporters, and editors to undermine the status of women in a society where broadcasting male-related news is more common.

Thus, the current study aims to explore the representation of women in print media of Pakistani English newspapers by conducting a corpus analysis and incorporating Corpus linguistic tools. A corpus is a computer recorded or electronically saved written or spoken data and can be used to verify the hypothesis regarding a language. Hence, this study employs CDA and Corpus linguistic analysis to understand the gendered language by studying collocations in the selected editorials.

## Research Methodology

A mixed method approach has been used to vitalize the process of data collection and analysis. Although editorials were selected randomly yet it was ensured that these editorials were taken from Pakistani English newspapers. Secondly, it was considered that the selected editorials discuss all types of issues and news beyond considering gendered phenomenon so that no biased information may be given and results may not be equivocal. The selected data was cleaned with the help of software fiXit and saved as text files for the further processing.

Antconc 3.5.9 corpus analysis software was used to analyze the targeted corpus. This software offered different corpus tools such as n-grams, word lists, concordance lines, collocate, and keywords in context. To analyze the data for this research in the first place we developed a list of words and later collocate man, men, women, and woman, and male, female, he and she have been analyzed with the help of Antconc. This tool enabled the researcher to find out all possible structures and frequency of certain words usage with all targeted words which were analyzed later with the help of concordance lines. The application of the corpus-driven or corpus-based analysis in corpus linguistics is basically based on the analytical technique. In corpus-based analysis, the researcher inquires about a problem with the help of the corpus linguistic technique and then the researcher answers the given questions. Whereas, in corpus-driven analysis, first the hypothesis is built on corpus linguistic data along with research questions that are based on the selected data. A corpus-driven analysis is an example of inductive study while corpus-based analysis is an instance of deductive research.

Later, thematic analysis was carried out to divide the data into different themes. Thematic analysis is a method of analyzing qualitative data. It is usually applied to a set of texts, such as an interview or transcripts. The researcher closely examined the data to identify common themes, topics, ideas, and patterns of meaning that come up repeatedly (Caulfield, [2019](#)).

In corpus linguistics (CL) the analysis of concordance lines, recurrent patterns, collocations, and keywords is done with the help of various types of software, where corpus tools are applied to analyze the corpora. Some of the basic methods include; the generation of word lists, concordances, collocates, graphs, and statistical measurements of these elements. As CL is primarily used for the quantification of data, it is majorly used in the



context of a particular situation (Sinclair, 1991). Hence, these methods are applied to analyze discourses and its features such as language features, recognition of a specific function, symbolizing a variety of languages, and plotting the occurrences of a feature by examining the whole text (Conrad, [2002](#)).

Corpus linguistics (CL) has been used to study critical discourse analysis (CDA) by Baker et al. ([2008](#)) initially, these approaches were combined to study the Pakistani English newspapers.. Language in CDA is a type of social practice where researchers try to uncover the embedded meaning, power relations, and false ideologies in any discourse (Kobayashi, [2019](#)). Significantly, CDA and CL are merged effectively to analyze a text where certain words and phrases, collocations, and syntactic patterns of text are analyzed with the help of corpus tools including the use of pronouns for a particular gender in language, transitivity, and agency (Orpin, [2005](#)).

### **Sampling Technique and Sample Size**

A sample of 270 editorials has been selected randomly through convenient sampling from three major Pakistani English newspapers. This gives the most relevant information and saved time. The researchers build the corpus of all newspapers published from 2010-2020. The corpus comprises 429295 words. The titles of these editorials are general in their nature so that no subjectivity could affect the results.

### **Interpretive Framework**

Post positivism has been used as a guiding map. This is metatheoretical critique approach that has been employed to investigate the selected editorials. Post positivists believed that all cause and effect are probability and there is no strict cause-effect. Their primary focus was to support quantitative analysis in order to justify the results with the help of scientific research methods. Baker's ([2018](#)) approach has been employed as a theoretical framework to conduct a thematic analysis in the current research.

### **Findings and Discussion**

This section includes what the current research has found from the data to conduct a quantitative analysis of data, the researcher incorporated data in the form of screenshots from the software, which provided a detailed discussion of qualitative findings made by the researchers. Here, a detailed analysis of women is presented to highlight the gendered vocabulary used

in selected newspapers editorials. It is hypothesized that women/woman carries negative connotations in both Pakistani culture and newspapers. Moreover, they are underrepresented or misrepresented by news media. These gender-specific words help researchers to find the very set patriarchal social system and the mindset which declares women as an inferior gender. Collocates are calculated to analyze the syntactic structure of words neighboring the target words. The following examples from the data showed the possible collocates of target words and helped the researcher to understand the intentions and intended meaning of the writers. The first objective of this study is to explore the most frequently occurring collocates of women, women, men, and men in Pakistani English newspapers.

The data showed that men/men collocate with words that show authority, respect, and nobility in this discourse. On the other hand, women are often collocated with negative verbs and adjectives in the selected texts.

### Collocations

In most of the examples, collocate of both genders *are given* on the right and left side are adjectives, verbs, nouns, and prepositions. Tables 2, 3, 4, and 5 show right and left collocates of *men and women* respectively.

**Table 1**

*First 20 Most Frequently Used Collocates of Man*

Sr. #	Collocate	Frequency	Sr. #	Collocate	Frequency
1	Vilify	10.13470	11	Respectable	10.13470
2	Testimony	10.13470	12	Vested	9.71967
3	Symbol	10.13470	13	Permitted	9.71967
4	Odds	10.13470	14	Myself	9.71967
5	Noble	10.13470	15	Literally	9.71967
6	Holey	10.13470	16	Lest	9.71967
7	Detain	10.13470	17	Dies	9.71967
8	Conversing	10.13470	18	Essentially	9.71967
9	Balancing	10.13470	19	Smiling	9.71967
10	Wealthy	10.13470	20	Dress	9.71967

There are 880 collocate types and 15090 collocate tokens of *man* in the target corpus. Table 1 shows that the most frequently occurring collocates of *man* have four major categories:

1. Nouns: symbols, myself, dress, odds, testimony

2. Adjectives: smiling, wealthy, holey, noble, respectable
3. Verbs: detain, vested, permitted, dies, balancing, conversing, vilify
4. Adverbs: literally, essentially

**Table 2***First 20 Most Frequently Used Collocate of Men*

Sr. #	Collocate	Frequency	Sr. #	collocate	Frequency
1	Subordinate	56	11	traveled	2
2	Overgrown	4	12	Torturing	2
3	Outnumbered	2	13	threw	4
4	Sick	54	14	tempted	1
5	Awards	56	15	Thoroughly	2
6	Forgiven	4	16	Tauntingly	1
7	Respond	56	17	Politically	3
8	Football	112	18	Violently	4
9	Engaging	30	19	Roamed	26
10	Defeated	30	20	Reflecting	28

There are 2348 collocate types and 47150 collocate tokens of *men* in the target corpus. Table 2 displays the main categories of collocates of *men* that can be classified as:

1. Nouns: subordinate, football, resource, awards
2. Adjectives: overgrown, smarter
3. Verbs: threw, stormed, sympathize, forgiven, defeated, enraging, respond, outnumbered, traveled, tempted, reflecting, roamed,
4. Adverbs: politically, violently, tauntingly, and thoroughly

**Table 3***First 20 Most Frequently Used Collocate of Woman*

Sr.#	Collocate	Frequency	Sr. #	Collocate	Frequency
1	Worship	2	11	Testimony	1
2	Weight	1	12	Surrenders	3
3	Wedlock	3	13	Stigma	2
4	Un-women	2	14	Stamina	1
5	Unopiniated	3	15	Sensible	2
6	Unfaithful	4	16	Savaging	0
7	Grocery	1	17	Ridiculing	2
8	Treats	2	18	Rescued	1
9	Widowed	1	19	Provoked	3
10	Nicely	1	20	Possesses	2

There are 845 collocate types and 5780 collocate tokens of *women* in the target corpus. Table 3 shows that the most frequently occurring collocates of *women* have four major categories:

1. Nouns: weight, wedlock, stigma, stamina, testimony, grocery
2. Adjectives: unfaithful, un-opinioned
3. Verbs: worship, treats, surrenders, ridiculing, rescued, provoked, possesses, savaging, threats, widowed
4. Adverbs: nicely

**Table 4**

*First 20 Most Frequently Used Collocate of Women*

Sr. #	Collocate	Frequency	Sr. #	Collocate	Frequency
1	Suppression	2	11	Elevated	1
2	Subsidized	2	12	Drive	1
3	Shunning	2	13	Conflating	1
4	Performances	2	14	Beliefs	2
5	Nourished	1	15	Struggled	2
6	Marching	6	16	Poorest	2
7	Learning	2	17	Advancing	2
8	Internationally	3	18	Deserve	2
9	Intake	1	19	Encouraged	2
10	Impediments	2	20	Woefully	2

There are 3680 collocate types and 32300 collocate tokens of *women* in the target corpus. Table 4 shows that the most frequently occurring collocates of *women* have four major categories:

1. Nouns: suppression, intake, impediments, beliefs, performances
2. Adjectives: subsidized, poorest,
3. Verbs: shunning, struggled, encouraged, deserve, nourished, marching, learning, elevated, drive, conflating, advancing
4. Adverbs: internationally, woefully

It was noticed that grammatical categories of collocates with *women/woman* have more verbs, and some of them depict that they are moving towards a better life in society. Moreover, the verbs that collocate *man/men* are more concerned with the authority and supremacy of *man/men*. However, the use of adjectives was interestingly significant as negative adjectives have been used with words like *woman/ women* e.g. unfaithful, poorest, etc. the use of nouns was quite significant to discuss

nouns like suppression, worship, and beliefs are found as collocates of *women/ woman* whereas, the nouns like weapons, office, etc. were used with *man/ men*.

### **Thematic Analysis of Concordance Lines**

A meticulous analysis of concordance lines has been carried out in order to investigate the above-mentioned research questions. The findings were quite astounding which showed changes in the trends recently. The inquiry resulted in the following answers to the proposed research objective; to deconstruct the patriarchal social system of Pakistan. Thematic analysis has been carried out in order to unpack the prevailing situation in newspapers.

According to Hirdman (1998) in the present era, the representation of men and women in the media is not only the representation but in actual the construction of femininity and masculinity. Armstrong (2004) emphasized the problem of gender imbalance in news as an outcome of conventional norms which degraded and marginalized women in society, trying to assert the idea that men are more important in a community. As a result, there is a profound threat to society as it will reduce the probability of the masses accepting women occupying any powerful position in society which ultimately triggers gender stereotypes.

Baker's concept (2018) has been selected as the study's framework to analyze the data. A corpus-driven thematic analysis of data resulted in the following categories (See Table 5).

Table 5 shows that inequality is observed in women's representation in Pakistani English news editorials. They are considered as less important and their issues are also given low weightage in newspapers as compared to men. The women are low-paid and are considered to be doing something alleged if they demand a divorce. But a man has every right to keep or divorce a woman without any valid reason. Hence, the data showed that women editorial writers are now changing the trends as they talk about successful women of Pakistan by narrating the success stories of business women, women working in the corporate sector, and women working at home. It is also evident that women from rural areas are much less unaware of their rights and struggling to live their lives with violence staying at a marginalized position. Last but not least is the patriarchal system of the society where men are not only given more respect, attention, and

importance but love also. They are a symbol of power, superiority, and authority over women.

**Table 5**

*Common Categories in All Three Major Newspapers*

Themes	Keywords in corpus	Frequency
Inequality	Against	936
	Poor	2891
	Divorce	2710
	Working	1246
	Society	1112
	Equality	243
Women's initiatives	Empower	2811
	Encouraged	2134
	Young	4005
	Corporate	48
	Politics	124
	Fashion	337
	Superwomen	34
	Industry	11
	Real	91
	Educate	14
Unawareness	First	144
	Rural	521
	Poor	270
Patriarchy	Power alleged, demand	873

### Concordance Lines

Concordance lines further explain the use of certain words in a context. These lines actually reveal the underrepresentation of women with negative collocates which leads the researcher to find interesting findings.

### *Concordances of Women*

The above-mentioned 10 concordance lines showed that 'women' occurred with certain adjectives e.g. young, many, rural, and nouns e.g. empowerment, education, independence, development, field, partners, challenges, profession, etc. Furthermore, it depicted that news media was talking about women who were either excelling in their fields or were trying to excel. While a few other types of researches pointed out that in the first-

place women were under-represented in news media and in the second place when they were represented they are insufficiently dressed and are referred to as stereotypical rules (Collins, [2011](#)).

### ***Concordances of Men***

Here are the first 10 concordance lines of ‘men’ highlighted that in the English language *men* always come before *women*. Words such as harassment, equal, opportunity, gap, wage, power, empowering, women, influential, divorce, marriage, leaders, internet, qualified, and jobs showed that news media represent men with more powerful words in the context and participate in the meaning-making process. *Men* occur more frequently than women in the corpus which indicates that modern media has changed its trends. Another research conducted in South East Europe shows that newspapers showed patriarchal ideology and power structures of a society where women are misrepresented and criticized for certain actions whereas, men face no criticism for the same action (Bamburać et al., [2006](#)).

### ***Concordances of Man***

A detailed analysis of concordance lines has been given below. Concordances are measured with man, men, woman, and women as node words. Adjacent and window collocates were analyzed thoroughly to show the structures of discourse used by writers to represent women. The concordance lines have been studied by keeping the context under consideration and focusing on the research questions formulated for this paper. Though the current study focuses on collocates of women it also identified collocates of man/ men to underpin the patriarchal and gender-related cognitive structures of writers of Pakistani society. In a study conducted in the United States or England over-representation of men in newspapers has been noticed (Ross, [2007](#)).

The concordance lines clearly depict that man is used with words such as great, fought, acquire, father, relationship, wealthy, power, honor, symbol, and society which shows that man has a special status in society and so is in news media where reporters while talking about women give more margin to men by using significant words in the context. In a study conducted in the UK and Ireland, it is suggested that men’s news is three times more important than those of women’s (Ross & Carter, [2011](#)).

### ***Concordances of Woman***

In concordance lines, the word ‘divorce’ has been used with a man to display his power and authority where the writer says ‘man divorced his wife’ but in the above-mentioned example the same word divorce is used in women's context as a negative connotation and it is conveyed that asking for a divorce is ‘alleged’ demand. Words like super, smiling, fragile, real and Pakistani show that women editorial writers appreciated the efforts of women and wanted to represent them on media and on other forums too. The corpus showed that all editorials described the deprivation and inequality in social structures which were being practiced for ages. A study uncovered that women are presented in a narrow-minded and conservative manner in Pakistani English Fiction. Moreover, in Pakistan men are given more priority in every sphere of life and are considered dominating figures. Women are exploited and are not given equal rights and are not free to live on their own (Mahmood et al., [2013](#)).

### **Conclusion**

To conclude, it has been hypothesized that women are underrepresented in Pakistani media. A corpus of three eminent Pakistani English newspapers has been selected for the purpose to collect data. The corpus has been developed and built through software; textfixer and corpus software Antconc (5.3.9) have been used to analyze the data. A corpus-driven analysis has been carried out to analyze the collocation patterns and concordances of women and men in a corpus of Pakistani English newspapers. The findings are quite significant which need to be mentioned; i) collocates of women show that they are underreported or misreported, ii) they are changing the old trends through their voice, and iii) the patriarchal system in Pakistan is the root cause of their underrepresentation. The findings recommended that there should be conducted a comprehensible study on women's representation not only in news but also in political discourse. Moreover, the media discourse needs a thoughtful analysis as it represents the social and power structures of society. The selected newspapers have a very low percentage of news covering women issues. Most of the news is about male members of society representing politicians, bureaucrats, military persons, and sportsmen. They are represented as individuals with reference to their job and other official activities whereas; women are discussed with reference to their families, their outlook, appearances, social status, and personal affairs. The data showed that a few



editorials regarding women were written by male writers and most of them were written by female writers, which again shows the discrimination regarding women-related issues and the common practice to take their issues as trivial and unimportant. It is suggested that a serious effort should be made to explore the gendered language in other media also.

### References

- Afzaal, M., Hu, K., Ilyas Chishti, M., & Khan, Z. (2019). Examining Pakistani news media discourses about China–Pakistan Economic Corridor: A corpus-based critical discourse analysis. *Cogent Social Sciences*, 5(1), e1683940. <https://doi.org/10.1080/23311886.2019.1683940>
- Armstrong, C. L. (2004). The influence of reporter gender on source selection in newspaper stories. *Journalism & Mass Communication Quarterly*, 81(1), 139-154. <https://doi.org/10.1177/107769900408100110>
- Peter, L., & Armstrong, C. L. (2004). The influence of reporter gender on source selection in newspaper stories. *Journalism & Mass Communication Quarterly*, 81(1), 139-154. <https://doi.org/10.1177/107769900408100110>
- Baker, P., Gabrielatos, C., Khosravini, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273-306. <https://doi.org/10.1177/09579265080888962>
- Baker, P., & Vessey, R. (2018). A corpus-driven comparison of English and French Islamist extremist texts. *International Journal of Corpus Linguistics*, 23(3), 255-278. <https://doi.org/10.1075/ijcl.17108.bak>
- Bamburać, N. M., Jusić, T., & Isanović, A. (Eds.). (2006). *Stereotyping: Representation of women in print media in South East Europe*. Mediacentar.
- Caulfield, J. (2022, July 21). *How to do thematic analysis | A step-by-step guide & examples*. Scriber. <https://www.scribbr.com/methodology/thematic-analysis/>

- Caldas-Coulthard, C. R. (1993). From discourse analysis to critical discourse analysis: Theoretical developments. *Trabalhos em Linguística Aplicada*, 21, 49-62.
- Cheng, M. (2006). Constructing a new political spectacle: Tactics of Chen Shui-bian's 2000 and 2004 inaugural speeches. *Discourse & Society*, 17(5), 583-608. <https://doi.org/10.1177/0957926506066629>
- Collins, R. L. (2011). Content analysis of gender roles in media: Where are we now and where should we go? *Sex Roles*, 64(3-4), 290-298. <https://doi.org/10.1007/s11199-010-9929-5>
- Conrad, D. L. (2002). Engagement, excitement, anxiety, and fear: Learners' experiences of starting an online course. *The American journal of distance education*, 16(4), 205-226. [https://doi.org/10.1207/S15389286AJDE1604\\_2](https://doi.org/10.1207/S15389286AJDE1604_2)
- De Beaugrande, R. (2001). Interpreting the discourse of HG Widdowson: A corpus-based critical discourse analysis. *Applied Linguistics*, 22(1), 104-121. <https://doi.org/10.1093/applin/22.1.104>
- Fairclough, N. (2002). *New labour, new language?* Routledge.
- Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional language. In *Discourse in the professions: Perspectives from corpus linguistics* (pp. 11-33). HKUST The Institutional Repository.
- Galasiński, D., & Marley, C. (1998). Agency in foreign news: A linguistic complement of a content analytical study. *Journal of Pragmatics*, 30(5), 565-587. [https://doi.org/10.1016/S0378-2166\(98\)00007-1](https://doi.org/10.1016/S0378-2166(98)00007-1)
- Gerbner, G., Gross, L., Morgan, M., & Signorielli, N. (1986). Living with television: The dynamics of the cultivation process. In J. Bryant & D. Zillman (Eds.), *Perspectives on media effects* (pp. 17-40). Lawrence Erlbaum Associates.
- Hirdman, A. (1998). Male norms and female forms: The visual representation of men and women in press images in 1925, 1955 and 1987. *Nordicom Review*, 19(1), 225-254.
- Jia, S., Lansdall-Welfare, T., Sudhakar, S., Carter, C., & Cristianini, N. (2016). Women are seen more than heard in online newspapers. *PloS One*, 11(2), e0148434. <https://doi.org/10.5061/dryad.p8s0j>

- Kobayashi, A. (2019). *International encyclopedia of human geography*. Elsevier.
- Krishnamurthy, R., & Teubert, W. (2007). *Corpus linguistics*. Routledge.
- O'Halloran, K. (2007). Critical discourse analysis and the corpus-informed interpretation of metaphor at the register level. *Applied Linguistics*, 28(1), 1-24. <https://doi.org/10.1093/applin/aml046>
- Mahmood, R., Parveen, S., Shah, S. K., & Batool, A. (2013). A corpus driven comparative study of representation of 'Man' and 'Woman' in Pakistani and British English fictions. *Research on Humanities and Social Sciences*, 3(11), 95-105.
- Norberg, A. (2017). From blended learning to learning onlife - ICTs, time and access to higher education [Doctoral thesis]. Umeå University.
- Orpin, D. (2005). Corpus linguistics and critical discourse analysis: Examining the ideology of sleaze. *International Journal of Corpus Linguistics*, 10(1), 37-61. <https://doi.org/10.1075/ijcl.10.1.03orp>
- Pearce, M. (2008). Investigating the collocational behaviour of MAN and WOMAN in the BNC using Sketch Engine. *Corpora*, 3(1), 1-29.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)* (Vol. 55). John Benjamins Publishing.
- Ross, K. (2007). The journalist, the housewife, the citizen and the press: Women and men as sources in local news narratives. *Journalism*, 8(4), 449-473. <https://doi.org/10.1177/146488490707865>
- Ross, K., & Carter, C. (2011). Women and news: A long and winding road. *Media, Culture & Society*, 33(8), 1148-1165. <https://doi.org/10.1177/0163443711418272>
- Sotillo, S. M., & Wang-Gempp, J. (2004). Using corpus linguistics to investigate class, ideology, and discursive practices in online political discussions. In *Applied corpus linguistics* (pp. 91-122). Brill Rodopi.
- Sadia, S., & Ghani, M. (2019). Modality in editorials of Pakistani English newspapers: A corpus based study. *International Journal of English Linguistics*, 9(1), 144-151.

- Simon-Vandenberg, A. M. (1997). Modal (un) certainty in political discourse: A functional account. *Language Sciences*, 19(4), 341-356. [https://doi.org/10.1016/S0388-0001\(96\)00068-X](https://doi.org/10.1016/S0388-0001(96)00068-X)
- Shor, E., van de Rijt, A., Ward, C., Askar, S., & Skiena, S. (2014). Is there a political bias? A computational analysis of female subjects' coverage in liberal and conservative newspapers. *Social Science Quarterly*, 95(5), 1213-1229. <https://doi.org/10.1111/ssqu.12091>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.