Pakistan Social Sciences Review
www.pssr.org.pk

RESEARCH PAPER

A Corpus-assisted Study of Linguistic Features of Pakistani English Newspaper Headlines

Aqsa [1] Dr. Tayyaba Yasmin*[2] Dr. Azhar Majeed Qureshi[3]

1. M. Phil Scholar, Department of English, University of Education Lahore, Punjab, Pakistan
2. Associate Professor, Department of English, University of Education, Lahore, Punjab, Pakistan.
3. Assistant Professor, Department of Science Education, Allama Iqbal Open University, Islamabad, Pakistan

PAPER INFO	ABSTRACT
	This study attempts to uncover the ideology of journalists and the image of Pakistani Media by examining the linguistic features of headlines related to keywords of Covid, Afghanistan, Facebook, harassment, and the Sialkot tragedy. Headlines on different topics are selected like sports, political, national, and international topics from The Dawn and The Express Tribune from August to December 2021 to create the corpus. To derive frequencies, AntConc 3.5.8 corpus analysis software was used, followed by Fairclough's (1995) three-dimensional model based on three stages of description, interpretation, and explanation. The findings revealed that adjectives, modal verbs, grammatical features, and positive and negative lexical items are employed by journalists in headlines to draw readers' attention and encourage them to read the complete story.

Introduction

Media plays a critical role in turning ordinary everyday events into the news. People's perceptions of the world are based on secondary information obtained through the media. Newspaper headlines are the most significant aspect of the publication since they give a quick form of communication. The words in the headline convey the content of the article and reveal the ideologies of these words. Yasmeen et al. (2014) conducted a study and discovered that ideology has a role in the development of newspaper headlines in both English and Urdu newspapers in Pakistan. Findings of their study explicate that aside from demonstrating the common graphological characteristics of these headlines, numerous ideologies are at work behind many of the apparently simple words.

Rumors about biasness of media are expected. Journalists choose such types of vocabulary and grammatical features that attract the readers' attention. The majority of the time, authors employ linguistics features to amplify the significance of the headlines (Grabe, Zhou, & Barnett, 2001). This study has employed Fairclough's CDA-based three-dimensional model (1995). In this approach, discourse is a three-

dimensional concept encompassing text, discourse practice, and social-cultural practice. The approach of critical discourse analysis is often used in linguistics to uncover the concealed ideologies of language (Norman Fairclough, 1995). This study intends to critically analyze the similarities and differences between linguistic features employed in Pakistani English newspaper headlines: *The Dawn* and *The Express Tribune*. The frequency of keywords used in both newspaper corpora is analyzed using AntConc 3.5.8 software for quantitative analysis. For qualitative analysis concordance and new sources of some specific keywords are utilized.

. This study provides the language learners a better understanding of how journalists employ specific words to transmit their ideologies. Findings can be used as a foundation for future studies regarding linguistics analysis of headlines in various fields. The present research emphasizes the significance of the language used in news headlines. Headlines for the recent study are delimited from August 2021 to December 2021, which are chosen from different genres like political news, war news, national and international news, fashion news, and sports news. Despite covering a wide range of media genres, the scope of the research was focused on analyzing only a few keywords like Covid, Afghanistan, Facebook, Tragedy, harassment for their description, interpretation, and explanation.

Literature Review

There are a lot of studies with various theoretical frameworks that aim to investigate how newspapers represent events and situations. Taiwo (2007) investigated the relationship of language and ideology with power dynamics. The 300 Nigerian newspaper headlines were studied using the CDA method for unusual vocabulary and rhetorical devices to discover the ideologies that underpin their creations. The findings demonstrated that the headlines conceal ideological connotations and represent the perspectives of those whose interests are served and those whose claims are undercut. He also finds that headlines are used to prolong conversations and affect the readers' opinions (Taiwo, 2007). Duanprakhon (2012) studied how journalists create a particular image of certain things in the reader's mind by considering the Thailand youth crime news headlines. It was observed that a reader's mind could conjure up an ideological vision of a criminal solely based on the wording of headlines regardless of the complete news story.

Using critical discourse analysis, Ulum (2016) investigated the ideology of newspapers from diverse cultures. He compared and contrasted the viewpoints on Syrian migrants arriving in Europe in European and Turkish publications. According to the study's findings, the Turkish publications focused on the migrants' plight and what they had to endure, European newspapers headlines were solely concerned with refugees' threats to Europe (Ulum, 2016).

After reviewing the previous research on newspaper headlines, it turns out that all studies focus on headlines relating to specific issues. But the current study deals with various linguistic features that are used for making headlines catchy, like lexical choices, modality, and grammatical features. The one similarity with previous studies is that it is also a CDA-based study. Moreover, this research study looks at the power dynamics and ideologies that underlie media discourse.

Theoretical Background

CDA, which is derived from Critical Linguistics, was initially referenced in Fairclough's Language and Power (1989). Critical Discourse Analysis is a Discourse Analysis process that reveals how discourses are employed daily for content, power relations, and the advancement of new information (Freebody, Luke, & Muspratt, 1997). Fairclough's technique is grounded on the idea that power influences and regulates how people produce and interpret language. Critical discourse analysis examines discourse as a social practice to comprehend the power of language in societies. According to Van Dijk (1993), Critical discourse analysis has the goal of critically analyzing the sociolinguistic element of discourse. CDA focuses on power relations, jurisprudence, hegemony, diction, contextual factors, and forms of social misuse of power (Van Dijk, 1993).
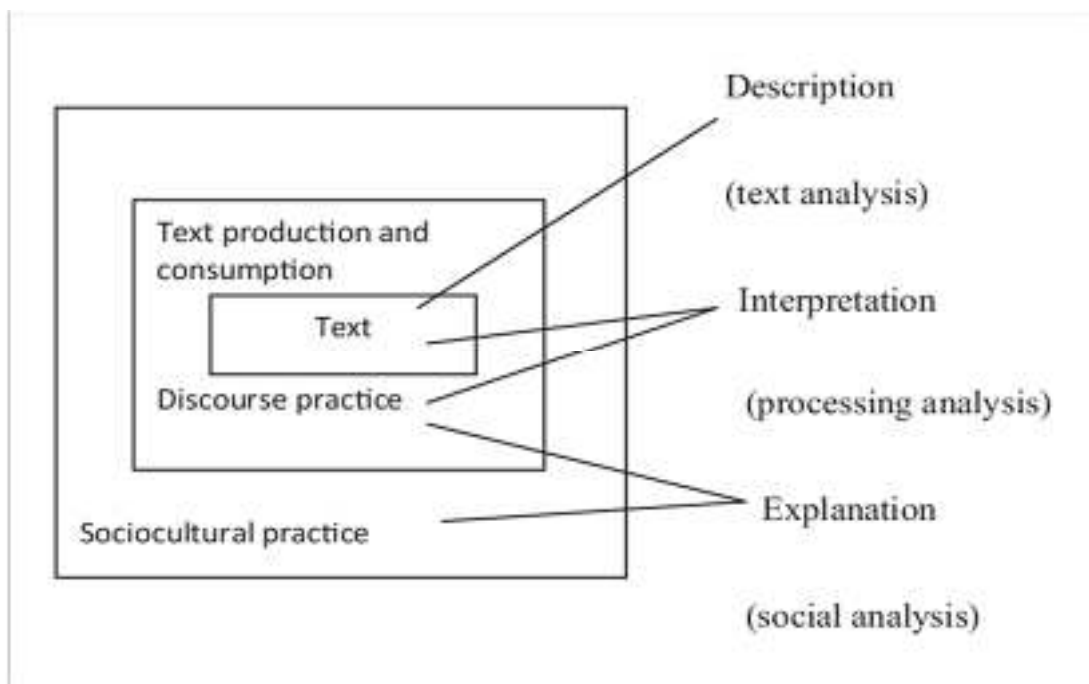

Figure.1 Fairclough model (1995)

The first dimension of the Fairclough model (1995) is text, the second is discourse practice, and the third is sociocultural practice. In this model, he offered three different phases for analyzing any discourse: description, interpretation, and explanation. According to Fairclough (1992), in the description stage, linguistic properties of text are being analyzed, like which type of vocabulary, grammar, wording, and transitivity are used in the text. This stage deals with the formal properties of text (Fairclough, 1992).

Respondents' text production procedures view the text as the output of a production process, whereas text interpretation considers the text as a source for understanding. Both are included in the interpretation stage. News sources and reporting modes can be examined when creating news reports. "Explanation is a process of viewing a discourse as a component of processes of social conflict, inside a matrix of relations of power," (Fairclough, 1989). Its goal is to explain how discourse

and social structure interact by showing speech as one component of a social process (Fairclough, 1989).

Material and Methods

From the start to the end of the research, two paradigms, qualitative and quantitative research methodologies, define the path of a research study. Data is collected from the most readable local Pakistani English newspapers, "*The Dawn*" and "*The Express Tribune*." Corpus of both newspapers was generated using news headlines published from August 2021 to December 2021. The number of words in *The Dawn* corpus is 24,000, and *The Express Tribune* contains 32,000. AntConc software version 3.5.8 has been applied for the analysis of corpora, and the findings are given through bar graphs. Fairclough three- dimensional model (1995) has been selected for the analysis of data.

As a result, this research study employed AntConc's "Keyword List" function to find out a numerically significant correlation between the target and reference corpora. For reference corpus, "Corpus of Pakistani English of various newspapers" is selected, a large corpus of 1,931,412 words. This study used the functions of "Concordance" to conduct a qualitative analysis of some specific keywords; Covid, Afghanistan, Facebook, Tragedy, and harassment in the keyword list to evaluate the effectiveness and function of these words in the construction of the discourse. The second level in Fairclough's (1995) three-dimensional framework is interpretation. This part examines the news sources of headlines like specific, semi-specific, and unspecific news sources to deal with the procedure of discourse production. The explanation stage aims to reveal hidden ideologies and cultural factors in the social context.

Data Analysis

This is a CDA-based study, and the core stream of this study was largely based on the Fairclough's Three-Dimensional model (1995), which is extremely useful for deciphering the ideology of newspaper headlines.

Textual Analysis (Description)

In this stage, the linguistic features of the text are analyzed and interpreted. For quantitative analysis, both newspapers' keyword lists and normalized frequencies are presented in Figures 2 and 3. For normalization, the raw frequencies are divided by the total number of words in the corpus and multiplied by a 10,000-norming base. Only the words which are used for further concordances analysis are given in figures 2 and 3.

Results and Discussion

The Express Tribune

The total keyword tokens of this file are 12671, and the whole numbers of keyword types are 435. Only specific words that link to the purpose of the study are selected; they are a total of 60 in number. The findings of the keyword list quantitative analysis of *The Express Tribune* are presented (see figure 2), proceeded by qualitative analysis utilizing the "Concordance" functions.
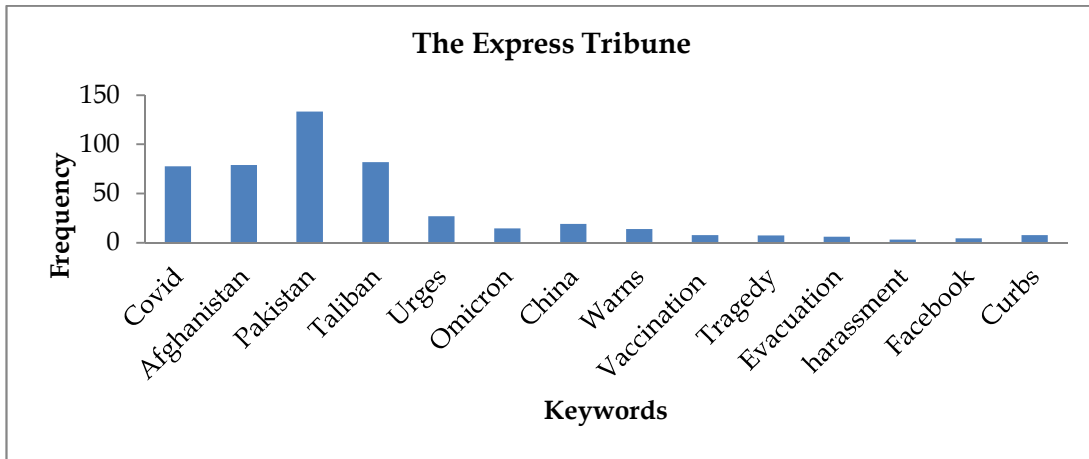
The Express Tribune

Figure 2 Graphical representation of frequently employed words in *The Express Tribune*

The Dawn

The keyword tokens of this file are 13372, and the total numbers of keyword types are 477. Only specific keywords shown in figure 2 are selected for a thorough analysis.
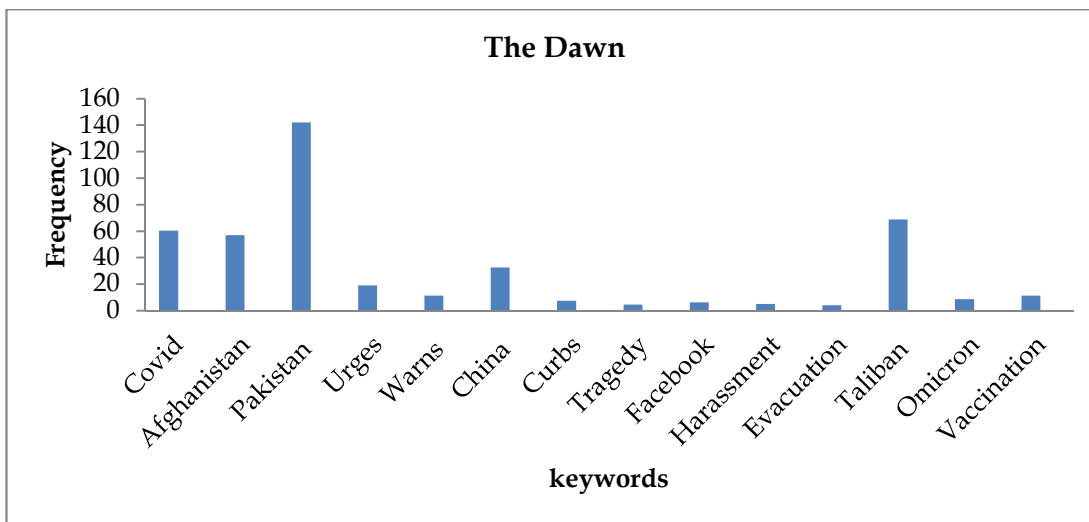
The Dawn

Figure 3 Graphical representation of frequently employed Words in *The Dawn*

Concordance Analysis

More in-depth information regarding keywords can only be discovered by studying concordance, thus enabling the analysis of keywords more accurately. Large numbers of keywords are selected in this study but for concordance, only a few of them have been focused.

(1)   The word "Covid" appears more frequently in *the Express Tribune* than in *The Dawn*. Some keywords are utilized concerning the keyword Covid for expressing the contrasting opinions of both newspapers' coverage, such as curbs, against, vaccine, protesters, and vaccination. The Governments of various countries have stiffened the restrictions in response to the new Covid variant, and they presume cooperation from citizens regarding the government's recent epidemic-related regulations. However, the use of the word "against" in Covid-related headlines indicates that individuals disagree

with the measures and restrictions implemented by various countries, as evidenced by their actions, such as people protesting against COVID curbs. Just as *The Express Tribune* reported in the following headlines.

Thousands protest in New Zealand against COVID-19 rules (The Express Tribune, Nov 09, 2021)

Violent protests against COVID curbs continue in Australian city (The Dawn, Sep 23, 2021)

Even though the countries were conscious that the epidemic was spreading and those constraints were required, instead of neglecting residents' needs and compulsions, they adopted a variety of efforts to offer all forms of necessities to people during the lockdown. It demonstrates that all countries place a high value on the needs of each individual who is unable to earn a living due to the restrictions. As a result, unlike *The Dawn*, this only reports on citizen protests against epidemic curbs. *The Express tribune* reports on constructive actions taken by countries to unite inhabitants from all areas for a common purpose of fighting against epidemics collectively.

(2)   The frequency of the word Afghanistan is higher in *The Express Tribune* than in *The Dawn*. It is repeatedly used to refer to the reaction of different countries to the evacuation of the US from Afghanistan.

UK lawmakers condemn PM Johnson and US President Biden over Afghanistan (The Express Tribune, August 18, 2021)

UK says US pullout from Afghanistan a "mistake" (The Express Tribune, August 13, 2021)

The rapidity with which the Taliban gained ground after the US forces evacuated most of their troops from Afghanistan astounded the West. The words "mistake and "condemn" are used to show the anger of UK lawmakers over the US pullout from Afghanistan. These headlines merely underscore the fact that the United States has placed all countries in danger due to the evacuation.

PM sees civil war if Taliban fail to form inclusive govt in Afghanistan (The Dawn, Sep 22, 2021)

India sponsored terrorism, invested $3b in Afghanistan to destabilize Pakistan (The Express Tribune, Sep 11, 2021)

If the Taliban fail to build an inclusive administration, Prime Minister Imran Khan has cautioned that Afghanistan may devolve into civil war. The more intense and robust words "civil war" and "terrorism" are used to grab readers' attention and make the headline sensational. These phrases frequently invoke terror in the minds of the general public. They instill panic and horror in the public sphere. Headlines of *The express Triune* also present the negative image of India, which is promoting terrorism and spoiling peace in Pakistan with the aid of Afghanistan.

We 'lost' the 20-year war in Afghanistan: top US general (The Express Tribune, Sep 30, 2021)

US forces leave Afghanistan after 20 years (The Dawn, August 31, 2021)

Both headlines are straightforward and effective in summarizing the incident of the defeat of the US in Afghanistan. The word "leave" represents the permanent withdrawal of the US from Afghanistan. News editors frequently utilize numbers in headlines to engage readers on both a practical and cognitive level to provide factual news to readers. The figure "20 years" headline offers the readers statistical information.

Respect for rights will lead to Afghanistan stability: PM (The Dawn, Nov 13, 2021)

Pakistan's efforts for inclusive govt in Afghanistan must be supported: Fawad (The express Tribune, August 30, 2021)

News writers employed modal verbs "will" and "must" to describe foreign minister Shah Mahmood Qureshi, Fawad Chaudhary, and PM comments regarding the practical steps that Pakistan will take in the future. These headlines show the solidarity of Pakistan with Afghanistan. Moreover, these headlines positively reflect Pakistan in this chaotic situation. In comparison to *The Dawn*, the terminology employed in *The Express Tribune* is more informative.

(3)   *The Express Tribune* uses the word "Facebook" more frequently than *The Dawn*. The findings of the keyword Facebook from both corpora are focused on Facebook's new glasses and the outage of Facebook, Instagram, and WhatsApp, which prevents users from assessing these social media apps. Whistleblower's accusations that Facebook is more concerned with profits than with people's safety are also highlighted in the headlines of both newspapers.

Whistleblower says Facebook pushed for higher profits, cavalier over safety (The Express Tribune, Oct 05, 2021)

Whistleblower says Facebook operating in the shadows (The Dawn, Oct 06, 2021)

These headlines demonstrate the people's distrust of Facebook, implying that it is willing to jeopardize public safety for profit. These headlines portray Facebook negatively in front of the general public. The whistleblower's use of phrases like "higher profit," "shadow," and the adjective "cavalier" demonstrates that they are attempting to persuade and convince people about Facebook's negligence concerning the safety of public data.

Facebook unveils its first "smart" glasses (The Express Tribune, Sep 09, 2021)

Facebook and Ray-Ban debut "smart" shades (The Dawn, Sep 10, 2021)

The Adjectives "first smart" and "new smart" are used to describe the digital advancement and development of glasses like the use of these smart glasses will provide a virtual environment. The monosyllabic word "first" shows that these glasses were first time introduced and launched as brand new in the digital market. *The Dawn* concentrates on Facebook's collaboration with Ray-Ban for the debut of smart glasses. At the same time, *the Express Tribune* merely mentions that Facebook has unveiled its smart glasses and does not mention the Ray-Ban association with Facebook in its primary headline.

(4)   The frequency of the occurrence of the word "harassment" is higher in *The Dawn* than in *The Express Tribune*. The sexual harassment of women is the focus of the concordance lines of "harassment" of both corpora. The findings demonstrate that

whether women work as industrial workers or as inspectors, they are harassed in every profession. The press coverage of a journalist's harassment case is also examined.

Tesla subjects women to nightmarish sexual harassment at factory (The Express Tribune, Nov 19, 2021)

IAEA complains to Iran over "harassment" of women inspectors (The Dawn, Sep 16, 2021)

The topic of women's harassment dominates the headlines. All women are in pain because of this ever-increasing issue. These headlines depict a pitiful condition with horrible circumstances. The language employed in headlines is full of emotions. The words like "sexual harassment" and the adjective "nightmarish" have a significant effect on the mind of readers. News writers employ emotive language to raise readers' emotions in response to these incidents. They attempt to portray a bleak mood through headlines, and they want safety measures to prevent such catastrophes.

32 journalist harassment cases filed in a year, Islamabad police inform Supreme Court (The Dawn, Nov 13, 2021)

SC constitutes five-member larger bench to hear journalists' harassment case (The Express Tribune, August 21, 2021)

The use of "32 journalists" and "a year" in the headline refers to a numeral reference to a number and incident. These words are used together to make the news more surprising and sensationalizing for the readers. Simple present tense "constitutes" is used to show the newness of the event. The verb "constitutes" indicates SC's action to establish a five-member bigger bench for hearing the harassment cases of journalists like shooting issues, assaults, hurt, death threats, etc.

(5)   The term "tragedy" appears more frequently in *The Express Tribune* than in *The Dawn*. The Sialkot tragedy, which shocked the entire Muslim world, is highlighted in the concordance results of both corpora. Different negative words are used to point out this issue, like "ashamed," "lynching," "condemns," and "ghastly murder."

Ashamed of Sialkot tragedy: Maulana Tariq (The Express Tribune, Dec 22, 2021)

Ghastly murder of Lankan man in Sialkot tragedy shames nation (The Dawn, Dec 04, 2021)

The word "ashamed" has a negative connotation, implying that something embarrassing has occurred for which we should feel ashamed. The adjective "ghastly" refers to something terrifying or disturbing. This headline depicts the news writer's attitude to the situation, which is that it is a national embarrassment. These statements elicit pity and empathy from the audience in response to the catastrophe. The tone of these headlines is emotive.

Both corpora's findings focus on political and important bodies' reactions to the disaster in Sialkot. The meaning of the word is apparent when it comes to persons who use the name of Islam to do such horrible acts. These occurrences are a primary cause of islamophobia, highlighting the urgent need to educate people about actual Islamic beliefs so that they are not duped in the name of Islam.

Analysis of Discourse Practice (Interpretation)

The discursive practice focuses on how texts are written, consumed, and interpreted in society. Discursive practice analysis focuses on the production and consumption of the text and how power relations are established. News sources will be analyzed in this stage to explore the process of text production. In examining news sources in news discourse, three methods have been used: specific, semi-specific, and unspecific news sources. Individuals who offer precise information about the news provider, such as their identity, position, and so on, are referred to as "specific sources." The semi-specific source offers general remarks about the source. The unspecific source is used when specific information is unknown. The overall numerical results of news sources for five keywords have been represented in Table 1.

The statistical results of table 1 have displayed the highest percentage from all sources, specifically news sources (with a 55.30% percentage). The highest usage of particular news sources indicates that both newspapers prefer to present balanced and impartial headlines free from subjectivity, which clarifies that the news writer is not giving his opinion but providing the statements of others. With 28.03% percent and 16.6% percent, respectively, the semi-specific and unspecific sources account for the second and third largest proportions. Some of the sentences that use specific, semi-specific, and unspecific sources may have ideological content. Here are instances to consider.

(1) 'Nation building' never a US goal in Afghanistan, **says Biden**

(2) **US study** suggests COVID-19 vaccines may be ineffective against Omicron without booster

(3) Facebook is selective in curbing hate speech against Indian Muslims: **study**

Example 1 shows that Biden strongly defended his decision to pull out US troops from Afghanistan. In this headline, the attitude of the news writer is balanced and unbiased by adding the words of President Joe Biden. This is a true example of a specific news source. Presenting the core of state ideology is example number two. Different studies are undertaken in every sector to guide the readers to a large extent. The validity of these judgments cannot be directly traced if the identification of these studies is not exposed. *The Express Tribune* is possibly utilizing its identities to convey its opinions.

Furthermore, the validity and reliability of these judgments must be examined further. According to a study conducted in the United States, the covid-19 vaccination is ineffective without a booster. Because of the country's name, which suggests that this study was conducted in the United States, there is some reliability and neutrality in this headline. This is an excellent example of a reliable semi-specific news source.

Headline (3) is an example of an unspecific news source. The reporter indicated that Facebook is curbing hate speech against Indian Muslims but didn't refer to the name of the study. This headline has an unspecific news source because the detailed information of the news provider is missing. There is only a mention that a study was conducted, but the name of the study is missing.

Table 1

News sources used in all headlines

Sources	Specific	Semi-specific	Unspecific	Total
Total	73	37	22	132
Percentage	55.30%	28.03%	16.6%	100%

Social Analysis (Explanation)

Discourse is regarded as a form of social behavior. It's vital to evaluate the institutional and societal context to determine what the reporter wants to say. This stage deals with the social analysis of keywords such as Covid, Afghanistan, Facebook, harassment, and tragedy. So, each keyword's analysis is conducted step by step. Because the current study includes general headlines that cover a variety of themes, rather than one single ideology, each topic has a different ideology. The goal of reporting on Covid-19, according to the headlines, is to raise public awareness about the pandemic and encourage people to engage in the government's immunization programme to combat it.

In this stage, the ideologies of media practitioners have discussed the issue of Afghanistan. It is apparent that Pakistan print media covers all aspects of the matters like the reaction of different countries regarding the evacuation of US forces. Still, the primary focus of media is Pakistan. It focuses on the positive steps that PM Imran Khan immediately took to provide aid to Afghanistan. He also asks other countries to provide assistance to Afghans and make the country stable to prevent any terrorist activities. Although social media apps are now ignoring children and adolescents' safety and mental health, as evidenced by recent unnamed whistleblowers. They shared a cache of business documents with Congress and The Wall Street Journal detailing how Facebook knew its products, particularly Instagram, were harming girls' self-image but ignored this fact.

By explaining headlines related to keyword harassment, it becomes clear that news media is trying to present the negative image of society that although we make progress in every field of life, we still live a conservative life where males are dominant. We still believe in a patriarchal system where women are not as accessible and powerful as men. Through the headline's news, reporters try to introduce women's empowerment that each society should work on the empowerment and safety of women.

The media is open in giving its opinions on various issues. Islamophobia is one of the universal issues in this era; people, by using the name of Islam, spread violence and barbarity in society. Sialkot tragedy is one of those issues which make the whole Muslim nation ashamed in front of the world. The central ideology of Pakistani print media behind this incident is to provide awareness to people that Islam does not allow such oppression and killing of anyone by using its name. Print media of a country builds a country's image. In the current times, the media is powerful enough to destabilize a country's government and its economy and can deteriorate its reputation. Pakistani print media is also doing the job of awareness creation about national and international events. It targets social issues with a perspective of educating the masses and thus providing means to eradicate these social evils.

Conclusion

The media authorities in Pakistan have a social obligation to contribute to social harmony and stability by playing a beneficial role. As a result, it is envisaged that Pakistan's media institutions will seek to promote acceptance, cooperation, peace, and prosperity by self-regulating their own media corporations. This research study focused on using linguistics features to represent different issues in headlines and the media's ideology toward these issues in *The Dawn* and *The Express Tribune* newspapers by examining the headlines using Norman Fairclough's three-dimensional model (1995) and the corpus analyzer AntConc.

In this analysis, the occurrence of keywords like Covid, Afghanistan, Facebook, harassment, and tragedy is significant. According to the results, *the Express Tribune* had a higher frequency of some of these terms than *The Dawn*. The findings further reinforced the notion that Critical Discourse Analysis may be used to track, discover, and make explicit the underlying ideologies of newspaper headlines that are usually hidden from the broader public. The results show that simple present tense, present participle, adjectives, modal verbs, and positive and negative lexical items are used frequently in both newspaper headlines. Moreover, results also prove that a specific source of news is employed mainly through newspapers.

Despite the thoroughness of the analysis, there are still certain restrictions. For starters, the data is limited to the analysis of five keywords, which is insufficient and can be expanded. If there is ample time, further keywords for concordance analysis should be included to make a more accurate and balanced conclusion. Furthermore, the analyses' views can be developed further, such as from the viewpoint of transitivity, modalities, and so forth. In terms of the study's future implications, the current research will aid forthcoming CDA researchers in Pakistan in understanding the importance, principles, and scope of CDA to explore its fundamental ability to introduce changes to society through spreading awareness.

References

Duanprakhon, P. (2012). *Critical Discourse Analysis of News Headlines: A Case of Youth Crime in Thailand* (thesis). National Institute of Development Administration.

Fairclough, N. (1989). *Language and Power*. London: Longman.

Fairclough, N. (1992). Discourse and text: Linguistic and intertextual analysis within discourse analysis. *Discourse & Society*, *3*(2), 193–217. https://doi.org/10.1177/0957926592003002004

Fairclough, N. (1995). *Critical Discourse Analysis: The Critical Study of Language*: Harlow: Pearson Education.

Grabe, M. E., & Zhou, S. (2001). Explicating sensationalism in television news: Content and the bells and whistles of form. *Journal of Broadcasting & Electronic Media*, *45*(4), 635–655.

Muspratt, S., Luke, A., & Freebody, P. (1998). *Constructing critical literacies: Teaching and learning textual practice*. Hampton Press.

Pornjan, D. (2012). Critical discourse analysis of news headlines: A case of youth crime in Thailand. *(Unpublished master's thesis). School of Language and Communication National Institute of Development Administration, Thailand*.

Taiwo, R. (2007). Language, Ideology and Power Relations in Nigerian Newspaper Headlines. *Nebula*, *4*(1), 218–245.

Ulum, O. G. (2016). Newspaper ideology: a critical discourse analysis of news headlines on Syrian refugees in published newspapers. *Turkish Studies*, *11*(5), 541-552.

Van Dijk, Teun A. (1993). Principles of critical discourse analysis. *Discourse & society*, *4*(2), 249-283.

Yasmeen, K., Mahmood, M. A., Mahmood, R., Shakir, A., & Sattar, A. (2014). A Critical Discourse Analysis of Newsworthiness in English and Urdu Newspaper Headlines In Pakistan. *New Media and Mass Communication*, *24*, 60–14.