

PAKISTAN JOURNAL OF LANGUAGE AND LITERARY STUDIES

Volume 2, Issue 1

Spring 2026

A Peer-Reviewed Journal of Linguistics, Literature and Language Studies

ISSN (Online): ***3137-0373***

Website: https://ojs.gscwu.edu.pk/index.php/pjlls

Email: pjlls@gscwu.edu.pk

Published by: Department of English, Faculty of Arts and Social Sciences, Government Sadiq College Women University, Bahawalpur, Pakistan

EDITORIAL TEAM

Patron in Chief	Vice Chancellor, Government Sadiq College Women University, Bahawalpur
Editor	**Prof. Dr. Musarrat Azher** Dean, Faculty of Arts and Social Sciences Govt. Sadiq College Women University Bahawalpur
Associate Editors	**Dr. Saba Sadia** Assistant Professor, Department of English Govt. Sadiq College Women University Bahawalpur **Dr. Breya Syal** Lecturer, Department of English Govt. Sadiq College Women University Bahawalpur **Dr. Nida Zahra** Lecturer, Department of English Govt. Sadiq College Women University Bahawalpur
Member from Dept	**Ms. Ayesha Durrani** Lecturer, Department of English Government Sadiq College Women University, Bahawalpur
Assistant Editor	**Engr. Rakib Mustafa** Assistant, Dean Faculty of Arts and Social Sciences Government Sadiq College Women University, Bahawalpur

Journal Information

ISSN (Online): 3137-0373
Frequency: Biannual (Spring and Autumn)
Mailing address: *Department of English, Faculty of Arts and Social Sciences, Government Sadiq College Women University, Bahawalpur, Pakistan*
Email: *pjlls@gscwu.edu.pk*
Website: *https://ojs.gscwu.edu.pk/index.php/pjlls*

Copyright: © 2026 Pakistan Journal of Language and Literary Studies. All rights reserved. Views expressed are those of the authors and do not necessarily reflect those of the editors or publisher.

ADVISORY BOARD

Foreign Members

No.	Name	Designation & Institution
1	***Prof. Dr. Dechao Li***	Professor Department of Chinese and Bilingual Studies The Hong Kong Polytechnic University *dechao.li@polyu.edu.hk*
2	***Prof. Dr. Jacklin Troy***	Professor The University of Sydney *jakelin.troy@sydney.edu.au*
3	***Dr Martin Weiser***	Visiting Professor University of Salzburg, Austria *weissermar@hotmail.com*
4	***Prof. Dr. Shobhana Chelliah***	Professor Department of Linguistics Indiana University Bloomington *schellia@iu.edu*
5	***Dr. Dongmei Cheng***	Associate Professor, Literature and Languages The Texas A&M University System *dongmei.Cheng@tamuc.edu*
6	***Dr. Flavia Belpoliti***	Associate Professor and Director of Spanish Grad Studies Literature and Languages The Texas A&M University System *flavia.Belpoliti@tamuc.edu*
7	***Dr. Muhammad Afzaal***	Associate Professor Research: Corpus Linguistics, Discourse Analysis Shanghai International Studies University, China *muhammad.afzaal1185@gmail.com*
8	***Dr. Tribhuwan Kumar***	Assistant Professor Department of English Language and Literature Prince Sattam bin Abdul Aziz University, Sulail, Riyadh *tribhuwan.literature@gmail.com | t.kumar@psau.edu.sa*

ADVISORY BOARD

Local Members

No.	Name	Designation & Institution
1	***Prof. Dr. Mamoona Ghani***	Professor Department of English Islamia University Bahawalpur *mamunaghani@hotmail.com | drghani2009@yahoo.com*
2	***Prof. Dr. Saiqa Imtiaz Asif***	Professor Department of English BZ University Multan *saiqaasif@yahoo.com*
3	***Prof. Dr. Asim Mehmood***	Professor, Chairman Linguistics Govt. College University, Faisalabad *masimmahmood@gcuf.edu.pk | masimrai@gmail.com*
4	***Prof. Dr. Muhammad Asif Khan***	Professor, Chairperson Department of English Literature Islamia University Bahawalpur *Asif.khan@iub.edu.pk*
5	***Dr. Muhammad Kamal***	Associate Professor, Department of English Allama Iqbal Open University, Islamabad *mkkamazai@hotmail.com*
6	***Dr. Tazanfal Tehseem***	Associate Professor, Department of English University of Sargodha *Tazanfal.tehseem@uos.edu.pk*
7	***Dr. Tehseen Zahra***	Associate Professor Department of Humanities and Social Sciences Bahria University *Tehseen.azhar@yahoo.com*
8	***Dr. Azhar Pervaiz***	Associate Professor, Department of English University of Sargodha *Azhar.pervaiz@uos.edu.pk*

CONTENTS

Front Matter ..... **i–ix**

Editorial Note ..... **vi**
Prof. Dr. Musarrat Azher

Framing Human Tragedy: A Semiotic Analysis of War Images in the Israel-Palestine Conflict **1–18**
Jawaria Sajid, Muhammad Arif & Muhammad Asaf Amir

The Role of Interactional Metadiscourse Markers in Portraying Female Identity: A Corpus-Based Study of Fantasy TV Characters **19-35**
Farhat Bano

Slavery, Trauma and Healing: A Corpus-Based Analysis of Temporal Markers in Selected African-American Neo-Slave Narratives. **36-51**
Mahnoor Fatima

Evaluating Forced Alignment in an Under-Documented Shakargarh Punjabi Speech Corpus **52-67**
Adnan Arif

Analyzing Social Disparity: A Critical Discourse Analysis of Hosseini's and the Mountains Echoed **68-80**
Arshad Hussain

EDITORIAL NOTE

It is a pleasure to present Volume 2, Issue 1 (Spring 2026) of the *Pakistan Journal of Language and Literary Studies*. The journal was founded to give scholars of language, literature and linguistics a rigorous, open space in which to share their work. This issue continues that commitment with 4 research articles that reflect the breadth of current scholarship in the field.

The contributions come from researchers at universities and colleges across Pakistan. They draw on a range of methods and frameworks, from discourse analysis and multimodal semiotics to literary criticism and language pedagogy. Several articles look at how meaning is made and contested across texts, images and social contexts. The opening article, for example, applies a semiotic model to news photographs, asking how visual signs construct meanings of loss, displacement and humanitarian crisis. Taken together, the papers show that questions of language and representation remain central to understanding the societies we live in.

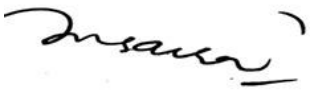
Every article here was assessed through a double-blind peer review process. We are grateful to our reviewers, whose careful and constructive feedback strengthened the work. We also thank the members of our Advisory and Editorial Boards for their guidance, and the authors for trusting us with their research.

We would like to see more interdisciplinary and comparative work, studies grounded in local linguistic and literary contexts, and the voices of emerging scholars. We invite researchers to submit original, unpublished manuscripts for forthcoming issues, following the guidelines on our website.

The views expressed in the articles are those of the individual authors and do not necessarily reflect those of the editors or the publisher.

We hope you find this issue informative and thought-provoking.

With best wishes,



Prof. Dr. Musarrat Azher
Editor, PJLLS
Dean, Faculty of Arts and Social Sciences
Govt. Sadiq College Women University Bahawalpur

CONTRIBUTORS

Framing Human Tragedy: A Semiotic Analysis of War Images in the Israel-Palestine Conflict

Jawaria Sajid
M.Phil., Department of Applied Linguistics, Government College University Faisalabad

Jawaria Sajid is currently pursuing a PhD in Applied Linguistics from Government College University Faisalabad. She holds an M.Phil. in Applied Linguistics from Government College University Faisalabad and a BS in English Literature and Linguistics from Government College Women University Faisalabad. Her professional experience includes English language teaching, IELTS training, and compliance instruction. She has co-authored several research articles published in national and international journals. Her research interests include Systemic Functional Linguistics, Discourse Analysis, Gender and language, Corpus linguistics, Multimodal discourse analysis, and language assessment. She was also selected for the competitive Amal Career-Prep Fellowship and has completed professional development courses offered by the British Council and the U.S. Department of State.

Email: *jawariasajid4126@gmail.com*

Muhammad Arif
Lecturer in English, Punjab Postgraduate College, Alipur, Punjab

Muhammad Arif is an English Lecturer at Punjab Group of Colleges, Alipur, and holds an MPhil in English. His scholarship brings radical feminist and postcolonial theory, together with discourse analysis, to bear on literature and linguistics, examining how misogyny and cultural power are encoded in narrative and language. Committed to critical self-reflexivity, he questions inherited assumptions, recovers marginalized voices, and writes to unsettle received readings of the literary canon. His areas of interest are Pakistani & Indian fiction in English, Literary Studies, and Discourse Stylistics.

Email: *arifniaazi@gmail.com*

Muhammad Asaf Amir
Associate Professor of English, Govt Associate College, Naushehra (Khushab)

Dr. Muhammad Asaf Amir is presently serving as an Associate Professor in the Higher Education Department, Government of the Punjab. He has vast experience in teaching and research. He holds a PhD in Applied Linguistics. His doctoral research analyzes the relationship between power and ideology in Pakistani English Fiction using a corpus-based discourse stylistic technique. His areas of interest include Corpus Linguistics, Critical Discourse Analysis, Stylistics, Postcolonial Studies, and Pakistani English Fiction.

Email: *masafamirmalik@gmail.com*

The Role of Interactional Metadiscourse Markers in Portraying Female Identity: A Corpus-Based Study of Fantasy TV Characters

Farhat Bano
BS Scholar Department of Linguistics and Social Sciences, The Begum Nusrat Bhutto Women University, Sukkur, Pakistan

Farhat Bano is an undergraduate student of BS English (Linguistics) at The Begum Nusrat Bhutto Women University, Sukkur, Pakistan. Her academic interests include corpus linguistics, discourse analysis, sociolinguistics, pragmatics, forensic linguistics, and language and gender. Her research

focuses on the application of corpus-based and discourse-analytic approaches to media, digital communication, and identity construction. During her research internship, she conducted this study on interactional metadiscourse and female identity in fantasy television characters. She has presented her research at academic conferences, including work on speech acts, political discourse, and mental health discourse. In 2024, she secured first position at the 3rd International Conference on Linguistics and Multidisciplinary Research (ICLMR).

Email: *farhatofficial6@gmail.com*

Slavery, Trauma and Healing: A Corpus-Based Analysis of Temporal Markers in Selected African-American Neo-Slave Narratives

Mahnoor Fatima
Research Scholar, Department of English Language and Literature, Government Fareed Bukhsh Ghousia Graduate Science College 333 GB , Toba Tek Singh

Mahnoor Fatima is a Bachelor of Science candidate in English Language and Literature at Government Fareed Bukhsh Ghousia Graduate Science College 333GB, affiliated with Government College University Faisalabad (GCUF), Pakistan. Currently, she is working as a Research Scholar. Her scholarly interests operate at the intersection of critical literary theory and empirical digital methodologies. Her primary research focus encompasses African-American studies, post-colonialism, Marxism, feminism, and eco-criticism, alongside a distinct computational focus on corpus linguistics and generative AI in online communication discourse. A recipient of the merit-based Honhar Scholarship, she also actively participates in the human rights and youth leadership initiatives.

Email: *harrara12345@gmail.com*

Evaluating Forced Alignment in an Under-Documented Shakargarh Punjabi Speech Corpus

Adnan Arif
PhD Scholar, School of English Language, Technology and Literature, Faculty of Languages and Literature, Minhaj University Lahore, Pakistan

Adnan Arif is a PhD Research Scholar in English Linguistics at the School of English Language, Technology and Literature, Faculty of Languages and Literature, Minhaj University Lahore. He holds an MPhil in Applied Linguistics from the University of Management and Technology. His research interests include Punjabi/Lahnda dialectology, corpus-driven morphosyntax, sociolinguistics, language variation, language documentation, orthography, natural language processing, and cross-linguistic research. His current doctoral work focuses on linguistic variation, documentation, and corpus-driven analysis of Punjabi varieties.

Email: *2023F-mulphd-eng-007@mul.edu.pk*

Analyzing Social Disparity: A Critical Discourse Analysis of Hosseini's and the Mountains Echoed

Arshad Hussain
MPhil scholar, Department of English, University of Sargodha, Sargodha

Arshad Hussain holds an M.Phil. in English from the University of Sargodha, Pakistan, and an M.A. in English. He has over three years of teaching experience and is currently serving as an English teacher at Comprehensive Model High School, Sargodha, Pakistan. His research interests include

Critical Discourse Analysis (CDA), the relationship between language and ideology, and the representation of social inequalities in literary texts. His M.Phil. research examined Khaled Hosseini's And the Mountains Echoed through Fairclough's three-dimensional model of Critical Discourse Analysis and social inequality theory. His scholarly interests focus on exploring how linguistic choices construct meaning, reflect power relations, and reproduce social disparities in literary discourse.

Email: *engl62bs21r007@gmail.com*

PAKISTAN JOURNAL OF LANGUAGE AND LITERARY STUDIES
Volume 2, Issue 1 (Spring 2026)
Pages: 1-18

Framing Human Tragedy: A Semiotic Analysis of War Images in the Israel-Palestine Conflict

Jawaria Sajid[1], Muhammad Arif[2] & Muhammad Asaf Amir[3]

ABSTRACT

The present study examines the human tragedy during Post-October 2023 Israel-Palestine conflict through the lens of Riera (2020)'s semiotic model, which consists of three levels: a) Concept (Sign, Context, and Meanings), b) Construct (semiosis), and c) Proposition (reader response). It focuses on how contextual elements, visual signs, and semiotic relationships build reflective and embedded meanings in war images. The study is qualitative with a sample of four war-related images, selected purposively from the Moscow Times news website for analysis. The findings show that there is a human tragedy in Palestine in the shape of social and cultural disturbance, as well as residential areas, healthcare facilities, religious spaces, and shelters. The selected images also show the unequal power relations such as vertical relations of the people being dependent on the war choices and bombing by IDF (Israeli Defense Force) from sky. In contrast, horizontal relations between people and places are abandoned, resulting in the destruction of houses, worship places, hospitals, and shelters. Despite making geopolitical statements, the study explores how these visual representations construct meanings of loss, displacement, and humanitarian crisis through semiotic processes. Humans have to suffer not only because of the actions of Israel but also due to the birth of human tragedy; supported by international powers with Israel. The current study contributes to visual semiotics and conflict discourse analysis by showing how war images construct different meanings about tragedy through semiotic processes.

Keywords: *Concept, Construct, Proposition, Human Tragedy, Semiotic Analysis*

[1] M.Phil. Dept of Applied Linguistics, Government College University Faisalabad
[2] Lecturer in English, Punjab Postgraduate College, Alipur, Punjab
[3] Associate Professor of English, Govt Associate College, Naushehra (Khushab)

Corresponding author:
Ms. Jawaria Sajid, M.Phil. Dept of Applied Linguistics, Government College University Faisalabad
Email: jawariasajid4126@gmail.com

1. Introduction

Human tragedy is a grave and disastrous occurrence that leads to a great deal of suffering, death, or destruction of human life. The phrase is often used to characterize circumstances like poverty, starvation, or disease outbreaks that cause great suffering and fatalities, as well as occurrences like war, natural catastrophes, terrorist attacks, or other acts of violence. Individual tragedies that affect an individual or a family are also considered as human tragedies. Human tragedies are often regarded as terrible and profound occurrences that significantly affect both individuals affected and the larger community. The term "human tragedy" typically refers to the emotional reactions of people and groups to the incident, in addition to the harm done, the loss, and the grieving process (Santos, 2023).

War, a human tragedy, involves large-scale violent conflict affecting the large number of people by putting an enormous toll on those who are involved in it (Sturma, 2019). The Israel-Palestine conflict is one of the major conflicts of the century. It has affected all the major surrounding countries due to conflict of ideology and culture between the two nations (Global Conflict Tracker, 2024). Hamas (the militant group of Palestinians) launched an operation against IDF (Israeli Defense Force) on October 07, 2023. The reason behind the attack was to drive the IDF back from the Gaza Strip (Global Conflict Tracker, 2024). Despite launching the ground operation against militants, IDF started airstrikes on Palestinians. About 34,000 deaths have occurred during the operation of IDF in the name of defense, including women, children, and the elderly (Global Conflict Tracker, 2024). About 1,200 people were killed by Hamas, and 250 people were kept as hostages (BBC, 2024). It was 1947 when UN (United Nations) split the land of Palestinians into Jewish and Palestine states, but Palestinians rejected it, and the conflict started (BBC, 2024). Thousands of people have been killed in this conflict while millions have been displaced (Alsaafin, 2023). The displacement of people has led to the birth of human tragedy.

United Nations Human Rights (2024) disclosed that about 1.9 million people have been displaced out of 2.3 million in the Israel-Palestine conflict. Moreover, the sanitary conditions are dreadful, and shelters are cramped after the IDF bombing (United Nations Human Rights, 2024). Humans suffer, and there is a massive scarcity of food, medicine, water, tents, and basic necessities (United Nations Human Rights, 2024). The international media broadly captured these humanitarian consequences and expressed them through news photographs and other forms of visual media. These visual representations not only present clash but also signify meanings of suffering, displacement, resilience, and war.

The occurrence of human tragedy results in severe human sufferings. In the recent conflict, 30,000 tons of bombs were dropped, and 13,000 people were killed, including approximately 5,600 children and 3,550 women (Tehran Times, 2023). Wars create their dynamics with language and images. The semiotics and the semantics are redefined in the war images to disclose the communication system, legitimization, and collective mass violence (Jacob, 2021). IDF has bombed all the Palestinian structures with their rage without considering the international law and ethical considerations. War images present the core aspects and the

semantics through different codes. The present study is conducted to reveal human tragedy through the war images due to IDF bombing.

The study adopted the semiotic model of Riera (2020), which consists of three aspects of analysis: a) Concept (Sign, Context, and Meanings), b) Construct (Semiosis), and c) Proposition (reader response). This framework provides a clear examination of how visual images construct distinct layers of meaning and outline audience interpretations of contemporary conflict imagery.

1.2. Statement of the Problem

Human tragedy is the broad term that encompasses human loss, displacement, and subjugation. The concepts of destruction present the hidden meanings of war, which results in the construction of human tragedy. Such human tragedy contributes to the disruption of human civilization, which can be better understood through sign construction.

1.3. Research Objectives

- To understand the human tragedy in Palestine through war Images.
- To explore the inlaid meanings in the signs of war images of Palestine.

1.4. Research Questions

- How do the concept, construct, and proposition present the human tragedy in Palestine?
- What intended meanings are constructed through representation of destruction in Palestine?

1.5. Significance of the Study

The study is descriptive, which helps to understand the core aspects of human tragedy in Palestine by Israel's heavy shelling. The present study can help to understand the inlaid meanings of signs, and their relations with the context. This study can also help to explore the recipient's response being the inhabitant of the planet.

2. Literature Review

International Humanitarian law is based on the Geneva Convention with the additional Protocols of 1977 including the protection of victims of war conflicts (ICRC, 2024). All the states have agreed to these laws, which include the safety and protection of those people and their goods which are not involved in the war as well. The major international treaties of the laws are as follows which are agreed by the international community:

- Convention for The Protection of Cultural Property in The Event of Armed Conflict, Plus Its Two Protocols (1954);
- Biological Weapons Convention (1972);
- Conventional Weapons Convention and Its Five Protocols (1980);
- Chemical Weapons Convention (1983);
- Ottawa Convention on Anti-Personnel Mines (1997);

- Optional Protocol to the Convention on the Rights of the Child on the Involvement of Children in Armed Conflict (2000).

All these laws and protocols are applied at the time of international conflict.

Brown (2007) explored "Tragedy, 'Tragic Choices' and Contemporary International Political Theory". The study mentioned that the fundamental idea of the tragic worldview is that human behavior occasionally, if not frequently, presents a choice between two drastically different but equally awful consequences, and thus, from one point of view, our actions in every given circumstance would be seen as improper. The study also suggests that tragedy is the significant element for understanding political and humanitarian consequences in international relations. However, Brown's research does not explore human tragedy through semiotic analysis and visual representation. The tragic vision is generally rejected by analytical political theory, and most of the current literature on humanitarian intervention and global distributive justice also opposes the idea that there are always clear-cut appropriate responses and acting is always bad. Refusing to acknowledge the sad aspect of human existence has negative effects on one's intellect as well as one's political standing (Brown, 2007).

Segell (2023) studied "An Alternative View of The Israel-Palestine Dispute: The Eschatological Dimension". This study used eschatology (a theological concept about end-of-age expectations that frequently provoke violent reactions) to analyze the Israel-Palestine conflict. The study described that comprehensive studies examining the function of eschatology in this setting are lacking despite a wealth of studies on the Israel-Palestine conflict. Extreme hatred and bloodshed are greatly fueled by religion in the Middle East, which includes Israel and Palestine. The impact of religious traditions, especially eschatology, which often elicits violent reactions, is the main topic of discussion. This was examined in five aspects of the study: daily religion, Jerusalem's function, eschatological prophecy, expansion beyond Jerusalem, and the message delivered. The study concluded that eschatology plays a major role in this struggle and that violent apocalyptic prophecies exacerbate it. Different interpretations of eschatological and apocalyptic prophesies, as well as an awareness of religion's role in the struggle, are necessary to resolve this dispute (Segell, 2023).

Assaiqeli (2021) studied "Palestine in Visual Representation: A Visual Semiotic Analysis of the Nakba". The study mentioned that Israel has made a concerted effort to erase historical Palestine from Arab and international public consciousness since its de facto creation. National themes are constructed and maintained in part through visual discourse. Through a critical discourse analysis of a corpus of visual representations, this paper sought to understand how visual representation memorializes and reconstructs national themes. In doing so, it can act as a counter-hegemonic discourse against attempts to mystify history and memorize the other. This social semiotic research examined the role of visual representations in the memorialization of important Palestinian national themes and, consequently, the reconstruction and preservation of historic Palestine by using Kress and van Leeuwen's (2006) grammar of visual design as a framework for visual analysis. The study highlighted the sociopolitical and emancipatory

function that graphically expresses national themes, such as ethnic cleansing and the Right of Return, in creating the collective consciousness of the Palestinian people about their past, present, and future. Further investigation into visual signals may provide more light on the role that multimodal and visual communication play in national emancipation and the maintenance of significant national themes (Assaiqeli, 2021).

Adawiyah et al. (2020) did a study "Multimodal Depiction of Israelis and Palestinians on Gaza Conflict in Two Middle Eastern Press". This study intended to investigate the portrayals of Israelis and Palestinians in the Gaza conflict by two Middle Eastern Press publications: Daily Sabah and Iran Daily. To identify media ideology, this study used multimodal critical discourse analysis. In other words, it emphasized semiotic resources manifested through visuals in addition to language qualities. Clauses containing Van Dijk's ideological square discourse are analyzed using Halliday's (2004) transitivity. Kress & van Leeuwen's (2006) symbolic function of the image was used to analyze the pictures included in the chosen articles. It was evident from the data that both media outlets portrayed Israelis as the "Other" group and Palestinians as the "Self" group. They depicted the "Other" as the agent of offensiveness, while the "Self" was seen as the victim and the righteous warrior (Adawiyah, Junining, & Tabiati, 2020).

A resent study done by Azher et al. (2025) "Media coverage of Israeli air strikes and aftermaths in Palestine: A visual semiotic analysis" compared two newspapers side by side that is Dawn (Pakistan), and the Guardian (UK) specifying how they portray the aftermath of Israeli airstrikes on Gaza through published pictures. They collected data from June and July 2024, characterizing the images by portrayals of deaths/ mourning, destruction of infrastructure, displacement, starvation, and military actions in Gaza, using Kress and Leeuwen's (2006) framework of visual semiotics, focusing on four major components of the framework namely gaze, social distance, camera angle, and modality. The other framework applied was Martin & White's (2005) Appraisal Theory of Emotions namely affect, judgement, and appreciation. 100 news images were analyzed, 50 from each newspaper. The findings uncover that Dawn newspaper covers more about displacement, deaths, and diseases, prompting more emotional engagement of the viewers. The Guardian focused on the images of wreckage, highlighting the landscape damage associated with humanitarian disasters, also stressing on some armed forces misconducts. The study aids to multimodal discourse investigation and focuses on the hidden principles and political schemes in the graphic depiction of the Israel-Palestine war. (Azher, Jawad, & Arshad, 2025)

Jawad, Azher, and Zehra (2025) did a study "Media Coverage of Israel-Induced Palestinian Starvation Crisis: A Visual Semiotic Analysis" that further used visual semiotics on the issue of Israel-Palestine conflict. They examined three newspapers namely Dawn, the express tribune, and the guardian, using images that portrayed hunger and starvation of Gaza from July 2025 issues. The data existed of 45 images, 15 images per newspaper. The study applied the framework of Kress and van Leeuwen, 2006, mostly focused on compositional markers namely

information value, salience, and framing along with Martin and White's Appraisal Theory (2005) with emotional components of affect, judgment, appreciation. The findings revealed that dawn and the express tribune used images that emphasized the struggle and reliance of people stuck in the disaster while using weak frames and crowded displays of aid distribution. On the other hand, the guardian used close images of malnourished children which portrayed their misery and naivety. The study stresses how print media plays a role in determining thoughts and inducing emotions with visual representation on Israel-Palestine conflict (Jawad, Azher, & Zehra, 2025).

Research Methodology

The current study is descriptive and qualitative in nature. Descriptive study discusses the current scenario of issues (Manjunatha, 2019). Qualitative study, on the other hand, explores the issues of the given situation and according to Gay et al. (2012), is based on the textual analysis (Gay, Mills, & Airasian, 2012). This study is based on the semiotic analysis in which the non-verbal aspects are explored with respect to the semiotic analysis. The study has adopted the semiotic analysis model of Riera (2020) which discusses the three aspects of analysis.

3.1. Riera (2020)'s Model of Semiotic Analysis

Riera (2020) proposed a new theory of semiotic analysis with respect to Berger and Luckmann's grand theory of social constructionism (1966), which involves three aspects:

- Concept (Sign, Context, and Meanings)
- Construct (Semiosis)
- Proposition (the signs that convey the precise context of the object, which can trigger the appropriate response from the recipient)

Thus, the model of semiotics theory defined as the sign provides the precise context as well as conveys its designed meaning to elicit the intended response.

3.1.1. Concept

The semiotic model consists of three main concepts: signs, context, and meanings.

The sign is the composition of the physical reality called the signifier and its conceiving and interpretation by the viewer (signified). Different things, including smells, body language, sounds, etc., manifest a sign.

Context is the provision of relevant and particular meanings to a particular exchange occurring (Bowcher, 2018). It helps the recipient interpret the proposed interactional sense and derive the intended meaning.

Meanings are the relationship between the recipients of a sign and the personal experience of the world around the recipients (Zlatev, 2018). The meanings are created by connecting and interacting with the sign and the surrounding reality.

3.1.2. Constructs (Semiosis)

Construct refers to the relationship between the sign, the context, and the meanings. Peirce (1873) proposed that semiosis involves the sign, including the production of the meaning. The interplay and the interaction are presented by semiosis between the concepts of sign, context, and meanings.

3.1.3. Proposition

The proposition is the response of the recipient in response to the sender's intended meanings of the social and cultural image. The main intentions of the sender are to enhance the comprehension and the response to the received information (Riera, 2020).

3.2. Population and Sample

Population is the set of all those participants on whom the research has to be generalized (Daniel, 2016). The population of the present study is all those war images which are of Israel-Palestine Conflict.

A sample is the representative part of any population which has certain characteristics of a population (Shorten & Moorley, 2014). A sample is selected to save time, resources, and energy (Shorten & Moorley, 2014). For the analysis of the current study, a sample of four images is selected through purposive sampling from the Russian agency, The Moscow Times. Using purposive sampling, four war-related images were selected from The Moscow Times. Purposive sampling is the process of selecting a sample based on choice (Thomas, 2022). The four images from The Moscow Times were selected purposively because they clearly portray humanitarian circumstances, destruction of civilian infrastructure, displacement, and other visual representations associated with the objectives of the study.

3.3. Process of Data Collection and Analysis

The data of four images were collected systematically from The Moscow Times. The analysis is based on the semiotic theory of Riera (2020) which has three aspects of analysis: Concept, Construct (Semiosis), and Proposition (reader response). The analysis included recognizing the major visual signs, construing their contextual meanings, examining the process of semiosis, and evaluating the responses of viewers. The findings were explained according to the research objectives and questions to ensure coherence and consistency.

4. Data Analysis

4.1. Image 1: Destruction in Gaza after Israeli Airstrikes



Source: (Smirnov, 2023)
https://www.themoscowtimes.com/2023/10/24/russian-propaganda-seizes-on-wests-double-standards-as-israeli-strikes-pound-gaza-a82850

4.1.1. Concept

Signs: Smoke, Demolished Building, Destroyed Streets, Scattered Material

Context: The war between Hamas and Israel continues while Israel bombed the residential buildings.

Meanings: The image describes the inhuman action of bombing, which not only destroyed the building but also proved that the places were not worth living. The image describes the destruction of life structures, inhibiting living and movement in the modern age

4.1.2. Construct

The sign presents the results of inhuman actions that the IDF does not believe in morality. The buildings are presenting the cruelty of bombing. There are vertical relations between the smoke and the building, which presents the strike of bombs to prove that it is not worth living the land. Smoke presents the cruelty of the IDF and rage of the force on the humans as well as on the buildings. Vertical relations from IDF in the shape of smoke and targeting the building present that inhuman actions are taken to prove supremacy, unethical practices, no love for humanity and land, are symbolic of savagery, and the construction of human tragedy. The vertical relations between the destroyed street and the scattered buildings relate to the human dispersal, which has made it not worth living. It is proved that the lands coming under the range of air strikes by IDF have no right of safety, security, and peace

4.1.3. Proposition

According to the image, social structures and cultural aspects are ruined alongside the violation of international law and values. The image presents the sufferings of humans and their tragic lives. Humans lose their homes and residential areas. Smoke presents that the shelling is made with inhumanity. The image conveys that the life of Palestinians is totally dependent on the choice and thinking of the IDF. The rage of the IDF shows the conversion of solid buildings into particles due to heavy bombing.

Though international law, ethical and religious norms give guidelines for conducting the war to make sure the safety of residential places and properties, but these rules are not followed during the conflict. The image asks the viewer to feel the pain and the sorrow to reshape the thinking that the international community is silent and does not consider the inhuman actions. It seems that the current policies observe that US and other European allies are supporting all the inhuman actions of Israel. A recipient can view that the violation of international laws is supported by the lawmakers, which reinforces the question of character and role. It illustrates that Israel is not only involved in producing human loss, suffering, and tragedies, yet the international community of powerful countries supports these inhuman actions

4.2. Image 2: The Aftermath of an Israeli Air Strike on the Gaza Strip



(The Moscow Times, 2023)

https://www.themoscowtimes.com/2023/10/12/400-russians-seek-evacuation-from-besieged-gaza-strip-a82748

4.2.1. Concept

Signs: Destroyed Ships, Destroyed Roads, Despaired Humans

Context: In response to the Hamas attack, civilians and their life structures are attacked through an air strike by Israel.
Meanings: The image presents that the life of a common man is made critical and threatening by Israel after the air strike to prove the superiority. Additionally, roads and shops are not saved. It presents a threat to the source of livelihoods and the life infrastructures

4.2.2. Construct

The sign presents that there are vertical relations where the invisible power is of the IDFs, and they have the right to demolish any Palestinian area where they wanted. Destroyed shops and streets present the loss of peaceful and safe places even in the modern age. Despaired humans can only show their rages but are unable to fight the forces. It looks like human loss and human tragedy are going to be produced due to the airstrikes by IDF on shops and streets. Destruction of the shops and markets presents that IDF not only believes in the destruction of the businesses but also believes in closing the movement. Horizontal relations are there between the despaired humans by watching their destroyed shops and streets that they are unable to rejoin and remake them. IDF relates the air strike to the targeting of the same ethnic group without the distribution of the enemy and the common man.

4.2.3. Proposition

Destruction of the business places and the roads reveals the inhuman actions of Israel. It expresses that all types of properties are not safe from the attacks of Israel. The international law protects the properties and the life structures of a common man, but these do not seem reliable and applicable. Human is losing its business places, minor shops, and roads. They are seeing their things desperately and are helpless to resume or rebirth their life. Business structures are presented as the loss, while the running fire presents the heavy shelling, proposing the high rage of Israel on the Palestinians. Though the war is between Israel and Hamas, the life structures of the common Palestinians are destroyed by Israel, giving the message that their life businesses are not important. Destruction of businesses and workplaces denotes the rebirth of starvation in society, but the war criminals neglect it. Though humanity is suffering and humans have to lose their property, working places, and life sources, the international community is silent too. The culture of war, the destruction of social structures, and the violation of international law presents that the life of the Palestinians is totally dependent on the choice of the Israelis. The birth of human tragedy is there when the business structures are not saved, while the heavy shelling in the shape of fire can be proved as a continuous threat and fear for the Palestinians. It can affect their brains and their thinking about life, too.

4.3. Image 3: Palestinians carry belongings as they leave al-Ahli Hospital, which they were using as a shelter, in Gaza City.



Source: (Buketov, 2023)

https://www.themoscowtimes.com/2023/10/20/russian-propagandists-are-dancing-on-the-bones-of-israel-hamas-victims-a82840

4.3.1. Concept

Signs: Destroyed Vehicle, Human Migration, Despaired People
Context: Al-Ahli Hospital was kept as the shelter and was targeted by the Israeli air strike, so the people are fleeing from the Hospital in search of new shelters.
Meanings: The image presents the lifeless and despairing feelings of the people in search of safety and shelters

4.3.2. Construct

Hopeless humans are not allowed even in a healthy place that has distinctive protection during the war. Human migration has direct relation with the inhumane bombing by the IDF to prove that their lives are dependent on the behavior and the policies of the IDF. On the other hand, destroyed vehicles show that means of communication are also destroyed by IDF to prove no relation with the immigrants. Immigrants holding minimal luggage and kids show that they keep the necessary things with them so that they can spend their terrible days straightforwardly. Leading male presents that the front person will face the danger and the issues. There is an invisible vertical relationship between the shelling to the lives and the health places, proving the supremacy and the cruelty of IDF to make people migrate. There are no horizontal relations between the places and the people as the people have to flee from the grounds. The grounds

and the health facilities are not worth living in. Even though the hospital is targeted, the surrounding areas are not safe to live in. It presents the detached structures of life when the common places are not worth living in for humans.

4.3.3. Proposition

The culture of the war and the inhuman actions present the lifeless and disguised situation of the people of Palestine. The movement of the people with minor things presents the idea that people are in a worse situation and helpless. The sufferings of the people are high level as they have no shelter, even in the hospitals too. Though international rules ask the warring countries to take care of the women, children, and health places so that the health facilities can continue to be offered, the IDF does not care about these aspects. The hospital has exemption from other places, too. Besides, the encouragement of the international powerful countries with Israel presents the interests of their own rather than of the people. Besides the hospital, the nearest places are asked to be empty, too. This presents the distress where the people cannot avail of the shelters as their homes are demolished. International war crimes committed in the shape of inhuman actions lead to the tragic suffering of children, women, and older people. All of it asks the reader about the parameters of humanity and the implementation of the rules. Still, these are lacking by the major international powers as they support their interests rather than the humans in the world. The culture of war and the imbalance of social power made the people suffer, and the image asks the recipients to revise their preferences set by the international powers.

4.4. Image 4: Smoke Rises Following an Israeli Airstrike in The Gaza Strip, As Seen from Southern Israel.



Source: (AFP, 2023)
https://www.themoscowtimes.com/2023/10/14/russian-diplomat-hopes-to-meet-hamas-for-talks-to-free-hostages-agencies-a82769

4.4.1. Concept

Signs: Smoke from Buildings, Demolished Fields

Context: Residential places around the mosque are struck by Israel to target its suspicious opponents.

Meanings: The image depicts that the worship places were not saved in the war and the surrounding areas were special targets. The image also presents the demolished grounds and fields.

4.4.2. Construct

Smoke shows the cruelty of IDF against the Palestinians and indicates that even worship places are not safe to live due to destruction. The horizontal relationship shows that the common places are not equally saved due to the attacks on worship places of the Muslims. Demolished fields and new smoke of bombing illustrate their vertical relations with IDF as the sky is not safe; any time the bombing can happen. On the other hand, worship places are not safe from the bombs of the IDF. Besides, the close relationship of the surrounding building with the worship places proves that the worship places are the main targets where the opponents can take shelter.

4.4.3. Proposition

There is a difference between war and crime. A reader can understand that the war is based on the ethics and moralities in which the worship and the common places are not attacked. The

common places, worship places, and movement places are not safe in the hands of Israel. Support from the international powers has given them the free hand to perform such actions that cannot be considered ethical or based on morality. The culture of war has destroyed the social structures. Being the reader of the planet, it can be said that the life of people living around religious places is not safe. Smoke relates to the dispersion of human thought, through which human tragedy can be better understood. No human is found in the given image, which shows the heavy shelling conducted against the Palestinians. Displacements, human loss, sufferings, and pain are much more evident. Still, the birth of the human tragedy in Palestine has asked the readers to revise their perception of international standards and the ways of working. Though international institutions assure the security of nations and the common people as well as world places, but all the rules are abandoned here, and inhabitant areas are especially targeted, which shows that humanity has lost the game. At the same time, the inhuman and the supporters have gained victory by giving birth to human loss and displacement.

5. Conclusion

Riera (2020)'s Model of Semiotic Analysis not only explored meanings and the context but also the relational meanings and the response of the reader due to the intended message of the image. The birth of human tragedy in Palestine is not a new phenomenon; it has its history of different stages. In recent times, huge bombing by the IDF on the social and cultural structures of Palestinian life gave birth to the human tragedy and the loss of lands in Palestine. Insecure feelings and destruction of life put the people in a trauma from which they cannot recover their previous state of stability.

5.1. Findings

- There are vertical relations between the smoke and the building, which presents the strike of bombs to prove that the land is not worth living on.
- Smoke presents the cruelty of the IDF and rage of the force on the humans as well as the buildings.
- Vertical relations from IDF in the shape of smoke and targeting the building indicates that inhuman actions are taken to prove supremacy, unethical practices, no love for humanity and land, are symbolic of savagery, and construct human tragedy.
- Moreover, the vertical relations between the destroyed street and the scattered buildings related to the human dispersed situation, which has made areas difficult to live.
- It is proved that the lands coming under the range of air strikes by IDF have no right to safety, security, and peace.
- Despaired humans can only show their rages but are unable to fight the forces.
- Human loss and human tragedy are going to be produced due to the air strikes by IDF on shops and streets.

- Destruction of the shops and markets presents that IDF not only believes in the destruction of the business but also believes in checking the movement.
- Horizontal relations are there between the despaired humans by watching their destroyed shops and streets that they are unable to rejoin and remake them.
- Burning fire relating the inhuman actions that the heavy shelling is conducted to demolish the minimal things coming in the range of the place.
- There is an invisible vertical relationship between the shelling to the lives and the health places, proving the supremacy and the cruelty of IDF to make people migrate.
- There are no horizontal relations between the places and the people as the people have to flee from the grounds.
- The grounds and the health facilities are not worth living in. Even though the hospital is targeted, the surrounding areas are not safe to live in.
- The horizontal relationship suggests that the common places around the worship places are also not safe.
- Demolished fields and new smoke of bombing illustrate their vertical relations with IDF as the sky is not safe. On the other hand, worship places are not safe from the bombs of the IDF.
- Besides, the close relationship presented of the surrounding building with the worship places proves that the worship places are the main targets where the opponents can take shelter.
- The images ask the viewer to feel the pain and sorrow to reshape the thinking that the international community is silent and does not consider the inhuman actions merciless.
- It seems that US and other European allies are supporting all the inhuman actions of Israel.
- A recipient can view that the violation of international laws is supported by the lawmakers, which reinforces the question of character and role.
- Both Israel and international community of powerful countries are responsible for producing human loss, suffering, and tragedies in Palestine.
- The rules and principles of conducting war are neglected during the conflict as Israel attacked on children, women, elderly people, religious and health places.
- Support from the international powers has given them the free hand to perform such actions that cannot be considered ethical or based on morality.
- The inhuman and destructive actions by Israel and its supportive parties on Palestine results in the construction of human tragedy that fills the heart of Palestinians with pain and sorrow.

The findings show that the selected war images represent several meanings through visual and contextual elements with the help of Riera's (2020) semiotic framework. Different visual signs such as damaged buildings, smoke, healthcare facilities, religious sites and abandoned streets construct themes of displacement, human sufferings, destruction and social disturbance. The

findings align with previous studies such as Assaiqeli (2021), who claimed that visual elements provide the themes of identity and displacement in accordance with semiotic resources. Likewise, the current study also supports Adawiyah et al. (2020), who proposed that public opinions are affected by visual representations by creating certain meanings. Inclusively, the findings indicate the significance of Riera's (2020) semiotic model in examining the visual representation of Israel-Palestine conflict.

5.2. Recommendations

- Riera (2020)'s Model of Semiotic Analysis not only explores the meanings of the signs yet it expanded the analysis with respect to the recipient's responses so it can be applied to explore the collective and personal meanings in the images.
- There is relationship between the signs, context and meanings so the new joint meanings can help a researcher to expand the meanings of the research with the help of Riera (2020)'s Model of Semiotic Analysis.
- A researcher can employ the model to understand the meanings of the signs both in given and in present time too

5.3. Gap for Further Studies

There can be future research on the images based on multimodal, pragmatics and the systemic functional analysis of the images. There can also be comparative research on the depiction of different media channels about the presentation of the Israel-Palestine conflict.

References

Azher, M., Javed, T., & Arshad, F. (2025). Media Coverage of Israeli Air Strikes and Aftermaths in Palestine: A Visual Semiotic Analysis. *Balochistan Journal of Linguistics*, *13*, 28-28.

Adawiyah, R., Junining, E., & Tabiati, S. E. (2020). Multimodal Depiction of Israelis and Palestinians on Gaza Conflict in Two Middle Eastern Press. Diglossia, 12(1), 20-35.

AFP. (2023, October 14). Russian Diplomat Hopes to Meet Hamas for Talks to Free Hostages – Agencies. Retrieved April 15, 2024, from https://www.themoscowtimes.com/2023/10/14/russian-diplomat-hopes-to-meet-hamas-for-talks-to-free-hostages-agencies-a82769

Alsaafin, L. (2023, October 9). What's the Israel-Palestine conflict about? A simple guide. Retrieved April 30, 2024, from https://www.aljazeera.com/news/2023/10/9/whats-the-israel-palestine-conflict-about-a-simple-guide

Assaiqeli, A. (2021). Palestine in Visual Representation: A Visual Semiotic Analysis of the Nakba. International Journal of Arabic-English Studies, 21(1), 99-126.

BBC. (2024, April 5). Israel Gaza war: History of the conflict explained. Retrieved April 30, 2024, from https://www.bbc.com/news/newsbeat-44124396

Berger, P. L., & Luckmann, T. (1966). The social construction of reality: A treatise in the sociology of knowledge. Anchor Books.

Bowcher, W. L. (2018). The semiotic sense of context vs the. Bowcher Functional Linguistics, 5(5), 1-19.

Brown, C. (2007). Tragedy, 'Tragic Choices' and Contemporary International Political Theory. International Relations, 21(1), 5-13.

Buketov, K. (2023, October 20). Russian Propagandists Are Dancing on the Bones of Israel-Hamas Victims. Retrieved October 15, 2024, from https://www.themoscowtimes.com/2023/10/20/russian-propagandists-are-dancing-on-the-bones-of-israel-hamas-victims-a82840

Daniel, E. (2016). The Usefulness of Qualitative and Quantitative Approaches and Methods in Researching Problem-Solving Ability in Science Education Curriculum. Journal of Education and Practice, 7(15), 91-100.

Gay, R. L., Mills, E. G., & Airasian, W. P. (2012). Educational Research: Competencies for Analysis and Application. Pearson, Upper Saddle River.

Halliday, M.A.K. & Matthiessen, C. M.I.M. (2004). An Introduction to Functional Grammar. London: Arnold. https://doi.org/10.4324/9780203783771

Global Conflict Tracker. (2024, April 24). Israeli-Palestinian Conflict. Retrieved April 30, 2024, from https://www.cfr.org/global-conflict-tracker/conflict/israeli-palestinian-conflict

Golzar, J., & Tajik, O. (2019). Sampling Method | Descriptive Research ; Convenience Sampling. IJELS, 1(2), 72-77.

ICRC. (2024). What is International Humanitarian Law? International Community of Red Cross.

Jawad, T., Azher, M., & Zehra, N. (2025). Media coverage of Israel-induced Palestinian starvation crisis: A visual semiotic analysis. *Research Journal of Language and Literary Studies, 1*(2), 1-52.

Jacob, F. (2021). War and Semiotics. Routlege.

Manjunatha, N. (2019). Descriptive Research. Journal of Emerging Technologies and Innovative Research, 6(6), 883-887.

Kress, G. and van Leeuwen, T. (2006) Reading Images: a Grammar of Visual Design, London: Routledge.

Peirce, C. S. (1873). Description of a notation for the logic of relatives, resulting from an amplification of the conceptions of Boole's calculus of logic. Memoirs of the American Academy of Arts and Sciences (New Series), 9(2), 317–378.

Riera, J. J. (2020). Semiotic Theory.

Riera, J. J. (2020). Semiotic Theory. Retrieved April 10, 2024, from https://opentext.wsu.edu/theoreticalmodelsforteachingandresearch/chapter/semiotic-theory/

Santos, M. (2023). What is human tragedy? Retrieved April 2024, 2024, from https://www.quora.com/What-is-human-tragedy

Segell, G. (2023). An Alternative View Of The Israel-Palestine Dispute: The Eschatological Dimension. Center of Middle Eastern Studies, 16(1), 1-11.

Shorten, A., & Moorley, C. (2014). Selecting the sample. Evid Based Nurs, 17(2), 32-33.

Smirnov, O. (2023, October 25). Russian Propaganda Seizes on West's 'Double Standards' as Israeli Strikes Pound Gaza. Retrieved April 15, 2024, from https://www.themoscowtimes.com/2023/10/24/russian-propaganda-seizes-on-wests-double-standards-as-israeli-strikes-pound-gaza-a82850

Sturma, A. (2019). Reconciliation: Rebuilding Trust After the Tragedy of Violent Conflict. United States Institute of Peace.

Tehran Times. (2023, November 21). Tragedy in Gaza is a great shame for humanity. Retrieved April 30, 2024, from https://www.tehrantimes.com/news/491676/Tragedy-in-Gaza-is-a-great-shame-for-humanity

The Mosco Times. (2023, October 12). 400 Russians Seek Evacuation From Besieged Gaza Strip. Retrieved April 15, 2024, from https://www.themoscowtimes.com/2023/10/12/400-russians-seek-evacuation-from-besieged-gaza-strip-a82748

Thomas, F. B. (2022). The Role of Purposive Sampling Technique as a Tool for Informal Choices in a Social Sciences in Research Method. Just Agricultural Multidisciplinary e-Newspaper, 2(5), 1-8.

United Nations Human Rights. (2024, January 30). “Gaza is a massive human rights crisis and a humanitarian disaster”. Retrieved April 30, 2024, from https://www.ohchr.org/en/stories/2024/01/gaza-massive-human-rights-crisis-and-humanitarian-disaster#:~:text=There%20are%201.9%20million%20people,tents%20and%20other%20basic%20necessities.

Zaltev, J. (2018). Meaning making from life to language, The Semiotic Hierarchy and phenomenology. Lund University.

PAKISTAN JOURNAL OF LANGUAGE AND LITERARY STUDIES
Volume 2, Issue 1 (Spring 2026)
Pages: 19-35

The Role of Interactional Metadiscourse Markers in Portraying Female Identity: A Corpus-Based Study of Fantasy TV Characters

Farhat Bano[1]

ABSTRACT

Human identity is constructed and expressed through language. This can be seen in the television industry through observations of female characters. Dialogue of female characters is crafted carefully to express their thoughts, emotions, and behaviors, providing a glimpse into their psyche. This study analyzed interactional metadiscourse (hedges, boosters, attitude markers, self-mentions, engagement markers) utilized by five female characters in Sci-Fi and fantasy television: Wednesday Addams, Wanda Maximoff, Lyra Belacqua, Ellie, and Eleven. This study employed Hyland's (2005) metadiscourse model and took a corpus-based approach, using AntConc to analyze a data set of 43 television episodes, totaling 69,908 tokens, and employing a sequential mixed-method research approach. Frequency data were compared and reported per 1,000 tokens for each of the female characters. The results showed that self-mentions were the most frequent metadiscourse marker for three of the five female television characters, with engagement markers ranking first for the remaining two. The results further showed the television character's emotional histories and personality traits by the presence or absence of hedges and boosters. This study posited that different and psychologically elaborate female characters were portrayed through interactional metadiscourse; this study advanced the field of metadiscourse by applying the concept to television dialogue and offered insights into the representation of female identity in contemporary popular culture.

Keywords: *Spoken Interactional Metadiscourse; Female Identity; Corpus-Based Analysis; Fantasy Television; Scripted Speech*

[1] BS Scholar Department of Linguistics and Social Sciences, The Begum Nusrat Bhutto Women University, Sukkur, Pakistan
Secondary Affiliation: Research Intern, Corpus Research Center, Minhaj University Lahore, Pakistan

Corresponding author:
Ms. Farhat Bano, BS Scholar Department of Linguistics and Social Sciences, The Begum Nusrat Bhutto Women University, Sukkur, Pakistan
Email: farhatofficial6@gmail.com

1. Introduction

1.1. Background of the study

Language facilitates more than communication. Language helps construct identity. It helps convey emotions, attitudes, and express stances. In linguistics, the facet of language that corresponds with this is referred to as metadiscourse. Metadiscourse refers to the language used by the speaker/writer to maintain their relationship with the audience while expressing their stances and personality (Hyland, 2005). When it comes to metadiscourse, the interactional dimension is the most pertinent to identity because it refers to the elements that enable the speaker to express uncertainty, confidence, and emotions, and to identify with the audience (Hyland & Tse, 2004).

One of the most salient findings of research has established that gender significantly influences the use of interactional metadiscourse. Analysis of a range of personal essays (Puspita & Suhandano, 2023), EFL student's argumentative writing (Al-Otaibi & Hussain, 2024), and motivational speeches (Azlia, 2022) revealed that language employed by females is richer in interactional metadiscourse than language used by males. These findings indicate that female speakers/writers are more attentive to emotional and personal elements in their discourse.

Television is also a medium through which female identities are constructed and transmitted to audiences worldwide (Wu & Fakhruddin, 2023; Montasseri et al, 2020). The dialogue printed in television shows is written. The response of a television character is ascribed to the configuration of that character's background, personality, and emotional configuration. This is even true for the Sci-Fi and fantasy genres. In these genres, female characters are often represented as multifaceted social beings who are ultra- empowered, socially isolated, and highly traumatized. The response of these characters, and even the features of the response, is indicative of the representation of female identities in the popular media. To the researcher's knowledge, the utilization of interactional metadiscourse in the analysis of scripted dialogues of female characters in fantasy television shows is novel. In this context, the analysis of this study focuses on the dialogues of five female characters in fantasy television series sketching the basic framework of the model proposed by Hyland (2005) and employing the corpus software, AntConc (Anthony, 2022).

1.2. Research Problem

Metadiscourse has been examined in a variety of genres. These include academic essays, TED talks, research articles, and television talk shows. However, there seems to be a dearth of studies that have utilized metadiscourse in the analysis of the scripted dialogues of fictional television characters (Mensah et al., 2025; Lasagabaster & Bier, 2025). Many of the studies on metadiscourse and gender have concentrated on real-world language users in formal contexts. Al-Otaibi and Hussain (2024) studied the argumentative essays of EFL students. Azlia (2022) studied speeches in TED Talks. Puspita and Suhandano (2023) examined personal essays. Wu and Fakhruddin (2023) show that Netflix TV dialogues reveal female identity ideologies but did not adopt a metadiscourse framework or corpus analysis.

This is a clear gap in their analysis. No one has studied metadiscourse in TV series dialogues and used corpus tools to identify these patterns across different TV series characters.

1.3.Research Gap

There are gaps in the literature which are contextual and methodological. Contextually, TV series dialogues are very different from the language of essays or speeches. They are scripted in a way that makes a character appear real, emotionally and psychologically consistent. Particularly in the Sci-Fi and fantasy genres, the female characters are emotionally complex, and the linguistic patterns reflect this. Although literature has a notable absence of studying these patterns, the absence of studying interactional metadiscourse is even more profound. Methodologically, TV series dialogues have not been studied using a corpus to do a cross-character analysis of metadiscourse. The speaking and writing pattern of discourse has been studied using a corpus and AntConc with Hyland's (2005) metadiscourse framework (Puspita & Suhandano, 2023; Azlia, 2022; Al-Otaibi & Hussain, 2024). Chou et al. (2023) revealed that different genres of discourse lead to different patterns of interactional metadiscourse and therefore illustrate the importance of context. The need for studying fictional dialogues as a distinct genre cannot be overstated. The current study aims to address these gaps in the literature.

1.4.Research Objectives

1. To identify the interactional metadiscourse features used in the dialogue of five selected female characters from Sci-Fi and fantasy TV series.
2. To compare the metadiscourse patterns used by five female characters across their respective series.

1.5.Research Questions

1. What interactional metadiscourse features appear in the dialogue of the five selected female characters?
2. How do the metadiscourse patterns differ or converge across the five characters?

1.6.Significance of the Study

This study makes an important contribution to metadiscourse research. Hyland (as cited in Mensah et al., 2025) indicated that metadiscourse analysis constricts research to a small variety of textual constructs and has not fully expanded research within different genres and contexts. Based on Hyland's (2005) model, this study will continue expanding existing research on fictional TV dialogues as a framework and show that metadiscourse analysis can derive results beyond academic research and formal spoken genres. Within media and gender studies, the results will justify the existence of female characters and the linguistic construction of female characters within fantasy TV. For language researchers, the study provides a new construction and methodology that can be used to study different characters and linguistic constructions in different genres and cultural contexts (Chou et al., 2023).

2. Literature Review

Language has the potential to construct identities and relationships, and to express feelings or attitudes beyond the verbal messages and the informational functions that may be encoded within the world. Metadiscourse (Hyland, 2005) is one of the linguistic tools that can

perform this task. At the core of metadiscourse analysis is the identification of linguistic resources that facilitate the management of discourse, expressions of attitudes towards the discourse, and audience engagement. The interactional aspects of discourse are most relevant to identity construction. These include the resources (linguistic and otherwise) that allow articulation of self and others in a given discourse context. This review explores research on interactional metadiscourse and gender, the construction of female identity in media, and the methodological approaches used in closely related studies, in order to frame the present study within the existing literature.

2.1.Hyland's (2005) Interactional Metadiscourse Framework

The main theory in this research is Hyland's (2005) model of metadiscourse and interpersonal relationships. Hyland splits metadiscourse into two; the first is interactive metadiscourse which organizes the text for the reader/listener. The second is interactional metadiscourse which organizes the personal and social aspects of communication. As Hyland and Tse (2004, p. 174) said, metadiscourse is "as important as the information," because it is how the reader/listener perceives and reacts to the writer/speaker. This model has been widely adopted in linguistics research, and it has been proven to be effective in both written and spoken discourse (Lasagabaster & Bier, 2025; Mensah et al., 2025).

The five sub-categories of interactional metadiscourse, which this research is focusing on, are as follows. Hedges are words that show an incomplete commitment to the proposition (Hyland, 2005). Example hedges are: maybe, perhaps, I think, and might. Boosters are words that show complete commitment to the proposition and, likewise, reject all other possibilities. Some examples of attitude markers are: unfortunately, surprisingly, and I hope. Self-mentions make the author present in the discourse (Puspita & Suhandano, 2023). Finally, engagement markers make the audience part of the discourse and words such as: you know, listen, look, and consider (Mensah et al., 2025).

A specific dimension of human identity is represented by each distinct element. Experts claim that hedges demonstrate emotional caution, while boosters indicate certain confidence. Attitude markers show emotional openness because self-mentions express personal power, and engagement markers reveal connections between people.

The reason for selecting the interactional over the interactive dimension is that the former is the one through which speakers position themselves emotionally and relationally, compared to which the latter is more concerned with the organization of the text for the reader. Each of the sub-categories was considered as an identity marker: hedges were regarded as signs of a character's emotional deliberation and uncertainty about the claim she is making; boosters as signs of a character's self-assurance and confidence; attitude markers as the character's evaluative and emotional stance; self-mentions as a character's assertion of her agency and of her centrality in the narrative; and engagement markers as a character's relational focus. The patterns in these categories were interpreted as providing a character's psychological profile rather than as simple counts of the usage of another linguistic sign.

2.2.Interactional Metadiscourse and Gender

Research on gender differences in interactional metadiscourse has highlighted and reinforced similar trends in different contexts and genres. Puspita and Suhandano (2023) examined 20 personal essays from Chicken Soup for the Soul using AntConc 3.5.9 and Hyland's (2005) model. Half the essays were authored by women and the other half by men. After normalizing their results per 1,000 words, the authors found that women exhibited a higher use of interactional discourse markers. Women had a significantly higher use of boosters (1.99 vs. 0.70), attitudes (0.66 vs. 0.30), and engagement markers (3.21 vs. 0.99). Self-mention was the highest category for both genders. The study found that women write in a more emotionally expressive and audience-centered style, whereas men write in a more event-centered style without emotional detail.

Similar research was conducted by Al-Otaibi and Hussain (2024), who focused on 46,453 tokens of Saudi EFL student's argumentative essays on a gender-sensitive issue. Using AntConc and Hyland's (2005) model, the authors found that female students had higher uses of attitude markers, boosters, hedges, and self-mentions than male students. Interestingly, female students who argued a position in favor of males used hedges, indicating that hedging was influenced by both gender and by the position taken. The authors attributed women's higher uses of boosters and attitude markers to their increasing assertiveness and suggested that social and cultural factors affect the use of metadiscourse.

Spoken discourse in TED Talk speeches from male and female motivational speakers was analysed by Azlia (2022) using a corpus of 43,160 female characters and 51,873 male characters. Many scholars believe that female speakers used more self-mention and engagement markers because these female speakers possessed more expressive and communicative qualities in their speech, while male speakers remained direct and content-centered. A consistent trend emerges across these studies. Female language users tend to be more interactionally rich, especially regarding self-references, engagement markers, and attitude markers. Hedging is more ambiguous, however. While Azlia (2022) and Al-Otaibi and Hussain (2024) observed that women hedge more, Puspita and Suhandano (2023) reported the contrary in personal essays. The results support the idea that hedging may not be a definitive gender phenomenon. Chou et. al. (2023) substantiated this view and conducted a corpus-based study of both translated and untranslated English texts. They reported that, within the various genres, hedges along with other interactional discourse markers were highly variable. The results confirmed that interactional metadiscourse is influenced by the communicative context in which it is found.

2.3.Female Identity in Television and Media

Television and film are seen as critical areas in the construction of gender identity and its circulation. A corpus-based analysis by Montasseri et. al. (2020) examined gender representation in American films through the lens of the COCA fiction corpus and Van Leeuwen's (2008) social actor model. Their analysis revealed that, in general, female characters were depicted as playing roles that were weak in power and low in status. The authors stated that movie scripts are, in general, ideological tools to shape audience views on

gender. Their analysis shows the value of analyzing scripted language as a representation of gender.

Wu and Fakhruddin (2023) arrived at a comparable result in their analysis using Systematic Functional Linguistics to study Netflix productions. They reported that Netflix female characters were created using westernized language choices that framed women's representation as submissive, high-maintenance, and objectified. Their study confirmed that TV dialogue constructs gender identity for a target audience at the global level. However, their study did not employ metadiscourse and corpus tools, which the current study addresses.

Mensah et al. (2025) employed Hyland (2005) to analyze five episodes of a Ghanaian TV talk show within the period of the 2020 elections. They found that participants of the talk show employed a considerable amount of interactional metadiscourse, with self-mentions and engagement markers being the two largest categories. They indicated that metadiscourse resources were ample in talk show discourse, including non-fictional talk show dialogue, and that the categories had distinct communicative functions. Mensah et al. (2025) importantly called for metadiscourse research in contexts outside the academy, noting the great need for metadiscourse resources in daily discourse and media discourse. Lasagabaster and Bier (2025) made a similar remark in their research on interactional metadiscourse in English-medium lectures of three disciplines, where they found that metadiscourse placement was determined by the speaker's contextual and institutional positioning. This is also equally relevent of fictional characters, whose created and scripted roles of discourse are likely to elicit the same metadiscourse response.

Through the academic works that were mentioned before, three major points were revealed. To begin, it has been observed that women who use language show higher communicative depth in interactional metadiscourse, especially when they use self-mentions, engagement markers, and attitude markers, even though hedging changes when the situation of communication changes. Another point shows that the identity of women is created on purpose when female language users make specific selections of words. As a third discovery, factors like the subject of discussion, the style of writing, or the environment of communication alter the interactional metadiscourse markers, which means that gender is not the only cause of these variations (Chou et al., 2023; Al-Otaibi & Hussain, 2024; Lasagabaster & Bier, 2025). Although these academic investigations have given valuable knowledge, researchers have not investigated interactional metadiscourse inside the prepared conversations of imaginary women who act in Sci-Fi and fantasy TV programs. Because this specific empty space in methodology and context exists, the present study addresses this empty space directly.

3. Research Methodology

A mixed method approach is adopted for this study. In this case, the researcher uses a quantitative framework for the first phase and follows it with a qualitative framework for the second phase. The design is appropriate for this study considering the nature of interactional metadiscourse. It necessitates the identification and counting of linguistic features to outline patterns, and subsequently, the interpretation of those patterns in context to assign their

significance (Puspita & Suhandano, 2023; Azlia, 2022; Al-Otaibi & Hussain, 2024). The first phase of the research focuses on quantitative analysis to ensure that fictional dialogues are not oversimplified to mere data. The second phase is a qualitative analysis that ensures that the results are not oversimplified. Measures established in the first phase of the research allow for qualitative, descriptive, and meaningful results.

3.1.Sampling Method and corpus Design

Purposive sampling is utilised in this scientific work because participants are chosen for a specific objective instead of being selected by chance (Mensah et al., 2025). Indeed, experts claim that the chosen characters analysed are Wednesday Addams from Wednesday (Season 1, 2022), Wanda Maximoff from WandaVision (Season 1, 2021), Lyra Belacqua from His Dark Materials (Season 1, 2019), Ellie from The Last of Us (Season 1, 2023), and Eleven from Stranger Things (Season 4, 2022). This study examines character's trauma, outsider identity, supernatural powers, and emotional complexity. These traits are believed to influence diverse and complex interactional metadiscourse. These traits are also uncommon in fictional characters, so random sampling was inappropriate.

The data includes all dialogue for the sampled characters across one full season, totaling 43 episodes. Sampling full seasons provides the opportunity to capture character's full emotional and communicative development throughout the series and improves the study's validity (Mensah et.al., (2025). Altogether, a sum of 69,908 tokens is gathered from all the selected characters, which represents the entire text size. Through closer inspection, it has been observed that Table 1 presents the specific numbers of these tokens for each individual character.

Table 1: Sub-Corpus Size by Character

Character	TV Series	Season / Year	Episodes	Total Tokens
Wednesday Addams	Wednesday	Season 1, 2022	8	12,358
Wanda Maximoff	WandaVision	Season 1, 2021	9	9,855
Lyra Belacqua	His Dark Materials	Season 1, 2019	8	16,573
Ellie	The Last of Us	Season 1, 2023	9	16,020
Eleven	Stranger Things	Season 4, 2022	9	15,102
Total			43	69,908

Note. Total tokens represent the combined word count of dialogue extracted for each character across the listed episodes.

3.2.Data Collection Procedure

Data collection came from publicly available transcripts of five TV series. These transcripts were collected from source websites (Springfield! Springfield!, Scraps from the Loft, and Simpleremix Scripts). The remaining data were collected from the focus character's dialogue. The speech of remaining focus characters, stage directions, scene descriptions, and the narrator's text were all discarded. The dialogue of each focus character was stored as a separate text document. Together, the five text documents constituted a multi-character corpus of this TV series and were imported to AntConc for analysis.

3.3.Data Analysis Procedure

The analysis followed four structured steps:

Step 1 (Corpus Preparation): Each character uploaded their text file into AntConc. Word frequency lists identified linguistic norms and checked that data was cleaned.
Step 2: Feature Identification. Using AntConc's Word List and Concordance, this study searched for interactional metadiscourse markers in each sub-corpus. The five markers by Hyland (2005) were as follows: hedges (maybe, perhaps, I think, I feel, I guess, might, could, possibly, probably, seem, sort of, and kind of), boosters (always, never, certainly, definitely, clearly, obviously, absolutely, of course, I know, and indeed), attitude markers (unfortunately, surprisingly, I hope, I wish, I am glad, I am sorry, terrible, amazing, and horrible), self-mentions (I, me, my, mine, myself, and we), and engagement markers (you know, listen, look, remember, you see, and consider). Each token was checked in its concordance context, as per the verification procedure of Al-Otaibi and Hussain (2024) and Lasagabaster and Bier (2025).
Step 3: (Quantitative Comparison) The unstandardized frequency of each category was recorded for each character. For each character, data was normalized to 1,000 tokens. This was the case as the five subcorpora presented disparate sizes, making raw numbers potentially unclear (Chou et al., 2023). The results were recorded in Microsoft Excel.
During Step 4 (Qualitative Interpretation), the analysis of quantitative data focused on character identities, trauma, and narrative roles. This analysis went beyond the description of the frequency of features and considered the reasons for the frequency.

3.4.Ethical Considerations

This research is an analysis of TV transcripts that the public can access, and no humans are involved in this research, so there are no ethical issues of privacy, consent, or confidentiality. The transcripts are cited, and this research follows the APA guidelines to eliminate plagiarism. The research is of a noncommercial nature, and the research adheres to the Corpus Research Center's data policy. The data stay jointly owned by the researcher and the Corpus Research Center.

4. Data Analysis

4.1.Overview of Data Analysis

This research examines the Spoken Dialogue of five female fictional characters from five Science Fiction and Fantasy television series. The focus for the study is on: Wednesday Addams (Wednesday, Season 1, 2022) who has a corpus of 12,358 tokens from 8 Episodes, Wanda Maximoff (WandaVision, Season 1, 2021) who has a corpus of 9,855 tokens from 9 Episodes, and Lyra Belacqua (His Dark Materials, Season 1, 2019) who has a corpus of 16,573 tokens from 8 Episodes. Ellie from The Last of Us (Season 1, 2023 provided 16,020 tokens across 9 episodes. Eleven from Stranger Things (Season 4, 2022) provided 15,102 tokens across 9 episodes. The total combined corpus is 69,908 tokens across 43 episodes. Transcripts were taken from Springfield! Springfield!, Scraps from the Loft, and Simpleremix Scripts. Because the investigation required accuracy, the researchers manually gathered and cleaned the dialogue of each fictional character while they removed other speaker's lines, stage instructions, and narration text. It has been observed that the collected data represents scripted spoken discourse which was created by authors to mimic realistic speech while building the persona, traits, and emotional state of each figure.

4.2.Method of Analysis

A sequential mixed-method design organises this scientific investigation. In the first quantitative phase, the interactional metadiscourse markers within the five sub-corpora were identified and computed using the software AntConc, which is consistent with Hyland's (2005) five-category model. When researchers compare language patterns across five different characters and text collections, using this methodology guarantees the evaluation is structured, objective, and repeatable, which is extremely important (Puspita & Suhandano, 2023; Al-Otaibi & Hussain, 2024). After counting, basic frequency numbers were adjusted to a standard of 1,000 tokens because this adjustment permits accurate comparisons to happen despite variations in the size of the text groups, which experts claim is a standard procedure in corpus-based metadiscourse research (Chou et al., 2023; Lasagabaster & Bier, 2025). Throughout the subsequent qualitative portion, the discovered language patterns were explained by linking them to the mental state, story function, and emotional background of each character, which connects statistical numbers to the true meaning that those numbers hold within the constructed identity of each figure.

4.3.Step-by-step Analytical Procedure

The same four-step procedure from the Methodology section (i.e. Corpus Preparation, Feature Identification, Quantitative Comparison, and Qualitative Interpretation) was conducted. Each metadiscourse marker was verified by context checking of the word, which concurs with the procedures of Al-Otaibi and Hussain (2024) and Puspita and Suhandano (2023).

4.4.Reliability and Trustworthiness

For the purpose of making the research analysis reliable, every found metadiscourse token was examined in its concordance context within AntConc because researchers wanted to confirm that the token was truly acting as a metadiscourse marker instead of a normal vocabulary word. Experts claim that this method of checking context is similar to the procedures which were utilized in previous similar investigations (Al-Otaibi & Hussain, 2024; Lasagabaster & Bier, 2025). Before the main analysis started, researchers completed a preliminary test on a small portion of every sub-corpus so that the system of coding would remain uniform for all five characters.

5. Findings

5.1.Overview of Findings

Hyland's (2005) interactional metadiscourse model was represented in every character's dialogue. The analysis showed a prevalence of Self-mentions, followed by Engagement Markers, Hedges, Boosters, and Attitude Markers. The distribution of the categories was Largely consistent across all five sub-corpora, with self-mentions ranking first for three characters (Wednesday, Lyra, and Eleven) and engagement markers ranking first for the remaining two (Wanda and Ellie). Thus, the most characteristic interactive feature for scripted female dialogue in this genre was the inclusion of Self-mentions and relational Engagement. Observed in Table 2 are the frequencies and rates per 1,000 tokens for all characters across the five categories

Table 2

Interactional Metadiscourse Frequencies and Normalised Rates per 1,000 Tokens by Character

Character	Tokens	Hedges (raw/1k)	Boosters (raw/1k)	Att. Markers (raw/1k)	Self-Mention (raw/1k)	Eng. Markers (raw/1k)	Total
Wednesday	12,358	130 / 10.52	79 / 6.39	7 / 0.57	946 / 76.55	650 / 52.60	1,812
Wanda	9,855	125 / 12.68	49 / 4.97	5 / 0.51	680 / 69.00	712 / 72.25	1,571
Lyra	16,573	209 / 12.61	56 / 3.38	2 / 0.12	1,216 / 73.37	1,034 / 62.39	2,517
Ellie	16,020	231 / 14.42	49 / 3.06	4 / 0.25	1,116 / 69.66	1,126 / 70.29	2,526
Eleven	15,102	184 / 12.18	44 / 2.91	7 / 0.46	1,155 / 76.48	971 / 64.30	2,361

Note. Values are presented as raw frequency / normalized rate per 1,000 tokens (raw/1k).

5.2. Finding for Research Objective 1: Interactional Metadiscourse Features in Female Character Dialogue

The first research objective investigated which interactional metadiscourse features were found in the dialogues of five selected characters. The analysis established that all five features were present in all five sub-corpora, although with variable frequencies.

Self-mentions were far and away the largest category. Wednesday generated 946 unadjusted self-mention tokens and, with Eleven (1,155), both achieved the highest counts, although Wednesday's count gave her a normalized figure of 76.55 per 1,000. The remaining normalized figures were: Lyra 73.37, Ellie 69.66, Wanda 69.00, and Eleven 76.48. The first-person singular pronoun 'I' was the dominant form of self-mention for all five characters. This corroborates the findings of Puspita and Suhandano (2023) and Al-Otaibi and Hussain (2024) regarding the prominence of self-mention in the respective corpora.

The second largest category was engagement markers. Wanda had the highest nominalized figure for engagement markers (72.25 per 1,000), and Ellie was next (70.29 per 1,000). Eleven was approximately 6 points behind (64.30 per 1,000), followed by Lyra (62.39 per 1,000), and Wednesday was last (52.60 per 1,000). Across all five characters, the second-person pronoun 'you' was the most frequent engagement marker. This was an engagement marker, and in a TV discourse analysis, Mensah et al. (2025) found that engagement markers were frequently the largest subset of interactional resources.

Table 2 shows the distribution of hedges amongst the five characters. A higher distribution of hedges was found for Ellie (14.42 per 1000, 231 raw tokens), Wanda (12.68 per 1000, 125 raw tokens), Lyra (12.61 per 1000, 209 raw tokens), Eleven (12.18 per 1000, 184 raw tokens), and Wednesday (10.52 per 1000, 130 raw tokens) in the order given. The highest used hedges in the five characters were maybe, might, I think, I guess, probably, and kind of. The

presence of hedges in the five characters shows that uncertain expressions are frequent markers of the scripted female speech in this genre. However, the presence of hedges in the speech of the five characters is not equal.

Boosters are also not common. Wednesday also shows the highest booster occurrence (79 raw, 6.39 per 1000), and the highest occurrence of boosters in total. Wanda had the next highest occurrence at (49 raw, 4.97 per 1000), Lyra (56 raw, 3.38), Ellie (49 raw, 3.06), and Eleven (44 raw, 2.91). The total number of attitude markers was also low. Wednesday and Eleven had 7 raw tokens, Wanda had 5, Ellie had 4 and Lyra had 2. This was lower than female writers in academic genres (Puspita and Suhandano, 2023; Al-Otaibi and Hussain, 2024). This also shows that attitude markers are less in scripted fictional dialogue than in academic written dialogue.

Some examples of the quantitative measures described above are as follows. They illustrate four of the marker types used in the samples of scripted dialogue. Each metadiscourse marker shown below is emphasized.

(1) Wednesday Addams - Booster

"It's definitely a match."

This booster reflects Wednesday's tendency to be assertive, which shows her refusal to soften her claims, and which is consistent with her having the highest booster rate of the 5 characters.

(2) Ellie - Hedge

"You seem capable enough."

This hedge reflects Ellie's guardedness and caution to others, which is consistent with her having the highest hedging rate of the 5 characters.

(3) Wanda Maximoff - Engagement Marker

"You're a real cut-up. You know, I always…"

This engagement marker directly calls upon the addressee to be part of the utterance, which is consistent with Wanda's high engagement and her narrative arc of being relational and connecting.

(4) Eleven - Self-Mention

"I'm here. I'm here."

The repeated self-mention focuses on Eleven's presence in the discourse, which is consistent with her being one of the characters with the highest self-mention rate.

(5) Lyra Belacqua - Attitude Marker

"Hopefully, we shall prove fast enough."

This attitude marker reflects Lyra's rare expression of evaluative stance through speech; her overall low attitude-marker frequency suggests that her emotions are more often conveyed through action than through explicit verbal evaluation.

These examples give quantitative measures when corresponding dialogue samples are provided.

5.3. Finding for Research Objective 2: Cross-Character Comparison of Metadiscourse Patterns

Research Objective 2 also includes the comparison of the metadiscourse patterns of the characters. However, the focus was on the differences and the convergence of patterns. The comparison of the characters showed that there are some common metadiscourse patterns for the five characters and some that were unique to an individual character.

In examining similarity, three of the five characters (Wednesday, Lyra, and Eleven) follow the categorical order: self-mentions, engagement markers, hedges, boosters, and attitude markers, while Wanda and Ellie show engagement markers ranking above self-mentions. The significance of the character's self-mention categorical order indicates that there is a pattern on the level of a genre, a way of constructing identities for women using language in Sci-Fi and fantasy TV. This pattern also supports findings across studies that women language users place a greater emphasis on the presence of the self and interpersonal engagement as opposed to the more direct or evaluative aspects of language (Puspita & Suhandano, 2023; Mensah et al., 2025).

The most notable variation is in the case of engagement markers. Wednesday's (52.60 per 1,000) is nearly 20 points lower than Wanda's (72.25 per 1,000); this is the largest disparity in any category. This speaks to each character's personality. Wednesday's character is designed to be emotionally and socially withdrawn; therefore, direct engagement with the audience is lower. In contrast, Wanda and Ellie are both characters with a designed strong need to engage and connect to the audience. Here, Wanda, who is trying to maintain her grief-constructed world, and Ellie trying to guard the ones she loves and express this in her high rates of engagement markers. This supports Al-Otaibi and Hussain's (2024) study on metadiscourse as relational positioning that goes beyond gender. In self-mentions, the gap between the highest and lowest rates is about 7.5 points (Wednesday at 76.55 vs Wanda at 69.00). Wednesday and Eleven produce the highest self-mention rates, which is consistent with both characters being written with deeply personal, identity-centred storylines. Wanda's lower rate may reflect her more team-centred narrative in WandaVision, where the story is as much about the people around her as it is about her.

In boosters, Wednesday clearly is the most assertive, at 6.39 per 1,000, which is more than double Wanda's rate and more than double Eleven's rate. This is consistent with her characterization as an intellectually self-assured and an assertive person. Conversely, Eleven, who is characterized as emotionally fragile and still healing from the aftermath of her traumatic past, yields the fewest boosters, at 2.91 per 1,000. Chou et al. (2023) observed that interactional markers, or boosters, are particularly sensitive to the speaker's context and frame, which is clearly visible in the contrast between Wednesday and Eleven.

In hedges, Ellie stands out as the only character with a significantly higher rate, at 14.42 per 1,000, while the other four characters showed values between 10.52 and 12.68 per 1,000. The higher hedging characterizes Ellie, who like a survivor in a dangerous context, is a teenager

in modern-day America where a high degree of care combined with a high degree of hedging is a psychosocially appropriate response. This is consistent with Azlia's (2022) conclusion that the female speaker's interactional markers indicate her engagement with the speaker's content.

In attitude markers, the numbers are sufficiently low that the ability to interpret with confidence is limited. However, Lyra's extremely low frequency of attitude markers, at 0.12 per 1,000, is significant and would appear to be consistent with the assertion that in the series, her emotions are primarily expressed through actions, and more direct forms of speech.

6. Discussion

The results from this research show that interactional metadiscourse displays an identifiable structure in the scripted female dialogue in Sci-Fi and fantasy television. The most outstanding result is that self-mentions in the five character's background have a broad scope of 69.00 to 76.55 self-mentions per 1,000 tokens. This pattern reveals characteristics of the writing process of those characters. Characters that predominantly use first-person self-referential language are written as being self-aware, agentive, and as being in the center of the narrative. The presence of this linguistic feature in the dialogue of the five characters is a deliberate writing decision to ensure that the five characters are the primary subjects of their narratives, as opposed to being secondary characters to someone else's narratives.

Confirmation of this reading is found in the next most dominant category in the metadiscursive engagement, which suggests that speakers and the addressees create a shared space, and as a result, the character is portrayed as being relational and emotionally present. Wanda and Ellie demonstrate the highest degree of engagement. This suggests that Wanda and Ellie are indicative of characters who are written to make relational engagement, connection, and presence, which are depicted to be potentially painful and even impossible in their circumstances. The engagement rate in Wednesday's dialogue is 52.60, which expresses her preference for a relatively distant style of emotional engagement. As Mensah et al. (2025) demonstrate, engagement markers are indicative of the speaker's relational intent, and this is depicted clearly in all the characters.

Regarding the identity of each person, the variation that is found in hedges and boosters among the different characters provides the most useful information. It has been observed that Ellie utilizes a very high rate of hedges (14.42 per 1,000), which demonstrates a speaking style that is defined by carefulness and uncertain words, and this language matches completely with her true self because she is a young person who is unable to give full trust to her world, her own feelings, or the people who live around her. Showing a great difference, the heavy booster rate of Wednesday (6.39 per 1,000), which is more than double the rate of any other character, reveals her strong mental confidence and her complete refusal to make her claims soft. Many scholars believe that these linguistic patterns of hedging and boosting in scripted dialogue are not simple accidents of style, but they are deliberate decisions that are made to make each character stable in their thoughts and familiar in their emotions.

Regarding the investigation, the lowest application of attitude markers was discovered across all five characters, which ranged from only 0.12 to 0.57 for every 1,000 tokens. These

numbers of attitude markers are extremely small when researchers compare the data with female academic writers working in scholarly environments (Al-Otaibi & Hussain, 2024; Puspita & Suhandano, 2023). Many people believe that the emotional state of characters in television drama is presented to the viewers through physical acting, background music, and screen visuals, which makes the presence of evaluative language less necessary. Individual characters can easily show deep sadness, terror, or happiness through these performance tools. Because of these visual and auditory tools, the high demand for attitude markers in written dialogue decreases when compared with the written text of an author.

6.1.Connection with Past Literature

These findings partially agree with existing literature on inter-related metadiscourse and gender, while also extending on some findings. The dominance of self-mentions and engagement markers is in line with Puspita and Suhandano (2023) who found self-mentioning was the most important interactional category in personal essays, and with Mensah et al. (2025) who described self-mentions and engagement markers as the most common interactional resources in television talk shows. However, in the present study, dominance of the two categories was more pronounced compared to the above mentioned real world examples. The likely reason for this is that scriptwriters enhance personal presence and direct relationalness to depict emotionally vivid characters. The variation across characters in the use of hedging supports Azlia's (2022) proposition that interactional markers used in spoken discourse reflect the speaker's relational context, and confirms Al-Otaibi and Hussain (2024), who found that hedging is influenced by attitude and social context, and is not gendered. Ellie's high hedge rate exemplifies what is situational for her, not a fixed gender way. Wednesday's booster contradicts the prevailing view that female speakers are generally less assertive than males. It aligns with Al-Otaibi and Hussain (2024), aiming for a more intricate framework, arguing that assertiveness in metadiscourse is contextually constructed identity and positioning, and is not necessarily bound by gender. The low attitude markers of the five characters in the current study are in stark contrast to Puspita and Suhandano (2023) and Al-Otaibi and Hussain (2024), who argue that female writers in an academic setting exhibit high levels of attitude markers. This, however, is a consequence of the difference in genre. Chou et al. (2023), in their multi-genre corpus study, argue that attitude markers and other stance features may vary across a wide range of genres, sub-genres, and suppress certain communicative contexts. Wu and Fakhruddin (2023) argue that Netflix, through a specific language framework, provides a construction for the gender ideology of its female characters. The current study provides evidence of this claim in the way that metadiscourse interactional features were employed to construct the identity of five female characters in a Non-Stereotypical way.

6.2.Theoretical Implications

Adding to metadiscourse theory, this study demonstrates that Hyland's (2005) interactional model can be further productively applied to scripted fictional dialogue. This extends the model's application from strictly academic writing and formal spoken genres to other domains. This finding helps answer Hyland's request (cited in Mensah et al., 2025) to expand

metadiscourse research from the few text types it has been restricted to. The fact that three of the five characters organise self-mentions, engagement markers, hedges, and boosters followed by attitude markers in the same hierarchy, while the remaining two show engagement markers ranking above self-mentions, offers a testable hypothesis about a metadiscourse framework for female characters in fantasy television. The hypothesis is certainly novel, but is also immediately testable, and as such can be improved further by other researchers.

6.3.Practical Implications

This study offers a systematic, evidence-based approach for media and gender researchers to examine the linguistic construction of female identity in popular television. For language teachers and corpus linguists, this study shows AntConc with Hyland's (2005) model can be applied to fictional media texts and beyond the academic texts. This broadens the types of text that can be included in corpus-based research and classroom pedagogy. Finally, screenwriters can learn how the use of hedging, boosting, self-mention, and engagement can enhance audience perceptions of the psychological and emotional reality of female characters.

6.4.Connection to Research Objectives

Both research objectives were fully accomplished. Analysis of the sub-corpora as a whole confirmed the presence of the metadiscourse features in each character's dialogue for all five categories and all five sub-corpora, allowing the researcher to achieve Research Objective 1. Normalized cross-character comparison permitted the researcher to achieve Research Objective 2, as it disclosed both the metadiscourse patterns categorical hierarchy that is shared by all five characters and the metadiscourse patterns individual differences that constitute all five characters' personalities and identities.

7. Conclusion

This study sought to portray female identity in Sci-Fi and fantasy television through the analysis of interactional metadiscourse in character dialogue, particularly focused on the characters of Wednesday Addams, Wanda Maximoff, Lyra Belacqua, Ellie, and Eleven. The study utilized a corpus of 69,908 tokens, compiled from 43 television episodes across five different series, and provided a large enough sample from which both qualitative and quantitative analyses to be conducted. Interactional features that were most prevalent across all five characters were self-mentions, followed by engagement markers, followed by hedges, and boosters and attitude markers, with self-mentions ranking first for three characters and engagement markers ranking first for Wanda and Ellie. The patterns that were identified for each character demonstrated that Sci-Fi and fantasy television genre features portray female characters who are relationally engaged. Metadiscourse made evident that each character's individual personality and emotional history and character role was distinguished through the use of hedging, boosting, and engagement.

This study presents three primary advancements. Theoretically, it broadens the scope of Hyland’s (2005) interactional model to scripted fictional dialogue. This highlights the model’s versatility and suggests that genres beyond the academic and formal prompts of spoken discourse may be appropriate. This development responds to the suggestions of

Mensah et al. (2025). Methodologically, it takes an initial step in developing a systematic, referable, and replicable framework for cross-character metadiscourse analysis, employing a corpus-based tool in AntConc. Empirically, it offers a resource of substantive and distinct engagement for five widely recognized fictional characters, which may be drawn upon for future inquiries into the interplay of language, identity, and representation within contemporary fictional constructs.

There are some limitations. The size of the corpus prevents reliable conclusions beyond the analysis of these characters, and the findings are restricted to these five characters and television series. Future studies may aim to incorporate a comparative analysis of male characters. One possibility may be applying this study's framework to the analysis of characters portrayed in television series that are produced in languages beyond English. It is also likely that a multimodal analysis of language and performance, sound, and visual elements may provide greater depth to the study of female identity. Although this study has many limitations, it may be the first of its kind to illustrate how language in a scripted medium constructs female identity within a highly reproductive sphere of popular culture.

Because this study lacks a male comparison group, it describes the interactional metadiscourse of the five female characters, rather than stating that such patterns are gendered, and claiming a gender difference would need a comparative male sample. Another limitation is regarding coding reliability: all data coding was done by one researcher and no check of inter-rater reliability was done. In order to improve coding quality, a second independent coder should classify the markers.

References

Al-Otaibi, G. M., & Hussain, A. A. (2024). The use of interactional metadiscourse markers by Saudi EFL male and female college students: The case of a gender-sensitive topic. Humanities and Social Sciences Communications, 11(1), 988. https://doi.org/10.1057/s41599-024-03506-3

Anthony, L. (2022). AntConc (Version 4.2.0) [Computer software]. Waseda University. https://www.laurenceanthony.net/software

Azlia, S. C. (2022). Interactional discourse of male and female motivational speech in TED Talks: A corpus-based study. Rainbow: Journal of Literature, Linguistics and Culture Studies, 11(1), 42–49. https://doi.org/10.15294/rainbow.v11i1.54777

Chou, I., Li, W., & Liu, K. (2023). Representation of interactional metadiscourse in translated and native English: A corpus-assisted study. PLOS ONE, 18(7), e0284849. https://doi.org/10.1371/journal.pone.0284849

Hyland, K. (2005). Metadiscourse: Exploring interaction in writing. Continuum.

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. Applied Linguistics, 25(2), 156–177. https://doi.org/10.1093/applin/25.2.156

Lasagabaster, D., & Bier, A. (2025). An examination of the use of spoken interactional metadiscourse markers in EMI lectures from different disciplines. English for Specific Purposes, 79, 137–151. https://doi.org/10.1016/j.esp.2025.05.003

Mensah, E. N. A., Hammond, C., & Wornyo, A. A. (2025). 'We mean more than what we say on air': Interactional metadiscourse resources in television talk shows. Cogent Social Sciences, 11(1), 2472917. https://doi.org/10.1080/23311886.2025.2472917

Montasseri, Z., Khaghaninejad, M. S., & Moloodi, A. (2020). Gender representation in American movies: A corpus-based analysis. The International Journal of Humanities, 27(4), 42–53. http://eijh.modares.ac.ir/article-27-30885-en.html

Puspita, S. R., & Suhandano. (2023). Gender-based analysis of interactional markers in personal essays. ELS Journal on Interdisciplinary Studies in Humanities, 6(3), 470–478. Gender-based Analysis of Interactional Markers in Personal Essays | ELS Journal on Interdisciplinary Studies in Humanities

Wu, L., & Wan Fakhruddin, W. F. W. B. W. (2023). Discourse analysis of female identity shaped by international media through Systematic Functional Linguistics. Eurasian Journal of Applied Linguistics, 9(1), 74–82. https://files.eric.ed.gov/fulltext/EJ1395143.pdf.

PAKISTAN JOURNAL OF LANGUAGE AND LITERARY STUDIES
Volume 2, Issue 1 (Spring 2026)
Pages: 36-51

Slavery, Trauma and Healing: A Corpus-Based Analysis of Temporal Markers in Selected African-American Neo-Slave Narratives.

Mahnoor Fatima[1]

ABSTRACT

The study explored the frequencies of temporal markers in selected African-American Neo-slave narratives, combining corpus linguistics with literary studies. These narratives focus on the historical experiences of slavery and its continuing psychological effects. While these narratives have been largely studied for their themes of trauma, history, identity and memory, limited attention has been given to linguistic features that shape them. The role of temporal markers and their contribution to representing trauma and healing remained under-explored. The study employed a sequential mixed-methods design, analyzing a corpus of five African-American Neo-slave narratives, Beloved, Homegoing, Kindred, Underground Railroad, and The Water dancer, consisting of 540,843 words. AntConc was used to generate frequency lists and concordance lines of temporal markers, while NVivo was used to code the extracted data according to Dominick LaCapra's trauma theory. The findings revealed that past-oriented temporal markers are abundant in the corpus. This presence supports LaCapra's trauma theory concept of Acting Out. Present and future-oriented markers revealed a progression toward healing corresponding to LaCapra's Working Through phase. This imbalanced distribution, with Acting Out coded 212 times against
146 for Working Through, showed the complex but progressive nature of healing within these narratives. It offered an empirical framework for understanding the relationship between language, trauma, and healing. The study provided insights into psychological research by highlighting how language simultaneously recorded and transformed the effects of historical trauma, suggesting that healing remains gradual or incomplete rather than fully resolved in trauma narratives. The study contributed to both corpus linguistics and trauma studies, suggesting that temporal markers can be measured as evidence of psychological state rather than only interpreted through close reading.

Keywords: *Temporal markers, Corpus linguistics, Neo-slave narratives, Trauma theory, Healing.*

[1] Research Scholar, Department of English Language and Literature, Government Fareed Bukhsh Ghousia Graduate Science College 333 GB , Toba Tek Singh, affiliated with GCUF.
Secondary Affiliation: Corpus Research Centre, Minhaj University Lahore (MUL), Lahore, Pakistan.

Corresponding author:
Ms. Mahnoor Fatima, Department of English Language and Literature, Government Fareed Bukhsh Ghousia Graduate Science College 333 GB , Toba Tek Singh, affiliated with GCUF
Email: harrara12345@gmail.com

1. Introduction

1.1. Background of the study

African-American Neo-slave narratives are literary works that center on the history of slavery and its enduring effects on African-American lives. These narratives are a collection of experiences from enslaved people that relate past to present realities. Neo-slave narratives are narratives by contemporary writers, as opposed to the original slave narratives, in which the author interweaves fictionalized elements with the history of slavery. Their works will serve to remember the history and track the intergenerational impact of slavery. Many Neo-slave narratives focus on trauma, history, identity, resistance, and freedom, highlighting the ways in which the brutality of slavery continues to affect the lives of individuals even after its abolition.

Trauma is central to this genre, shaping not only what is narrated but how it is narrated. Memories of painful events interrupt present experiences and create shifts between present and past. Trauma is in these narratives, not just a theme, but a narration itself. The genre is structured by fragmented structures, non-linear chronology and recurring flashbacks, which emphasize the fact that traumatic memory is unconfined to past.

1.2. Research Problem

African-American Neo-slave narratives hold a prominent position in literary history because of the impact they have on history of slavery and its psychological, social and cultural impact. Existing research has largely approached these narratives through thematic analysis, focusing on memory, identity, oppression, history, and feminist perspectives. However, these narratives are fundamentally linguistic constructs in which meaning is shaped, trauma is portrayed, and healing is signaled through language itself. In particular, the way language simultaneously functions as both the representation of trauma and the possibility of healing remains theoretically overlooked. There is insufficient understanding of how linguistic structures, specifically the way temporal experience is encoded and shapes readers understanding of trauma and healing within these narratives. Hence, further empirical and systematic research is required to examine how the linguistic patterns specifically temporal markers shift between traumatic past and present experiences.

1.3.Research Gap

While existing studies have approached Neo-slave narratives through literary criticism, post-colonial theory, and cultural analysis (Friedman, 2021), systemic investigation of these texts remains overlooked. Caruth (1996) emphasizes trauma as a phenomenon that resists integration into non-linear memory, appearing through repetition and temporal disruption. Crespo and Fernandez-Lansac (2016) trace this pattern empirically, finding that tense shifts and disrupted temporal context correlate with severity and healing outcomes.

Specifically, no corpus-based study has examined how temporal markers shift across traumatic and healing-oriented narrative moments in this literary tradition. This represents both a methodological and analytical gap; the linguistic mechanisms through which these narratives move from trauma to healing have not been empirically examined. The study, therefore, responds to this gap by identifying temporal markers and their frequencies in

selected African-American Neo-slave narratives. This analysis reveals language as a double-edged sword, simultaneously destructive and therapeutic in the movement from trauma to healing.

1.4.Objectives

1. To identify the frequencies of temporal markers in selected African-American Neo-slave narratives..
2. To examine the temporal markers representing the transition from trauma to healing across selected African-American Neo-slave narratives.

1.5.Research Questions

1. What are the frequencies of temporal markers in selected African-American Neo-slave narratives?
2. How do the temporal markers represent the transition from trauma to healing across selected African-American Neo-slave narratives?

1.6.Significance of the Study

The study is significant in various dimensions. Firstly, it is significant to the field of literary linguistics because it moves away from thematic analysis to systematic analysis of linguistics for African-American Neo-slave narratives. This approach offers a way of thinking about the functioning of language in the process of building trauma and healing. Secondly, the study is of methodologically significance because it introduces a corpus-based analysis to a genre which has traditionally been studied in a qualitative literary critical style. This empirical approach enables to analyze the temporal markers systematically, which further enhances the analysis conducted on these narratives.

Thirdly, the study has important inter-disciplinary value. It combines literature and corpus linguistics, indicating that this kind of narratives can only be fully understood through the examination of the words used and their linguistic features. This inter-disciplinary approach has broader implications for how researchers engage with other literary traditions beyond the African-American context. Lastly, the study carries social and cultural significance by drawing attention to the expressive ability of African-American writers whose linguistic choices reflect deep psychological and historical realities of slavery.

2. Literature Review

Language plays a significant role in expressing pain, memory, and healing in literary narratives. Temporal markers are words or phrases such as before, after, then, and now that indicate transitions in a narrative. In trauma narratives, non-linear or recursive markers often reveal “psychological mutilation” produced by traumatic experience (Habli, 2023). African-American Neo-slave narratives have been studied for their reconstruction of slave history and their representation of the psychological effects of enslavement (Edwards-White, 2016). Edwards-White focuses on the lasting psychological effects of enslavement on the individual, Guettaia (2025) analyzed Edward P.Jones’ “The Known World” and Morrison’s “A Mercy, broaden the frame to show these narratives connect past experiences of slavery to present-day identity, memory, and resistance, shifting the focus of analysis from the individual psyche to collective identity formation. This review synthesizes literary-critical and corpus-linguistic

scholarship on trauma and temporality in order to situate the present corpus-based study of temporal markers within both traditions.

The study draws on Dominick LaCapra's (2001) trauma theory, particularly the concepts of Acting Out and Working Through. "Acting Out" appears when the traumatic past continues to haunt individuals and the narrative remains stuck in a cycle; linguistically, this shift is often signaled by recursive markers and disrupted chronological sequencing. In contrast, "Working Through" is a process in which individuals organize their experiences into a linear narrative and move toward reconstructing identity; linguistically, this shift is often marked by forward-oriented, resolving temporal markers that restore chronological progression. Scholars have applied LaCapra's framework directly to Neo-slave narratives, arguing that texts like Beloved and Kindred use narrative techniques such as haunting and time travel to achieve therapeutic witnessing of slavery's traumatic history (Madsen, 2010), treating Working Through as a psychological state internal to the protagonist. Chicholewski (2020), in contrast, reframes the same narrative techniques as "Chronopolitical interventions" that challenge linear history and foreground slavery's socio-political consequences, relocating Working Through from individual psychology to collective historical assessment. Examined through a critical lens, these two accounts suggest that Acting Out and Working Through operate on more than one scale at once, which is part of the motivation for analyzing temporal markers empirically rather than assuming a single, linear progression.

One of the recurring themes in this scholarship is the temporality of the Neo-slave narrative: the rejection of the temporal linearity, as a way to present trauma as a continuing process not something finished. Both, Habli (2023) and Abdulsahib (2026) examined Morrison's Beloved as an enactment of this temporal mutilation through repetition and displacement, arguing that this fragmentation is a window into the way of traumatic memory that obscures the past and present. Lee (2015) extends this focus on structure to the language itself, demonstrating how protagonist of Dessa Rose, Kindred and Corregidora oscillate between trauma and healing in different linguistic modes. These linguistic forms are oral testimony, words and songs, and suggest that language is a major means of oppression and resistance in these narratives.

Mu (2026) argues that trauma is not confined to the past but continues to persist in present. Lee's medium-based account and Edwards-White and Mu's identity-based account describe the same underlying interplay between Acting Out and Working Through, but locate its evidence in different contextual features: form of utterance versus content of identity. Examining temporal markers offer a way out of this conceptual split. Because these markers operate within both formal structures and expressions of identity, tracking them provides a single, unified method to analyze both dimensions simultaneously.

A second strand of this scholarship shifts from narrative structure to the social and communal functions of language. Focusing on relationship between language, trauma, emotion, race, and space in contemporary African-American literature, Mikic's (2025) emotion-race framework shows how racial and spatial oppression produce emotions such as fear, shame, guilt, anger, and grief, and how Black place-making practices convert these emotions into resistance and community-based healing.

Boularas (2024) and Viscuso (2024) narrow this communal lens to testimony, specifically, Boularas frames testimony as a tool for challenging racist stereotypes and reconstructing Black identity. In contrast, Viscuso treats individual traumatic moments as turning points that reshape identity, behavior and relationships across generations. Where Mikic operates at the level of individual utterance and life-event, suggesting that Working Through in these narratives is rarely a private achievement but can be achieved by communal structures.

Fontes (2021) and Ortega-Williams et al. (2024), both drawing on the concept of Post-Traumatic Slave syndrome, locate healing in collective rather than individual processes. This extends Mikic's and Boularas communal framing but shifts the focus from spatial practice or testimony to inter-generational transmission, suggesting that Working Through in Neo-slave narratives is collective.

Regardless of this literary-critical engagement with trauma and temporality, the field has examined almost entirely through qualitative and thematic methods. Salama et al. (2016) and Tazzioli (2021) rely on close reading and thematic analysis; Abood (2024) applies systematic functional grammar, but only to nineteenth century narratives. In contrast, Rodriguez (2021), Akter (2025), and Nelson (2026) foreground memory, identity and historical recovery without analyzing linguistic features at all. However, Nelson's treatment of language confined to eighteenth and nineteenth century texts. Using different analytical frameworks, all the stories remained focus on the idea that slave narratives do not follow a linear storyline.

However, corpus-linguistic engagement with this literary-critical tradition remains incomplete. Mohammed (2017) used frequency and percentage counts of lexico-syntactic features, sentence types, African-American Vernacular English markers, and cohesive devices to characterize narrative style of Morrison in Beloved and Paradise. This study finds that Neo-slave narratives are amenable to frequency-based analysis. Elmerot et al. (2026) apply a different corpus-based method to older material, building an annotated corpus of autobiographies and interviews of formerly enslaved persons, comparing key noun usage across sub-corpora. By using Linguistic Inquiry and Word Count on clinical trauma narratives, Ryan (2023) shifts the objective of study, finding that trauma leaves a measurable, lexical signature. Taken together, these studies differ in text type and in method but converge on a single point that frequency-based analysis provides evidences that close reading does not offer. This is a methodological warrant for treating temporal markers themselves as measurable evidence of the shift between Acting Out and Working Through.

In summary the reviewed literature establishes non-linear temporality, testimony, and communal healing as central to Neo-slave narratives. It also highlights the role of narrative structures in representing trauma, memory, pain, and identity in African-American Neo-slave narratives. However, a smaller body of corpus-linguistic research confirms the empirical viability of frequency-based method for trauma more broadly. No study bridges nineteenth century and contemporary Neo-slave narratives to trace a transition from trauma to healing. This gap underscores the need for an empirical investigation of how language operates in African-American Neo-slave narratives, representing transitions between Acting Out and Working Through.

3. Research Methodology

The study used mixed-methods sequential approach, whereby the quantitative analysis of the number of temporal markers was followed by a qualitative interpretation of the meaning of the temporal markers. In first step, quantitative approach is employed to determine frequencies of temporal markers in the selected African-American Neo-slave narratives. This quantitative method is most appropriate for the research problem since it allows the researcher to use objective frequency measurement and also conduct thorough statistical analysis of the data. The qualitative interpretation is guided by Dominick LaCapra's trauma theory, specifically his distinction between Acting Out and Working Through. Acting Out refers to a repetitive reliving of traumatic past in which the boundary between past and present collapses. In contrast, Working Through refers to the gradual progression from trauma to agency and engagement with life. These two concepts, in this study function as qualitative coding categories. Each temporal marker identified in the corpus is examined in the context and classified as either Acting Out or Working Through. This methodology best suits the present study because it will provide insights into how language functions in trauma narratives.

3.1.Sampling Method and size

The targeted population of this study comprises of African-American Neo-slave narratives, centered on the experiences of slavery, memory, trauma, and identity. The study uses non-probability sampling, specifically purposive sampling by selecting five narratives (N5) are selected as the primary dataset. The selected narratives include Beloved (1987) by Toni Morrison, Homegoing (2016) by Yaa Gyasi, The Underground Railroad (2017) by Colson Whitehead, Kindred (1979) by Octavia Butler, and The Water Dancer (2019) by Ta-Nehisi Coates. The entire corpus comprised of 540, 843 words. Each novel was selected on the basis of its standing within the Neo-slave narrative tradition and confirming a recurring use of temporal markers shifting between Acting Out and Working Through. Analyzing complete narratives rather than excerpts, ensures comprehensive identification of temporal markers.

Fifteen temporal markers were selected on the basis of their categorization into present, past, and future category for the frequency and concordance analysis: then, now, still, always, suddenly, finally, again, soon, after, before, while, since, until, and at last. These were chosen because they recur across Neo-slave narratives as linguistic signals of temporal orientation, giving a consistent dataset across all five African-American Neo-slave narratives.

3.2.Data Collection Tools

AntConc was used for data extraction, with the cleaned corpus loaded. Each of the fifteen temporal markers was searched individually, one marker at a time in order to generate its frequency list and concordance lines, rather than running a single combined search. This one-by-one procedure made it possible to identify each markers frequency, range, normalized frequency separately, and to read every occurrence in its surrounding context before moving to qualitative coding phase.

Nvivo used for qualitative coding. Concordance lines of each temporal marker were imported to NVivo and coded into one of two nodes: Acting Out and Working Through. In Nvivo, a

marker was coded Acting Out when its context showed the past interrupting or repeating within the narrative present in the form of recurring flashbacks, cyclic phrasing, and collapse of then and now. While, it was coded Working Through when its context showed resolution or forward movement in the form of explicit before or after contrasts, and narrative linearity. Applying this scheme across the corpus produced 212 coding references under Acting Out and 146 Under Working Through.

3.3.Data Analysis Plan

The present study analyzed data by employing a mixed-methods sequential approach. The quantitative aspect focuses on the identification of the frequency of temporal markers, while qualitative aspect examines how these temporal markers mark transition from trauma to healing in selected African-American neo-slave narratives. Firstly, all the narratives are prepared and loaded into AntConc, providing frequency lists and concordance lines of temporal markers. For qualitative analysis, NVivo is used to organize and categorize the extracted patterns following LaCapras's concepts of Acting Out and Working Through. This enables markers to be analyzed in terms of repetitive trauma or progression towards healing, making the analysis systematic and theoretically sound. The combination of both tools allows for a comprehensive analysis of narratives. This data analysis plan directly supports the research objectives and ensures both research questions are fully answered.

3.4.Ethical Considerations

The study was based on secondary data, selected African-American Neo-slave narrative, no human participants were directly involved in this study. So, issues regarding informed consent were not applied. However, the research strictly followed ethical standards of academic integrity and use of data. To maintain ethical responsibility, all the sources were properly cited and referenced following standard academic requirement. This study further ensured honest and accurate representation of data on the basis of evidence avoiding bias or prejudice. It also employed anti-plagiarism and citation compliance to ensure balanced, objective interpretation of literary data. The study consists of published literary works so, the analysis remains respectful and objective. All the collected data was stored in digital format, no authorized sharing or misuse of data occurred. Data collected during this internship remained the joint intellectual property of the Corpus Research Center (CRC), Minhaj University, Lahore, and the research to ensure transparency and ethical compliance throughout the internship.

4. Data Analysis and Findings

The study analyzed five African-American Neo-slave narratives as a primary corpus, comprising of 540,843 words. The selected narratives include Beloved (1987) by Toni Morrison, Homegoing (2016) by Yaa Gyasi, The Underground Railroad (2017) by Colson Whitehead, Kindred (1979) by Octavia Butler, and The Water Dancer (2019) by Ta-Nehisi Coates. For quantitative part, AntConc was used to generate frequency lists and concordance lines of temporal markers, facilitated the extraction of statistical data for further analysis. This method was based upon the principles of corpus linguistics, as described by Biber et al.

(1998) and the principles of frequency and concordance analysis described by McEnery and Hardie (2012).

For qualitative part, thematic analysis following principles of Braun and Clarke (2006) was applied, supporting identification and interpretation of recurring linguistic patterns and meanings across selected narratives. The dataset consisted of approximately 22 pages of purposively selected textual instances for further detailed analysis in NVivo. By employing Dominick LaCapra's (2001) trauma theory, particularly the concepts of Acting Out and Working Through, data was coded and analyzed. The integration of both qualitative and quantitative approaches ensured a comprehensive analysis of temporal markers within selected African-American neo-slave narratives.

The analysis followed seven steps: firstly, each narrative is cleaned in Notepad to remove footnotes and page numbers. Then, five cleaned novels compiled and uploaded in AntConc. After that, each of the fifteen temporal markers is searched individually in AntConc to generate frequency lists, concordance lines, and N-Gram patterns of size 3. Qualitative part includes importing the concordance lines into NVivo. In NVivo, each line was coded as per the definitions of LaCapra's trauma theory concepts of Acting Out and Working Through. Then, integrating the frequency list data (step 2) with the coding results (step 5) to trace how the balance between Acting Out and Working Through shifts across each narrative, answering both quantitative and qualitative research questions in a single, linked analysis.

To ensure reliability of the analysis, instrumental reliability was first achieved through AntConc and NVivo, which reduced the risk of inconsistency in automated frequency counting, supplemented by manual tabulation for verification and cross-checking of key findings. The contextual patterns were assessed under consistent criteria across the corpus. This computational approach was complemented by institutional oversight, Corpus Research Center, Minhaj University, Lahore, providing procedural supervision. The study was carried out over a planned period of 2 months, allowing the findings to be cross checked by the periodic experts. This combination of software counting, and institutional supervision was aimed to limit prejudice and strengthen the overall reliability of findings.

4.1.Findings Rol:

The analysis revealed that temporal markers serve as key linguistic features in African-American Neo-slave narratives, structuring narrative time, and shaping how trauma is expressed and processed. Corpus-Based analysis identified frequent recurrence of markers such as then, when, before and now, which recur across the five narratives, revealing that past-oriented markers are abundant as compared to that of present-oriented markers. The recurrences of markers organize memory construction and narrative continuity.

Figure 1

Values represent normalize frequency per (10,000) words of each temporal markers highest frequency trigram across five narratives. Bars are ordered by descending frequency.

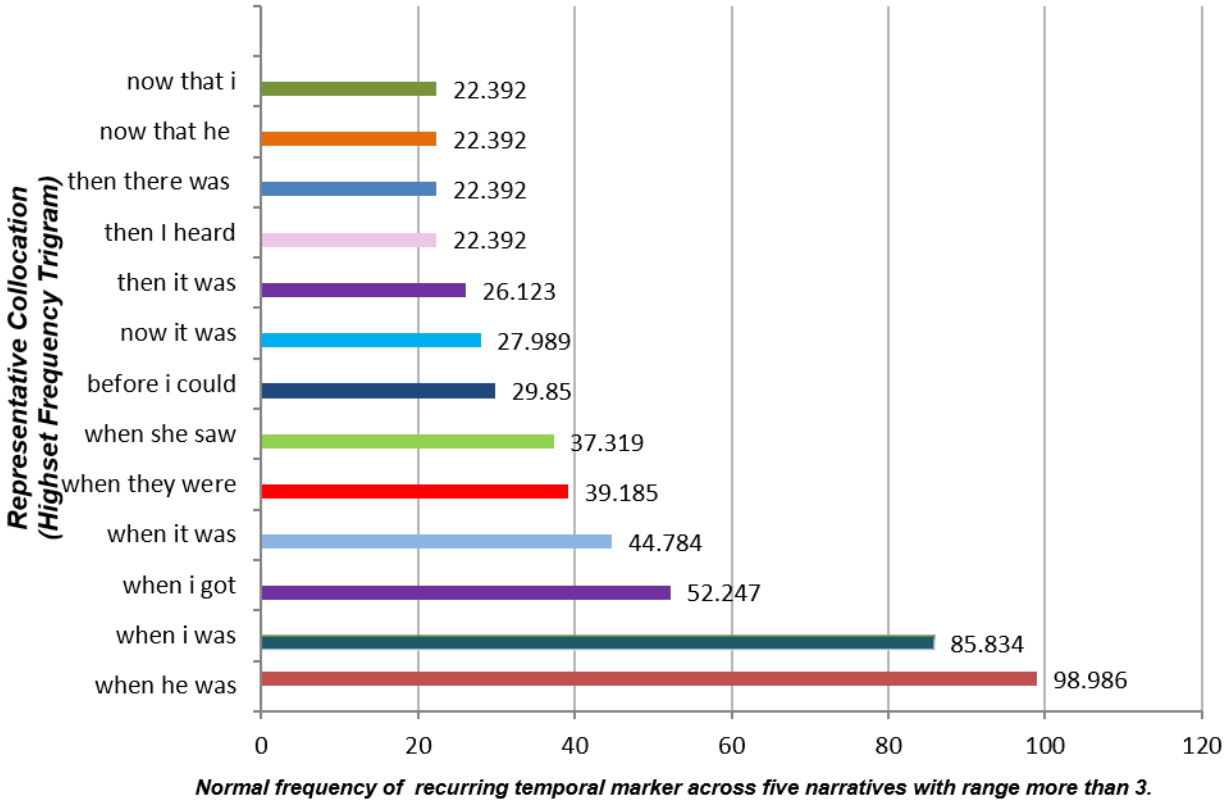


Table 1

No.	**Temporal Marker**	**Representative Collocation** (Highest Frequency Trigram)	***f*** (Collocation)	**Range** (Narratives)	**Total f** (Marker)	**Norm. *f***
1	When	*"When he was"*	53	5/5	1248	98.986
2	Then	*"Then I was"*	14	2/5	634	26.123
3	Now	*"Now and Then"*	16	4/5	463	29.855
4	Before	*"Before I could"*	16	3/5	304	29.855
5	After	*"After a while"*	21	2/5	235	39.185
6	Until	*"Until he was"*	10	4/5	150	18.66
7	Again	*"Again and again"*	9	4/5	147	16.794
8	Still	*"Still in the"*	7	4/5	132	13.062
9	While	*"While I was"*	15	2/5	104	27.289
10	Since	*"Since the day"*	15	3/5	76	27.989
11	Soon	*"Soon as I"*	11	4/5	62	20.526
12	Always	*"Always said that"*	4	1/5	39	7.464

13	At Last	*"At last I"*	5	1/5	16	9.33
14	Suddenly	*"Suddenly I was"*	5	2/5	13	9.33
15	Finally	*"Finally I got"*	3	1/5	11	5.598

The table shows the high normal frequency of each temporal markers appeared in AntConc with their frequency of the collocation patterns and clusters/N-Grams.
Function;Range(/5)= number of narratives, out of five in which the pattern occurred. Norm. f=Normal frequency per 10,000 words.

Frequency analysis (figure:1 and table:1) shows that past-oriented markers, particularly then, now and when occurred frequently than present-oriented forms as now (When: 1248 total occurrences, 5/5 narratives, Then: 634 total occurrences 4/5; Now: 463, 5/5). This numeric dominance of past-oriented temporal markers is the evidence of the narrative's orientation toward memory toward memory and unresolved trauma. A pattern consistent with LaCapra (2001) concept of Acting Out in which trauma is relived rather than fully resolved. The comparatively lower but still consistently distributed frequency of now signal a point of character's agency and psychological stability. The remaining temporal markers such as after, until, again, always, today, finally, suddenly, at last, while, again, since, and soon occurred with comparatively lowered but consistently distributed frequencies (Table:1), reinforcing the corpus-based tendency toward past-oriented temporal markers.

Concordance Lines for the Temporal Marker "Now" Generated via AntConc's KWIC Function.

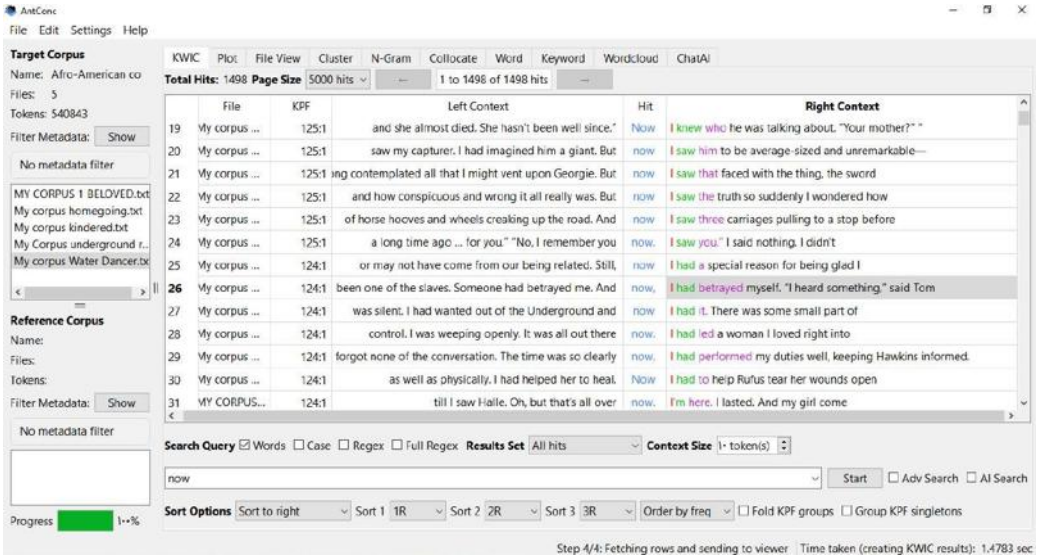


Figure 2

Concordance lines for the temporal markers generated using AntConc's KWIC function (Figure:2). This inspection of now in its immediate context reveals that the markers frequency co-occurs with verbs of cognition and perception (I knew, I saw, I had), situating in within moments of retrospection realization rather than present tense narration. It indicates its function extends beyond simple chronological marking to reflect deeper shifts in narrative perspectives.

These findings support the corpus linguistic framework of McEnery & Hardie (2012), emphasizing that frequency and contextual repetition help to uncover important meanings in discourse.

4.2.Findings RO2

Temporal markers organize and shape overall coherence of trauma and healing based experiences. Thematic and interpretive analysis of these markers revealed that these markers appear in the contexts of fear, harm, emotional sufferings, and healing. In Kindred, the line "Man cursing and lec-turing as he beats me" (Butler, 1979, p. 107), showed the moments of physical pain and suffering. In Beloved, "Then she heard close thunder crawling upon the stairs" (Morrison, 1987, p. 103), showed that temporal markers frame moments of fear and psychological tension. Overall, these instances demonstrated that how trauma repeatedly disrupts narrative consciousness, compelling characters to relive and process painful experiences. These extracts represented traumatic repetition, fear, and emotional suffering. This aligned with LaCapra's concept of Acting Out, where trauma is relived through fragmented memory and emotional pain.

In contrast, these Neo-slave narratives also carry the instances of gradual healing. In Water Dancer, "I tried to crawl, and I stood" (Coates, 2019, p. 45), showed gradual progression toward healing. In Kindred, "Then I got back under the cover, I am healing" (Butler, 1979, chap. 2, para. 4). These extracts aligned with LaCapra's trauma theory, particularly the concept of Working Through, where characters begin to move forward toward emotional stability and healing.

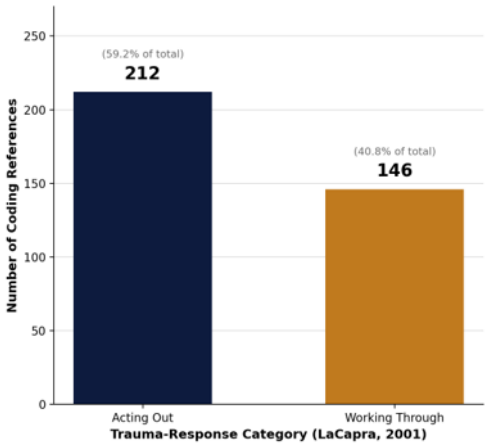


Figure 3

The coding references for the two phases: Acting Out identified in 212 instances, while Working Through appeared in 146 instances. Temporal markers related to trauma were found in code, Acting Out, while, healing related markers appeared in code; Working Through, created as per LaCapra's trauma theory definitions of Acting Out and Working Through.

Table 2

Marker	Concordance Line	Source Text	Coded Function

now	"...how conspicuous and wrong it all really was. But now I saw the truth so suddenly..."	*Kindred*	Transition pivot — sudden insight following prior distortion.
now	"...I had wanted out of the Underground and now. I had it."	*Kindred*	Healing-marking — arrival at a previously unattainable state.
now	"...I had helped her to heal." Now I had to help Rufus tear her wounds open—	*Kindred*	Counter-evidence , 'now' marking regression; healing framed as non-linear.
today	"...it is hard on us. It exhausts us. Even today, I could only do this with the aid of my brothers."	*The Water Dancer*	Healing-marking — present endurance framed as relational, not as individual.
today	"...one of the police officers... shook his head, and said, 'Do that and you'll die today.'"	*Homegoing*	Trauma-marking — acute crisis-point that later leads tochange.
then	"Pain again shot through my ankle... And then I stood. I took one step. It was agony. I fell. And again..."	*Beloved*	Transitional pivot — sequences incremental, faltering physical recovery.
then	"...I am determined to get Thena out... Then she leaned over and kissed me lightly on the cheek."	*The Water Dancer*	Healing-marking — relational tenderness following trauma-testimony.

The representative concordance lines for the temporal markers (Now, Then, Today) in Table 2, extracted from the corpus to provide qualitative evidence to show transition from trauma to healing. These coded instances provided evidence of varied functions performed by temporal markers

4.3.Discussions

The findings revealed that temporal markers are indicators of trauma, memory, and healing in African-American Neo-slave narratives. Their occurrence reflects how characters continuously move between past and present. Past-oriented markers such as then and when represent traumatic memory, return of suffering, and narrative fragmentation. However, present-oriented markers such as now and today indicated self-awareness, emotional stability and narrative progression. The findings further suggested that fragmented linguistic structures reflect unstable and non-linear nature of traumatic experiences.

4.4.RO1

The findings of the study supported previous research on trauma, memory, temporality, and African-American Neo-slave narratives. The frequent use of temporal markers such as then, now, when supported Habli (2023), argued that trauma narratives disrupt linear time through repetition and past memories. This numeric dominance also supported Madsen (2010) and

Chicholewski (2020), who both read repetition and non-linearity in in Neo-slave narratives as central to trauma representation, though their analysis was relied heavily on close reading. Similarly, the present study found that past-oriented temporal markers repeatedly occur in contexts related to emotional fragmentation, suffering of past, and flashbacks. However, the present study differs from the previous study such as Viscuso (2024), (Boularas, 2024), and Fontes (2021) which primarily focused on historical, psychological, and thematic representation of trauma. These studies discussed trauma, memory, and healing conceptually, paying limited attention to corpus-based analysis of linguistic patterns specifically, temporal markers. In this regard, the present findings extend the work of Mohammed (2017) and Elmerot et al. (2026), moving beyond stylistic or lexical frequency counts to reveal that temporal markers can be traced as measurable evidence of trauma and healing across an entire corpus rather than a single narrative

4.5.RO2

Temporal markers actively trace the psychological movement between trauma and healing supporting LaCapra's (2001) thematic distinction between Acting Out and Working Through. The dominance of Acting Out (212 coded instance) over Working Through (146 coded instances), aligns with Habli (2023) and Abdulsahib (2026), who argued that traumatic memory in Neo-slave narratives resists closure and continues to reappear in the present. However, the presence of a considerable number of Working Through (146), supports Madsens (2010) reading of Kindred and Beloved as therapeutic witnessing, in which healing gradually emerges alongside trauma rather than emerging as final narrative resolution.

Unlike Fontes (2021) and Ortega-Williams et al. (2024), who located healing primarily in collective or intergenerational processes, the present findings of RO2 suggests that healing is also traceable at the level of individual choice of language within a single narrative. The findings strongly supported Dominick LaCapra's trauma theory, particularly concepts of Acting Out and Working Through. The use of temporal markers such as then, before, and when, highlighted fragmented narration, non-linear progression, and emotional pain, corresponding to LaCapra's concept of Acting Out. In contrast, the markers such as now, after, and today highlighted emotional relief, gradual healing, and narrative linearity, corresponding to Working Through. Researchers, scholars, and psychologists can use temporal markers analysis to understand how language represents trauma, memory, and healing in Neo-slave narratives. The study may facilitate the identification of psychological and emotional patterns within literary texts.

The study successfully achieved both objectives. The first objective: To identify the frequency of temporal markers in selected African-American Neo-slave narratives, achieved through frequency lists and concordance lines analysis in AntConc. The second objective: To investigate temporal markers that indicate transitions from trauma to healing in selected African-American stories using KWIC analysis and coding in NVivo. The temporal markers used for the healing purposes were not concentrated on the endings but spread throughout the narratives. This distribution has shown that healing happens slowly over time in the aftermath of trauma, not just at the end of each story

5. Conclusion

In conclusion, temporal markers play a significant role not only in representing trauma but also healing within African-American Neo-slave narratives. The trauma related temporal markers recurrence highlights that trauma is not confined to past, instead it continues to haunt the individuals in present. Healing gradually becomes possible through testimony and communal voice, allowing for the reconstruction of the self and psychological stability. The integration of corpus-based tools with LaCapra's trauma theory helps to understand the role of language in trauma narratives.

The present study is limited to five narratives, while future research may analyze ten or more novels for broader findings. Additionally, researchers may employ the same corpus-based framework for narratives written by authors from distinct ethnic backgrounds, such as Caribbean and African diasporic literature. This comparative approach would reveal whether the linguistic patterns, more specifically the temporal markers are genre specific or appear in trauma literature globally. Future research is also encouraged to employ more diverse corpus tools such as Sketch Engine or Python, advancing frequency and collocation analysis. This would strengthen corpus-based studies in terms of both reliability and the depth of corpus-based studies.

Future research may benefit for combining corpus analysis with psychological assessments, examining how readers emotionally process temporal markers in trauma narratives. Such inter-disciplinary approaches would further bridge the gap between linguistic analysis, literary studies, and psychological research.

References

Abdulrafiu, T. (2025). Reimagining slave narratives in contemporary African American and Caribbean literature: A postcolonial and intersectional analysis of trauma, memory, and resistance. GRS Journal of Multidisciplinary Research and Studies, 2(9). https://doi.org/10.5281/zenodo.17119939

Abdulsahib, Z. A. (2026). Power, voice, and representation in Toni Morrison's Beloved: A critical study of narrative discourse. International Journal of Education Humanities and Social Science, 9(02), 317–327. https://doi.org/10.54922/IJEHSS.2026.1245

Abood, H. H. (2024). Systemic functional linguistics to stylistics analysis of Beloved novel by Toni Morrison. Journal of Language Studies, 8(11), 319–343. https://doi.org/10.25130/Lang.8.11.20

Akter, M. (2025). Exploring the works of contemporary writers inspired by Toni Morrison through the lens of intersectionality: Slavery, trauma and dehumanisation (Bachelor's thesis, BRAC University). http://hdl.handle.net/10361/26527

Biber, D., Conrad, S., & Reppen, R. (1998). Corpus linguistics: Investigating language structure and use. Cambridge University Press. https://doi.org/10.1017/CBO9780511804489

Boularas, S. (2024). Collective memory and the recasting of post-civil war African-American identity in selected Toni Morrison's novels. Theory and Practice in Language Studies, 14(10), 3125–3130. https://doi.org/10.17507/tpls.1410.14

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Butler, O. E. (1979). Kindred. Doubleday.

Caruth, C. (1996). Unclaimed experience: Trauma, narrative, and history. Johns Hopkins University Press.

Chicholewski, A. (2020). Chronopolitical Interventions in the Afterlife of Slavery: Forms and functions of temporal disruptions in contemporary speculative neo-slave narratives. [Doctoral dissertation, Carl Von Ossietzky University of Oldenburg]. https://oops.uni-oldenburg.de/id/eprint/5181

Coates, T. N. (2019). The Water Dancer. Random House.

Crespo, M., & Fernández-Lansac, V. (2016). Memory and narrative of traumatic events: A literature review. Psychological Trauma: Theory, Research, Practice, and Policy, 8(2), 149–156. https://doi.org/10.1037/tra0000041

Edwards-White, M. M. (2016). Critical exploration of the neo-slave narrative in selected Walker and Morrison novels. http://opus.govst.edu/thesis/83

Elmerot, I., Olsson, L.-J., & Rönnbäck, K. (2026). Voices of formerly enslaved: A new text corpus of narratives by formerly enslaved persons. Scientific Data, 13, Article 682. https://www.nature.com/articles/s41597-026-07340-x

Fontes, F. M. da R. (2021). Memory, language and identity construction in Toni Morrison's Song of Solomon and Beloved (Master's thesis, Universidade Federal de Minas Gerais). https://hdl.handle.net/1843/39468

Friedman, G. (2021). Unsentimental historicizing: The neo-slave narrative tradition and the refusal of feeling. American Literature, 93(1), 115–143. https://doi.org/10.1215/00029831-8878542

Guettaia, S. E. (2025). Unveiling the continuum: Analyzing slave narratives and neo-slave narratives through Edward P. Jones' The Known World and Toni Morrison's A Mercy. Dirasat Insaniyyah wa Ijtimaiyyah [Journal of Social and Human Science Studies]500–491 ,(1)14 ,. https://asjp.cerist.dz/en/article/262335

Gyasi, Y. (2016). Homegoing. Knopf.

Habli, R. (2023). Beloved & the erotics of temporal mutilation. BAU Journal - Society, Culture and Human Behavior, 4(2). https://doi.org/10.54729/2789-8296.1126

Lee, J. C. (2015). Formations of matrilineal memory: Sites of trauma within twentieth-century neo-slave narratives (Master's thesis, University of California, Los Angeles). https://escholarship.org/uc/item/0rg6d92w

Madsen, D. L. (2010). Teaching trauma: (Neo-)slave narratives and cultural (re-)memory. In G. Wisker (Ed.), Teaching African American women’s writing (pp. 60–74). Palgrave Macmillan. https://archive-ouverte.unige.ch/unige:87694

McEnery, T., & Hardie, A. (2012). Corpus linguistics: Method, theory and practice. Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

Mikić, M. (2025). Emotion, race, and space in contemporary African American literature. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-85795-9

Mohammed, H. (2017). A linguistic stylistic analysis of Toni Morrison's Beloved and Paradise [Unpublished doctoral dissertation]. Ahmadu Bello University.

Morrison, T. (1987). Beloved. Plume.

Mu, S. (2026). An analysis of neo-slave narrative in The Underground Railroad. Scientific Journal of Humanities and Social Sciences, 8(4), 232–240.

Nelson, A. (2026). "The past was waiting for me": Removing the veil in the neo-slave narrative (Publication No. 32700790) [Master's thesis, California State University, Bakersfield]. http://hdl.handle.net/20.500.12680/2j62sf516

Ortega-Williams, A., Stephens, T., & Henderson, Z. (2024). Black intergenerational healing and well-being: Reimagining posttraumatic growth. American Psychologist, 79(8), 1171–1184. https://doi.org/10.1037/amp0001302

Rodriguez, G. (2021). The presence of past: Experimental modes of representation in neo-slave narratives [Doctoral dissertation, The University of Texas at Austin]. Texas Scholarship Repository. https://hdl.handle.net/2152/89461

Ryan,R. A. (2023). Trauma and linguistic markers for symptom severity in prolonged exposure for primary care, as observed through the use of narrative review (Doctoral research project, Florida Institute of Technology). https://repository.fit.edu/etd

Salama, E. G. (2016). Afro-American literature: A systemic functional analysis of discourse in selected Afro-American slave narratives [Master's thesis, Banha University]. https://doi.org/10.13140/RG.2.2.25957.60642

Tazzioli, F. (2021). Literature and trauma: The neo-slave narrative in the contemporary British context. In M. Ascari & G. Imposti (Eds.), Quaderni del Dottorato: Letteratura e altri mondi: Generi, politica, società (pp. 63–76). University of Bologna. https://doi.org/10.6092/unibo/amsacta/6793

Udoette, M. (2023). Re-inventing the past as "re-memory": Trauma, motherhood and history in Toni Morrison's Beloved. Sapientia Global Journal of Arts, Humanities and Development Studies, 6(3), 207–216.

Viscuso, C. (2024). Understanding racial encounters in African American literature: Trauma, identity, and literary analysis in African American narrative. International Review of Literary Studies, 6(2), 27–41.

PAKISTAN JOURNAL OF LANGUAGE AND LITERARY STUDIES
Volume 2, Issue 1 (Spring 2026)
Pages: 52-67

Evaluating Forced Alignment in an Under-Documented Shakargarh Punjabi Speech Corpus

Adnan Arif[1]

ABSTRACT

This study evaluated the applicability and usefulness of forced alignment for developing a searchable speech corpus of under-documented Shakargarh Punjabi, a local Punjabi variety spoken in Narowal District, Punjab, Pakistan. Positioned at the intersection of corpus linguistics, speech technology, and language documentation, the study addressed the problem that low-resource speech recordings often remain insufficiently transcribed, time-aligned, searchable, and reusable for linguistic analysis. The design adopted was a mixed method that consisted of quantitative descriptive evaluation and qualitative error analysis. Primary spoken data were generated from 24 recordings of 12 native speakers using semi-structured interviews, natural conversations and short personal accounts that produced 120 rated utterances (19.79 minutes, 1,565 words). The recordings were subsequently transcribed in Shahmukhi Punjabi; with a consistent Roman transliteration based on Montreal Forced Aligner, annotated in ELAN and manually verified in Praat. Successful utterances were those in which the ABDev remained under 50ms and the segment mismatch was not present. 74 (61.7%) utterances were successfully aligned initially and 96 utterances (80.0%) were successfully aligned after project specific pronunciation-lexicon refinement. The number of OOV cases dropped from 62 to 18 instances while the OOV rate dropped from 3.96 to 1.15 OOV/100 words. The remaining problems primarily were those items which were not present in the standard language, connected speech, and vowel reduction. The findings show that semi-automatic alignment, supported by local lexicon refinement and manual validation, can materially improve the searchability and documentary value of Shakargarh Punjabi speech data.

Keywords: *forced alignment; Shakargarh Punjabi; speech corpus; language documentation; low-resource languages*

[1] PhD Scholar, School of English Language, Technology and Literature, Faculty of Languages and Literature, Minhaj University Lahore, Pakistan

Corresponding author:
Mr. Adnan Arif, PhD Scholar English Linguistics, Minhaj University Lahore, Pakistan
Email: 2023F-mulphd-eng-007@mul.edu.pk

1. Introduction

The interface between language documentation and corpus linguistics has assumed a prominent position in recent scholarly literature. Today, however, the documentation of a modern language is not merely marking the moment of recording on a phonetic device and archiving it for posterity; what is increasingly emphasized is the transformation of language data into analyzable resources, ones that are easy to reuse and search. Beyond the large written corpora, too, corpus linguistics has grown to now embrace high-quality transcribed and sequentially annotated spoken corpora—copies of the spoken text correlated with the audio file. This association plays a much bigger role for under-documented and low-resource speech varieties, in which there are recordings, but their lack of systematic documentation and accessibility prevents them from being used to their full potential.

Thus, forced alignment is a crucial methodology in speech corpus construction. Forced alignment is the procedure of connecting an audio recording with its written transcriptions and for tagging phonetic or word segments with time boundaries. This process makes it possible to convert raw audio data to a searchable corpus, offering the basis for further phonetic investigation as well as investigation at the lexical and discourse levels. Technological interventions in the creation of resources, in annotation and mobilization of resources are noted in the literature to be important, especially for communities with fewer corpora or digital systems, such as speech communities (Berez-Kroeker et al., 2023). But also corpus-based projects show that documentation collection is of limited use and that a documentation collection that is retrieved and temporally aligned, can be used to systematically retrieve, compare and analyze linguistic data (Janssen & Seifart, 2025; Paschen et al., 2020).

Shakargarh is a tehsil of the district Narowal in Punjab province of Pakistan, where the present study was done. Narowal is a border district and Shakargarh is one of the tehsils in it. The geopolitically and historically important area of Shakargarh Tehsil borders Pakistan and India and was historically a part of Gurdaspur before being separated from it during the partition and handed over to the Pakistani territory. Official sources further mention that Kartarpur is also an important locality of Shakargarh which further emphasizes the cultural and linguistic significance of this border area with complex historical development (District Narowal, Government of the Punjab, n.d.-a, n.d.-b, n.d.-c).

In the present research, Shakargarh was taken as a specific speech-community area besides its geographical position. The Punjabi spoken locally remains poorly represented in searchable corpus form. This scarcity creates scope for community-based research grounded in actual speech data rather than generalized descriptions of Punjabi. Such work is relevant to applied linguistics and corpus linguistics because it enables investigation of spoken language as it is used within a particular community. At the same time, local varieties may not correspond closely to existing speech models, pronunciation lexicons, or alignment systems developed for better-resourced languages.

The study therefore examined forced alignment in an under-documented Shakargarh Punjabi speech corpus. It first established the research problem, gap, objectives, and questions, and

then evaluated a completed pilot corpus through a mixed quantitative and qualitative workflow. Particular attention was given to the extent to which local pronunciation, dialect-specific vocabulary, connected speech, and pronunciation-lexicon refinement affected alignment performance and the documentary usability of the resulting corpus.

2. Research Problem

Low-resource and under-documented speech varieties are often recorded, but most of the time these recordings do not develop into searchable and analyzable corpora. As a result, useful speech data still remain difficult to access, segment, and analyze in a systematic way. Forced alignment provides an effective tool for matching speech recordings to transcripts at the word or segment level, which facilitates corpus building, phonetic research, and long-term language documentation. However, in low-resource settings, it tends to be less effective because of small datasets, natural speech recording conditions, and mismatches between local speech varieties and the models on which alignment tools rely. The case of the Shakargarh Punjabi variety is especially important because this variety is not sufficiently documented, and it has not been developed into a searchable corpus for linguistic research.

3. Research Gap

Forced alignment has been demonstrated to be an important tool for increasing the resources for under-resourced languages, such as in corpus compilation and speech annotations (Leinonen et al., 2022; Tosolini & Bowern, 2025). However, much of current literature is focused on the general aspect of low-resource or cross-linguistic settings – and not specifically on Shakargarh Punjabi. The researcher is not aware of any published work which has attempted to quantify the forced-alignment results on any Punjabi speech corpus from Shakargarh, or consider manual evaluation or lexicon refinement as methods that may help record the local variety. There is thus a gap between existing low-resource forced-alignment work and its empirically assessed use with this poorly-documented Punjabi-speaking community.

4. Research Objectives

- To identify the challenges affecting alignment accuracy in this under-documented Shakargarh Punjabi variety.
- To evaluate the usefulness of forced alignment for corpus development and language documentation of this selected variety of Punjabi.
- To assess the performance of forced alignment on a Shakargarh Punjabi speech corpus.

5. Research Questions

- What are the main challenges affecting alignment accuracy in this under-documented Shakargarh Punjabi variety?
- How effectively does forced alignment perform on a Shakargarh Punjabi speech corpus?
- How useful is forced alignment for corpus development and language documentation of Shakargarh Punjabi?

6. Literature Review

In the current study into language documentation, documentation is no longer taken to mean simply the collection of linguistic materials, but as a more inclusive process of recording, annotating, preserving, making accessible and re-using language materials. Berez-Kroeker et al. (2023) believe that in the last few years the process of writing a document has been closely linked to technologies of creating and mobilizing resources. This view is applicable because documentation is considered a process going back to the recorded language events and reconstructing them to make them useful for documentation. It applies directly to the present study, since the Shakargarh Punjabi corpus was designed not merely as a collection of recordings but as a searchable and analyzable speech resource. Documentation becomes especially valuable when linguistic data are organized for long-term academic and community use.

A similar trend can be observed in time-aligned documentation corpora. In their analysis of the DoReCo project, Paschen et al. (2020) illustrate how the data presented in language documentation can be converted into a time-aligned cross-linguistic reference corpus. It is also worth mentioning that their contribution demonstrates that the value of documentation corpora can be increased considerably when audio, transcription, and annotation are organized in a manner that facilitates retrieval and comparison. Despite the fact that DoReCo is a huge collaborative enterprise and the present research is conducted on a considerably smaller scale, the principle is exactly the same: unprocessed recordings can be much more useful once they are aligned and systematized. This observation justifies the need to study forced alignment as part of corpus development in Shakargarh Punjabi.

The question of searchability has also been further developed in another work by Janssen and Seifart (2025), who show that documentary corpora can be made searchable with the help of TEITOK implementation through the integration of audio, transcription, translation, and interlinear information. They are of particular relevance because their study emphasizes the role of accessibility and search functions in language documentation. This is to say that documentation deals not only with the preservation of speech, but also with making it retrievable and processable. In the case of the current study, the literature helps explain the importance of alignment; forced alignment is valuable not only because it saves time but also because it can transform a local collection of recordings into a more usable corpus that can be used in future analysis and documentation.

Recent research on forced alignment as such provides further support for its relevance. Leinonen et al. (2022) examine the application of semiautomatic speech alignment in under-resourced languages and find that alignment can indeed be used beneficially even in low-data situations. Their work is especially useful since it does not present alignment as a flawless or entirely automated solution; instead, it demonstrates that cross-language forced alignment can be improved when researchers combine alignment tools with domain knowledge and validation. This is especially applicable to the situation of an under-documented form such as Shakargarh Punjabi, where performance can be affected by gaps in transcription, vocabulary

peculiar to the dialect, and local phonological features. Their study, therefore, justifies the careful and critical use of forced alignment rather than unreflective technological application.

This argument is further reinforced by Ahn et al. (2024), who conducted a study of phone alignment in the case of Panãra. They show that the results of the alignment process may differ considerably depending on whether researchers approach it through language-specific methods or cross-linguistic methods. According to their findings, the structure of the phone inventory and the modeling strategy are important. This is particularly relevant to the present study because it implies that the effectiveness of forced alignment on Shakargarh Punjabi cannot be assumed in advance. Although a tool may be available, its outcomes will depend on the suitability of the acoustic and phonological assumptions of the model to the target variety. That is, forced alignment is not only a technical process, but also a linguistic issue.

A closely related insight is offered by Tosolini and Bowern (2025), when they compare multilingual and cross-lingual approaches using Montreal Forced Aligner for different low-resource related languages. Their findings indicate that a larger baseline model can still be applied conveniently to previously unseen languages, though they assert that performance has to be evaluated carefully. This investigation is particularly relevant in that it offers a direct answer to the question of the conditions under which forced alignment can be applied in low-resource settings. Most importantly, it is relevant to the current study as it indicates that the usefulness of forced alignment is limited not only by the software itself but also by the way models are trained, adapted, and evaluated in relation to real speech data.

Further evidence comes from current research on minority and non-dominant speech communities. Passmore et al. (2025) discuss forced alignment in two English-based Pacific Creoles and demonstrate that even in minority-language conditions, alignment may be helpful when tested critically against the linguistic and acoustic peculiarities of the corpus. Their contribution leads the discussion outside a particular language family and indicates that the issue is broad: although forced alignment can help streamline the process of corpus generation and annotation, the output should be approached with care in relation to the reality of the speech community. This directly relates to the case of Shakargarh Punjabi, where local variation and under-documentation suggest that alignment cannot be taken at face value.

There is also literature that helps to indicate a wider applied-linguistic relevance to the current study. In language pedagogy, pronunciation study, discourse analysis, and sociolinguistic description, searchable speech corpora can also be useful. This could then be followed by a corpus of Shakargarh Punjabi speech that would enable comparative analysis of various varieties of Punjabi, language maintenance in the local area, and the creation of future educational or descriptive materials. In this sense, the present study is not purely technical; it stands at the intersection of corpus linguistics, speech technology, and community-based applied linguistics, with the primary issue being how to make under-documented types of speech more visible and analyzable with the help of tools.

Collectively, the literature points to four key points. To start with, searchability, accessibility, and reuse are all becoming important in language documentation. Second, time-aligned corpora are extremely useful in comparison with unstructured recordings. Third, forced alignment has obvious potential in low-resource contexts, although the key to its success is the condition of the data, linguistic adaptation, and evaluation procedures. Fourth, regardless of the increased research in the area of low-resource alignment, no apparent study has investigated forced alignment in a Shakargarh Punjabi speech corpus. These findings are relevant to the current research in that they demonstrate the relevance of the present study in the context of modern research on corpus development and documentation technologies.

The study was guided by a documentation-as-reusable-resource perspective. Himmelmann (2006) conceptualizes language documentation as the creation of lasting multipurpose records of linguistic practices, while more recent work emphasizes accessibility, mobilization, and technological reuse (Berez-Kroeker et al., 2023). In corpus terms, this orientation means that documentary recordings gain additional research value when transcription, annotation, and temporal alignment make them searchable and comparable (Paschen et al., 2020). Forced alignment was consequently not considered to be an ultimate goal, but as a corpus-enabling procedure, which to be of benefit, had to enhance access to, and analysis of, authentic community speech. This theoretical orientation led to the sampling, the documentation value annotation and the error analysis and interpretation in the successive sections.

7. Research Methodology

The sequential mixed-methods design was used in this study where quantitative evaluation of alignment performance is followed by qualitative analysis of the error sources in linguistic and technical aspects. The quantitative strand measured successful and problematic utterances, out-of-vocabulary (OOV) instances, boundary deviation, segment mismatch, and change after pronunciation-lexicon refinement. The qualitative strand interpreted recurrent alignment problems in relation to dialect-specific vocabulary, connected speech, vowel reduction, overlap, recording noise, and transcription inconsistency. This design matched the research questions because the study required both measurable evaluation of the aligner and linguistically informed explanation of why particular forms succeeded or failed (Leinonen et al., 2022; Tosolini & Bowern, 2025). Figure 1 summarizes the workflow from community-based participant selection and recording through ELAN transcription, MFA alignment, Praat verification, error analysis, and assessment of documentation value. The figure also highlights that manual checking was integrated into, rather than separated from, the computational workflow.

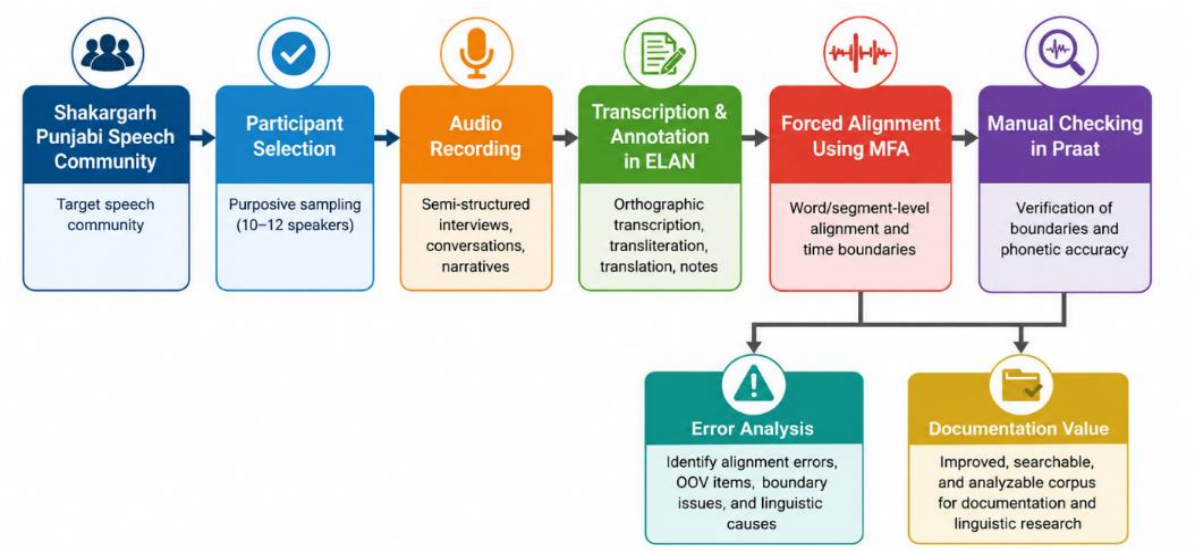


Figure 1

Research workflow for evaluating forced alignment in a Shakargarh Punjabi speech corpus

8. Sampling Method and Size

Non-probability purposive sampling was applied to draw the sample of native speakers of Adult Shakargarh Punjabi, who were long time residents of Shakargarh and areas close to it, in the Narowal District, and who regularly spoke the language. Six women and 6 men, 22-63 years old, and from both urban and rural areas were included in the sample. There were 24 recording files, each from two different speakers. The design was based on the central importance of detailed transcription, forced alignment, development of the pronunciation-lexicon and manual boundary verification, which led to the decision to keep the corpus at pilot scale. Signal wording was assessed on 120 sentences split evenly into semi-structured interviews, natural conversation and personal accounts (40 utterances per speech context). In documentation with limited resources, targeted sampling can be justified when the size of the sample depends either on the lack of a large, consistent corpus, or on the lack of consistent transcription and verification of a large corpus. (Berez-Kroeker et al., 2023; Leinonen et al., 2022).

9. Data Collection Tools

Primary spoken data was gathered using semi-structured interviews, short personal narratives and natural conversations. The use of this combination not only enabled the guided speech but also relatively spontaneous speech so that forced alignment was able to be evaluated under different speech condition. Speaker code, age group, gender, residence, speech context and recording conditions were documented on a participant information sheet and a consent form and on a metadata sheet. All recordings were transferred to WAV format to retain best quality of recording for transcription and alignment. Time-aligned multi-tier transcription/annotation of ELAN was conducted (Max Planck Institute for Psycholinguistics, 2026). Montreal Forced Aligner (MFA), a trainable alignment system based on Kaldi, was used to align audio with transcripts (McAuliffe et al., 2017). Manual inspection of word and

segment boundaries and verification of decisions regarding phonetic alignment were carried out using Praat (Boersma & Weenink, 2026). These tools facilitated both corpus compilation and alignment, verification and later, error analysis

10.Data Analysis Plan

The analysis combined quantitative descriptive statistics with qualitative error analysis. The 120 evaluated utterances represented 19.79 minutes of speech and 1,565 running words. Shahmukhi transcripts were paired with standardized Roman transliteration, and a project-specific pronunciation lexicon was developed for local forms that were absent or inadequately represented in existing resources. MFA output was manually inspected in Praat. An utterance was classified as successful when the absolute boundary deviation was 50 ms or less and no segment mismatch remained; an utterance was classified as problematic when the boundary deviation exceeded 50 ms and/or a segment mismatch was present. This criterion was applied consistently before and after lexicon refinement. The successful utterances, problematic utterances, frequency of OOVs, rate of OOVs, types of residual errors, and success levels of utterances per speech context were calculated through quantitative analysis. Finally, the qualitative stage translated the linguistic sources of the error marks left out secondary (dialect variation, under-completion, connected speech, overlap, background noise, transcript-audio mismatch) into English. This procedure ensured that performance was numerically captured, while maintaining the need to keep the linguistic interpretation for corpus-based tasks in low resource settings (Ahn et al., 2024; Leinonen et al., 2022; Tosolini & Bowern, 2025).

11.Ethical Considerations

Ethical guidelines for conducting community based linguistic research were followed. Before recording, the participants were told the reason of the research and informed consent was obtained both in written form and verbally. Participants chose to participate, and confidentiality was ensured without using names and instead using speaker codes. The audio files, transcripts, metadata and annotation files were maintained securely, and only utilized for suitable and legitimate academic purposes. Local forms of speech were expressed carefully and verified with speakers, if needed, in order to avoid arbitrating the interpretations made on the data gathered from the community into a standardized form. Both proper citation and original analysis and responsible archiving was observed throughout the study. Both the attention given to documenting with care and the location-by-sort system of controlled access and responsible reuse with others align with the current practice of language documentation (Berez-Kroeker et al., 2023).

12.Data Analysis

12.1. Overview of Data Analyzed

A total of 12 speakers took part in the completed pilot, with 24 recordings and 120 evaluated utterances of a total duration of 19.79 mins and 1,565 words. The source of the data was equally from semi structured interviews and for short personal narratives from natural conversation. Table of the material used was transcribed in Shahmukhi Punjabi and subsequently translated into Roman characters. Dialect-specific vocabulary, reduced/non-standard pronunciation, connected speech, overlap and other features typical of spontaneous speech were included in the dataset to assess the forced alignment in this low-resource environment. After developing a community-specific pronunciation-lexicon, the experts reviewed the model generating the local pronunciation-lexicon and made adjustments to its pronunciation rules. The aligned result of the refined pass after community-specific pronunciation-lexicon development was compared with the result of the initial pass, enabling the effectiveness of local community-specific pronunciation-lexicon development on alignment performance to be evaluated.

A sample excerpt from the corpus illustrates the linguistic nature of the collected data:

”اوہ پائی، کیہہ کرے کرنا ایں یار؟ ایتھے کون کسے دا ہُندا اے؟ کوئی کسے دا نئیں، پا جی۔“

Romanized form:

“Oh paai, keeh karay karna ain yaar? Aithay kaun kisay da hunda ae? Koi kisay da nai, paa ji.”

The excerpt represents naturally occurring Shakargarh Punjabi rather than standardized Punjabi and therefore provides a realistic test of alignment performance. The recurring manuscript forms were standardized as پائی / paai, ایں / ain, نئیں / nai, and نِکھاکھرا / nikhakhra so that the same lexical items could be traced consistently across the analysis, findings, discussion, and conclusion.

12.2. Method of Analysis

Analytical framework was a mix between the quantitative descriptive analysis and the qualitative error analysis. The quantitative measures consisted of: utterance-level alignment status, absolute boundary deviation, segment mismatch, OOV frequency, and performance before and after pronunciation-lexicon refinement. A qualitative analysis was used in order to determine and explain the linguistic and technical sources of the problematic alignments observed and the effect of local phonology, vocabulary, pace and talk-in on output. A numerical accuracy rating by itself would not capture the reason for failure or success of an aligner with an under-documented variety.

Past low-resource alignment research also indicates that the quality of the alignment could vary considerably with certain factors; for example, due to dialect variation, pronunciation differences, phone inventories, or spontaneous speech conditions (Ahn et al., 2024; Leinonen et al., 2022; Tosolini & Bowern, 2025). Descriptive statistics and linguistic interpretation were used, however, as complementary elements of the evidence in the present analysis.

12.3. Analytical Procedure

The analysis was done in three interrelated steps. Initially, recordings were sorted and made into WAV format. The transcription of each utterance was done manually in Shahmukhi Punjabi script and a standard Roman transliteration was prepared for pronunciation tracking and for constructing the lexicon. The Utterance boundaries, speaker turns were then set up for time-aligned annotation in ELAN.

Secondly, a project specific pronunciation lexicon was created and evolved from common locally recurrent forms. Special consideration was paid to words that were differently pronounced or differently spelt in relation to the standard, such as پائی / paai, ایں / ain, نئیں / nai, نِکھاکھرا / nikhakhra, ایتھے / aithay, and combinations like اَپنا تن ای / apna tan ee. The local lexical entries included in the working lexicon were verified with the participation of native speakers.

These kinds of forms directly tested for the effect of local lexical alignment. For instance, the item نِکھاکھرا / nikhakhra was sensitive to OOV, and the connected phrase اَپنا تن ای / apna tan ee was sensitive both in terms of the lexicon and reduction. Thus the lexicon was a controlled version of an adaptation layer on top of the community transcription than replacement for the original speech data onto the general purpose aligner.

Thirdly, the corpus was processed using MFA and the boundaries checked in Praat. Absolute boundary deviation and segment mismatch was noted for every single utterance. Critically one segment mismatch, OR a boundary deviation of more than 50 ms, was considered problematic for a successful alignment. This operational definition enabled the manual process of verification to be repeated and provided for comparison of initial and refined alignment passes directly.

The quantitative analysis tabulated successful and problematic utterances, OOV occurrences, rate of incorrect OOV per 100 words, and error categories that were not explained by the OOV categories. This was followed by a qualitative analysis of the language sources of residual problems. Of particular interest was the handling of dialect-specific/out-of-vocabulary items, boundary ambiguity of connected speech, vowel reduction/non-standard pronunciation, overlapping speech, background noise, and transcription inconsistency.

The corpus included successful and problematic micro-level examples. In UTT007, the sequence of connected words (apna tan ee kam aanda ae) moved from an absolute BD of 135 ms to a final BD of 14 ms after lexicon refinement without any remaining mismatch and hence was classified as successful. In contrast, UTT008, bearing the sequence paani neeviñ thaañ ee charrhda ae, had an issue that arose even after the process of refinement: the boundary deviation went up from 59 to 90ms while a segment mismatch remained. The corresponding Shahmukhi forms were then kept in the corpus annotation and verified against the original audio. The reasons why numerical thresholds were needed to determine alignment, along with linguistic interpretation of local connected speech, are revealed in these paired examples.

12.4. Reliability and Trustworthiness

Repeat listening, standardized Roman forms, the systematic use of the 50-ms boundary criterion and transcript checking were used for reliability. The output for each utterance was examined using Praat, and the diagnoses of errors made were analyzed with regard to the audio material, the transcripts and the local pronunciation evidence. The consistency of status assignment was done both prior to and following the lexicon refinement process, minimizing the likelihood that improvements were due to variable manual rulings.

The trustworthiness was still further supported by comparing the Shahmukhi with the Romanized forms, and by having the forms used by native speakers confirmed within the projects. The interpretive checking process was still needed because the material in the corpus was community speech that was not put out in a controlled laboratory setting, but for expressions of local meaning or discourse forms or reductions.

12.5. Ethical Handling During Analysis

Ethical procedures outlined in section 11 were adhered to during analysis and archiving, with codes for speakers retained in all of the analysis files while no personally identifiable information was included in the examples.

13.Findings

13.1. Overview of Findings

Forced alignment resulted in a measurable improvement following pronunciation-lexicon refinement. A total of 74 utterances of a possible 120 were successful (61.7%) and 46 problematical (38.3%) during the initial pass for MFA. Following the refinement process 96 utterances (80.0%) fulfilled the success criterion while 24 utterances (20.0%) represented problematic ones (18.3 percentage points more fulfilling the success criterion). The number of OOV went down from 62 to 18 and the rate of OOV went down from 3.96 to 1.15 number of OOV per 100 words (down by 71.0%). The results demonstrate that local lexicon refinement substantially enhanced corpus-level alignment, but some linguistic and acoustic issues still remain. Initial and refined performance measures are summarized in Table 1.

Table 1. Forced-alignment performance before and after pronunciation-lexicon refinement

Indicator	Initial	After refinement	Change
Successful utterances	74 (61.7%)	96 (80.0%)	+18.3 percentage points
Problematic utterances	46 (38.3%)	24 (20.0%)	−18.3 percentage points
OOV instances	62	18	−71.0%
OOV rate per 100 words	3.96	1.15	−2.81

13.2. Findings Related to Objective 1

The goal of the first objective was to determine the main issues that impact alignment accuracy. After the refinement process, 24 utterances remained problematic, of which the most frequent error types were dialect-specific/OOV items (29.2%) and connected-speech boundary ambiguity (29.2%). Five (20.8%) were listed as due to vowel reduction and/or non-standard pronunciation, three (12.5%) as due to overlapping speech, and one (4.2%) each due to either background noise or transcription inconsistency. This distribution illustrates that the other errors were not caused by a lack of one specific technical skill only, but by a

combination of local lexical structure, spontaneous speech and recording conditions. The distribution of the residual is given in Table 2.

Table 2. Residual primary misalignment errors after lexicon refinement

Primary error type	Frequency	% of problematic utterances
Dialect-specific/OOV item	7	29.2%
Connected speech/boundary ambiguity	7	29.2%
Vowel reduction/non-standard pronunciation	5	20.8%
Overlapping speech	3	12.5%
Background noise	1	4.2%
Transcription inconsistency	1	4.2%

These are exemplified in the transcribed examples. Alignments like پائی / paai, ایں / ain, نئیں / nai, and نِکھاکھرا / nikhakhra were mentioned several times in connection with OOV-sensitive and/or dialect-sensitive alignments. Similarly, connected sequences like اَپنا تن ای کم آندا اے / apna tan ee kam aanda ae also exhibited reduction and compression of boundaries; rapid delivery in conversations was also complicating boundaries, as was overlap. The patterns account for the failure of the pronunciation-lexicon refinement to account for many occurrences while still maintaining errors due to connected speech and acoustic competition.

13.3. Findings Related to Objective 2

Overall performance and variation by speech context were evaluated in the second objective. Semi-structured interview utterances had the highest success rate at 90.0% (36/40) followed by personal utterances 85.0% (34/40). Only 65.0% of natural-conversation utterances were successfully aligned (26/40), indicating that natural conversation was quite difficult to align. Qualitative observation testifies to the fact that spontaneous, connected speech with varying pacing and overlap etc., decreases the clarity of boundaries; guided or more carefully paced speech increases the clarity. The refined performance by speech context is detailed in table 3.

Table 3. Refined alignment performance by speech context

Speech context	Evaluated n	Successful n	Success rate
Semi-structured interview	40	36	90.0%
Natural conversation	40	26	65.0%
Personal narrative	40	34	85.0%

The impact of lexicon refinement could also be observed at utterance level. UTT007 initially had a problem with boundary deviation but after refinement it was reduced to 14ms from 135ms and was classified as successfully aligned. On the other hand, UTT008 was still a problem at 90 ms, and exhibited a segment mismatch, indicating that lexical adaptation can help resolve certain local-form issues, but not fully offset connected-speech and acoustic complexity. Manual checking in Praat was therefore necessary to identify true improvement from residual alignment errors.

13.4. Findings Related to Objective 3

The third objective explored usefulness for corpus development and for language documentation. The rise in percentage of successful utterances from 61.7% to 80.0% and the 71.0% decrease in the number of OOV instances was evidence that forced alignment materially facilitated the ability to organize speech into searchable time-aligned units. All the other 20.0% problematic utterances were also analytically productive since errors could be explicitly written down, rather than simply left as "alignment" output.

So it created more than automated segmentation workflow. It established a structured record of the audiotapes, the Shahmukhi transcription, the Roman transliteration of the text, entries of local pronunciation of the text, alignment status, and diagnosis of the error. This hierarchical structure of organization contributes to making the corpus useful in future phonetic, lexical, morphosyntactic and sociolinguistic studies, at the same time maintaining the relationship with the documentary record of authentic community discourse.

14.Discussion

The results illustrate that forced alignment can be useful to develop an under-documented Shakargarh Punjabi speech corpus, however, the performance of forced alignment is strongly dependent on linguistically and speech-contextually sensitive factors. The results of successful utterances indicate that the quality of alignment was not determined by the quality of the baseline model, but by the fact that lexicon was refined as well, the ratio of successful utterances has increased by 18.3 percentage points after that process. Local pronunciation and the use of local vocabulary explicitly mattered, and there appears to have been an additional difficulty with connected and spontaneous speech which is indicated by a lower success rate for natural conversation (65.0%) than for interviews (90.0%) or narratives (85.0%).

One of the very obvious reasons for difficulty was the use of dialect specific words. Even with general resources of pronunciation, forms like پائی / paai, ایں / ain, نئیں / nai, and نِکھاکھرا / nikhakhra were not consistently represented and dialect-specific/OOV items remained high (29.2%). Characteristic of adaptation and validation in under-resourced alignment (Leinonen et al., 2022), and in line with Tosolini and Bowern (2025), who demonstrate that multilingual models might still need to be language- or variety-adjusted. The reduction in the number of OOV instances (from 62 to 18) offers an immediate, quantifiable proof of that interpretation.

Connected speech was also crucial. Another 29.2% of the remaining problematic utterances were due to connected-speech boundary ambiguity, with vowel reduction/non-standard pronunciation making up 20.8%. This is consistent with the results of Ahn et al. (2024), who find that the phonological structure and the strategy of modeling influence the results of alignment. The present results in a context-specific context support this, as even native speech, which was processed with lexicon refinement, proved to be more difficult than guided speech to make an appropriate alignment of its boundaries, which suggests that beyond lexicon refinement there are still other aspects of boundary ambiguity that need to be resolved.

Forced alignment, despite these constraints, was undoubtedly documentary. Four of the five evaluated utterances fulfilled the operational success criterion, providing a considerably larger amount of material that also can be searched, displayed, and used in a time-aligned manner. This is part of a larger argument in support of greater long-term usability of documentary resources through aligned corpora (Berez-Kroeker et al., 2023; Paschen et al., 2020). Thus, we can see that forced alignment was not just a useful technical tool, but also a production process as part of the construction of a corpus that allowed for an increase in the amount of recorded community speech.

The results also make it clear the pronunciation lexicon's role and manual verification. Lexical refinement, specific to the project, lowered the OOV rate by 71.0%, but seven utterances (29.2%) remained problematic because of OOV items, which means that a fully automatic workflow would have swallowed important remaining errors. These, therefore, had to be combined for methodologically equivalent use, as was the combination of local lexicon development and Praat verification. By following a trainable approach for creating alignments for the popular MFA framework, this recent pilot shows why it is essential for users to adapt to their local domain and not rely on a given “baseline” dictionary.

The study also affirms the connection of the corpus linguistics and language documentation. Language documentation aims at long-term, multi-purpose documentation of practices while corpus linguistics brings systematic organization, retrieval and analysis (Himmelmann, 2006). These objectives were partially met by using forced alignment in order to convert the oral recordings of Shakargarh Punjabi into structured, time referenced corpus data, but retaining the connection to the original recordings (oral baseline). This is especially true of a local variant, which is largely spoken and which has hardly appeared in retrievable digital materials.

Overall, the three research objectives were met. The study established the main alignment error sources, provided quantitative results of the performance with and without the pronunciation-lexicon refinements, and presented the usefulness of the documents by considering a semi-automatic workflow. The evidence suggests forced alignment is effective, but it requires a combination of local linguistic knowledge, consistent transcription, explicit errors and manual confirmation to obtain the best results.

15.Conclusion

Based on this study the conclusion is drawn that with local knowledge of the Punjabi pronunciation as well as manual validation, forced alignment could play a valuable role in the development of a searchable Shakargarh Punjabi speech corpus. After pronunciation-lexicon refinement, the successful alignment rate jumped to 80.0% from 61.7%, and the number of OOVs was reduced by 71.0%. However, it is not a purely automatic process as evidenced by residual errors involving dialect-specific forms, connected speech, vowel reduction and overlap. The most effective method was thus to have semi-automatic procedures (Shahmukhi transcription and standardized Roman transliteration based on the MFA procedures and project-specific lexicon) followed by verification carried out by Praat software, with a 50-ms boundary criterion. This workflow resulted in a more systematic, searchable and reusable corpus resource from a pilot collection of community speech, and offers an empirical foundation for low-resource Punjabi documentation.

16.Study Limitations

The study had its limitations on the pilot level – 12 speakers, 24 recordings, 120 evaluated utterances, 19.79 minutes of speech. Though this data set was adequate to make controlled comparison between initial and refined alignment, the results should not be generalized for all

speech situations and speakers of Shakargarh Punjabi. The analysis also relied on manual transcription, Roman transliteration, as well as verification by researchers and online alignment tools were not available for Shakargarh Punjabi. Such restrictions are crucial since performance could differ in the case of larger corpora, another, more variable set of speakers, or different recording settings.

17.Recommendations and Future Directions

Especially future studies should include more speakers, age groups, villages and interactional situations, with the same explicit verification criteria. Project-specific pronunciation lexicon should be expanded by ongoing native-speaker validation along with systematic notation of pronunciation variations. Comparative testing with other acoustic or multilingual alignment model hypotheses would be useful to explore whether the achievements shown here are unique to the MFA pipeline discussed here or whether they are generalizable to other systems. An aligned corpus, with more data, can then inform phonetic analyzes, lexical profiling, morphosyntactic analyzes, sociolinguistic variation analyzes, and community-oriented digital documentation of Shakargarh Punjabi.

References

Ahn, E. P., Chodroff, E., Lapierre, M., & Levow, G.-A. (2024). The use of phone categories and cross-language modeling for phone alignment of Panãra. In Proceedings of Interspeech 2024 (pp. 1505–1509). ISCA.

Berez-Kroeker, A. L., Gabber, S., & Slayton, A. (2023). Recent advances in technologies for resource creation and mobilization in language documentation. Annual Review of Linguistics, 9, 195–214.

Boersma, P., & Weenink, D. (2026). Praat: Doing phonetics by computer [Computer software]. University of Amsterdam.

District Narowal, Government of the Punjab. (n.d.-a). District profile.

District Narowal, Government of the Punjab. (n.d.-b). Our history.

District Narowal, Government of the Punjab. (n.d.-c). Important places.

Himmelmann, N. P. (2006). Language documentation: What is it and what is it good for? In J. Gippert, N. P. Himmelmann, & U. Mosel (Eds.), Essentials of language documentation (pp. 1–30). Mouton de Gruyter.

Janssen, M., & Seifart, F. (2025). Searchable language documentation corpora: DoReCo meets TEITOK. In É. Le Ferrand, E. Klyachko, A. Postnikova, T. Shavrina, O. Serikov, E. Voloshina, & E. Vylomova (Eds.), Proceedings of the Fourth Workshop on NLP Applications to Field Linguistics (pp. 58–64). Association for Computational Linguistics.

Leinonen, J., Partanen, N., Virpioja, S., & Kurimo, M. (2022). Semiautomatic speech alignment for under-resourced languages. In A. K. Ojha, S. Ahmadi, C.-H. Liu, & J. P. McCrae (Eds.), Proceedings of the Workshop on Resources and Technologies for Indigenous, Endangered and Lesser-Resourced Languages in Eurasia within the 13th Language Resources and Evaluation Conference (pp. 17–21). European Language Resources Association.

Max Planck Institute for Psycholinguistics. (2026). ELAN (Version 7.1) [Computer software]. Author.

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. Proceedings of Interspeech 2017, 498–502. https://doi.org/10.21437/Interspeech.2017-1386

Paschen, L., Delafontaine, F., Draxler, C., Fuchs, S., Stave, M., & Seifart, F. (2020). Building a time-aligned cross-linguistic reference corpus from language documentation data (DoReCo). In Proceedings of the Twelfth Language Resources and Evaluation Conference (pp. 2657–2666). European Language Resources Association.

Passmore, S., Roque, L. S., Gillespie, K., Nath, S. K., Davey, K., Mullan, K., Cawley, T., Biggs, J., Billington, R., Evans, B., Thieberger, N., Evans, N., & Barth, D. (2025). English-based acoustic models perform well in the forced alignment of two English-based Pacific Creoles. In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 31172–31183). Association for Computational Linguistics.

Tosolini, A., & Bowern, C. (2025). Multilingual MFA: Forced alignment on low-resource related languages. In J. Lachler, G. Agyapong, A. Arppe, S. Moeller, A. Chaudhary, S. Rijhwani, & D. Rosenblum (Eds.), Proceedings of the Eighth Workshop on the Use of Computational Methods in the Study of Endangered Languages (pp. 100–109). Association for Computational Linguistics.

PAKISTAN JOURNAL OF LANGUAGE AND LITERARY STUDIES
Volume 2, Issue 1 (Spring 2026)
Pages: 68-80

Analyzing Social Disparity: A Critical Discourse Analysis of Hosseini's And the Mountains Echoed

Arshad Hussain[1]

ABSTRACT

Hosseini's fiction deals with the themes of displacement, human sufferings and inequality which make his narratives the valuable texts for analyzing linguistic choices. Although a few studies have been conducted to explore his texts linguistically, yet certain areas like CDA are still unexplored. This study aims at investigating how language has been exploited and manipulated to portray social disparity in Hosseini's And the Mountains Echoed. The selected text is explored through the lens of Fairclough's Tri-Dimensional Model (1989) and Ridgeway's (2014) Theory of Social Inequality to delineate three different stages for analyzing a text, each facilitating a systemic evaluation, analysis, and interpretation of the text. The first stage, description, engages with the grammatical level of text, offering a deep understanding of text's lexical, structural and syntactic features. The second stage, interpretation, focuses on to provide an analytical acumen in interpreting the text according to its linguistic feature and author's ideology behind constructing it. The third stage, explanation, encompasses an extensive discipline of explanation that considers the social norms and ideologies aligned with the text. The results of the study reveal that the author has highlighted the dominant issues of society like gender inequality, economic inequality, ideological inequality, status inequality and healthcare inequality. Furthermore, it also exposes the reasons behind inequality such as gender differences and social stratification. The study seeks to present a meaningful understanding about how language can be exploited and manipulated to construct power and ideology.

Keywords: *Critical Discourse Analysis, Social Disparity, Tri-Dimensional Model, Social Inequality Theory, And the Mountains Echoed, Khaled Hosseini*

[1] MPhil scholar, Department of English, University of Sargodha, Sargodha

Corresponding author:
Mr. Arshad Hussain, MPhil scholar, Department of English, University of Sargodha, Sargodha
Email: engl62bs21r007@gmail.com

1. Introduction

Social disparity remains a significant social concern, as unequal access to resources, opportunities, and social privileges can shape individuals' experiences and relationships within society. These inequalities have been explored by numerous scholars and literary writers who examine their effects on human lives and social structures. Khaled Hosseini, an Afghan-born American novelist, is among the writers whose works explore social inequalities and their consequences for individuals and communities. Through his literary narratives, Hosseini highlights the ways in which differences in social status, wealth, and privilege influence human relationships and individual experiences. This critical discourse based study has been conducted to examine the linguistic exploitation in Hosseini's And the Mountains Echoed. The main focus of the study is to investigate how language has been used to portray different characters belonging to different gender or social class. In order to draw a qualitative descriptive analysis, the study adopts Hosseini's And the Mountains Echoed as primary data and utilizes Fairclough's Three Dimensional Model (1989) as methodological framework in addition to Ridgeway's Social Inequality Theory (2014). Findings of the research reveal that the text is overloaded with the ideologies related to social and gender disparities.

2. Research Questions

The study focuses on to answer the following questions:

- How does the writer construct the text to portray the theme of social disparity?
- Which kind of explanation of above mentioned theme can be given through CDA?
- What ideologies does the author portray in the text?

3. Research Objectives

The study focuses on to answer the following questions:

- To explore how the given text has been constructed to portray the theme of social inequality.
- To uncover the hidden ideology of the author behind constructing the text that offers diversity of socio-cultural identities.
- To draw an interpretation of bracketed theme contextualized in the text.

4. Methodology

The study focuses on examining the text through Critical Discourse Analysis (CDA) perspective in order to find out the embedded theme of social inequality. Therefore, the study employs Fairclough's Tri-Dimensional model from the broad field of Critical Discourse Analysis (CDA) as its theoretical framework, alongside inequality theory, with purpose of analyzing and examining social inequality within the text. Moreover, in order to support the arguments related to social inequality, Ridgeway's ideas about inequality have been employed. Cecilia L. Ridgeway is an American psychologist and a professor of social sciences. Her fields of interest are to underscore the dominant issues like inequality in society, gender issues and status problems.

5. Theoretical Framework

In this study, Fairclough's Three Dimensional Model and Ridgeway's Social Inequality have been used a theoretical frame work to analyze and explore the selected text. Fairclough is a prominent scholar among the pioneers of Critical Discourse Analysis. His works are notable as theoretical frameworks for researchers of CDA. His Tri-Dimensional model is much popular among the researchers. Tri-Dimensional model was introduced in his work Language and Power (1989). In this book, the author discusses discourse and its relationship with power, identity, social practices and ideology. Furthermore, a comprehensive theoretical framework has been developed in the book for analyzing and interpreting any discourse for exploring embedded power relations, identity and ideology in the language. The model provides three dimensions of the text for interpretation.

- **Description**: At the level of description the model suggests to explore linguistic features, grammar and vocabulary of the text at micro level.
- **Interpretation**: At the second level of analysis, the model focuses to investigate how a discourse can be produced, distributed, consumed and interpreted.
- **Explanation**: This dimension of the model aims to uncover the context of the discourse at broader level. It looks how discourse shapes or shaped by social and power relations

5.1. Social Inequality

Fitria and Danu (2024) explain that social inequality can occur due to gender, race and economic imbalance among people. Ridgeway (2014) critically examines the causes and effects of social inequality and concludes that along with numerous other reasons, social status is also a crucial factor behind social inequality. She focuses on how status matters and effects social inequality. In her views, status is the main cause of social inequality.

6. Literature Review

In order to achieve a fruitful analysis of the study, previous literature has been explored and reviewed methodologically. Bloor and Bloor (2013) investigated the practice and study of Critical Discourse Analysis. They argued that this practice mainly focused on exploring social goals and their link with language and discourse. Furthermore, they elaborated that it provides numerous methods, techniques, and tools in order to explore social relation, societal structures, human psychology, cognition and behavior. Similarly, Ahmad and Khan (2020) conducted a study to explore the ideology embedded in Hosseini's And the Mountains Echoed which challenges and opposes the norms of patriarchal society. Additionally, they highlighted the ideology of liberal feminism through the character of Nila Wahdati. The study revealed how she resisted her father and opposed the idea of male dominancy. It also revealed how the author opposed the norms and tradition of the patriarchal society. The cited study was conducted through Critical Stylistic perspective with the purpose of exploring how women showed their power through the use of language. Linguistics tools like noun, pronoun, modality and use of clauses were explored to find out the clashes between feminist and patriarchal ideologies.

Fareed et al. (2022) argued that the text has been composed for justifying and promoting American invasion in Afghanistan. The researchers elucidated that the author portrayed Afghanistan as a deserted and abandon place where western intervention is necessary to trigger novelty and independence in the society. Through qualitative subjective analysis, the study explores that the author represented Afghan society and Islam as pitiable and miserable. In this way, he made the way for Western scholars to portray an erroneous identity of the culture and religion of Afghanistan. The study argued that the depicted Afghan society in the text is totally different from that of the real one and the author's representation is based on biasness.

Gender roles and power dynamics are commonly observed across various societies. Literature often highlights the social issues embedded within these structures and helps to expose the so-called civilized façade of seemingly well-developed societies. In this regard, Christopher et al. (2018) explored and analyzed the power relations and gender issues in Hosseini's And the Mountains Echoed. The study focused on to examine how gender role and power were embedded in the text through the lens of patriarchy and masculinity. The study revealed how male dominancy and female limitations were implanted in the text. The results of the study exposed that patriarchy has strong impact on both men and women and it influences the gender role in the given society.

Hosseini's literature not only encapsulates human problems but is also relates humanity to ecosystem in order to convey aesthetic sense. Shah & Kumari (2019) conducted a study to examine the ecocritical perspective of Hosseini's And the Mountains Echoed. It is the study of relationship between literature and concrete environment. The researcher utilized Glotfelty's views about Ecocriticism, in order to find out how the author has connected human beings with natural environment. Their study described Hosseini as a representative of Afghan society with its relation to nature and in his works he depicts the ecological issues. The representation of Oak tree and its pain has been made so realistically as if writer himself is feeling the pain of the tree. Furthermore the author also highlights the environmental problems faced by the people of Afghan society. The text starts with an allegory that narrates the story of natural effects on humans and throughout the plot different symbols of natural influence on human beings have been shown. The novel also represents the liberal feminist and patriarchal ideologies.

Literature always explores various human problems and inequality is one of them. Multiple studies have been conducted to analyze such literature. Difa & Tarihoran (2020) conducted qualitative descriptive study to explore inequality in The Adventures of Tom Sawyer to analyze how principal characters strive to wipe out social inequality. Structuralism theory has been used as theoretical framework in order to analyze the text objectively, independently and separately from its context. It has been explored that the main character of the novel Tom adventurously struggles for equality and he does not discriminate people. Moreover, the researchers also focused on to explore which type of social inequality is found the text and came up with the findings that inequality based on race, social strata and slavery is found in the text.

Likewise, another study has been conducted to examine A Thousand Splendid Suns, through Critical Discourse Analysis perspective, in order to investigate the hidden powers and ideologies behind discourse (Khan, 2016). The researcher employs Fairclough's three dimensional model from the field of Critical Discourse Analysis for examining and analyzing discourses. Furthermore, he also adopts Feminist Critical Discourse Analysis perspective with the purpose of interpreting power exercise against women. The analysis has been conducted through socio-political and socio-cultural milieu of Afghanistan. The study sums up the analysis with the results that power depends upon the context. It is not merely the matter of gender but it can be associated with wealth or political power. The text reveals that Taliban have the power and dominancy upon other people while Rasheed has power and tries to suppress his wives. At the end, his wives resist and take control over power.

Marginalization of women is also a common problem in different societies which has always been unveiled by various writers to highlight this critical issue. Shakeel and Shaheen (2024) examined the translated texts of Khadija Mastur The Women's Courtyard (2018) and A Promised Land (2019) through the lens of Fairclough's three dimensional model. Their study uncovered the way female characters of the novels resist typical female role and challenge the stereotypical norms of the society through language choices. Furthermore, use of language is considered the bedrock of social interaction which unveils the social status of different characters. In this regard certain studies have been explored. Firstly, Shafak's Three Daughters of Eve has been examined through the lens of Fairclough's three dimensional model. The study aims at exploring how the linguistic choices of different characters reveal their religion, identity and the power within the text (Bilal et al., 2021). Secondly, Hadi's The Black Eyes has also been critically analyzed through the lens of Fairclough's three dimensional model as theoretical framework in order to investigate the role they have to play in the society and the hurdles they have to face (Rahim, 2024). All the above discussed studies reveal two relating aspects of this particular study. The first one is that various studies have been conducted to explore nearly similar objectives by using similar model, but in different texts. Secondly, it has been explored that up till now, no study has been conducted to this particular text using same model to explore same objectives.

7. Analysis and Discussion

The study utilizes six excerpts chosen from the text in order to analyze how language has been exploited to portray the theme of social inequality. From the six elected excerpts, four have been analyzed comparatively in form of pairs. Excerpts 3, 4, 5, and 6 have been respectively analyzed in pairs which represent status disparity. While the other two excerpts 1 and 2 have been analyzed separately and they represent gender disparity.

7.1. Excerpt 1

"[…] not accustomed to talk back […] had rows […] what I wear, […] went, […] said, how I said it, who I said it to. I had turned bold and adventurous, […] he […] ascetic and emotionally austere. We […] natural opponents […]" (Hosseini, 2013, p. 119).

7.1.1. Description

This excerpt underscores a clash between a conservative father and a modern, bold minded daughter who has rebellious nature and stands against the stereotypical customs of society. She is bold, brave and strong enough to stand against patriarchy and male chauvinism. The lines are articulated by Nila Wahdati and she is elucidating the ups and downs of her relationships with her father. She describes that her father had aggressive nature and was not accustomed to listen talk back instead he only used to speak in authoritative and orderly manners. Pronouns like 'he", and 'I' are used which represents different identities. 'He' is used by the narrator for her father which shows her deficient relations with her father while 'I' is used to portray empowerment and independency. Logical Connectors "Over what", "where", "how" have been used to develop a strong cohesive relation and sequence in the language of the excerpt. Ideological words such as "bold" and "adventurous" are used to describe the novelty and modernization of the society in which a female character is portrayed as bold and adventurous unlike to the typical traditions of the society. Mode and tone: Declarative mode of the narration makes the tone descriptive and informing. Tense: Past tense has been used to describe that in past there were conflicts and clashes between stereotypical and modern people. Syntactic structure: Compound sentences reveal the empowered and independent nature of the character

7.1.2. Interpretation and explanation

The extracted part of the text is ideologically loaded with the problematic condition of gender discrimination and male chauvinism in the society. It proclaims that when the main character of the novel, Nila Wahdati, was young, her father divorced her mother and forced her to live with him but she did not possess sufficient capability to put up with his father's ferocious behavior. Consequently, she stood against his will and started to retaliate against societal barriers imposed on her. Moreover, she explains how her father was careful about her activities. He craved to be informed about everything she did or said and who to say it. On the contrary, she did not expect any interference in her personal life. Metaphorically, she represents the new generation which is striving for liberty and personal freedom in congested and stereotypical society. In addition, the implementation of ideologically loaded words like, 'bold' and 'adventurous' represent her powerful and strong will. These words are usually associated with men but in this particular context the narrator portrays the character of Nila as a bold and adventurous woman. On the other hand, she uses the words 'ascetic' and 'austere' for her father which represent his aggressive and stern behavior. It also depicts certain negative aspect of men and portrays them as inexorable and impatient. Deployment of pronominal reference is another aspect of the text which portrays the authoritative and confident modality of the speaker. She is frequently using pronoun 'I' which depicts her authority, confidence, boldness and empowerment. Moreover, instead of saying 'my father'

or according to social and cultural ideology ‘baba’, she employs third person pronoun “he” for her father. It represents her hidden hatred for her father. Furthermore, the employment of advance vocabulary characterizes her character strong willed, and socially and economically powerful. Words like, ‘ascetic’ and ‘austere’ represent the social status of Nila Wahdati that she is bold and strong enough to use advance and complex vocabulary as compared to the other socially lower characters of the novel. As noted in the Literature Review, this linguistic defiance directly supports the findings of Ahmad and Khan (2020), who highlighted Nila's resistance to patriarchal constraints through her expressive choices. This structural polarization between her modern vocabulary and her father's traditional stance mirrors the foundational concepts of description and interpretation outlined in Fairclough’s (1989) model established in the methodology.

7.2. Excerpt 2

“Father had sold […] his muse […] Gone […] Vanished […] nothing left […] the finger cut, to save the hand […] panicking. […] shrieking […]” (Hosseini, 2013, p. 33).

The excerpt has been drawn from the second chapter of the text which narrates the pathetic story of love empathy between Abdullah and his sister Pari. In the selected text, Abdullah’s feelings have been described after when his sister has been sold by his father to a wealthy family

7.2.1. Description

At the level of description, Fairclough’s three dimensional model provides a deep insight in the used vocabulary, grammar and the textual structure of the text. Tense: Past simple and past perfect tense symbolically reveal that the author is trying to expose the bitter reality of destruction that has happened in Afghanistan during numerous foreign invasions. Mode and tone: Declarative and exclamatory sentences reveal the wistfulness of the narrator which makes the tone of the text gloomy and melancholic. Sentence Structure: Simple, Compound and Complex sentences make the situation perplex, chaotic and pathetic. Ideologically loaded Words/Phrases like “Div”, “Muse”, and “The finger cut, to save the hand” depict the influence of supernatural entities on human life. Symbolically, ‘Div’ represents powerful and wealthy people who exploit the poor and loot their resources. Logical connectors such as “but’, ‘and’, ‘that’, and ‘so” show the narrator’s reluctance and resilience towards a particular ideology. Pronouns like “he and she” have been frequently used to reveal that the author is much concerned about third persons who are facing different challenges.

7.2.2. Interpretation and Explanation

The excerpt illustrates the profound and unwavering bond of love between brother and sister. Abdullah and Pari belong to lower class family and both love each other passionately and cannot live without each other. But their father is facing financial difficulties and sufferings therefore he has to sell her daughter ‘Pari’ to a wealthy family. The narrative vividly depicts the dire circumstances of an impoverished family and their exploitation by an affluent and opulent family of Kabul named ‘Wahdatis’. The author uses the metaphoric expression of “Div” for cruel creature that abducts the daughter of a poor father. This metaphor is used for Wahdatis who purchase Pari from her father. Moreover, this event also exposes the unequal

treatment of females who are sold and purchased as an object of business. Another metaphor of "Muse" has been used to emphasize the innocence and integrity of Pari. Furthermore, the constraints of the poor have been described critically and logically in order to reveal social inequality among Afghans. The author uses an ideologically teeming phrase, "The finger cut, to save the hand" to expose how poor people manage their problems and issues in critical circumstances. They have to sell their each and everything, even the most beloved, children, in order to earn their bread and butter. In addition, the author symbolically associates 'Pari' with a finger that has to be cut for protection and preservation of hand. The hand is also used as symbol for the impoverished family of 'Pari'. In this way, the author embeds the idea of social inequality found in Afghan society. This systemic commodification of female characters underscores the broader societal hierarchies investigated by Christopher et al. (2018) regarding deeply rooted masculine dominance. Furthermore, this economic desperation demonstrates the micro-level grammatical 'description' phase of the Tri-Dimensional framework, pointing back to the core research objectives to uncover how socio-cultural identities are structurally constrained by material poverty.

7.3. Excerpt 3 & 4

- **Excerpt 3** "[…] the beautiful, graceful creature […] big house. […] lit a cigarette and smoked without hurry […] bewitching grace […] cupping her hand before her mouth […] slender wrist […] a long-lashed woman with flowing dark hair […] pale delicate fingers […]" (Hosseini, 2013, p. 51).
- **Excerpt 4** "[…] a thickshouldered […] rough-skinned […] stubby fingers […] puffed, rounded face […] possessed none of the grace […]" (Hosseini, 2013, p. 18).

7.3.1. Description

In the process of description, the researcher uses the proposed theoretical framework based on Fairclough's three dimensional model (Fairclough, 1995). In these extracted parts of the text, Parwana and Nila, two major characters of the novel, have been introduced. Nila's character is unveiled when she is smoking and enjoying her life freely, while Parwana is introduced when she is making bread on tandoor and paying her duties as a working woman. As she belongs to lower class family she is being portrayed through informal and simple language. Mostly, the narrator uses negative adjectives for the portrayal of Pari's character. Quite the opposite, Nila's character has been drawn from the elite class, so she is portrayed with positive adjectives and through attractive, embellished and formal language

In this part of the analysis, two different extracts from the text are being analyzed comparatively. One is related to the description of Nila's character while in the other text Parwana's character has been introduced. The language that has been used to display Parwana's character is far different from that which is used to portray Nila's character. It is obvious that the author uses absolutely different language for portraying different characters from different strata. The discrimination of language will be explored in this analysis

- **Description of the text 3: Ideologically loaded words and phrases** such as "beautiful", "graceful", "whore", "bewitching grace", "slender wrist", "pale delicate fingers", and "Big house" reveal a different identity of a woman who belongs to upper

class. She is portrayed as elegant, exquisite and enchanting lady. **Mode and tone:** The use of Simple and declarative mode of the sentences with polished and standard language makes the tone of the narration sensual, attractive and charming. **Tense:** The use of present tense portrays the intentions of the writer behind constructing a different ideology. It shows his obsession with a particular image of Afghan society and its people. **Syntactic Structure:** The use of compound and complex sentence makes the narration attractive and cheerful. **Pronouns** such as **'I', '**my' and 'me' reveal that the author is more declined toward a particular ideology.

- **Description of the text 4: Ideologically loaded words and phrases** such as "daily naan"**,** "thick shouldered"**,** "rough-skinned"**,** "stubby fingers" and "tandoor' all reveal the culture and identity of the society. It is also embedded with the ideology about the sufferings and difficulties of Afghan women in rural areas. **Mode and Tone:** Declarative mode of the selected text makes the tone reflective, observant and gloomy. **Tense:** The use of past tense in this particular excerpt shows author intentions behind the discourse. It seems he wants to turn his back on the stereotypical and traditional role of women. **Syntactic Structure:** Compound and complex sentences make the situation serious and critical which show the significance of language that imparts the hidden ideologies. **Vocabulary:** The use of simple and informal vocabulary reveals the fact that the excerpt has been narrated by a lower class person, resulting into the development of identity. Moreover, the frequent use of **pronoun 'her'** reveals the supremacy, authority and empowerment of the character

7.3.2. Interpretation and Explanation

In the handpicked excerpts, narrator's embedded ideology of power and identity is quite apparent. Both excerpts introduce two major characters but from different social stratum. It is evident that the language which has been leveraged for the portrayal of lower stratum character contains different ideological perspective; while on the other hand, the language which is leveraged to introduce upper class character contains different ideological perspective. Different adjectives have been incorporated for both characters according to their social stratum. Ideologically loaded words also make it evident that there is manifested discrimination between the two characters. Nila Wahdati belongs to upper class and spending a luxurious and lavish life. She has been introduced as heroin and presented as valiant and audacious lady. While introducing her, the author uses the adjectives like "beautiful, graceful creature" which embellish her identity. Additionally, these adjectives make her more attractive and socially high standard woman who has the vigor of grace and beauty to impress, attract and dominate others. It shows her dominancy and importance in social strata. Moreover, the phrase like "big house" makes her personality more formidable and exceptional and it also denotes that she ascribes to upper echelon. Furthermore, she has been praised through certain more vivid adjectives like "long-lashed woman", "dark hair" and "delicate fingers" which represent her identity as beautiful, elegance and suitable for suitors. Moreover, these words make her a perfect woman as these qualities are considered necessary for women. Parwana belongs to lower class. She has been described with negative adjectives.

Her identity has been constructed less attractive and remarkable as compare to that of Nila. The author uses adjectives; "rough-skinned, puffed face and stubby fingers" to portray her character. She possesses no other qualities beyond those that portray her elegant and exquisite and esteemed. Consequently, her identity has been constructed as she has no concern with being beautiful or graceful like Nila. She is really a working women belonging to lower class and has no concern with her outer look. Moreover, she has been introduced while making daily 'Naan' on 'Tandoor'. However, Nila has been introduced while smoking and enjoying her life. It shows a massive discrimination between the two characters. This stark linguistic divergence in characterization serves as a prime empirical example of Ridgeway's (2014) status inequality theory discussed earlier, where material and social standing dictate perceived societal worth. It also provides a direct answer to Research Question 1, showing exactly how the writer constructs textual descriptions to institutionalize class boundaries.

7.4.Excerpt 5 & 6

- **Excerpt 5** "[…] thin waist […] pretty mouth […] arched eyebrows […] pink toenails […] lipstick. […] peach dress without sleeves […] dark sunglasses […] She smiled […] high-heeled shoes […]" (Hosseini, 2013, p. 28).
- **Excerpt 6** "Parwana […] pregnant […] shrouded figure […] sitting […] stiff silence, shriveled up into a ball. […] shoulders gathered, feet tucked beneath her swollen belly […] disappear into the wall. […] shielded […] soiled veil. […] shame rising from her, like steam, the embarrassment, how small she felt, […]" (Hosseini, 2013, p. 28).

The above mentioned extracted parts of the text represent Abdullah's views about the two major characters of the novel; Nila and Parwana. Both characters represent different social stratum and the author usually uses distinctive language for portraying either of the two. In the previous texts, it has been investigated how both are introduced to the audience but here it is explored the way they behave to each other and how much difference is there in both. Moreover, it has also been tried to explore how they encounter to each other.

7.4.1. Description

At the description level, it is evident that both characters have been portrayed through different language. The author uses ideologically loaded words with different perspectives to display the exact situation of the characters.

- **Description of the text 5:** Ideologically loaded words such as "thin waist", "pretty mouth", "arched eyebrows", "pink toenails", "lipstick", "sunglasses" and "high-heeled shoes" are the words that reveal the ideology about beauty, attraction, modernity, liberty, and empowerment. Use of such fascinating and reflective words for an elite class woman reveals the discriminated ideology about portraying different characters. Syntactic Structure: Complex sentences make the narration more embellished for a character that belongs to upper class. Mode and tone: By using declarative sentences, polite, admiring and praising has been developed to portray the character sophisticated and commendable. Pronouns: Frequent use of the pronouns 'her' and 'she' for a rich lady uncovers her authority and empowerment.

- **Description of the text 6:** Ideologically loaded words like “pregnant”, “shrouded figure”, “stiff silence”, “shriveled”, “shoulders gathered”, “feet tucked”, “swollen belly” and “veil” encapsulate the ideology behind a specific characterization. Mode and tone: The use of declarative sentences and pitiable words make the tone of the narration pathetic and gloomy and invoke empathy in readers for the character. The use of figurative language such as ‘like’ highlights the status of the character that is always compared with different things that are socially and cultural less important and are considered as the symbol of rough and tough. Additionally, logical connecters like ‘how’, ‘and’, ‘but’ are used as cohesive devices to construct clash, conflict and relationship in between the sentences and ideas.

7.4.2. Interpretation and Explanation

The excerpts are embedded with numerous ideologically embedded words that reflect the social system of the given society. Conjointly, the texts depict contrastive culture and society in which two female characters are living a life that is totally opposite to one another. First extract represents the upper echelon in which Nila’s character has been presented with highly prestigious status. The words like, “thin waist”, “pretty mouth” and “arched eyebrows” represent her physical beauty, remarkable look and standardized social status. Thin waist and pretty mouth are regarded as the crucial parts of women’s beauty and attraction. Moreover, arched eyebrows add much more to polish it. Whereas, the words like, “pink toenails”, “lipstick”, “sunglasses” and “high-heeled shoes” represent her economic standard. She possesses proficient wealth to apply these ornaments. Moreover, these things represent liberty, authority, superiority, autonomy, and boldness of Nila. In addition, these things are also regarded as mandatory for women. In many societies they are suppose to apply these ornaments. In the second extracted part of the text a lower class woman has been portrayed that is no longer enjoying prestigious and standardized social status. She belongs to a village where women are suppose to work all the day and take care of their children, Furthermore, the people of that village are not much wealthier to provide any kind of facilities to women. Therefore, women have no chance to embellish their beauteous features, look more attractive and utilize ornaments. The words like, “Pregnant”, “shrouded figure” and “swollen belly” depict the supposed role of women in villages particularly when they belong to lower stratum. These words reflect Parwana’s social role and her poor condition. Moreover, she represents all the women of her village and stratum. She has no concern with her beauty and outer look. Whether she is carless about it or not, it is equivocal, but the way the narrator ignores to portray her beauty, shows that lower class women are deprived to enjoy the right to look elegant and exquisite and are bound to live woeful and poignant life. Likewise, the words, “stiff silence”, “shoulders gathered”, “feet tucked”, and “veil” depict Parwana’s cultural obligations. She has covered herself with veil. Her cultural and social stratum does not allow her to unveil her face. Moreover, the sense of inferiority before an alluring, pulchritudinous and affluent lady makes her more submissive and fragile. Likewise, her silence adds more to her socially and economically docile and deferential condition. Throughout the scene, she remains silent in the presence of an opulent and enchanting woman. She lacks the supplies to

decorate herself and look more attractive and exquisite, so she seems unpretentious and modest. This submissive silence and spatial marginalization align heavily with Khan's (2016) observations on context-dependent power dynamics, where wealth explicitly determines discursive dominance. Geographically and culturally, this matches the macro-level 'explanation' stage of Fairclough's model, satisfying the final research objective by illustrating how the text reproduces localized stratification within the broader post-war Afghan milieu.

8. Conclusion

The main focus of the study is to explore social inequality through the lens of Fairclough's Three Dimensional Model and Ridgeway's Social Inequality Theory in Hosseini's And the Mountains Echoed and the analysis reveals that various types of social inequality have been portrayed through the language. Moreover, various causes behind social inequality have also been explored. Following types of social inequality and their causes are evident in the text:

Types of social inequality found in the text: The analysis of the text from the perspective of Tri-Dimensional model reveals that there are two major types of social inequality that have been portrayed in the given text. These types are being elaborated in the following section:

- **Gender Inequality**: A burning issue of every society, gender inequality, has been explored from the given text. The author has portrayed the unequal treatment of different female characters in the text. The major female character, Nila, faces various difficulties in her life and due to her worse treatment and for the sake of her daughter's secure future she flees to France. She stands against the so called norms and traditions of the society which makes her an iron lady. Excerpt 1 and 2 reveal that she has been treated unequally by her father and other 'Patriarchs' of her society. All these aspects of her life have been portrayed through the use of specific language that reveals her strong and determined willpower and empowerment. Moreover, another character can also be seen as the victim of gender inequality that is 'Pari' who is sold for the sake of money. Pari's pathetic story has been revealed through her brother's character that uses unequivocally different linguistic choices to describe the challenges faced by his sister. The phrase 'finger cut to save the hand' reveals the how peripheral she is for her family that she has been sold only to get some money.
- **Status Inequality**: Inequality on the basis of social status and class difference is also evident in the given text. Characters from different social status adopt different linguistic choice in their communication to each other. The excerpts 3, 4, 5, and 6 expose the fact the class distinction and status issues exist in the given society. Through the use of language the author has embedded the ideology that lower class people are always misbehaved and marginalized by the wealthier and the upper crust.

References

Ahmad, A., & Khan, R. N. (2020). Conflicting and Challenging Patriarchal and Liberal Feminist Ideologies and Norms in Afghanistan: Critical Stylistic Study of Khaled Hosseini's And the Mountains Echoed.

Bilal, H. A., Abbas, S., Kiran, A. R., & Mufti, A. W. (2021). Shafak's Treatment of Divinity and Religion in Three Daughters of Eve: A Critical Perspective. Psychology and Education, 58(4), 136–144.

Bloor, M., & Bloor, T. (2007). The Practice of Critical Discourse Analysis: An Introduction (1st ed.). Routledge.

Christopher, G., Sunday, J., & Loveth, A. (2018). Gender and Power Dynamics in Khaled Hosseini's And the Mountains Echoed: An Analysis of Social Hierarchies and Their Legitimization.

Difa, R. T., & Tarihoran, R. K. (2020). An Analysis of Inequality as Reflected in The Adventures of Tom Sawyer's novel. Journal of Linguistics, English Teaching and Education, 1(3), 86-90.

Fairclough, N. (1989). Language and Power. Longman.

Fareed, I., Khan, W. A., & Ahmed, A. (2022). Othering of Afghans: A Critical Study of And the Mountains Echoed. The Dialogue, 17(2), 108–122.

Fitria Deviana, & Danu Wahyono. (2024). Social Inequality on Women in Khaled Hosseini's "A Thousand Splendid Suns" Novel. Austronesian: Journal of Language Science & Literature, 3(2), 75–88. https://doi.org/10.59011/austronesian.3.2.2024.75-88

Hosseini, K. (2013). And the Mountains Echoed. Riverhead Books.

Khan, R. N. (2016). A Critical Study of Discursive Power in the Selected Passages of Khaled Hosseini's A Thousand Splendid Suns. NUML Journal of Critical Inquiry, 14(1), 87–113.

Rahim, N. (2024). A Critical Discourse Analysis of Women's Representation in Maysaloon Hadi's Novel The Black Eyes. Arab World English Journal, 15(3), 395–410.

Ridgeway, C. L. (2014). Why Status Matters for Inequality. American Sociological Review, 79(1), 1-16.

Shah, S., & Kumari, P. (2019). An Eco-critical Study of Khaled Hosseini's And The Mountains Echoed. An International Peer-Reviewed Open Access Journal of Humanities, 2(3), 1–7.

Shakeel, S., & Shaheen, U. (2024). Challenging Oppressive Norms to Reclaim Agency: A Feminist Critical Discourse Analysis of Mastur's The Women's Courtyard and A Promised Land. Jahan e Tahqeeq, 7(2), 149–170.