

A Corpus-Based Analysis of the Perspectives of Chinese and American Mainstream Media in the Israeli-Palestinian Conflict

Guang Yang

School of Foreign Languages, Guangdong University of Technology, Guangzhou Guangdong,
510000, China

862148083@qq.com

Abstract. Based on the research method of integrating corpus linguistics and critical discourse analysis (CDA), this study systematically compares the linguistic features, reporting positions, and propaganda strategies of mainstream Chinese and American media in reporting on the Israeli-Palestinian conflict since 2023, using a self-built corpus of 100 news articles from each country. The findings reveal that American media actively construct an image of Israel as a sovereign state in their reporting, linking the Palestinian side to violent attacks to form a semantic framework of sovereign states confronting non-state violent organizations, reflecting their pro-Israel stance and anti-terrorism narrative orientation. In contrast, Chinese media emphasize the geopolitical and historical roots of the conflict, highlighting multilateral rules and civilian protection through a humanitarian lens, echoing China's official support for the "two-state solution" and its role as a responsible major power in global governance. The research exposes the ideological differences and power relations behind language choices in international news reporting, filling an academic gap in the systematic comparison of Chinese and American media coverage of the Israeli-Palestinian conflict.

Keywords: The Israeli-Palestinian Conflict; Media Comparison; Critical Discourse Analysis; Corpus Linguistics.

1. Introduction

The armed conflict between Palestine and Israel stands as one of the longest-running and most complex regional disputes in the Middle East. On October 7, 2023, during the Jewish festival of Sukkot, Hamas launched attacks on two Israeli kibbutz settlements-Be'eray and Reem-targeting civilians. At the time, the Netanyahu-led Israeli government, representing the right-wing faction, had carried out numerous extreme retaliatory strikes against both Hamas and civilians in the Gaza Strip. Through the viral spread of firsthand videos on social media, this Israeli-Palestinian conflict captured global attention.

This sudden and violent clash, amplified through social media's viral effect, has escalated the Middle East regional conflict into a global ideological debate. In the digital age, social media and news platforms have become core tools for ideological collectives (including nations, international organizations, and NGOs) to shape identities, disseminate ideas, and achieve political objectives. In this highly complex and sensitive international issue, news media are not only conveyors of information but also constructors of reality, shapers of public opinion, and reflectors of national positions. As two major global powers with significant influence, how the media of the United States and China report on the Israeli-Palestinian conflict not only reflects their respective journalistic professionalism, ideological positions, and foreign policy orientations, but also profoundly influences the perception and attitudes of domestic and international audiences toward the conflict.

Currently, although a considerable number of media studies have been analyzed through various theoretical frameworks and statistical methods, on the one hand, the analysis of U.S. and Western media reports has formed a mature, theoretical, and methodologically diverse research field with substantial accumulated findings. On the other hand, specialized studies focusing on China's media coverage of this conflict remain in their infancy and fragmented, while systematic and comparative academic research directly comparing Chinese and American media is even rarer, leaving a significant academic gap.

Therefore, the author uses the theory of critical discourse analysis to analyze the two self-built corpora of mainstream media in China and the United States reporting the Israeli-Palestinian conflict, in order to realize the analysis of mainstream news media in China and the United States.

2. Literature Review

2.1 Overview of Media Coverage of the Israeli-Palestinian Conflict

Geopolitical positioning serves as a pivotal factor in shaping media narratives. Kaur and Arora (2025) demonstrate that Anglo-American and Middle Eastern media often construct diametrically opposed frameworks: the former tends to portray Hamas as the “aggressor” and Israel as the “victim,” while the latter adopts the reverse stance. These adversarial narrative differences are subtly conveyed through linguistic strategies that reflect the respective media outlets and nations’ positions. Further research by Sarwar and Qasim (2025) reveals that Western media predominantly employ transitive processes to highlight Israel’s proactive and defensive narratives, whereas media from Islamic and Global South countries tend to emphasize the connection between Palestinian suffering and moral frameworks, as well as the existential processes of these narratives.

Even within Western media, reporting lacks uniform standards. Arikat (2011) was among the first to highlight this issue, revealing that Western outlets-particularly U.S. media like FOX News-display clear Orientalist biases in covering the Israeli-Palestinian conflict, often portraying Palestinians as irrational ‘others.’ In contrast, BBC and Al Jazeera’s coverage appears more balanced. Kareem and Najm (2024) further confirmed through their YouTube platform research that Western media portray Palestinians negatively as “armed militants” or “terrorists” using explicit or implicit language strategies, while simultaneously reinforcing Israel’s positive image. This representation profoundly influences public perception and judgment. However, most reports focus solely on media in the Western world and the Middle East, with few studies examining the attitudes of relevant media toward China, an increasingly important link in the multi-tiered world. One of the goals of this study is to fill this academic gap.

2.2 Critical Discourse Analysis

Critical Discourse Analysis is a theoretical methodology for revealing the hidden ideology, concepts, and power imbalance in the textual frame. It is a methodology aiming to reveal the hidden ideology, concepts, and power imbalance in the text framework (Wodak and Meyer, 2009; Fairclough, 2013; Van Dijk, 2015). As mentioned above, different media reports on the Israeli-Palestinian conflict show antagonism in the content involving Israel and Palestine, respectively, and lack a unified standard. Such differences in expression lead to a tendency to compete for the position choice of information recipients in shaping the legitimacy of the behavior of both sides.

The CDA theory views discourse as a form of social practice, which scholars have endorsed for its ability to establish power relations among different actors. “An important perspective in CDA related to the notion of ‘power’ is that it is very rare that a text is the work of any one person. In texts, discursive differences are negotiated; they are governed by differences in power that are in part encoded in and determined by discourse and by genre” (Ramanathan and Hoon, 2015: 66). The media behind the news reports are a collection of different ideologies, representing their interests, so the language choice in the news reports is more or less biased, and there is an unequal description of the power relationship between Palestine and Israel. In CDA, the use of media language is regarded as a form of social practice: “the use of language is a ‘social practice’ which is both determined by social structure and contributes to stabilizing and changing that structure simultaneously” (Wodak and Meyer 2009:7). “The use of language is a ‘social practice’ that is constrained by social structure and at the same time affects the stability and change of that structure.” Moreover, ideological positions shape media news tendencies, which in turn lead to structural conceptual changes in the recipients of media narratives, “which provide power in discourse and signifies various world views from different perspectives” (Ramanathan and Hoon, 2015).

Therefore, given the limitations of using Critical Discourse Analysis (CDA) alone in processing large-scale text data, such as insufficient sample representativeness and researcher subjectivity, this study adopts an integrated approach combining Critical Linguistics (CL) and CDA by leveraging CL's systematic and objective nature to enhance the reliability and validity of CDA analysis, while employing CDA's critical theoretical perspective to deepen the interpretation of the social significance underlying CL data.

3. Methodology

Current research primarily focuses on analyzing the framing, discourse, and biases in American media, with key theories including critical discourse analysis and framing theory. However, when confronted with the massive volume of news data in the internet age, critical discourse analysis alone proves inadequate for effective processing.

3.1 Research Questions

(1) What are the features of the high-frequency lexical words and the co-occurrence network in the news reports on the Israeli-Palestinian Conflict in the Chinese relevant media corpus and the American relevant media corpus?

(2) What stances or attitudes the media take from their national backgrounds are constructed in the news texts of the corpus?

(3) What narrative strategies are used by mainstream media in both countries' mainstream media to present their stance and ideology?

3.2 Analytical Software

3.2.1 AntConc

AntConc is a powerful and user-friendly corpus analysis tool developed by Anthony (2005). It is widely used for analyzing large text corpora by using functions resemble: (1) Concordance, this feature displays search terms in their contextual context using the 'Keyword in Context' (KWIC) format, helping researchers observe the actual usage of words; (2) Concordance Plot, a bar chart that visually displays the distribution of search results within the text, revealing the patterns of lexical distribution; (3) File View, users are enabled to directly view the original content of individual text files in the corpus for comparative analysis; (4) Clusters and N-Gram, which help to identify recurring lexical patterns (e.g., fixed collocations, idioms) to reveal recurring phrase structures in the text; (5) Collocates, with this set, words that co-occur with the target term within a specific window can be revealed their semantic relationships between lexical items; (6) Word List (Frequency List), this set can generate a frequency ranking list of all words in the corpus, supporting statistics by word form, entry, and other methods; (7) Keyword List, by comparing the target corpus with the reference corpus, this method identifies words that occur significantly more or less frequently in the target corpus, which is used for term extraction or stylistic feature analysis.

3.2.2 KH Coder

KH Coder is a free software for quantitative text analysis developed by Koichi Higuchi and first released in 2001 (Higuchi, 2016). Supporting Windows, Linux, and Macintosh operating systems, it helps researchers conduct systematic quantitative analysis of textual data. Its core design philosophy combines quantitative and qualitative analysis methods, deepening data understanding through iterative application of both approaches. KH Coder offers comprehensive text analysis features including: (1) Preprocessing tools supporting word extraction, part-of-speech tagging, and encoding rule configuration for Japanese, English, and other languages; (2) Lexical analysis including word frequency lists, descriptive statistics, KWIC contextual retrieval, and word association analysis, with visualization methods such as correspondence analysis, multidimensional scaling, hierarchical clustering, co-occurrence network analysis, and self-organizing maps; (3) Document analysis

supporting document clustering, Naive Bayes classification, and document-word matrix export. Encoding analysis features include frequency statistics, cross-tabulation analysis, and similarity matrix calculations.

3.3 Corpus Design and Construction

To construct a representative and topic-related corpus for this study, this paper selects influential international media in China and the United States from the NexisLexis database to obtain as extensive a range of domestic media attitudes as possible in order to explore the collective ideological attitudes of the two countries toward the Israeli-Palestinian conflict. The search was conducted using a targeted set of keywords resembling “Israel”, “Palestine”, “conflict”, “ Hamas” and “people”. This design was set to capture a broad range of media reports related to the Israel-Palestine conflict. Moreover, the search was restricted to the top 300 newest articles published in English from 2023.10 to 2025.8 based on the database and algorithm of the system.

Having attained basic data sources, a manual exclusion process was applied to eliminate non-relevant articles, 100 most relevant articles were chosen separately as materials for each corpus. Each selected article was downloaded in Microsoft Word (.docx) or PDF format. During preprocessing, non-essential metadata such as headlines, author names, publication dates, and URLs were systematically removed. Only the BODY section, which contains the main textual content of each article, was retained. Then, all the bytes are collected in two UFL-8 text files, one representing Chinese media and the other American media. These files were then imported into AntConc4.3.1, the corpus analytical tools that generate wordlists and collocation patterns.

These two corpora, the Chinese media corpus mainly includes 100 news reports mainly from China Daily, the most influential English portal in China, while the American media corpus includes 100 news reports from the prominent leading media companies such as ABC, CNN, Washington Post, Fox, etc. The American media corpus has 154,892 tokens in total, and the Chinese media corpus has 80,702 tokens in total separately.

4. Results

4.1 Keywords List and Frequency Analysis

Table 1. Chinese media top 20 high-frequency lexical words.

Rank	Word	Frequency	Rank	Word	Frequency
1	Israel	857	11	military	193
2	gaza	580	12	people	191
3	hamas	515	13	middle east	153
4	Israeli	408	14	attack	146
5	palestinian	361	15	netanyahu	133
6	conflict	298	16	humanitarian	120
7	war	297	17	minister	119
8	China	279	18	peace	115
9	palestine	266	19	killed	104
10	international	215	20	government	78

According to Upton (2009:3), “The advantages of a corpus approach for the study of discourse, lexis, and grammatical variation include the emphasis on the representativeness of the text sample, and the computational tools for investigating distributional patterns across discourse contexts.” Therefore, beyond the traditional analysis of phrase and clause structures, the basic themes from the lexical dimension and the language features can be modified by high-frequency words of the corpus. Because word lists suggest key areas for investigation and highlight problem areas in a corpus (Anthony, 2005), the analysis starts with two word lists of the top 20 high-frequency lexical words to

identify the basic features of the two constructed corpora. Table 1 and Table 2 show the top 20 high-frequency lexical words of Chinese and American media corpora, respectively.

Table 2. American media top 20 high-frequent lexical words.

Rank	Word	Frequency	Rank	Word	Frequency
1	Israel	1272	11	West	222
2	Israeli	812	12	government	219
3	gaza	680	13	peace	195
4	palestinian	569	14	killed	185
5	hamas	546	15	attack	178
6	people	473	16	resolution	174
7	war	302	17	support	164
8	jewish	291	18	jews	164
9	palestine	280	19	jerusalem	158
10	conflict	229	20	arab	152

4.1.1 Chinese Corpus Features

The Chinese media corpus features distinctive high-frequency terms including: “conflict”, “China”, “international”, “Middle East”, “military”; “netanyahu”, “humanitarian”, and “minister”. These terms can be categorized into three semantic dimensions: multilateral and global perspectives, humanitarian and governance frameworks, and the dynamics of power and responsibility.

From a multilateral and global perspective, terms like “international”, “Middle East”, and “China” frame the conflict within the framework of “global governance”. By using geographical and institutional labels such as “international” and “Middle East”, the Israeli-Palestinian conflict is constructed as an international public issue requiring multilateral intervention. From the perspective of humanitarianism and governance, “humanitarian” abstracts the consequences of conflict into a humanitarian crisis, thereby transforming the critique of political violence into a neutral issue of rescue and governance. This discourse choice focuses on rescue-oriented expressions such as disaster relief and civilian protection, endowing China with moral high ground and avoiding the politicization of stance statements. From the perspective of power and responsibility, by focusing on official actions and leaders’ decisions through terms like “military”, “netanyahu”, and “minister”, the military operation in the Gaza Strip differs from terrorist attacks.

4.1.2 American Corpus Features

The American media corpus features distinctive high-frequency terms: West, resolution, support, Jews, city, and Arab. The terms “West” (Western bloc), “resolution”, “support”, “Jews”, “city”, and “Arab” are particularly prevalent. These terms function semantically to establish two dimensions: stance presupposition and identity manipulation, ultimately shaping the framework of news reporting presuppositions.

In constructing their ideological narrative, American media predominantly adopt an event-driven approach, pre-setting a pro-Western and pro-Israel reporting stance. They employ terms like “West”, “resolution”, and “support” to frame conflicts within a depoliticized and dehistoricised technical framework, steering public discourse toward specific interpretations. This implicitly positions Israel through the lens of an “Israeli-Palestinian” identity dichotomy, suggesting that their narratives consistently align with the Western alliance’s stance toward Israel.

The frequent use of “Israel” and “Israeli” indicates that American narratives primarily focus on Israel's interests, while the term “Palestinian” appears with comparable frequency alongside Hamas. This demonstrates how U.S. media ideology equates “Palestinian” with Hamas, constituting a form of propaganda-driven power suppression. Meanwhile, the high frequency of “Arab” and “city” in reports highlights the conflict’s geographical and cultural context. “Arab” is used as a generalized racial-cultural symbol, reducing diverse Arab populations to homogeneous others. In U.S. media, “Arab” often co-occurs with terms like “terrorist” and “conflict”, implying an inherent link between Arab identity and violence. This epitomizes the U.S.-advocated “rules-based international order,

as a long-term, structural “conflict” rooted in the “Middle East” region and centered around the core political issue of “Palestine” statehood.

Furthermore, “Israel”, “Gaza”, “strip” and “Hamass” as the most prominent core nodes occupy the absolute center of the network, primarily focusing the narrative on the “Israel-Hamas” conflict and indicating that the core location of the conflict is in the Gaza Strip. Additionally, it is noteworthy that “Palestine” and “Palestinians” as nodes are visually one level lower than Israel, suggesting that Chinese media’s reporting weakens their core status. This presentation creates a narrative disparity in perspective, with one side being clearly stated as the main actor, while the other side is more associated with geographical regions, non-state armed organizations, Hamas, and ethnic groups, Palestinians. This selective emphasis on Hamas while failing to provide an equivalent Palestinian political entity to Israel’s official one implies, on one hand, that Palestine has failed to establish a state due to historical reasons, with its land occupied by Israel and its settlements controlled and segregated. On the other hand, it characterizes the conflict as a confrontation between a state and an armed organization, rather than a political struggle between two nations or political entities.

Subgraphs 01 and 03 relationship chains collectively construct a semantic field centered on international norms and moral obligations. Starting with war, through the connections of Netanyahu and the Prime Minister, it points to foreign and China’s diplomatic spokesperson, forming a narrative path of “conflict origin→Israeli leadership→China’s diplomatic solution.” The chain of “International-community-law-humanitarian-aid” constructs a semantic field of rule-based international humanitarian order. The strong association of the terms “international-community”, “law” and “humanitarian-aid” firmly anchors the discussion of conflict on the high ground of international rule of law and humanitarian values.

4.2.2 American Corpus Concordance Analysis

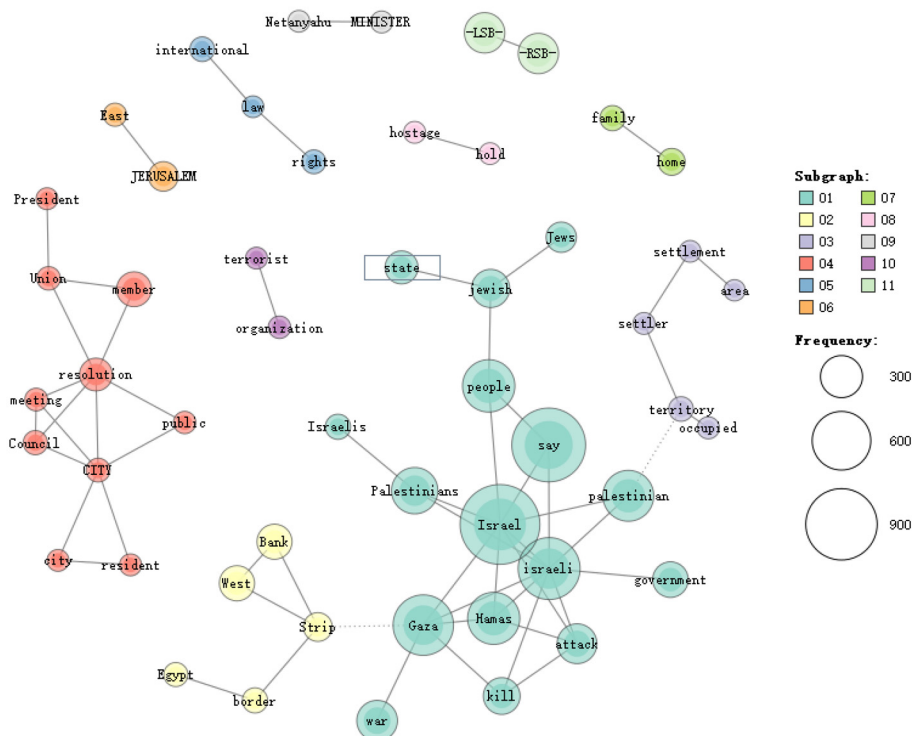


Figure 2. US media Concordance Networks

The textual structure reveals how U.S. media strategically nationalizes and officializes Israel through lexical choices, while simultaneously safeguarding and terrorizing the Palestinian narrative, thereby establishing a preliminary framework. The most striking feature of the chart is the overwhelming dominance of “Israel” and its derivative term “Israeli” in node size, indicating the

media's intense focus on Israel's perspective, which grants it inherent "subjectivity" and "centrality" in the narrative. This frequency imbalance inherently creates a "dominance of salience", subtly guiding audiences to perceive the conflict primarily through Israel's lens. In contrast, nodes representing the opposing political entity "Palestinians" appear disproportionately small, while "Hamas" may be highlighted. By framing the territorial conflict between states as a confrontation between sovereign states and non-state armed organizations through media narratives prioritizing Israel vs. Hamas over Israel vs. Palestine, this lexical strategy obscures the semantic distinction between the Palestinian people and armed groups.

Moreover, according to online reports, the image fails to include Palestine as a state node, instead featuring the geographical entity Gaza and the political organization Hamas. This lexical substitution strategy effectively achieves a semantic shift, reinterpreting ethnic conflicts or territorial disputes under international law as sovereign states confronting non-state violent actors, thereby establishing the logical foundation for subsequent counter-terrorism narratives.

In addition, at the level of discourse practice, the American media makes the above text structure appear objective and natural through the selection of sources and the juxtaposition of frames. The American media nationalizes and officializes Israel while safely terrorizing the Palestinian side through the choice of words, thus laying the initial foundation for the narrative framework. The network's core architecture and source selection jointly reinforce the unique strategic relationship between the United States and Israel. This framework constructs Israel as part of us (the Western democratic bloc) while labeling Hamas as the other (terrorism). Such binary opposition solidifies collective identity and provides discursive justification for policy. By emphasizing relatively mild terms like "settlement" and "rights" while avoiding critical concepts such as occupation and apartheid, the media confines discussions to politically acceptable parameters, thereby safeguarding the existing hegemonic order from fundamental challenges.

5. Discussion

Based on the framework of Critical Discourse Analysis, this paper makes a layered comparison of the language features, reporting positions, and communication strategies of Chinese and American media in the coverage of the Israeli-Palestinian conflict.

In China's media reports, "Israel" serves as the core node, slightly larger than the opposing entity "Hamas", while the conflict is framed within the long-term geopolitical and historical context through clusters like "conflict, Middle East, Palestine." Humanitarian terms such as "humanitarian aid" "international community", and "law" form a strongly correlated subgraph in the co-occurrence graph, reflecting the media's emphasis on international rule of law and humanitarian frameworks.

In U.S. media coverage, the terms "Israel" and "Israeli" dominate co-occurrence networks, with their node frequency far exceeding that of "Palestinians" or "Hamas", creating an unequal semantic field of state organization. Additionally, lexical binding occurs. "Israel/Israeli" frequently co-occurs with "President", "state" and "say", shaping Israel's image as a formal state actor. Conversely, "Palestinians" is tightly associated with words like "Hamas", "attack", "war", framing the Palestinian side as security-focused and terrorist. Chinese media emphasized the structural roots of the conflict, tracing back to the historical territorial dispute between Palestine and Israel, and proposed the state and two-state solution, echoing China's official political stance of supporting the UN's two-state solution in history. While Chinese media pointed out the violence between Hamas and Israel, it focused more on condemning the civilian casualties and humanitarian crisis caused by both sides rather than criticizing the violent conflicts, reflecting moral concern for vulnerable groups and human rights protection.

By portraying Israel as a sovereign state and Hamas and the Palestinian as a terrorist organization, American media implicitly endorse Israel's right to self-defense while aligning with Washington's counterterrorism narrative. To avoid critical narratives about Israel, American media employ

relatively neutral terms. This strategy serves to maintain the existing hegemonic order and protect Israel's international image as a strategic ally.

Chinese media embeds the current conflict into the grand narrative of the long-term Middle East conflict and the issue of Palestinian statehood, avoiding simplifying the conflict as sporadic violence while emphasizing its political essence. Leveraging international law and multilateral discourse, China positions itself as an advocate for international rules and peaceful resolution, enhancing its global governance discourse power.

American media reinforce the identity boundaries between the Western democratic camp and terrorism by creating semantic oppositions among conflict parties, providing discursive support for the legitimacy of Israeli government policies. In describing Israeli military actions, American media presuppose the nature of the conflict as counter-terrorism rather than national liberation or territorial disputes, leading readers to unconsciously accept this presupposition.

6. Conclusion

This study utilizes a self-built corpus of Chinese and American media coverage on the Israeli-Palestinian conflict, employing corpus linguistics and critical discourse analysis to systematically compare differences in lexical choices, semantic frameworks, and narrative strategies between mainstream media outlets. Chinese media, while also using "Israel," "Hammas", and "Gaza" as core nodes, placed the conflict within the long-term geopolitical and historical context through chains like "conflict-Middle-East-Palestine-state," emphasizing its structural roots. Meanwhile, the strong association of humanitarian terms such as "international-community-law-humanitarian-aid" highlighted the framework of international rule of law and multilateral solutions. In contrast, American media constructed Israel's official image as a formal state actor through the high-frequency co-occurrence of terms like "Israel/Israeli", "President", "state", and "official". At the same time, they associated the Palestinian side with words such as "attack", "terrorist", and "kill", completing their "safety" and "terrorism" labeling. Additionally, the semantic substitution of "Gaza" and "Hammas" for "Palestine" further eroded Palestinian national subjectivity.

The theoretical significance of this study lies in filling the academic gap of comparative research on media coverage of the Israeli-Palestinian conflict between China and the United States. By integrating corpus-driven methods with critical discourse analysis, the research enhances the objectivity and systematicness of textual analysis. The differences in media reporting between China and the U.S. not only reflect their distinct foreign policies and values but also subtly shape public perception and moral judgments about the conflict through the repeated use of discourse. Future research could expand the temporal scope and media range of linguistic data by incorporating multimodal data such as social media and visual discourse, thereby deepening our understanding of the dynamics of discursive power in cross-cultural communication.

References

- [1] Anthony, L. (2005, July). AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. In IPCC 2005. Proceedings. International Professional Communication Conference, 2005. (pp. 729-737). IEEE.
- [2] Arikat, H. (2011). Eye on the Media: The Israel-Palestine conflict.
- [3] Fairclough, N. (2013). Critical Discourse Analysis: The Critical Study of Language. Routledge.
- [4] Higuchi, K. (2016). KH Coder 3 reference manual. Kyoto (Japan): Ritsumeikan University.
- [5] Kaur, A., & Arora, A. (2025). Beyond the Battlefield: Framing Analysis of Media Coverage in Conflict Reporting. arXiv preprint arXiv:2506.10421.
- [6] Kareem, A. H., & Najm, Y. M. (2024). A critical discourse analysis of the biased role of western media in the Israeli-Palestinian conflict. *Journal of Language Studies*, 8(6), 200-215.

- [7] Ramanathan, R., & Hoon, T. B. (2015). Application of Critical Discourse Analysis in Media Discourse Studies. *3L: Southeast Asian Journal of English Language Studies*, 21(3).
- [8] Sarwar, M., & Qasim, H. M. (2025). FRAMING THE ISRAEL-PALESTINE CONFLICT IN MEDIA HEADLINES: A COMPARATIVE ANALYSIS OF WESTERN AND EASTERN PERSPECTIVES USING APPRAISAL AND TRANSITIVITY FRAMEWORKS. *Journal of Applied Linguistics and TESOL (JALT)*, 8(3), 1919-1931.
- [9] Upton, T. A., & Cohen, M. A. (2009). An approach to corpus-based discourse analysis: The move analysis as example. *Discourse studies*, 11(5), 585-605.
- [10] VAN DIJK, T. A. (2015). 22 Critical Discourse Analysis. *Discourse Analysis*, 466.
- [11] Wodak, R., & Meyer, M. (2009). Critical discourse analysis: History, agenda, theory and methodology. *Methods of critical discourse analysis*, 2(1), 1-33.