

How Arabic Media Construct Armed Conflict: A Corpus-Driven Concordance Analysis

M Indra Mulyadi¹, Chairunnisa Ahsana Amalan Shaliha², Abzari Jakfar³, Azhari⁴, Akmal Fajir⁵

^{1,2,3}Department of Arabic Language and Literature, Faculty of Adab and Humanities, UIN Ar-Raniry Banda Aceh

⁴University of Pannonia, Hungary

⁵Univeristas Al Washliyah Darussalam Banda Aceh

Email: 190502047@student.ar-raniry.ac.id

Article Info

Article history:

Accepted 12/28, 2025

Keywords:

Arabic News Discourse
Corpus Linguistics
Framing
Concordance Analysis

ABSTRACT

This study analyzes the construction of armed conflict in Arabic news discourse through the nuanced application of four lexical anchors in Al Jazeera Arabic reporting: ḥarb (war), širā' (conflict), silāh (weapons), and ḍaḥīya (victim). The issue at hand is that assessments of responsibility, humanitarian impact, and political legitimacy are frequently embedded in local lexical selections that are insufficiently analyzed in limited Arabic corpora. The goal is to figure out what these words mean in context and how they shape the way we talk about violence, competition, ability, and agency. The research utilizes a concordance-only corpus linguistic design through Sketch Engine. Ten hard news articles are collected into a custom corpus, and every Key Word in Context line for the four target words is read and coded, keeping clause-level co-text like intensifiers, numeric strings, attribution verbs, and source nouns. The findings show that there are different roles in the discourse. Ḥarb co selects with more intense descriptors and more precise numbers for humanitarian harm, such as detailed casualty and economic figures. This moralizes and measures war. Širā' records competition between multiple actors and structural conflict, often placing current conflicts in the context of longer historical patterns. Silāh manifests in claims concerning military action that are evidentially circumscribed, as well as in material records detailing logistics, funding, and safe havens, while simultaneously leveraging public opinion data. ḍaḥīya functions as a transition from passive victimhood to an agentive identity through value-laden lexicon and the legitimization of surveys. Implications encompass the efficacy of concordance-based deep analysis for small Arabic corpora, practical monitoring criteria for newsrooms regarding numeric displays and source attribution, and groundwork for cross-outlet comparisons and Arabic-sensitive quantitative layering in subsequent research.

Penelitian ini menganalisis konstruksi konflik bersenjata dalam wacana pemberitaan Arab melalui penerapan nuansa empat jangkar leksikal dalam laporan Al Jazeera Arabic, yaitu ḥarb (perang), širā' (konflik), silāh (senjata), dan ḍaḥīya (korban). Permasalahan utama dalam kajian ini adalah bahwa penilaian mengenai tanggung jawab, dampak kemanusiaan, dan legitimasi politik sering kali tersemat dalam pilihan leksikal lokal yang belum dianalisis secara memadai dalam korpus bahasa Arab yang terbatas. Penelitian ini bertujuan untuk mengungkap makna kata-kata tersebut dalam konteks penggunaannya serta bagaimana kata-kata tersebut membentuk cara wacana membicarakan kekerasan, persaingan, kapasitas, dan agensi. Penelitian ini menggunakan desain linguistik korpus berbasis konkordansi melalui Sketch Engine. Sepuluh artikel berita keras (hard news) dikumpulkan ke dalam sebuah korpus khusus, dan setiap baris Key Word in Context (KWIC) untuk keempat kata sasaran dibaca dan dikodekan dengan mempertahankan ko-tekst tingkat klausa, seperti penguat makna, deretan angka, verba atribusi, dan nomina sumber. Hasil penelitian menunjukkan bahwa masing-masing leksikon memainkan peran diskursif yang berbeda. Kata ḥarb berkolokasi dengan deskriptor yang lebih intens serta angka yang lebih presisi terkait dampak kemanusiaan, seperti rincian jumlah korban dan kerugian ekonomi, sehingga berfungsi untuk memoraliskan sekaligus mengukur perang. Kata širā' merekam persaingan antara berbagai aktor dan konflik struktural, serta sering menempatkan konflik aktual dalam konteks pola sejarah yang lebih

panjang. Kata *silāh* muncul dalam klaim-klaim terkait tindakan militer yang dibatasi secara evidensial, sekaligus dalam catatan material mengenai logistik, pendanaan, dan tempat perlindungan, sembari memanfaatkan data opini publik. Sementara itu, kata *ḍahīya* berfungsi sebagai transisi dari korban yang pasif menuju identitas yang memiliki agensi melalui leksikon bermuatan nilai dan legitimasi survei. Implikasi penelitian ini mencakup efektivitas analisis mendalam berbasis konkordansi untuk korpus bahasa Arab berukuran kecil, kriteria pemantauan praktis bagi ruang redaksi terkait penyajian data numerik dan atribusi sumber, serta landasan bagi perbandingan lintas media dan pengembangan pelapisan kuantitatif yang sensitif terhadap bahasa Arab dalam penelitian selanjutnya.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. INTRODUCTION

News discourse is very important for how people understand armed conflict, how information flows, and how ideologically important stories are made up about people, events, and responsibilities. Research in critical discourse studies has demonstrated that linguistic selections in the media are not impartial; they convey judgments, emphasize or obscure agency, and utilize lexical and grammatical resources that correspond with institutional stances and socio-political agendas (Fairclough, 1995; van Dijk, 1998). In media studies, framing theory posits that journalists and media outlets choose specific problem definitions, causal attributions, moral judgments, and treatment suggestions that shape audience perceptions of conflict (Entman, 1993). Simultaneously, studies on news values reveal consistent biases favoring negativity, conflict, prominent figures, proximity, and immediacy, which influence the reporting and prominence of events (Galtung & Ruge, 1965; Bednarek & Caple, 2017). These viewpoints converge on the assertion that the language of war reporting constitutes a domain of discursive power, wherein lexical selection, collocational patterns, citation practices, and grammatical structures collaboratively shape the narrative of violence, validate or invalidate participants, and allocate blame and accountability.

Corpus linguistics offers robust methodologies to implement and evaluate these insights across extensive or limited text collections. Methods like frequency profiling, keyword analysis, collocation, concordance inspection, and dispersion measures help researchers find systematic linguistic patterns that aren't obvious when reading for fun (Stubbs, 2001; Partington et al., 2013). Corpus-assisted discourse studies have shown that semantic prosody around certain words can show evaluative tendencies and alignments. They have also shown that patterns of agency suppression are linked to passivization, nominalization, and elision of actors in news reporting (Baker, 2006; Baker, Gabrielatos, & McEnery, 2013). Methodological advancements and instruments, Sketch Engine and its Corpus Query Language, enable reproducible, detailed analyses, facilitating lemma-based querying, part-of-speech filtering, and the systematic extraction of collocational profiles from media corpora (Kilgarriff et al., 2014). These methodologies have been utilized in the examination of subjects including migration, terrorism, and social groups, demonstrating that differential labeling and co-occurrence patterns are pivotal in the construction of social reality within news discourse (Baker et al., 2013; Bednarek & Caple, 2017)

Even with this strong set of methods and theories, there isn't much research on Arabic-language media discourse using corpus methods in international journals. A significant portion of corpus-based discourse scholarship has concentrated on English-language media in Western contexts, whereas Arabic-language reporting has frequently been analyzed through qualitative case studies, content analysis, or political communication frameworks (Pintak, 2014; Seib, 2005). Comparative studies of Arabic and English media have provided significant insights into transnational news flows and the politics of representation; however, comprehensive corpus analysis of the Arabic lexicon and grammar in conflict reporting remains relatively uncommon in mainstream applied linguistics literature (Figenschou, 2010; Hamdy & Gomaa, 2012). This gap is significant due to Arabic's unique morphological and syntactic characteristics, including complex derivational morphology, prevalent nominalization, and variability in definite marking, which affect the encoding of agency and evaluation in text. Research in Arabic natural language processing and morphology has underscored the necessity for meticulous lemma normalization and tag-aware querying in the analysis of Arabic corpora (Habash, 2010; Darwish & Mubarak, 2014). In the realm of war reporting, these linguistic characteristics may consistently influence the equilibrium

between active and passive constructions, the allocation of responsibility, and the evaluative prosody associated with principal figures.

Reporting on armed conflict in the Arabic-speaking world is also affected by complicated regional politics, media environments, and institutional ties. Pan-Arab networks and national broadcasters navigate professional standards, governmental influences, and audience anticipations that can affect lexical choices, the visibility of humanitarian perspectives, and the procurement of information regarding casualties and military activities (Seib, 2005; Pintak, 2014). Under these circumstances, granular corpus analysis is adept at pinpointing the deviations of discourse patterns from generalized Anglo-centric models and elucidating language-specific mechanisms for framing armed conflict. Moreover, the emphasis on reproducibility and query transparency in contemporary corpus linguistics contributes to the mitigation of subjectivity and the enhancement of methodological rigor, especially when examining ideologically sensitive topics that necessitate interpretive claims to be substantiated by systematic evidence (Partington et al., 2013; Kilgarriff et al., 2014).

There exists a distinct research deficiency at the convergence of corpus linguistics, Arabic news discourse, and conflict reporting. First, the existing literature is deficient in concentrated, replicable analyses of Arabic-language war reporting that integrate corpus methodologies with theoretical frameworks from framing and critical discourse analysis. Second, studies seldom record lemma-informed and tag-aware query protocols specifically designed for Arabic morphology to investigate agency suppression, event-target structures, and evaluative labeling. Third, the majority of corpus-assisted media studies have focused on English-language news or bilingual corpora, which may obscure the unique linguistic and discursive characteristics of Arabic text. Consequently, our comprehension of how Arabic news media portray armed conflict through lexical selection, grammatical frameworks, and sourcing methodologies remains insufficiently detailed and empirically substantiated. Furthermore, the interplay of these elements in signaling legitimization, delegitimization, or humanitarian emphasis in brief intervals surrounding key terms is not well understood (Baker et al., 2013; Bednarek & Caple, 2017).

This study fills this gap by doing a corpus linguistic analysis of ten Arabic-language news articles about armed conflict that were published on aljazeera.net. This focused corpus, though small, allows for a thorough and methodologically clear study of discursive patterns through Sketch Engine's Corpus Query Language. The study implements four constructs that are pivotal in discourse and framing research and are linguistically manageable in Arabic. These constructs encompass actor representation and labeling, event-target frameworks for civilian harm, agency suppression via passivization and nominalization, as well as evidentiality and modality in casualty reporting. Actor representation is analyzed via lemma-based inquiries concerning state forces, non-state armed groups, and civilian classifications, supplemented by lexical categories that signify legitimizing or delegitimizing labels. Event-target structures are put into action by linking conflict verbs to civilian or protected entities in short time frames. This makes it possible to find co-selections that highlight humanitarian impact. Periphrastic passive constructions, agentless casualty expressions, and nominalizations that lessen clear responsibility assignments all show how agency suppression works. Quotation verbs, source nouns, and hedging markers used with casualty figures and operational updates are used to look at evidentiality and modality (Entman, 1993; Partington et al., 2013; Kilgarriff et al., 2014).

The research possesses two interconnected aims. The first goal is to record and examine the language strategies that Al Jazeera uses when reporting on armed conflict. This includes looking at how actor labels, grammatical choices, and sourcing practices work together to create frames of responsibility, humanitarian concern, and legitimacy. The second goal is methodological. The study offers a clear and reproducible corpus protocol that utilizes lemma-based CQL, part-of-speech restrictions, and windowed co-occurrence patterns specifically designed for Arabic morphology and syntax. The study illustrates that corpus-assisted methods can produce significant insights even from a limited dataset by presenting counts and proportions appropriate for a small corpus and by correlating quantitative trends with concordance analysis. This dual orientation corresponds with the demands in applied linguistics and discourse studies for the amalgamation of theory-driven constructs with operationalizable, reproducible methodologies that are attuned to language-specific characteristics (Stubbs, 2001; Partington et al., 2013; Habash, 2010).

The study's contribution is threefold. It improves our understanding of Arabic-language conflict reporting by finding systematic patterns in Al Jazeera's coverage of labeling, agency management, and humanitarian framing. Methodologically, it presents a well-defined CQL-based protocol that can be modified for larger datasets and additional Arabic sources, thereby facilitating cumulative research and comparative analysis. It empirically demonstrates that even a modest, meticulously curated corpus can reveal significant patterns in evaluative prosody, event-target co-occurrence, and evidential sourcing that are directly pertinent to discussions in media discourse analysis. In practical terms, the methodology can be applied to multi-outlet corpora to analyze ideological orientations among pan-Arab channels and state-

aligned broadcasters, as well as to monitor shifts in framing throughout various phases of conflict and news cycles. The findings contribute to theoretical discourse regarding the construction of social realities of war and violence through language, as well as the influence of institutional contexts and audience expectations on the allocation of agency and evaluation in news reporting (Fairclough, 1995; van Dijk, 1998; Entman, 1993).

The study contextualizes Arabic news discourse within established frameworks of framing and critical discourse analysis, employs a corpus linguistic methodology tailored to the structural characteristics of Arabic, and offers empirically substantiated insights into the language of conflict reporting. By emphasizing reproducibility, linguistic accuracy, and theoretically informed operationalization, the analysis enhances an evidence-based comprehension of how Arabic news portrays armed conflict and how language influences the interplay among events, actors, and interpretations. The study addresses the necessity for corpus-based research in Arabic media discourse and presents a framework for future endeavors that amalgamate stringent methodology with theory-informed investigation in applied linguistics and discourse analysis.

Research on news discourse concerning armed conflict utilizes a cohesive framework of theories that elucidate the ways in which language encodes ideology, structures information, and shapes public interpretation. Critical Discourse Analysis views discourse as a social practice, illustrating how lexical selection, transitivity, and nominalization can either emphasize or diminish agency, thus legitimizing or delegitimizing actors and actions in the news (Fairclough, 1995; van Dijk, 1998). Framing theory delineates how media texts select certain elements of a perceived reality and highlight them through problem definitions, causal attributions, moral evaluations, and treatment recommendations, collectively shaping audience inference and judgment (Entman, 1993; Scheufele & Tewksbury, 2007). Systemic Functional Linguistics offers analytical instruments for delineating ideational and interpersonal significances via transitivity patterns, voice variations, evaluative resources, and modality, all of which are pivotal in the reporting of conflict events and accountability (Halliday & Matthiessen, 2014; Martin & White, 2005). In corpus linguistics, techniques such as keyword analysis, collocation, and semantic prosody uncover systematic evaluative tendencies and co-selection patterns that are not readily apparent through intuition alone (Stubbs, 2001; Partington, Duguid, & Taylor, 2013). Research on political violence news finds that there are certain patterns that show stance and ideology, such as actor labeling, event-target pairing, passivization and agent deletion, quotation patterns and sourcing, and the balance between humanitarian and strategic frames (Baker, Gabrielatos, & McEnery, 2013; Bednarek & Caple, 2017). These works collectively indicate a set of factors pertinent to Arabic conflict reporting: the lexical range for designating armed groups, the grammatical allocation of agency and responsibility, the evaluative prosody surrounding principal actors, and the evidential framework that substantiates assertions regarding casualties and operations (van Leeuwen, 2008; Richardson, 2007).

Within these traditions, scholars have formulated complementary methodologies and frameworks that can be aligned with quantifiable aspects, dimensions, and indicators for empirical investigation. According to framing theory, the shared dimensions encompass responsibility, conflict, human interest, morality, and economic repercussions, which can be implemented via lexical indicators, event schemas, and sourcing methodologies (Entman, 1993; Semetko & Valkenburg, 2000). The social actor model from critical discourse analysis has categories like activation vs. passivation, personalization vs. impersonalization, and inclusion vs. exclusion. You can find these categories by looking at transitivity choices, nominalization, and reference chains (van Leeuwen, 2008; Fairclough, 1995). Systemic Functional Linguistics provides dimensions of transitivity types, voice, modality, and appraisal subsystems, including attitude, engagement, and graduation, which can be associated with clause-level and phrase-level indicators in news texts (Halliday & Matthiessen, 2014; Martin & White, 2005). Corpus-assisted discourse studies convert these dimensions into queryable indicators: the keyness of labels, the collocational profiles of actors and events, the windowed co-occurrence of event verbs with civilian targets, the frequencies of periphrastic passive constructions, and the patterns of quotation verbs with source nouns (Baker, 2006; Partington et al., 2013). In the current research context, these methodologies are tailored to the Arabic language by focusing on lemma-based querying, acknowledging morphological variation, and being attuned to Arabic-specific passive and nominal forms that directly influence agency and evaluation (Habash, 2010; Darwish & Mubarak, 2014).

This study employs an integrated framework that originates from the convergence of corpus linguistics and discourse-analytic theories, commonly referred to as Corpus-Assisted Discourse Studies. Initial corpus research demonstrated the significance of frequency and concordance evidence in ideology studies, whereas Critical Discourse Analysis (CDA) and framing provided theoretically grounded constructs that could be transformed into quantifiable patterns (Stubbs, 2001; Baker, 2006). Later research on media discourse combined these ideas to look at touchy subjects like migration, terrorism, and Islam.

They used collocation networks, semantic preference, and prosody to figure out how different outlets and times felt about these issues (Baker et al., 2013; Bednarek & Caple, 2017). In the case of Arabic, advances in tokenization, lemmatization, and POS tagging made it possible to more reliably interrogate corpora, which is necessary for replicable analyses of languages with rich morphology (Habash, 2010; Kilgarriff et al., 2014). Researchers have previously integrated frameworks for pragmatic purposes: Critical Discourse Analysis (CDA) serves as a critical lens for interpreting agency and actor representation, framing theory offers a typology of news functions, Systemic Functional Linguistics (SFL) provides clause-level diagnostics, and corpus methods deliver quantitative leverage and transparency (Partington et al., 2013; Richardson, 2007). In similar contexts to this study, researchers have modified models to accommodate the linguistic characteristics of Arabic by emphasizing lemma-normalized patterns, acknowledging periphrastic passives like the tam construction with masdar, and considering the significance of nominalization and definite marking in actor reference (Darwish & Mubarak, 2014; Habash, 2010).

The integrated approach is suitable for analyzing the material object of Arabic conflict news as it correlates formal constructs with discernible linguistic signals in texts. Framing dimensions can be retrieved through systematically queryable proxies; for instance, responsibility frames can be identified by the prevalence of passives, nominalizations, and attribution verbs that allocate or diffuse blame; human-interest frames can be discerned through collocations that connect event verbs to civilian or protected entities; and conflict frames can be analyzed through escalation lexicon and clusters of battle events (Entman, 1993; Semetko & Valkenburg, 2000). The social actor model and SFL transitivity focus on the grammar of agency, which lets us look at activation and passivation through voice changes and role configurations in clauses. This is possible with concordance sampling and CQL constraints on verb lemmas and nearby noun phrases (van Leeuwen, 2008; Halliday & Matthiessen, 2014). Corpus methods are especially good for testing these ideas because they present you ways to count, compare, and put patterns in context across texts that can be repeated. Concordance lines, on the other hand, keep the qualitative nuance needed for discourse interpretation (Baker, 2006; Partington et al., 2013). To put it simply, the theoretical framework fits with the linguistic interface in a way that makes it possible to measure and understand Arabic news texts. This is especially true when there is strong linguistic preprocessing and query languages like those used in Sketch Engine (Kilgarriff et al., 2014; Habash, 2010).

The current study examines Arabic-language news articles regarding armed conflict disseminated by Al Jazeera. Al Jazeera is a pan-Arab news source that serves both regional and global audiences. It has a unique place in the media landscape because it has to balance professional standards with political sensitivities when reporting on war, civilian casualties, and political figures (Seib, 2005; Figenschou, 2010). Reporting on armed conflict usually includes quick updates, changing numbers of dead and wounded, both official and unofficial sources, and the need to find a balance between being quick and being sure. These pressures influence linguistic selections, such as the frequent use of hedging expressions for numerical data, the dependence on quotation formulas citing military or humanitarian sources, and the employment of grammatical structures that either intensify or mitigate responsibility (Bednarek & Caple, 2017; Richardson, 2007). In Arabic, other linguistic features are more important. These encompass intricate derivational morphology that generates nominalizations from verbal roots, periphrastic passives utilizing auxiliary-like verbs such as "tam," and patterns of definiteness that influence referential specificity and actor tracking throughout sentences. These characteristics generate various avenues for depicting events and participants, subsequently influencing readers' interpretations of agency, legitimacy, and humanitarian significance in conflict reporting (Habash, 2010; Darwish & Mubarak, 2014).

News stories about conflicts usually go through several stages, from the beginning of the event to later stages. These include breaking news stories that focus on immediate facts and uncertainty, consolidation pieces that combine official statements and eyewitness accounts, and analytical features that put operations, casualties, and diplomatic responses in context. Language choices change as these stages go on. Early reports frequently utilize modality and hedging to convey uncertainty in casualty figures, whereas subsequent reporting may solidify terminology and attributions as official narratives converge (Galtung & Ruge, 1965; Bednarek & Caple, 2017). The dynamics of sources also change; at first, people rely on local or activist sources, but later, they rely on confirmations from the government, military, or NGOs. In Arabic conflict reporting, the shift between active and passive constructions often reflects these changes in knowledge. Agentless passives and nominalizations emerge in narratives of harm when responsibility is ambiguous or disputed, whereas the activation of specific actors is more prevalent when official sources offer attributions. When violent event verbs are used near civilians or protected sites, the humanitarian aspect becomes more important. This creates co-selection patterns that act as cues for human-interest framing. Over time, the style of the outlet and the editorial lines affect the balance of labels used for armed groups. These labels can be neutral or ideologically charged, with positive or negative prosodies (Baker et al., 2013; Richardson, 2007).

Previous corpus-assisted discourse studies on conflict and security have primarily concentrated on English-language media, particularly regarding the representation of Islam, terrorism, and migration, and have illustrated the efficacy of collocation, semantic preference, and quotation patterns in elucidating ideological positioning (Baker et al., 2013; Partington et al., 2013). Research examining Arabic media frequently employs qualitative discourse analysis or political communication methodologies, yet seldom records reproducible, lemma-based corpus protocols that are attuned to Arabic morphology and syntax (Pintak, 2014; Seib, 2005). This study concentrates on Arabic-language reporting from a single prominent outlet and establishes a clear, replicable framework for examining four interconnected areas: actor labeling, event-target pairing, agency suppression, and evidentiality and modality in casualty and operational reporting. The specificity resides in the congruence of theory-driven constructs with Arabic-aware corpus indicators and in the illustration that a meticulously constructed small corpus can produce substantial insights when methodologies are transparent and concordance validation is essential. This is urgent because there are still armed conflicts going on in the area, and Arabic news is very important to people in the area and in the diaspora. To base discussions about media responsibility, ideology, and humanitarian reporting on empirical evidence that can be compared across outlets and time periods, we need better methodological clarity and language-sensitive analysis (Entman, 1993; Kilgarriff et al., 2014).

2. METHOD

Research Design

This study employs a qualitative corpus linguistic framework, utilizing a concordance-based analysis of specific lexical items associated with armed conflict. Concordance analysis offers direct insight into word usage within context, enabling researchers to investigate patterns of meaning, framing, and evaluation as they manifest in genuine discourse (Baker, 2006; Partington, Duguid, & Taylor, 2013). The design is well-suited for small to medium-sized corpora and is particularly suitable for Arabic news texts, where morphological richness and syntactic flexibility necessitate contextual examination for accurate interpretation (Habash, 2010). Instead of using statistical keywording or collocation measures, this study looks at how four chosen lexical anchors show conflict by closely reading concordance lines.

Data Source

The data are Arabic news stories that were published on the Al Jazeera Arabic website (aljazeera.net). Al Jazeera is one of the most important pan-Arab broadcast networks. It is often cited in studies of media in the Middle East and around the world (Seib, 2005; Figenschou, 2010). The chosen articles are all about armed conflict. They were chosen because they all have repeated references to important semantic domains that are important for reporting on conflict. This collection of texts makes up a coherent dataset that lets researchers look at lexical framing and evaluative language in a single journalistic voice.

Data Collection

Ten hard news articles were manually selected and obtained from the Al Jazeera Arabic online archive. There were two criteria for selection. First, the article must talk about a recent conflict event, like an increase in military activity, a political fight, or harm to civilians. Second, the article must have at least one of the four target words in it. Manual retrieval made sure that the content was relevant and the topics were consistent. This is important for concordance-based analysis, where the quality of KWIC lines depends on how well the source texts fit (Baker, 2006). Each article was saved as plain text and made ready to be uploaded to the corpus.

Instrument

All analyses were executed utilizing Sketch Engine, a corpus analysis platform extensively employed in computational linguistics and lexicography (Kilgarriff et al., 2014). Sketch Engine supports Arabic tokenization and has a reliable concordance feature, which was the main tool used in this study. The platform enables accurate retrieval of each occurrence of a target word, presents it within its textual context, and retains the surrounding co-text essential for interpreting meaning, stance, and evaluative load. To keep the qualitative design focused, only the concordance tool was used. Collocation, word sketching, and keyness statistics were not used.

Data Collection Procedures

Sketch Engine put the prepared texts together into a custom user corpus. The system used its built-in Arabic models to automatically tokenize the Arabic text and add part-of-speech tags. After making the corpus, the Concordance: word function was used to look up each of the four node words (حرب، صراع،) (سلاح، ضحية) one at a time. The interface gave back Key Word in Context (KWIC) lines that showed the co-text to the left and right of each term. To make sure the documentation was correct and to allow for repeated close reading, all concordance lines were exported as CSV files. Because Arabic has a lot of different forms, each line was checked by hand to make sure that the example was the right lemma and that any noise from retrieval didn't change the meaning. This method is in line with what is already known

in Arabic NLP and corpus work, where manual validation is necessary because of spelling differences and the difficulty of morphology (Habash, 2010; Darwish & Mubarak, 2014).

Measurement

In this study, measurement entailed qualitative categorization instead of numerical frequency analysis. Every concordance line functioned as an analytical unit. Measurement concentrated on: (1) the contextual function of each lexical item, for instance, whether حرب indicated destruction, escalation, or humanitarian loss. (2) Semantic framing, looking at whether صراع meant political rivalry or multi-actor structural contention. (3) Evidential status, pinpointing hedges, attribution markers, or source disclaimers in lines that include سلاح. (4) Role framing, particularly in victim lines, which were analyzed for patterns of victimization, agency, or moral judgment. The focus was on recognizing recurrent semantic patterns across KWIC lines instead of measuring their frequency. This is in line with the best practices for qualitative corpus linguistics when working with small datasets (Baker, 2006).

Data Analysis

The analysis of the data used ideas from corpus-assisted discourse studies, which mix contextual linguistic interpretation with systematic retrieval from a corpus (Partington et al., 2013; Baker, 2006). The analysis progressed through several phases: (1) Initial familiarization, during which all concordance lines were read repeatedly to discern preliminary semantic themes. We looked at how each word worked in both the story and the situation. (2) Thematic coding: Interpretive categories were made from groups of concordance lines. For حرب, the categories were destructive intensity, measured harm, and political effects. For صراع, there were two main types: political multipolarity and historical continuity. For سلاح, verification issues, transnational logistics, and public opinion were some of the categories. For ضحية, categories focused on victimhood reversal and identity formation. (3) Microanalysis of lexical framing: Each KWIC line was scrutinized for essential textual indicators, including modifiers, numerical expressions, attribution verbs, value-laden vocabulary, and ideological markers, in accordance with framing theory and discourse analysis (Entman, 1993; van Leeuwen, 2008). (4) Integrative interpretation: The four lexical domains were analyzed to discern overarching discourse patterns throughout the corpus. This phase sought to correlate linguistic insights with established characteristics of conflict reporting, such as responsibility allocation, humanitarian focus, and identity politics (Richardson, 2007; Bednarek & Caple, 2017). During the process, interpretations were based solely on the retrieved concordance lines, adhering to the principle that corpus evidence should inform discourse claims (Stubbs, 2001).

3. RESULTS AND DISCUSSION

a. Results

This part reveals the findings of a concordance-based research of four major words that Al Jazeera Arabic uses while talking about armed conflict: حرب (war), صراع (conflict), سلاح (arms), and ضحية (victim). Each term was examined in relation to its co-text within the corpus, enabling a sophisticated understanding of its contextual roles. The results are organized by theme, which is based on the semantic roles and discourse patterns that go with each node term (Table 1 and Figure 1).

The concordance lines for حرب always illustrate that war is a bad thing that happens to a lot of people. Some KWIC samples use حرب in sentences that are critical or make things worse, such as "حرب تدميرية" and "حرب إبادة" and "واسعة". These phrases show how big and morally heavy war is, making it seem like a planned event that goes beyond typical military fighting. There are also sentences that talk about activities like "شن حرب شاملة", which make the idea of deliberate escalation even stronger.

Next to the war, humanitarian effects are very crucial. One line of concordance gives explicit numbers of deaths, like "١٠٠ ألفا و 67 شهيد", as well as further details like "ألف طفل 20" and "٥٠٠ امرأة 12". It also states "مفقود 9500" and "٥٠٠ مصاب 196". The presence of these persons in the immediate environment suggests that war is articulated as measurable by its human toll. Quantified suffering exemplifies prevailing tendencies in conflict reporting that emphasize deaths as indicators of severity and legitimacy concerns (Bednarek & Caple, 2017).

Aside from deaths, حرب also talks about material and financial expenses. One-line claims that losses are "١٠٠ مليار دولار", while other lines talk about the societal ramifications, like "هجرة عكسية". These patterns depict war not merely as a military event but also as a catalyst for societal change, a categorization consistent with the analysis of war discourse in worldwide media (Richardson, 2007).

While حرب is about bad outcomes, the concordances for صراع are about political rivalry and conflict across structures. One sentence talks about a "صراع ثلاثي" that includes the Yemeni government, a separatist council, and tribal groupings. This three-part statement indicates that conflict is complicated and that political goals can be the same. Another KWIC phrase is "صراع عسكري على النفوذ والسيطرة على الحقول", which means that the struggle is about who has the most authority and control over critical resources. This aligns with the literature that connects conflict frames to the quest for political legitimacy and economic advantage (Entman, 1993).

There is also a paragraph in the dataset that talks about a "صراع مفتوح" that is part of a long history. It talks about how the population changed, noting that the number of settlers went from "280 ألف" after the Oslo Accords to "950 ألف" by 2023. These concordances illustrate that *صراع* helps us understand today's conflicts in terms of longer eras of political upheaval. This historical layering corresponds with media methods that incorporate current conflicts into broad narratives of grievance and global change (van Leeuwen, 2008).

There are two basic sorts of discursive patterns for *سلاح*: statements about military activity and mentions of logistical and financial infrastructure. First, some lines talk about airstrikes. For instance, one KWIC says, "سلاح الجو السعودي قصف... مواقع لقوات النخبة الحضرية". But the same phrase swiftly makes it clear that there is no verification: "لم يصدر حتى الآن أي بيان رسمي" and calls the report "رواية من طرف واحد". This combination shows that the word *سلاح* is regularly used in news items that show skepticism or lack of full proof. Hedging like this is in line with journalistic principles that say that information about combat is often disputed (Fowler, 1991).

Second, some lines put *سلاح* into a long narrative about how military logistics function between countries. There are talks of anti-tank missiles, cooperating with "ميليشيات عراقية مرتبطة بـ إيران" paying "300 ألف دولار" to probable fighters, and buying communication gear that costs "136 ألف دولار". One sentence also talks about safe houses built for former Syrian pilots "ليكونوا جاهزين لأي تصعيد عسكري محتمل". These patterns show that *سلاح* is used to strengthen stories about the material and financial systems that support armed groups and state organizations.

Finally, this lexical environment displays statistics on public opinion. One line indicates that "77% of Palestinians refuse to disarm Hamas" in exchange for a ceasefire. Another line says that "57%" are happy with Hamas, while "23%" are happy with the Palestinian Authority. These data suggest that *سلاح* is also a part of the political discussion about how legitimate armed resistance is and how much people believe it.

Even though *ضحية* only shows up in one concordance line, its rhetorical purpose is evident. The word "ضحية دائمة" talks about the predicament of the Palestinians; however, the material around it disputes this by suggesting that Palestinians are no longer passive victims. The sentence shows that Palestinians can take the lead, and it uses words like "الكرامة" and "التضحية" to support this claim. This transformation from victim to agent is substantiated by opinion polls indicating that "72%" of individuals endorse the strategic decision articulated in the article.

The amalgamation of *ضحية* with values, agency, and survey evidence signifies a deliberate reframing technique wherein victimization evolves into a narrative pivot rather than remaining a fixed identity. Rhetorical reversals such as these are well-documented in media studies of conflict discourse, where victimization and agency frequently coexist in complex manners (van Dijk, 1998).

When you bring all the concordances together, they reveal that conflict is always built in different ways: (1). War is all about death, destruction, and how it affects society as a whole. (2). *صراع* stresses how different politics are and how hard it is to get power throughout time. (3). *سلاح* illustrates both how armed operations work in real life and how information is often challenged during combat. (4). The victim influences identity and agency, signifying shifts in the moral and political stances of individuals.

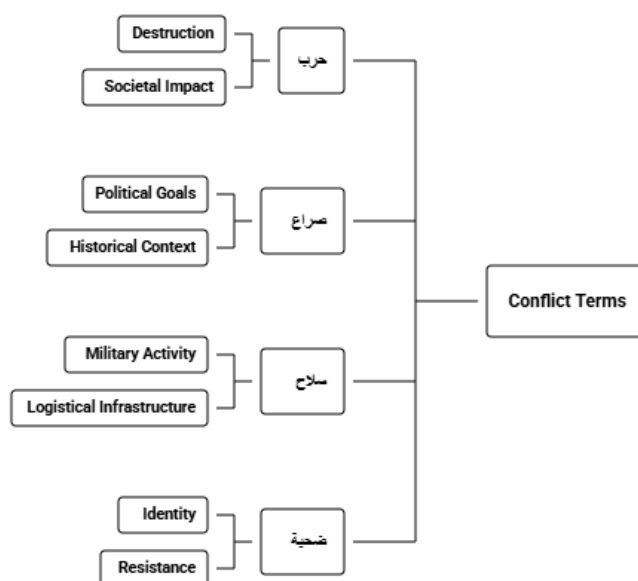
The four lexical items together create a clear discourse structure that measures and moralizes war, places conflict in historical and political context, situates weapons within transnational networks and verification practices, and transforms victimhood into a means of asserting identity and resistance. This trend aligns with prior findings in corpus-assisted research on conflict reporting, indicating that word choice and contextual framing are crucial in shaping public understanding (Baker, 2006; Partington et al., 2013).

Table 1. Concordance Analysis of Four Lexical Items

Lexical Item	Concordance Size	Typical KWIC Examples	Discursive Function	Emergent Themes
حرب (War)	6 lines	حرب تدميرية واسعة؛ "حرب إبادة"; "شحن حرب شاملة" casualty figures: "ألفا و ١٠٠" "67 ٢٠" ألف طفل "١٢ ألفا و ٥٠٠ امرأة" "٩٥٠٠٠ مفقود" "١٩٦ ألفا و ٥٠٠ مصاب" economic cost "مليار دولار 100"	Frames war as destructive, wide-scale, and morally charged; foregrounds measurable humanitarian loss	<ul style="list-style-type: none"> • Large-scale destruction • Quantified casualties • Economic and social disruption • Escalation and responsibility • Humanitarian framing

صراع (Conflict)	3 lines	صراع ثلاثي؛ “صراع عسكري على النفوذ والسيطرة على الحقول النفطية”؛ “صراع مفتوح... ٢٨٠ “ألفا → ٩٥٠ ألف	Represents political and structural contestation; embeds conflict within long historical arcs	<ul style="list-style-type: none"> • Multi-actor rivalry • Struggles for authority and territory • Resource competition • Historical continuity • Structural drivers of conflict
سلاح (Weapons)	3 lines	سلاح الجو السعودي قصف... لم يصدر أي بيان رسمي... رواية من “طرف واحد ألف 300” funding “تولار”; equipment “ألف دولار 136” ليكونوا “ جاهزين لأي تصعيد “عسكري محتمل public opinion مقابل 57%”، “77%” ٢٣%”	Highlights weapons as both a physical and political resource; reveals verification gaps, cross-border logistics, and public attitudes	<ul style="list-style-type: none"> • Airstrike claims and verification gaps • Transnational military logistics • Funding flows • Safe-haven networks • Political legitimacy and disarmament debates
ضحية (Victim)	1 line	ضحية دائمة... ليثبت أنه فاعل قادر... الكرامة والتضحية والدفاع عن القدس... ٧٢%”	Reframes Palestinians from passive victims to active agents; pairs moral values with numeric legitimation	<ul style="list-style-type: none"> • Victimhood reversal • Moral valorization • Agency construction • Survey-based legitimacy • Identity and resistance narratives

Figure 1. Discourse Analysis of Conflict Terms in Arabic



b. Discussion

The concordance-based analysis of four lexical anchors in Al Jazeera Arabic conflict reporting demonstrates a coherent structure of news discourse at a granular level. First, حرب shows up in situations that make damage and moral weight worse, along with the immediate co-selection of measured humanitarian harm and material costs. Minimal KWIC fragments such as “حرب إبادة”، “حرب تدميرية واسعة”، “حرب ١٠٠ ألف شهيد، ٢٠ ألف طفل، ١٢ ألفاً و٥٠٠ امرأة، ٦٧ ٩٥٠٠” occur with detailed casualty tallies “حرب شاملة” and macroeconomic references like “نحو ١٠٠ مليار دولار” and “مفقود، ١٩٦ ألفاً و٥٠٠ مصاب”. Second, صراع shows rivalry between multiple actors and structural contention. For example, “صراع ثلاثي” in Hadramawt and “صراع عسكري على النفوذ والسيطرة على الحقول النفطية”. It also shows how current conflicts fit into longer ones, like “صراع مفتوح” which will see a demographic change from “٢٨٠ ألفاً” to “٩٥٠ ألف” by 2023. Third, سلاح works in two ways: it makes claims about military action that are clearly hedged and source marked, like “سلاح الجو السعودي قصف... لم يصدر حتى الآن أي بيان رسمي... رواية من طرف واحد” and it gives details about logistics, funding, equipment, and safe havens (“١٣٦ ألف دولار، ملاذات آمنة 300”) along with opinion figures “٥٧٪، ٧٧٪، ٢٣٪ مقابل” that show how people feel about disarmament and performance. Fourth, الكرامة، التضحية، “ضحية” is used to shift from passive victimhood (“ضحية دائمة”) to agency framed by values (“الدفاع عن القدس”) and supported by survey data (“72%”).

These lexical behaviors demonstrate that war is discursively quantified and moralized, conflict is historically and politically layered, weapons discourse is both evidential and infrastructural, and victimhood is narratively transformed into agentive identity. The patterns come from reading each KWIC line and looking at the immediate context, not from using summary statistics or automated reduction. The functions identified for حرب correspond with recognized frameworks of news values and evaluation in conflict reporting. Bednarek and Caple (2017) assert that hard news emphasizes negativity, impact, and magnitude, frequently employing quantification and evaluative language; the recurrent association of حرب with casualty figures and substantial economic statistics exemplifies this inclination. Richardson (2007) contends that media discourse creates moral framing surrounding violence by amalgamating intensified descriptors with quantifiable harm. Our detailed KWIC analysis reveals these mechanisms functioning at the clause level in Arabic, substantiating that the combination of heightened war terminology with casualty reporting is a universal characteristic of conflict news across languages, despite the specific lexicon being language-dependent (Bednarek & Caple, 2017; Richardson, 2007).

The conduct of صراع aligns with framing theory's focus on the delineation of problems and causal attributions over time (Entman, 1993). The triadic and resource-centric applications in Yemen, alongside the historically entrenched Israeli Palestinian instances, correspond to van Leeuwen's social actor model, wherein activation, institutional roles, and processes of inclusion or exclusion are allocated among competing actors (van Leeuwen, 2008). Previous corpus-assisted studies have demonstrated that conflict coverage frequently distinguishes between episodic battle language and structural rivalry language; our KWIC evidence indicates that Arabic texts similarly differentiate lexically, employing صراع to encompass institutional and historical context while حرب focuses on destructive episodics (Baker, 2006; Partington, Duguid, & Taylor, 2013).

The two registers of سلاح align with the existing literature on evidentiality and verification in journalism. The clear labeling of unverified claims and one-sided sourcing is similar to the careful methods described by Fowler (1991) and Richardson (2007), in which the responsibility for statements is spread out through quoting and attributing. Simultaneously, the infrastructural enumeration of funding, equipment costs, and secure locations corresponds with the portrayal of the materiality of military capability, a characteristic emphasized in corpus-assisted discourse analyses that integrate logistics and finance into narrative frameworks (Partington et al., 2013; Baker, 2006). The inclusion of opinion figures in the سلاح KWIC lines introduces a layer of political legitimacy, aligning with research that utilizes quantitative public sentiment to either reinforce or challenge positions on resistance and disarmament (Bednarek & Caple, 2017; Richardson, 2007).

Finally, the ضحية pivot aligns with van Dijk's assertion that discourse oscillates between victimization and agency to navigate identity and legitimacy in conflict narratives (van Dijk, 1998). Appraisal theory posits that the co-selection of positive values occurs when a text recontextualizes actors from passive suffering to moral agency (Martin & White, 2005). Our detailed KWIC reading corroborates that Arabic reporting executes this reversal within the identical local context by correlating values and survey figures to substantiate agency.

The research illustrates the significance of concordance in-depth analysis for small, specialized Arabic corpora. Working at the level of KWIC lines avoids the risks of automated reduction that might hide important contextual clues, especially in Arabic, where morphology, definiteness, and nominalization interact with evaluation and agency (Habash, 2010; Darwish & Mubarak, 2014). By reading and coding each line, we maintain micro-rhetorical elements like intensifiers, numeric strings, attribution verbs, and

source nouns, which are crucial for framing and accountability but may be obscured in aggregated metrics. This method goes along with, rather than replaces, statistical tools. It is especially useful when researchers want to understand meaning in a way that is sensitive to language.

Lexical selection becomes a fundamental mechanism of frame construction. حرب acts as a focal point for intensification and quantification that supports discussions of responsibility and humanitarian importance, while صراع provides institutional and historical context that helps us understand the event. سلاح shows two sides of the story: evidence and infrastructure. These two things help people make claims about action and ability. ضحية shows how identity framing can be used strategically to reassert agency. These observations indicate that editorial decisions at the lexical level significantly influence public comprehension of armed conflict, corroborating previous assertions that language shapes social reality in news discourse (Fairclough, 1995; Entman, 1993).

The results underscore Arabic-specific resources, such as productive nominalization, periphrastic passives, and a value-laden moral lexicon, that engage with quantification and attribution. Even though the current analysis didn't look at passive rates, the hedging and source marking seen around سلاح and the agentless casualty reporting seen around حرب fit with patterns of agency management seen in Arabic and other languages (Habash, 2010; van Leeuwen, 2008). Editors and analysts can use these insights to make rules that make sure there is a balance between immediacy and verification, and to keep an eye on how moral language and numbers are spread throughout the coverage.

There are three things that should be noted. First, the corpus is small and only comes from one source. While concordance-only reading allows for in-depth analysis, broader generalization necessitates multi-outlet comparison and elevated token counts. Second, the research focuses on four-node words. Conflict discourse relies on event verbs (قتل، أصيب، اعتقل، قصف، استهدف) and civilian site nouns (مستشفى، مدرسة، مخيم، منزل) that are semantically important. Third, the analysis does not calculate association statistics or dispersion metrics. Qualitative regularities were robust; however, triangulation with collocation profiles and frequency normalization would augment their robustness, contingent upon adequate Arabic preprocessing quality (Kilgarriff et al., 2014; Habash, 2010).

The current findings yield several extensions. Broaden lexical scope and grammatical assessments. Use event predicates and protected entity nouns to connect systematic events to target pairing and agency suppression. This can be accomplished while maintaining a concordance-first methodology and incorporating restricted collocation windows to measure co-selection strength (Partington et al., 2013; Baker, 2006). Comparison across outlets. Create parallel subcorpora from pan-Arab and state-aligned media to assess whether حرب continues to serve as the focal point of quantification and whether صراع maintains structural functions across editorial boundaries. Comparing multiple sources is important for figuring out ideological differences (Richardson, 2007; Bednarek & Caple, 2017). Time dispersion and narrative stages. Observe the evolution of lexical framing throughout various reporting phases, ranging from breaking news updates to analytical segments. Investigate the reduction of hedging and attribution concerning سلاح as official narratives become more stable, and whether ضحية pivots align with particular political occurrences or shifts in public opinion (Entman, 1993; Martin & White, 2005). Quantitative layering in Arabic. If the quality of preprocessing allows it, add careful frequency normalization and LogDice collocations to line-by-line reading. This mixed method would keep interpretive fidelity while giving Arabic corpus linguistics (Kilgarriff et al., 2014; Habash, 2010) scalable metrics for cumulative research. Studies on how audiences react. Connect lexical framing to audience interpretations via survey experiments or focus groups, examining whether numeric displays and moral lexis in حرب lines yield enhanced humanitarian salience and whether صراع lines activate structural attributions of responsibility (Bednarek & Caple, 2017; van Dijk, 1998).

The current in-depth examination reveals that Arabic news coverage of armed conflict employs intricate lexical strategies to formulate concepts of violence, rivalry, capability, and identity. By examining each concordance line for ضحية، سلاح، صراع، حرب and scrutinizing the co-text with utmost precision, we reveal the local mechanisms that define framing in practice. These findings confirm the efficacy of concordance-only methodologies for small Arabic corpora and delineate distinct avenues for cumulative, linguistically informed research that amalgamates qualitative rigor with judicious quantitative stratification.

4. CONCLUSION

This study has demonstrated, through concordance-based reading at a fine granularity, that four lexical anchors in Al Jazeera Arabic conflict reporting serve distinct yet complementary discourse functions. War consistently coexists with intensified descriptors and quantified humanitarian harm, positioning war as a devastating event with quantifiable social and economic repercussions. صراع frames multi-actor rivalry and structural competition, embedding current confrontations in longer historical arcs and institutional dynamics. سلاح functions as both a material register, outlining logistics, funding, and

capacity, and as an evidential register, indicating verification status and source attribution. ضحية serves as a transition from passive suffering to an active identity, combining value-laden language with opinion figures to validate a redefined position. These patterns show how word choice and immediate context shape responsibility, humanitarian importance, and political meaning in Arabic news discourse.

There are three main effects. First, for corpus linguistics, the results show that concordance only deep analysis is useful for small, purpose-built Arabic corpora. Reading every KWIC line keeps small rhetorical features that can be lost in automated aggregation, especially in a language with a lot of different forms. Second, the results of the media and framing analysis show that word-level editorial choices affect how people understand violence and accountability: war is moralized and quantified, conflict is put in context through structure, weapons are talked about through evidence and infrastructure, and victimhood is turned into agency. Third, for practice, the patterns show newsrooms and analysts where to look for clear signs, such as how many numbers are used in war stories, how hedging and source marking are used around claims of military action, and how moral lexis is used with survey data when identity is at stake.

The corpus is small and only comes from one source, but the method gives strong insights that can be used on a larger scale. Future research may broaden the lexical scope to encompass event verbs and protected entities, juxtapose outlets with varying editorial alignments, and judiciously integrate quantitative metrics once the quality of Arabic preprocessing is guaranteed. Mixed method designs that integrate concordance reading with collocation profiles and audience studies would elucidate the interplay between lexical framing, reception, and policy discourse.

The study elucidates how Arabic news formulates the language of armed conflict through meticulous lexical selections that structure destruction, rivalry, capacity, and identity. By anchoring interpretation in every concordance line for ضحية, سلاح، صراع، حرب، the analysis illustrates that meticulous attention to local context can reveal the mechanisms through which news discourse influences public comprehension. The following reflection is straightforward and pragmatic: when the stakes involve warfare and human lives, linguistic precision transcends mere stylistic choice and becomes a moral obligation.

REFERENCES

- Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.
- Bednarek, M., & Caple, H. (2017). *The discourse of news values: How news organizations create newsworthiness*. Oxford University Press.
- Darwish, K., & Mubarak, H. (2014). Arabic NLP: Challenges and solutions. In *Proceedings of the EMNLP Workshops* (pp. 1–7).
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58.
- Fairclough, N. (1995). *Media discourse*. Edward Arnold.
- Fowler, R. (1991). *Language in the news: Discourse and ideology in the press*. Routledge.
- Habash, N. (2010). *Introduction to Arabic natural language processing*. Morgan & Claypool.
- Kilgarrieff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36.
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies*. John Benjamins.
- Richardson, J. E. (2007). *Analysing newspapers: An approach from critical discourse analysis*. Palgrave.
- van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. Sage.
- van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford University Press.
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). Sketching Muslims: A corpus-driven analysis of representations of Islam in the British press. *Applied Linguistics*, 34(3), 255–278. <https://doi.org/10.1093/applin/ams048>
- Bednarek, M., & Caple, H. (2017). *The discourse of news values: How news organizations create newsworthiness*. Oxford University Press.
- Darwish, K., & Mubarak, H. (2014). Arabic NLP: Challenges and solutions. *Proceedings of the EMNLP Workshops*, 1–7.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>

- Fairclough, N. (1995). *Media discourse*. Edward Arnold.
- Figenschou, T. U. (2010). Young, female, Western: The portrayal of Al Jazeera English journalists. *Journalism Studies*, 11(2), 196–210. <https://doi.org/10.1080/14616700903436064>
- Galtung, J., & Ruge, M. H. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research*, 2(1), 64–90.
- Habash, N. (2010). *Introduction to Arabic natural language processing*. Morgan & Claypool. <https://doi.org/10.2200/S00277ED1V01Y201008HLT007>
- Halliday, M. A. K., & Matthiessen, C. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.
- Hamdy, N., & Gomaa, E. H. (2012). Framing the Egyptian uprising in Arabic language newspapers and social media. *Journal of Communication*, 62(2), 195–211. <https://doi.org/10.1111/j.1460-2466.2012.01637.x>
- Kilgariff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovár, V., Michelfeit, J., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. John Benjamins.
- Pintak, L. (2014). *The new Arab journalist: Mission and identity in a time of turmoil*. I.B. Tauris.
- Richardson, J. E. (2007). *Analysing newspapers: An approach from critical discourse analysis*. Palgrave.
- Scheufele, D. A., & Tewksbury, D. (2007). Framing, agenda setting, and priming: The evolution of three media effects models. *Journal of Communication*, 57(1), 9–20. <https://doi.org/10.1111/j.1460-2466.2006.00326.x>
- Seib, P. (2005). Hegemonic no more: Western media, the rise of Al-Jazeera, and the influence of diverse voices. *International Studies Review*, 7(4), 601–615. <https://doi.org/10.1111/j.1468-2486.2005.00531.x>
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Blackwell.
- van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. Sage.
- Van Leeuwen, T. (2008). *Discourse and practice: New tools for critical discourse analysis*. Oxford University Press.