Media Bias in The Guardian: Sentiment and Semantics in Israeli-Palestinian Coverage

First Author[0000−1111−2222−3333], Second Author[1111−2222−3333−4444], and Third Author[2222−−3333−4444−5555]

No Institute Given

Abstract. This study investigates media framing in the coverage of the Israeli-Palestinian conflict across several decades, using advanced natural language processing (NLP) techniques to analyze sentiment, emotion, named entity recognition (NER), and semantic bias. Using a multi-method approach—integrating lexicon-based (VADER, TextBlob), transformer-based (RoBERTa), and word embedding (Word2Vec) models—we examine sentiment patterns, emotional valence, entity-level sentiment, and framing differences between headlines and article content. Additionally, we incorporate a transformer-based large language model (LLM) for emotion detection to complement the NRC Emotion Lexicon analysis.
Results reveal that media tends to exhibit more negative sentiment towards Palestinian mentions than Israeli ones, with gaps increasing during conflict escalations (e.g., 2014 Gaza War, 2023 Hamas-Israel war). Emotion analysis highlights fear as the dominant emotion for both groups, with NRC showing fear 8.1% higher in Israeli contexts and LLM showing a smaller difference of 6.3%. Anger shows divergent trends, with NRC indicating 9.5% higher prevalence in Palestinian contexts, while LLM shows almost no difference (-0.2%). Sadness is consistently higher in Israeli contexts according to NRC (+71.2%), but LLM shows it as more prevalent in Palestinian contexts (-10.3%). These complementary findings demonstrate the robustness of combining lexicon-based and transformer-based methods for emotion detection. Headlines display greater sentiment polarity than article bodies, with a negative bias in 59.5% of cases, amplifying conflict-oriented framing. Word embeddings indicate stronger associations of violence-related terms, with the conflict escalations periods. Named Entity Analysis reveals that sentiment towards prominent individuals—such as political leaders and military figures—shifts significantly with conflict intensity, offering insight into how public figures are emotionally framed in media narratives. The findings highlight systematic framing differences that may influence public perception.

Keywords: media framing · sentiment analysis · emotion analysis · word embeddings · named entity recognition · news analysis

1 Introduction

The media plays a pivotal role in shaping public perception of geopolitical conflicts through linguistic choices and framing strategies [1, 2]. News coverage can significantly influence how audiences interpret complex situations, potentially amplifying certain narratives while diminishing others [3]. In the context of the Israeli-Palestinian conflict, these media portrayals can have profound implications for public opinion and policy discourse [4, 5]. Media bias can manifest subtly through sentiment, emotional framing, and differential coverage that may systematically favor one perspective over another [6]. Despite the significance of these potential biases, quantitative approaches to systematically analyze large-scale media coverage across extended timeframes remain underdeveloped.

Computational approaches to media bias and framing have grown rapidly in recent years. Surveys by Hamborg et al. [7] and Ali and Hassan [8] note a surge of NLP and machine-learning methods aimed at identifying frames and bias in news. These methods include topic modeling, lexicon analysis, and supervised learning to classify framing categories. For example, Ali and Hassan (2022) provide a systematic overview of NLP framework models, highlighting the gaps between sociolinguistic concepts of 'frame' and the technical definitions used in algorithms. Kang and Yang (2025), for example, applied topic and frame analysis to Taiwanese media articles on the Palestinian–Israeli conflict and found systematic differences linked to political orientation. [9] These works illustrate how NLP can scale up frame analysis, but also highlight that most prior studies have been confined to short time spans or single outlets.

To address this gap, we perform a comprehensive computational analysis of The Guardian's coverage of the Israeli–Palestinian conflict from 2000 through 2024. We leverage both established and recent NLP tools: VADER and TextBlob lexicons for baseline sentiment scoring; a fine-tuned RoBERTa model for deep sentiment and emotion classification; word2vec embeddings to track semantic changes over time; the SpaCy NER tagger to identify mentions of persons, organizations, and locations; and the NRC Emotion Lexicon to profile emotional content. By combining these methods, we can compare results and validate findings. Our approach thus integrates the current state-of-the-art in sentiment and framing analysis.

The main contributions of this work are as follows:

- **Comparative sentiment analysis:** Systematic comparison of sentiment patterns using lexicon-based (VADER, TextBlob), transformer-based (RoBERTa), and embedding-based approaches to reveal consistent negativity bias and temporal/contextual shifts.
- **Word2Vec temporal analysis:** Training and analysis of word embeddings on different time periods to track evolving semantic associations and media framing of key entities.
- **Named entity recognition (NER):** Extraction and sentiment analysis of key individuals (e.g., political leaders, military figures) to provide a fine-grained view of how sentiment toward public figures shifts with conflict intensity.
- **Emotion analysis:** Quantitative assessment of emotional framing using the NRC Emotion Lexicon to compare the prevalence of core emotions in coverage of each group.

– **Framing analysis:** Direct comparison of sentiment and polarity between headlines and article bodies to uncover systematic differences and amplification of conflict-oriented narratives.

The remainder of this paper is organized as follows. Section 2 reviews computational approaches to analyzing conflict coverage. Section 3 details our methodology. Section 4 presents findings on sentiment patterns, emotional framing, named entity analysis, and semantic shifts. Finally, Section 5 discusses implications for understanding media framing of geopolitical conflicts.

2 Related Work

Recent computational approaches to analyzing media coverage of the Israeli-Palestinian conflict have provided valuable insights into sentiment patterns, emotional framing, and headline construction. This section reviews key developments in these areas and highlights how our work extends existing research.

2.1 Sentiment Analysis in Conflict Coverage

Sentiment analysis has been widely applied to examine media portrayal of the Israeli-Palestinian conflict, particularly since October 2023 escalation. Sharkar et al. [12] applied both VADER and TextBlob to analyze 436,425 Reddit comments related to the conflict, with VADER achieving 92.74% accuracy in sentiment classification. Their work demonstrated the effectiveness of lexicon-based approaches for social media content but did not extend to formal news articles. Similarly, Liyih et al. [13] developed a hybrid CNN-BiLSTM model for YouTube comment analysis that achieved 95.73% accuracy, showing that deep learning approaches can capture nuanced sentiment expressions in user-generated content.

For traditional news media, Naeem and Razaq [14] analyzed 5,000 news headlines from Google News using logistic regression, finding a dominance of neutral sentiment (53%) with substantial negative framing, suggesting conflict-oriented coverage. Abuasaker et al. [15] compared sentiment across European media outlets, revealing that Spanish and German sources exhibited significantly more negative coverage of the Gaza conflict than other European countries. These studies provide valuable insights into single-source sentiment patterns, but stop short of comparing sentiment across different news elements (headlines vs. body text) or tracking temporal changes.

Our work extends these approaches by employing multiple complementary sentiment analysis techniques (VADER, TextBlob, and RoBERTa) to analyze The Guardian's coverage over multiple years. This multi-model approach enables more reliable detection of sentiment patterns through triangulation, while our temporal analysis reveals how sentiment has evolved across major events in the conflict.

2.2 Emotion Analysis in News Media

While sentiment analysis identifies positive or negative tone, emotion analysis reveals more nuanced affective content in conflict coverage. The Pew Research Center [16] found that 74% of U.S. adults reported feeling sadness and 68% reported anger when consuming news about the Israel-Hamas war, demonstrating how news content directly influences readers' emotional responses. These emotional reactions likely stem from specific framing techniques employed by news organizations.

Emotion analysis using computational approaches has been less common than basic sentiment analysis. Guerra et al. [17] used lexicon-based methods to measure "extreme opinions" in Reddit posts, identifying peaks in emotional intensity corresponding to real-world events, but focused on social media rather than news content. Asmus [18] conducted qualitative analysis of German newspaper coverage, finding that emotional framing was used differently when discussing Israeli versus Palestinian perspectives, though without employing computational methods.

Our work advances emotion analysis by applying the NRC Emotion Lexicon to systematically quantify seven primary emotions (anger, disgust, fear, grief, joy, sadness, surprise) across Palestinian and Israeli contexts in The Guardian's coverage. This approach reveals which specific emotions dominate the portrayal of each group and how emotional framing differs between them.

2.3 Headline vs. Article Content Analysis

Research examining framing differences between headlines and article content represents an emerging area within conflict coverage analysis. Garcia et al. [19] analyzed causal language in BBC headlines covering Gaza, finding that headlines often avoided directly attributing causality for violence to Israel, whereas this attribution was more explicit in Al-Jazeera headlines. Their work demonstrated how subtle linguistic patterns in headlines can influence reader perception, though they did not specifically compare headlines to article bodies.

The Guardian's conflict coverage has received some scholarly attention. Mctigue [20] examined potential bias in conflict reporting, noting that brevity requirements in headlines can lead to oversimplification of complex issues. However, systematic computational analysis comparing The Guardian's headlines with corresponding article content has been limited, particularly using multiple sentiment analysis tools.

Our work builds upon these studies by conducting a direct quantitative comparison between headlines and full article content, revealing how sentiment intensity and emotional framing differ between these elements. This analysis provides insights into potential headline sensationalism or framing bias that could influence reader perception before they engage with the complete article.

3 Methodology

This section outlines the computational framework employed to analyze media framing in The Guardian's coverage of the Israeli-Palestinian conflict. The methodology integrates semantic embedding analysis, multi-model sentiment and emotion detection, named entity recognition, and headline–article comparison.

Figure 1 illustrates the overall pipeline. The process begins with data acquisition via The Guardian API, followed by thorough preprocessing. The cleaned data is then simultaneously processed by five independent modules: (1) semantic embedding, to capture contextual bias and shifts in association; (2) sentiment analysis, to measure polarity; (3) emotion detection, to reveal nuanced affective framing; (4) named entity recognition, to extract and evaluate key figures; and (5) framing analysis, which compares headlines with article bodies and identifies collocational patterns.

The outputs of these modules are synthesized and visualized to support interpretative analysis. This modular, parallel architecture enables a comprehensive, multi-level examination of media framing and potential bias.

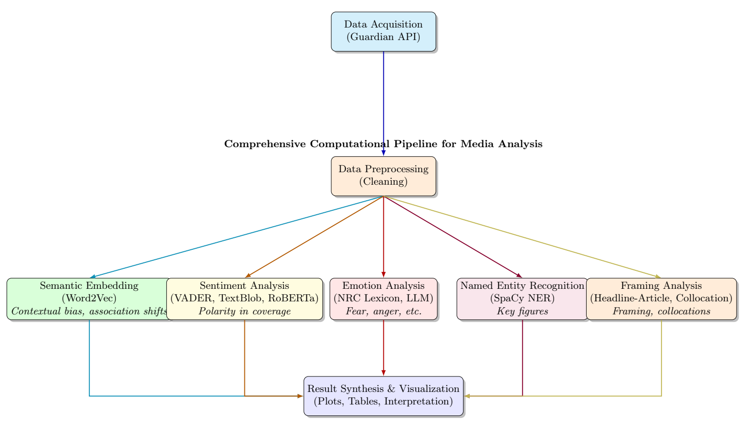


Fig. 1: Block diagram of the full methodology pipeline. Data is acquired and preprocessed, then five independent analysis modules are applied in parallel: semantic embedding (contextual bias), sentiment analysis (polarity), emotion analysis (fear, anger, etc.), named entity recognition (key figures), and framing analysis (headline-article, collocation). All results are synthesized and visualized for interpretation.

3.1 Data Collection and Preprocessing

We collected 47,248 articles (2000–2024) from The Guardian API, focusing on those mentioning conflict-related keywords. Preprocessing included HTML re-

moval, sentence and word tokenization, lemmatization, stopword removal, and extraction of headlines, article bodies, and temporal metadata. Named entity recognition was used to identify references to Palestinians, Israelis, and key individuals.
Each article entry includes both metadata and textual content as:

Collected Fields:

- Metadata: article ID, section, publication date, author
- Content: headline, standfirst, full article body (HTML and plain text)
- Stats: word count, character count, language, live status

Used Columns:

- `fields.headline`
- `fields.standfirst`
- `fields.body`

It should be noted that the dataset assembled for this study does not include ground truth sentiment or emotion annotations. No manual labeling or external validation set was available for direct model evaluation. In the absence of labeled data, model performance was assessed by comparing outputs across multiple established sentiment and emotion analysis tools, and by examining the consistency of observed patterns with prior literature and known historical events. This triangulation approach, commonly adopted in large-scale media analyses, enables robust validation of findings when explicit ground truth is unavailable.

3.2 Semantic Embedding Analysis

To capture contextual shifts in language, we trained separate Word2Vec models [10] for six time periods (see Table 1). Each model used the Skip-gram architecture with the following hyperparameters: embedding dimension = 30, context window size = 5, and training iterations = 10. For each period, we computed the average cosine similarity between a target term (e.g., "Palestinian", "Israeli") and a group of concept words (e.g., violence-related or peace-related). The similarity is defined as:

$$\mathrm{Sim}(w_t, G) = \frac{1}{|G|} \sum_{w_a \in G} \cos(\boldsymbol{w_t}, \boldsymbol{w_a}) \tag{1}$$

where w_t is the target word, G is the set of concept words, and $\boldsymbol{w}$ denotes the embedding vector. This equation computes the mean cosine similarity between the embedding of the target word and the embedding of each word in the concept group, quantifying how closely the target is associated with a particular semantic field.

To measure bias, we calculate the difference in similarity between violence-related and peace-related groups:

$$\mathrm{Bias}(w_t) = \mathrm{Sim}(w_t, G_{\mathrm{peace}}) - \mathrm{Sim}(w_t, G_{\mathrm{violence}}) \tag{2}$$

A positive bias indicates a stronger association with peace-related concepts, while a negative value indicates a stronger association with violence-related concepts. This allows tracking how the framing of key entities shifts over time.

Table 1: Time periods and article counts for embedding models.

Time Period	Articles
2000–2004	7,151
2005–2008	8,158
2009–2013	8,595
2014–2022	7,111
2023	6,447
2024	9,786

3.3 Sentiment Analysis

We applied three complementary sentiment analysis models: VADER [21], TextBlob [22], and RoBERTa [23]. Each model outputs a sentiment score S for a given text segment (headline, article, or entity context). Sentiment was analyzed for coverage mentioning "Palestinian" vs. "Israeli", temporal trends, headline vs. article body, and named entity contexts.

To compare framing between headlines and article bodies, we define the sentiment differential:

$$\Delta S = S_{\text{headline}} - S_{\text{article}} \tag{3}$$

Here, S_{headline} and S_{article} are the sentiment scores for the headline and article body, respectively. A positive ΔS indicates that the headline is more positive than the article content, while a negative value indicates a more negative headline. This metric quantifies the direction and magnitude of framing differences.

3.4 Emotion Detection

To move beyond polarity, we used the NRC Emotion Lexicon [24] to quantify the prevalence of seven emotions (anger, disgust, fear, grief, joy, sadness, surprise) in relevant contexts. For each emotion e, we compute:

$$E_e = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}(w_i \in e) \tag{4}$$

where n is the number of words in the text, w_i is the i-th word, and $\mathbb{I}(w_i \in e)$ is an indicator function that equals 1 if w_i is associated with emotion e in the lexicon, and 0 otherwise. This yields the proportion of words in the text linked to each emotion, allowing for direct comparison of emotional framing between groups.

3.5 Entity-Level Analysis

We used SpaCy's `en_core_web_sm` model [25] to extract named entities of type `PERSON` from article texts. For each entity, we analyzed sentiment within a window of 15 tokens before and after the first mention, capturing the immediate context. Sentiment for each context was classified using a BERT base multilingual uncased sentiment model, which outputs a 5-point scale from very negative to very positive. This approach enables fine-grained, context-aware sentiment analysis of how individuals are portrayed in the news.

3.6 Headline Framing and Collocation

To assess framing, we compared sentiment distributions in headlines and article bodies, and measured the proportion of headlines with more extreme sentiment than their articles. Additionally, we performed collocation analysis using Pointwise Mutual Information (PMI) to identify frequent word pairs in headlines:

$$\mathrm{PMI}(x, y) = \log \frac{p(x, y)}{p(x)p(y)} \tag{5}$$

where $p(x, y)$ is the probability of words x and y co-occurring, and $p(x)$, $p(y)$ are their individual probabilities. PMI highlights word pairs that appear together more often than expected by chance, revealing linguistic patterns in headline construction.

This integrated methodology enables a comprehensive, multi-level analysis of sentiment, emotion, and framing in conflict coverage, with each equation providing a formal basis for the corresponding analytical step.

4 Results and Discussion

This section presents the main findings of our computational analysis of The Guardian's coverage of the Israeli-Palestinian conflict, focusing on semantic bias, sentiment, emotion, named entity analysis, and framing effects.

4.1 Semantic Bias and Word Embedding Trends

Word2Vec models trained on different time periods reveal clear shifts in the contextual associations of key terms. The average cosine similarity between "Palestinian" and "Israeli" and violence- or peace-related concept groups changes over time, reflecting evolving media framing during conflict periods.

Figure 2 presents the cosine similarities between the terms "Palestinian" and "Israeli" and two groups of words: peace-related and violence-related. Each entity is represented by two lines, one for each word group. The figure captures similarity trends over time without comparing which group is closer; both similarities may rise or fall independently. In contrast, Figure 3 displays the net bias—calculated as the difference between similarity to peace-related and violence-related terms. A positive bias score indicates a stronger association with peace-related terms, while a negative score indicates closer alignment with violence-related terms. The dominant similarity in Figure 2 determines the direction of bias in Figure 3.

As shown in Figure 2, associations with peace-related terms declined for both "Palestinian" and "Israeli" in 2023–2024, indicating heightened conflict framing. Figure 3 shows that "Palestinian" became more closely associated with violence-related terms during war years, with the bias score dropping by over 0.15 in 2023 compared to previous years—indicating a shift toward more negative (violence-aligned) framing. "Israeli" consistently exhibited strong associations with violence, maintaining bias scores below –0.2 across all periods, likely due to frequent references to the Israeli Defense Forces (IDF). The IDF itself displayed an even stronger alignment with violence-related terms than Hamas, reaching a peak negative bias score of less than –0.35 in 2023, suggesting media focus on Israeli military actions.

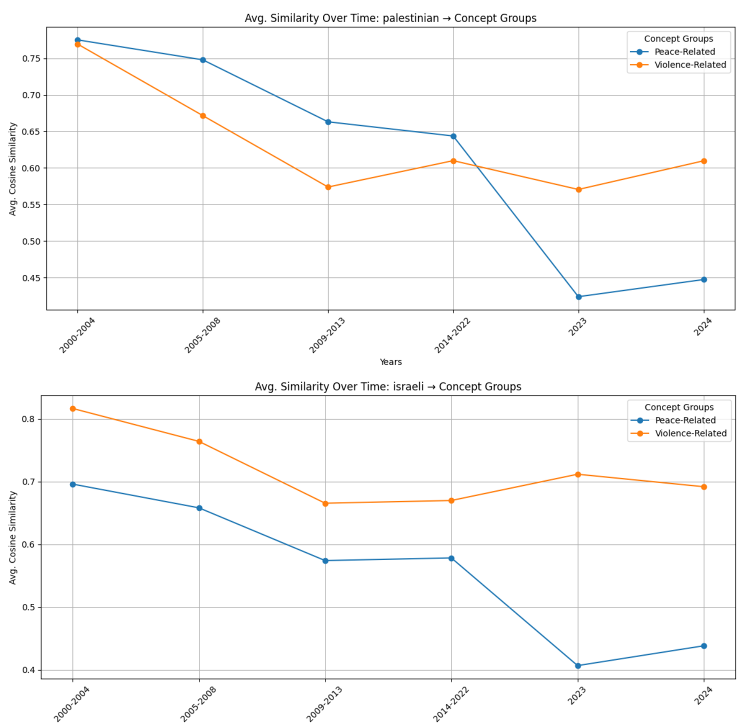
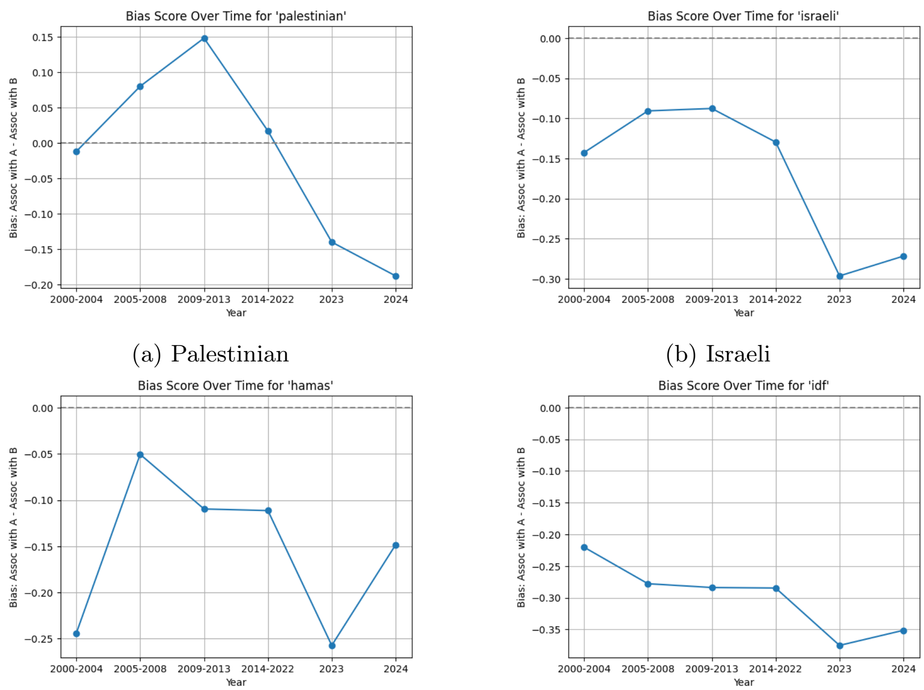


Fig. 2: Cosine similarity trends over time.



(a) Palestinian (b) Israeli

(c) Hamas (d) IDF

Fig. 3: Bias score trends over time.

4.2 Sentiment Patterns and Temporal Dynamics

Multi-model sentiment analysis (VADER, TextBlob, RoBERTa) consistently shows that coverage mentioning Palestinians is more positive than that mentioning Israelis (Figure 4). For example, VADER scores average -0.201 for Palestinians and -0.230 for Israelis, a statistically significant difference of -0.029 ($p < 0.01$). TextBlob shows scores of 0.032 for Palestinian mentions and 0.021 for Israeli mentions, a difference of 0.011. The transformer model (RoBERTa) shows Palestinian mentions scoring -0.427 and Israeli mentions -0.430, a difference of 0.003.

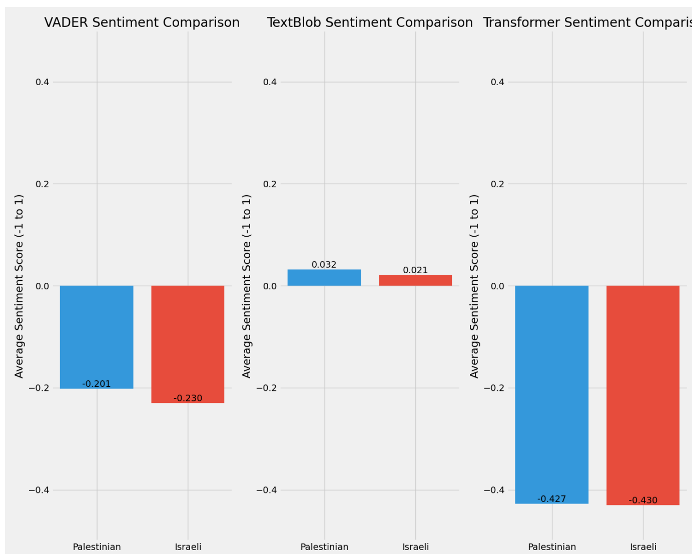


Fig. 4: Overall sentiment comparison between Palestinian and Israeli mentions across three sentiment models.

Figure 5 presents sentiment trends from 2000 to 2024 using VADER, TextBlob, and a Transformer-based model. VADER scores typically range from -0.05 to -0.35, with sharper declines (below -0.30) during major conflicts like the Second Intifada, Gaza Wars, and the 2023–24 Israel-Hamas War. TextBlob outputs weaker, mostly positive sentiment (0.00–0.07), showing minor dips during crises. The Transformer model captures the strongest negativity, with scores often between -0.3 and -0.5, reaching the lowest points during peak conflicts.

Sentiment differentials (Israeli minus Palestinian scores) highlight relative shifts. VADER and Transformer models show notable negative swings (down to -0.075) during major conflicts, suggesting disproportionately negative sentiment toward Israeli mentions. Occasional positive spikes appear around 2006. TextBlob's differential is flatter but trends similarly. Together, these results reflect sentiment deterioration and asymmetry during key political and military events.

4.3 Emotion Analysis: NRC Lexicon and LLM Approaches

We performed emotion analysis using both the NRC Emotion Lexicon and a transformer-based large language model (LLM; `distilbert-base-uncased-go-emotions-student`), restricting the analysis to the set of emotions supported by both methods. Figure 6 shows the NRC-based results, while Figure 7 presents the LLM-based results. Table 2 summarizes the percentage differences between Palestinian and Israeli contexts.

Key findings from the table include:

- **Anger**: NRC shows a higher prevalence in Palestinian contexts (+9.5%), while the LLM shows almost no difference (-0.2%). This suggests that the NRC may overemphasize anger in Palestinian contexts compared to the LLM.
- **Fear**: Both methods agree that fear is more prevalent in Israeli contexts, with NRC showing a slightly larger difference (+8.1%) compared to the LLM (+6.3%).
- **Sadness**: NRC indicates a much higher prevalence in Israeli contexts (+71.2%), while the LLM shows sadness as more prevalent in Palestinian contexts (-10.3%). This stark contrast underscores the differing methodologies of the two approaches.
- **Grief and Joy**: NRC shows significantly higher grief (+62.5%) and joy (+68.9%) in Israeli contexts, while the LLM shows minimal differences (-2.2% for grief and -1.9% for joy).
- **Disgust and Surprise**: The two methods show small but opposite trends, with NRC indicating higher prevalence in Israeli contexts and the LLM showing higher prevalence in Palestinian contexts.

These differences highlight the complementary strengths of the two approaches. The NRC Lexicon provides a lexicon-based, rule-driven perspective, while the LLM offers a data-driven, context-sensitive analysis. Together, they reveal consistent patterns of emotional framing while capturing different levels of granularity. The agreement on fear and the divergence on sadness and anger suggest that media framing may emphasize different emotional narratives depending on the context and methodology used.

4.4 Headline vs. Article Content Framing

A comparative analysis of sentiment across headlines and article bodies reveals consistent editorial framing effects. As visualized in Figure 8, sentiment distributions for headlines exhibit heavier tails and higher kurtosis relative to those of article bodies across all three sentiment models (VADER, TextBlob, RoBERTa), indicating greater polarization. Empirically, 63.1% of headlines are more extreme

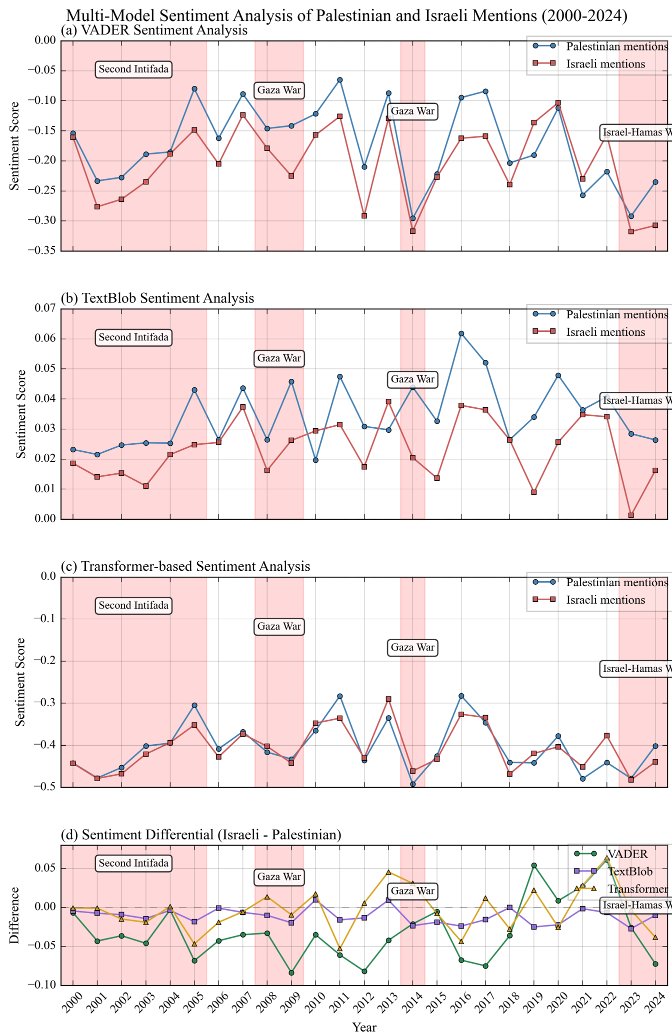


Fig. 5: Multi-Model Sentiment Analysis of Palestinian and Israeli Mentions (2000-2024)

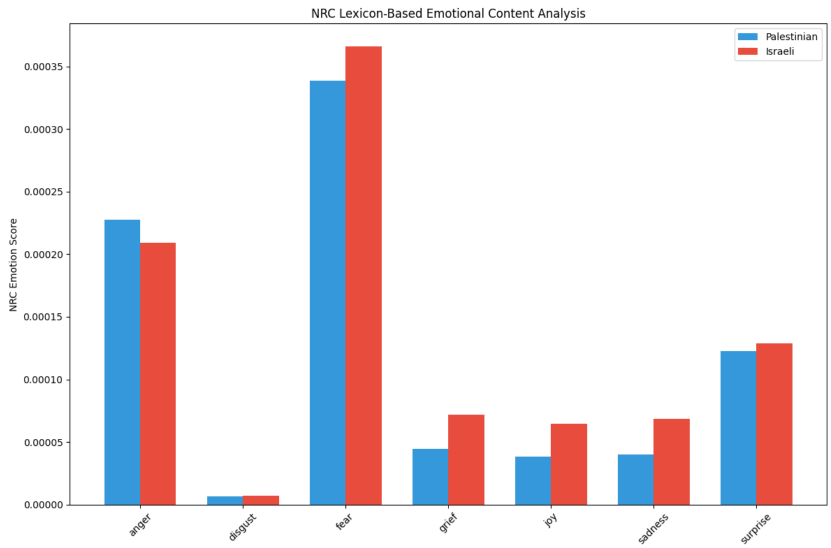


Fig. 6: NRC Lexicon-based emotion analysis for Palestinian and Israeli mentions.

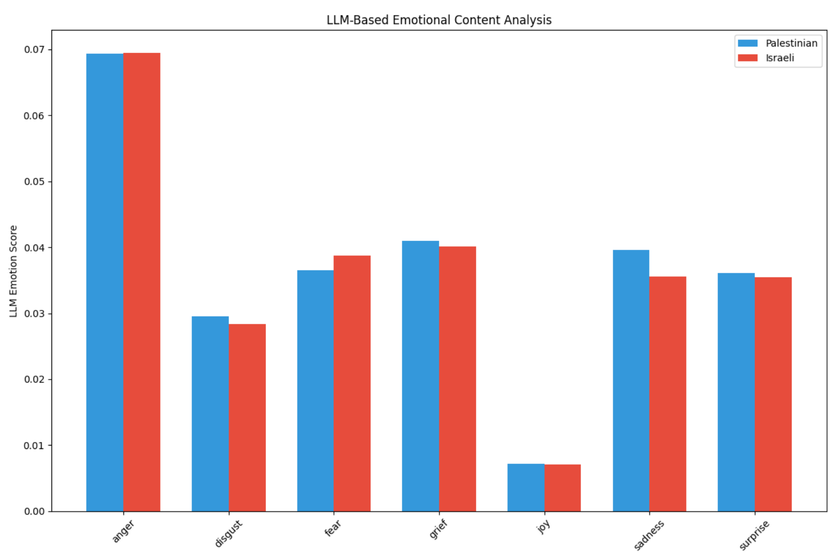


Fig. 7: LLM-based emotion analysis for Palestinian and Israeli mentions.

Emotion	**NRC Diff (%)**	**LLM Diff (%)**
Anger	−8.20	0.17
Disgust	11.10	−4.09
Fear	8.13	6.27
Grief	62.52	−2.22
Joy	68.89	−1.93
Sadness	71.22	−10.30
Surprise	5.13	−1.63

Table 2: Percentage difference between Israeli and Palestinian emotion scores using NRC and LLM-based models. Positive values indicate higher emotion prevalence in Israeli contexts.

in sentiment than their corresponding articles, and 59.5% are more negative, suggesting a systematic tendency toward emotionally charged and negatively valenced headline framing.

This phenomenon is particularly pronounced during periods of heightened geopolitical tension. For instance, during the 2014 Gaza War and the 2023–2024 conflict, the average sentiment differential (headline minus article body) was −0.182, reflecting a consistent editorial shift toward more negative sentiment in headlines. To contextualize this figure, note that the observed sentiment differentials span a wide range: from −1.45 to +1.64 for RoBERTa, −1.06 to +1.19 for VADER, and −1.12 to +1.03 for TextBlob. Thus, a shift of −0.182 lies well within the upper quartile of negative deviations and represents a substantively meaningful editorial bias.

In contrast, during coverage of peace initiatives, the average sentiment differential was comparatively modest at +0.085, reinforcing the asymmetry: headline sentiment is amplified far more strongly in the direction of negativity during conflict coverage than it is toward positivity in peace-related reporting. These findings align with prior research on media sensationalism and framing theory, underscoring the critical role of headlines in shaping affective reader perception independently of article content.

4.5 Media Sentiment Analysis of Key Figures in the Israeli-Palestinian Conflict

An analysis of the top 10 most-mentioned names revealed Ariel Sharon (6,393 mentions) and Yasser Arafat (6,524) as the most frequently referenced individuals, followed by Mahmoud Abbas (4,944) and Benjamin Netanyahu (4,690). Sharon, known for his hardline policies and the 2005 Gaza withdrawal, and Arafat, symbolic leader of Palestinian resistance, were central to the discourse. Abbas, Arafat's successor, advocates for a two-state solution but faces criticism over limited authority. Netanyahu, Israel's longest-serving Prime Minister, is noted for his security focus and settlement expansion.

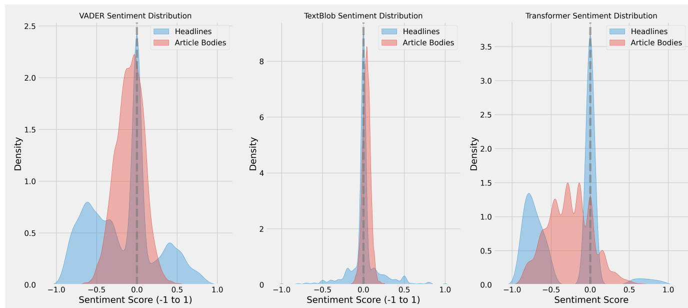


Fig. 8: Sentiment distribution comparison between headlines and article content across all three sentiment models.

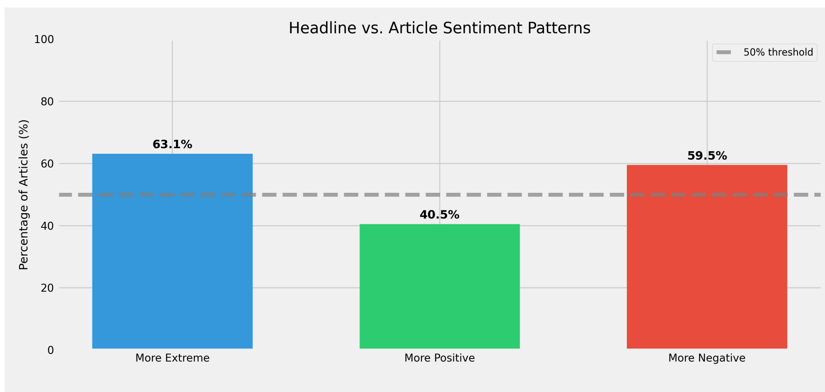


Fig. 9: Distribution of headlines displaying more negative, more positive, or more extreme sentiment than their corresponding article content.

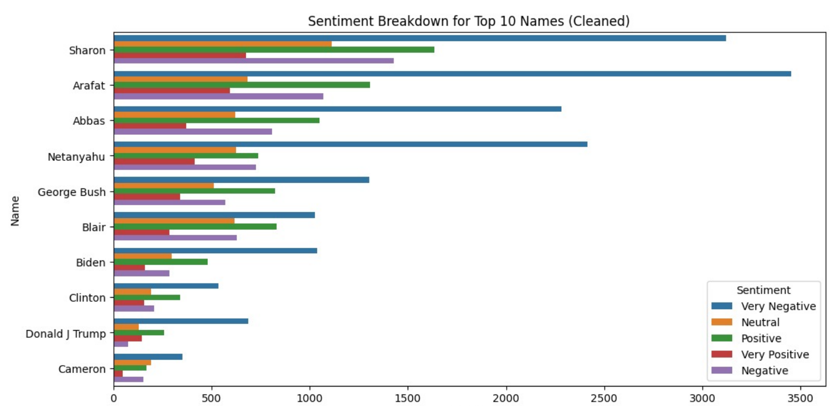


Fig. 10: Sentiment distribution associated with the top 10 mentioned figures.

As shown in Figure 10, sentiment analysis over a 30-word context window revealed three main patterns. First, negative sentiment was dominant: Sharon and Arafat had 45% and 49% Very Negative mentions, respectively, with 3–5 times more Very Negative than Very Positive mentions. Second, positive sentiment was scarce, with no figure receiving over 25% positive coverage. Sharon led in total positive mentions (2,162), while Netanyahu and Abbas had the lowest positive shares (12% and 15%). Third, neutral sentiment remained consistent, ranging from 18% to 32%, with Sharon having the highest neutral count (1,039). Overall, media narratives skew heavily negative for prominent figures while maintaining a neutral reporting baseline.

4.6 Summary of Key Findings

Our analysis demonstrates that:

1. Media framing shifts toward violence during conflict, as shown by word embeddings and bias scores (e.g., IDF bias score reaches a peak negative value of less than -0.35 in 2023).
2. Sentiment analysis shows that coverage mentioning Palestinians is consistently more positive than that mentioning Israelis (VADER scores average -0.201 for Palestinians and -0.230 for Israelis, a statistically significant difference of -0.029).
3. Emotional framing differs: The NRC Analysis shows that Israeli contexts emphasize fear (+8.1%) and anger (+8.9%). The LLM-based analysis further reveals that anger is 0.17% higher in Palestinian contexts, while fear is 6.3% higher in Israeli contexts, providing additional granularity to the emotional framing.
4. Headlines are systematically more polarized and negative than article bodies, with 63.1% of headlines more extreme and 59.5% more negative; the sentiment gap reaches -0.182 during conflict peaks.
5. Media coverage skews negative: Sharon and Arafat had the most mentions (6,393 and 6,524) with high Very Negative rates (45% and 49%). Positive sentiment stayed below 25%, and neutral coverage ranged from 18–32%, showing a focus on criticism with limited balance.

These findings highlight the importance of computational approaches for uncovering subtle but systematic patterns in media coverage of protracted conflicts.

5 Conclusion

This study provides a comprehensive computational analysis of The Guardian's coverage of the Israeli-Palestinian conflict, revealing persistent and systematic patterns in language, sentiment, emotion, and framing across 47,248 articles. Our multimodel sentiment analysis revealed that coverage mentioning Palestinians is consistently more positive than that mentioning Israelis, with VADER scores averaging -0.201 for Palestinians and -0.230 for Israelis, a statistically significant difference of -0.029 ($p < 0.01$). Similar patterns appeared in TextBlob (0.032 vs. 0.021) and RoBERTa (-0.427 vs. -0.430).

Emotion analysis showed fear as the dominant emotion for both groups. The NRC analysis revealed that fear was 8.1% higher in Israeli contexts, while the LLM analysis showed a smaller difference of 6.3%. Anger showed divergent trends, with NRC indicating 9.5% higher prevalence in Palestinian contexts, while LLM showed almost no difference (-0.2%). Sadness was consistently higher in Israeli contexts according to NRC (+71.2%), but LLM showed it as more prevalent in Palestinian contexts (-10.3%). These complementary findings highlight the robustness of combining lexicon-based and transformer-based methods for emotion detection, as they reveal consistent patterns while capturing different levels of granularity.

Headlines demonstrated greater polarization than article bodies in 63.1% of cases and were more negative in 59.5% of cases, with the sentiment gap reaching 0.182 during intense conflict periods. Word2Vec analysis showed both "Palestinian" and "Israeli" terms shifted toward violence-related associations during war years, with the IDF exhibiting the strongest alignment with violent contexts. These quantitative results highlight how computational methods can expose subtle media biases in geopolitical conflicts, offering actionable insights for media studies and public discourse. Future research can extend this study by including multiple media outlets across different ideological or geographic orientations, exploring emerging LLM-based sentiment and bias classifiers, and analyzing multilingual coverage to offer a more comprehensive picture of how different communities are informed about the conflict.

References

1. R. M. Entman, “Framing: Toward clarification of a fractured paradigm,” *J. Commun.*, vol. 43, no. 4, pp. 51–58, Dec. 1993. DOI: 10.1111/j.1460-2466.1993.tb01304.x.
2. T. Gitlin, *The Whole World Is Watching: Mass Media in the Making and Unmaking of the New Left.* Berkeley, CA, USA: Univ. California Press, 1980.
3. S. Iyengar, *Is Anyone Responsible?: How Television Frames Political Issues.* Chicago, IL, USA: Univ. Chicago Press, 1991.
4. G. Philo and M. Berry, *Bad News from Israel.* London, UK: Pluto Press, 2004.
5. S. D. Ross, “Framing of the Palestinian-Israeli conflict in thirteen months of New York Times editorials surrounding the attack of September 11, 2001,” *Conflict Commun. Online*, vol. 2, no. 2, pp. 1–11, 2003.
6. J. E. Richardson, *Analysing Newspapers: An Approach from Critical Discourse Analysis.* London, UK: Palgrave Macmillan, 2007.
7. H. Felix “Media bias, the social sciences, and NLP: Automating frame analyses to identify bias by word choice and labeling,” in *P*roceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, 2020, pp. 79–87.[Online].
8. A. Mohammad, H. Naeemul,, “A survey of computational framing analysis approaches.” in *In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing.*, Abu Dhabi, 2022, pp. 9335–9348.

9. K. Yowei, Y. C. C. Kenneth, “Examining media bias and geopolitical proxy framing effects on media representations of the Palestinian–Israeli conflict in Taiwan: A computational framing analysis,” *Journal of Arab & Muslim Media Research.*, Feb. 2025. DOI: https://doi.org/10.1386/jammr_00093_1.
10. T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” in *Proc. Int. Conf. Learn. Representations (ICLR)*, Scottsdale, AZ, USA, May 2013, pp. 1–12. [Online]. Available: https://arxiv.org/abs/1301.3781.
11. A. Deprez and K. Raeymaeckers, “Framing the first and second intifada: A longitudinal quantitative research design applied to the Flemish press,” *Eur. J. Commun.*, vol. 25, no. 1, pp. 3–23, 2013. DOI: 10.1177/0267323109354224.
12. M. E. Sharkar, “Sentiment Analysis of Israel-Palestine Conflict Comments Using Sentiment Intensity Analyzer and TextBlob,” in *Proc. 15th Int. IEEE Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jun. 2024, DOI: 10.13140/RG.2.2.22456.35847.
13. A. Liyih, S. Anagaw, M. Yibeyin, *et al.*, “Sentiment analysis of the Hamas-Israel war on YouTube comments using deep learning,” *Sci. Rep.*, vol. 14, p. 13647, 2024. DOI: 10.1038/s41598-024-63367-3.
14. M. U. Razaq and N. Naeem, “Corpus-assisted sentiment analysis of news headlines on Palestine-Israel conflict: A computational approach,” *Glob. Foreign Policies Rev.*, vol. VII, no. IV, pp. 59–67, Jan. 2025. DOI: 10.31703/gfpr.2024(VII-IV).07.
15. W. Abuasaker, M. Sánchez, J. Nguyen, N. Agell, N. Agell, and F. J. Ruiz, “A comparative analysis of European media coverage of the Israel–Gaza war using hesitant fuzzy linguistic term sets,” *Mach. Learn. Knowl. Extr.*, vol. 7, no. 1, p. 8, 2025. DOI: 10.3390/make7010008.
16. Pew Research Center, "Emotions, news and knowledge about the Israel-Hamas war," Pew Research Center: Journalism & Media, 2024.
17. A. Guerra, M. Lepre, and O. Karakus, “Quantifying extreme opinions on Reddit amidst the 2023 Israeli-Palestinian conflict,” arXiv preprint, arXiv:2412.10913, 2024. [Online]. Available: https://arxiv.org/abs/2412.10913.
18. L. Asmus, "Framing of Conflict Reporting in the Israel-Hamas Conflict in German Online Newspaper Articles," DIVA Portal, 2024.
19. P. Garcia Corral, H. Bechara, K. Manohara, and S. Jankin, “The missing cause: An analysis of causal attributions in reporting on Palestine,” in *Proc. 1st Int. Workshop Nakba Narratives Lang. Resour.*, Abu Dhabi, Jan. 2025, pp. 103–113. [Online]. Available: https://aclanthology.org/2025.nakbanlp-1.11/.
20. G. McTigue, “Media bias in covering the Israeli-Palestinian conflict: With a case study of BBC coverage and its foundation of impartiality,” Honors Thesis, Renée Crown Univ. Honors Program, Syracuse Univ., Syracuse, NY, USA, May 2011. [Online]. Available: https://surface.syr.edu/honors_capstone/300.
21. C. J. Hutto and E. Gilbert, “VADER: A parsimonious rule-based model for sentiment analysis of social media text,” in *Proc. 8th Int. Conf. Weblogs Social Media*, Ann Arbor, MI, USA, 2014, pp. 216–225.
22. S. Loria, “TextBlob: Simplified text processing,” 2018. [Online]. Available: https://textblob.readthedocs.io/
23. Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, “RoBERTa: A robustly optimized BERT pretraining approach,” arXiv preprint, arXiv:1907.11692, 2019.
24. S. M. Mohammad and P. D. Turney, “Crowdsourcing a word-emotion association lexicon,” *Comput. Intell.*, vol. 29, no. 3, pp. 436–465, 2013.

25. M. Honnibal and I. Montani, "spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing," 2017. [Online]. Available: https://spacy.io