

# China's Stance on Rohingya Refugees Issues in The Local Newspaper Through Corpus Sentiment Classification

Minjie Chen<sup>1</sup>, Wei Lun Wong<sup>2</sup>, Warid Mihat<sup>3</sup>, Henri E. Lemana II<sup>4</sup> & Jing Liu<sup>1</sup>

<sup>1</sup> School of Foreign Languages, Mianyang Teachers' College, Sichuan, P. R. China

<sup>2</sup> Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia

<sup>3</sup> Academy of Language Studies, Universiti Teknologi MARA, Kelantan, Malaysia

<sup>4</sup> School of Languages and General Education, Walailak University, Thailand

Correspondence: Wei Lun Wong, Faculty of Education, Universiti Kebangsaan Malaysia, Selangor, Malaysia. E-mail: colin\_me-and-you@live.com.my

Received: December 11, 2024

Accepted: March 7, 2025

Online Published: April 25, 2025

doi:10.5430/wjel.v15n7p42

URL: <https://doi.org/10.5430/wjel.v15n7p42>

## Abstract

The predicament of the Rohingya refugees has garnered significant attention in China. The present study scrutinises the newspaper's language alignment to the representation of China's stance on the Rohingya refugees. The primary objectives are twofold: firstly, to analyse the top ten salient nouns, verbs, and adjectives, and secondly, to evaluate the sentimentality of the aforementioned salient vocabulary and phrases. The present study employs a mixed-method approach with a corpus-driven research design. The corpus was procured from China Daily, comprising a total of 78 newspaper reports with 24,769 words and 1,013 sentences. The utilisation of Sketchengine and Atlas.ti, was selected. A wordlist of vocabulary has been produced for top ten salient nouns, verbs, and adjectives. The vocabulary was compared with the British National Corpus for keywords. Subsequently, the keywords were employed for trigrams and qualgrams. The presented materials consisted of concordances pertaining to phrases. The results were analysed sentimentally. The top ten salient nouns, verbs, and adjectives included Myanmar, people, Rohingya, be, have, say, more, human, and international. Then, keywords were compared to the reference corpus to select significant trigrams and qualgrams produced by Myanmar, Rohingya, migrants, Bangladesh and humanitarian. Sentimental analysis was performed on 60 linguistics items. Referring to the findings, the nation of China had a neutral stance. The findings suggest that scholars and politicians may benefit from a more empirical approach to analysing a nation's stance, as opposed to relying solely on subjective interviews, as reported language can serve as a factual basis.

**Keywords:** China, corpus-driven, Rohingya refugees, sentimentality, vocabulary

## 1. Introduction

As millions of individuals have been compelled to vacate their abodes and seek asylum and sanctuary in foreign territories, the contemporary global refugee crisis has engendered fervent debates and discussions among state leaders around the world (Field et al., 2021). Amidst the multifarious demographic cohorts impacted by the ongoing crisis, the Rohingya refugees hailing from Myanmar have been subjected to exceptionally dire predicaments. Against the aforementioned backdrop, the stance and policies adopted by various nations with respect to the Rohingya crisis carry significant weight (Kironksa & Peng, 2021).

China, being a significant global force and a pivotal regional actor in Southeast Asia, wields considerable sway in moulding the dialogue and consequences pertaining to the Rohingya predicament (Yoshikawa, 2022). Although China is committed to resolving matters related to refugees, little is known about how this information is disseminated to the public through newspapers. Therefore, it is suggested that a study be conducted to investigate the interconnection between China's current refugee policy and the sentimentality portrayed by tabloids or newspapers in the public sphere.

To this end, the present study employs a corpus linguistics methodology with a corpus-driven research design to elucidate the salient vocabulary, phrases and sentiment utilised by China in connection with the Rohingya refugee crisis. Corpus linguistics is a research methodology that involves the systematic analysis of extensive collections of written or spoken language, commonly referred to as corpora (Alsuhaibani, 2022; Wong et al., 2023a). This approach aims to uncover valuable insights into the patterns, usage, and underlying ideologies of language. Through this approach, this study hopes to understand how words in the newspapers are used to enforce or reject the policy practised by China government. The research questions were:

1. What are the top ten salient nouns, verbs, and adjectives obtained from the collected corpus?
2. What is the sentimentality of the aforementioned salient vocabulary and phrases?

## 2. Literature Review

### 2.1 Rohingya Refugee Crisis

The year 2017 witnessed a marked increase in the level of violence, which in turn acted as a trigger for the most noteworthy surge of Rohingya displacement in contemporary times. The Myanmar military's severe crackdown, which has been marked by extensive accounts of extrajudicial killings, sexual violence, and arson attacks on Rohingya villages, has resulted in a mass migration of refugees towards neighbouring countries, particularly Bangladesh. The velocity and magnitude of the displacement posed a formidable humanitarian predicament, exerting immense pressure on the resources and capabilities of the host nations and global relief agencies (Ware et al., 2022).

The ongoing Rohingya refugee crisis has engendered pressing inquiries pertaining to the domains of human rights, ethnic strife, and statelessness (Kironka & Peng, 2021). The issue at hand presents notable complexities with regards to expeditiously furnishing humanitarian aid, guaranteeing the safeguarding of displaced persons, and tackling the underlying origins of the predicament (Ware et al., 2022). Moreover, the crisis bears significant implications for the stability of the region, given the neighbouring countries' struggles with accommodating the substantial influx of displaced Rohingya, while Myanmar confronts global censure for its handling of the Rohingya populace (Tan, 2021).

China's apprehension regarding the Rohingya crisis is substantiated by a multitude of factors. China's acknowledgment of the Rohingya crisis as a subject of concern emanates from its dedication as a signatory to the 1951 Refugee Convention. In accordance with international humanitarian principles, it is incumbent upon China to attend to the pressing predicament confronting refugees. Significantly, it is worth noting that the ongoing crisis in Myanmar and Bangladesh holds significance that extends beyond its local contexts. This is primarily due to the geographical proximity of these nations to China, thereby giving rise to regional implications that demand attention and cannot be disregarded. The substantial influx of Rohingya refugees and the consequent destabilisation in the region may give rise to diverse ramifications, encompassing potential security vulnerabilities and disruptions to trade and economic endeavours. In light of China's burgeoning influence and substantial investments in the region, it is imperative for the nation to assume a proactive stance in effectively addressing humanitarian crises, thereby safeguarding its esteemed reputation and credibility on the global stage. Moreover, in its capacity as a steadfast constituent of the United Nations Security Council, China bears the onus of fostering tranquilly, equilibrium, and the safeguarding of human rights, thereby accentuating the significance of its constructive involvement in the resolution of the Rohingya predicament.

### 2.2 Recent Rohingya Studies in China

Between 2013 and 2023, there have been around 34 studies conducted and published on this topic in China. Generally, those studies can be categorized into four themes with different objectives. Table 1 below elucidates the categories.

Table 1. Past themes

Past themes	Researchers
Myanmar's geopolitics and politics	Han, 2021; Mostafanezhad et al., 2022; Xue, 2021; Morada, 2021; Mallick, 2020; Han, 2020; Houston, 2020; Kelemen & Turcsányi, 2020; Park, 2020; Zin, 2020; Hussain, 2019; Taufiq, 2019; Casolari, 2018
Human rights and democracy	Rasul, 2022; Tan, 2021; Tieri, 2021b; Ghosal, 2013
Rohingya Crisis	Hossain, 2023; Martel & Glas, 2022; Shi et al., 2022; Kamal & Kamal, 2022; Zahed, 2022; Karim, 2021; Tieri, 2021a; Uesugi et al., 2020; Palakodety et al., 2020a; Fink, 2019; Kipgen, 2019; Fumagalli, 2018; Mallampalli, 2023; Malji, 2021
Rohingya's Economy	Bennett & Faxon, 2021; Møller, 2021; Islam et al., 2018

This study provided a detailed examination of several pivotal themes relating to Myanmar, employing a critical analysis of existing research to deepen the understanding of each area. These themes included Myanmar's geopolitics and politics, human rights and democracy, the Rohingya crisis, and the economic conditions affecting the Rohingya population. The geopolitical and political landscape of Myanmar has been rigorously analysed by researchers, who have explored the complex interplay between national sovereignty and international pressure, the persistent ethnic conflicts, and Myanmar's strategic relationships with both neighboring countries and global powers. Notable scholars such as Han, (2021), Mostafanezhad et al. (2022); Xue (2021), Morada (2021), Mallick (2020), Han (2020), Houston (2020) Kelemen & Turcsányi (2020) Park (2020), Zin (2020), Hussain (2019), Taufiq (2019), and Casolari (2018) have likely discussed the balancing act between influences from China and pressures from Western nations. These contributions provide a nuanced view of Myanmar's international positioning, though they may not fully represent the perspectives of local ethnic groups, including the Rohingya, within these political discussions.

The second theme addressed the serious challenges and criticisms concerning human rights abuses and democratic deficits in Myanmar. Works by authors like Rasul (2022) and Ghosal (2013) might emphasise the systemic violations and the responses from the international community. Tan (2021) and Tieri (2021b) could critique the effectiveness of international human rights frameworks in changing the realities on the ground in Myanmar. While these studies underscore the urgent need for robust democratic institutions, they might also overlook the complexities involved in implementing democratic reforms within a country traditionally dominated by military governance.

Regarding the Rohingya crisis, this remains a critical humanitarian issue, with scholars such as Hossain (2023), Martel and Glas (2022), Shi

et al. (2022), Kamal and Kamal (2022), Zahed (2022), Karim (2021), Tieri (2021a), Uesugi et al. (2020), Palakodety et al. (2020a), Fink (2019), Kipgen (2019), Fumagalli (2018), Mallampalli (2023), and Malji (2021) likely discussing the immediate human rights violations and the international legal ramifications.

Finally, the economic dimensions of the Rohingya's life are explored, focusing on aspects such as labor market integration and the economic contributions of the Rohingya in exile. Researchers like Bennett & Faxon (2021) and Møller (2021) have studied these issues, and Islam et al. (2018) might detail the economic impacts of displacement on both the Rohingya and their host communities. These analyses are vital for understanding the economic challenges faced by refugees, yet they frequently fail to fully capture the economic potential of the Rohingya, focusing more on immediate survival strategies rather than long-term economic integration.

Table 2. Research methods employed

Research Instruments	Researchers
(Participant) Observation	Mostafanezhad et al., 2022, Kamal & Kamal, 2022,
Geospatial Data	Bennett & Faxon, 2021
Survey	Shi et al., 2022
Interviews	Martel & Glas, 2022; Mostafanezhad et al., 2022; Kamal & Kamal, 2022; Zahed, 2022; Tan, 2021
Document Analyses	Martel & Glas, 2022; Mallampalli, 2023; Hossain, 2022; Rasul, 2022; Tan, 2021; Bennett & Faxon, 2021; Han, 2021; Karim, 2021, Tieri, 2021a; Bennett & Faxon, 2021; Xue, 2021; Malji, 2021; Morada, 2021; Mallick, 2020; Han, 2020; Park, 2020; Tieri, 2021b; Møller, 2021; Kelemen & Turcsányi, 2020; Uesugi et al., 2020; Zin, 2020; Palakodety et al., 2020a; Fink, 2019; Kipgen, 2019; Hussain, 2019; Taufiq, 2019; Fumagalli, 2018; Casolari, 2018; Ghoshal, 2013
Open Ended Questionnaire	Islam et al., 2018

The decision to utilise the corpus method over other methodologies listed in Table 2 of the provided document necessitates a critical evaluation of the unique advantages corpus linguistics brings, particularly in the context of media discourse analysis. Observational methods, including direct and participant observation used by researchers such as Mostafanezhad et al. (2022) and Kamal & Kamal (2022), offer qualitative insights into behaviours and interactions within natural settings but suffer from observer bias and challenges in replicating conditions. In contrast, the corpus method allows for a more detached and replicable analysis across a vast array of documented interactions, making it ideal for systematic examination of media discourse.

Other methods such as geospatial data analysis, employed by Bennett & Faxon (2021), provide powerful spatial insights but fail to capture the linguistic nuances crucial for understanding media narratives and their impact on public perception and policy. Similarly, surveys and open-ended questionnaires, utilised by Shi et al. (2022) and Islam et al. (2018), though effective for gathering a broad range of responses, rely heavily on self-reporting and lack the linguistic depth necessary for media text analysis. Interviews, offering in-depth individual perceptions and utilized by researchers like Martel & Glas (2022), focus more on subjective experiences rather than the objective textual analysis essential for studying media language on a broader scale.

Document analysis, another common method, explores written materials to discern themes and patterns but often lacks the quantitative rigor and scalability afforded by corpus methods, which can systematically analyse larger volumes of text. Corpus linguistics not only provides this scalability but also introduces a quantitative rigor through statistical language analysis, enhancing the detection of patterns and trends that smaller-scale or manual studies might miss.

Additionally, corpus methods stand out for their replicability, reducing subjective interpretation and enhancing the reliability of findings, crucial for verifying research across studies. The integration with linguistic software tools like SketchEngine and Atlas.ti further facilitates detailed analyses such as keyword comparisons and sentiment analysis, providing deeper insights into how language frames issues. In conclusion, the corpus method was selected for its systematic, scalable, and replicable approach to language analysis across substantial text volumes, offering broad yet precise linguistic insights. This approach is particularly valuable in analyzing media representations, where the subtleties and patterns of language use are critical for interpreting media influence on public opinion and policy formulation.

### 2.3 Corpus Linguistics, Sentiment Analysis and Rohingya Refugee

Corpus linguistics, as a research methodology, encompasses the meticulous examination of expansive compilations of written or spoken texts, commonly referred to as corpora (Brandenburg-Weeks & Abalkheel, 2021; Wong et al., 2023b; Wong et al., 2024). The primary focus of this field of study pertains to the examination and analysis of language patterns, frequencies, and usage. The overarching objective is to acquire a deeper understanding of linguistic phenomena and discourse through meticulous investigation.

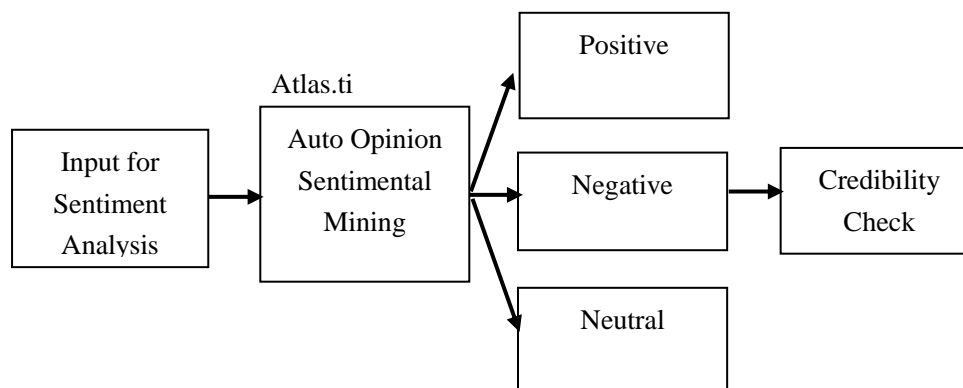


Figure 1. Theoretical framework (sentiment analysis)

Sentiment analysis, also known as sentiment classification, is a methodological approach employed within the domains of natural language processing, computational linguistics, and text mining. Its primary objective is to ascertain and discern the underlying sentiment or emotional tone that is implicitly conveyed within a given textual corpus. In recent years, a multitude of sentiment analysis models have been put forth by scholars in the field of human-computer interaction (Xu et al., 2023; D’Aniello et al., 2022; Tsai et al., 2022), information retrieval, and the analysis of multimodal signals (Cambria et al., 2017) across various domains, including but not limited to tourism, finance, and shopping. The study conducted by Hu et al. (2022) elucidates the viability and logicity of concurrently modelling emotion and cognition from a psychological perspective.

In accordance with the theoretical framework presented in Figure 1, the input data was subjected to sentiment analysis. In the context of this particular study, the salient vocabulary and phrases were duly identified and examined. Utilising the software tool Atlas.ti, automated sentiment analysis was performed on lexical expressions, classifying them into distinct categories of positivity, negativity, or neutrality. Prior to conducting a credibility assessment, the researchers engaged in a preliminary phase of manual opinion sentiment analysis. This involved scrutinizing the semantic connotations of the vocabulary and phrases that had been automatically categorised, thereby aiming to enhance the precision and accuracy of the subsequent findings. Finally, it is worth noting that these findings underwent a rigorous process of verification, as they were thoroughly examined and evaluated by a panel of three distinguished linguistics scholars, renowned for their expertise in the field.

In the context of the Rohingya issue, the application of corpus linguistics has been employed in two scholarly investigations (these studies however were not conducted in China as discussed in 2.2) to analyse the language and discourse pertaining to the aforementioned humanitarian crisis. The studies have made significant contributions to our comprehension of the manner in which the Rohingya crisis is portrayed, structured, and deliberated upon within diverse sources.

Houston (2020) undertakes a corpus analysis of the official discourse propagated by the government of Myanmar pertaining to the conflict in the region of Rakhine. The scholarly analysis posits that Myanmar can be characterised as a transitional state grappling with a tenuous grasp on political legitimacy. Furthermore, it asserts that the government's official narrative concerning the conflict in Rakhine ultimately falls short in effectively establishing a sincere departure from the preceding military regime. The present discourse posits that the governmental endeavours to reconcile national and international pressures, while concurrently assuming a neutral stance as an arbiter, manifest a betrayal of its accountability for the predicament faced by the Rohingya population. Furthermore, it contends that the efficacy of external pressure in influencing Myanmar's conduct shall be constrained.

In their scholarly endeavour, Palakodety et al. (2020a) posit an innovative methodology for the identification and discernment of comments that espouse support for the Rohingyas within the context of YouTube videos pertaining to the ongoing refugee crisis using corpus linguistic. The authors espouse the utilisation of artificial intelligence in discerning benevolent discourse with the aim of fortifying the internet's security for communities that have been historically marginalised. The paper introduces a considerable collection of YouTube video comments and employs a fusion of various Active Learning methodologies alongside an innovative active sampling strategy grounded in nearest-neighbours within the comment-embedding space. This approach is utilised to develop a classifier capable of identifying comments that advocate for the Rohingyas amidst a larger volume of derogatory and impartial comments. The aforementioned study culminates in the assertion that the suggested methodology possesses the potential to identify instances of assistance-oriented discourse in alternative scenarios.

#### 2.4 Connecting Government Stance and Sentiment Played in the Public

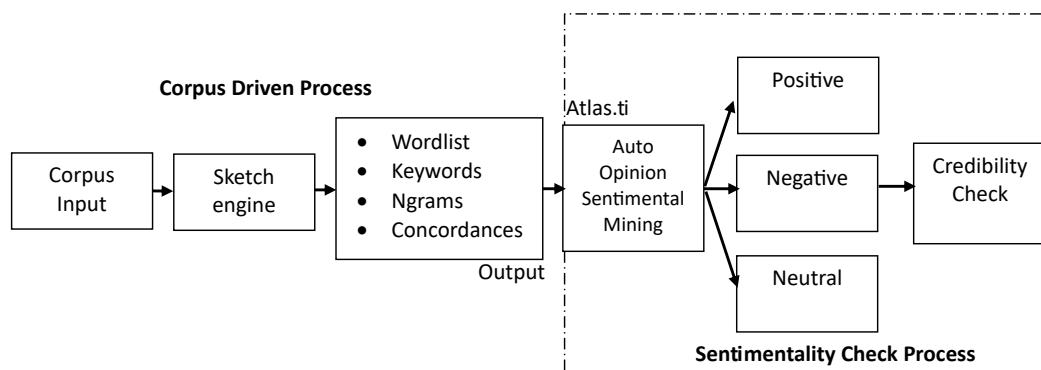


Figure 2. Conceptual framework

In order to establish a correlation between words and policy via sentiment analysis, a series of pivotal procedures must be executed. To commence, it is imperative for researchers to identify the source of the corpus under investigation. This input could derive from newspaper articles, letters, magazines, just to name a few. Once the corpus data have been identified, they will be utilised as the fundamental basis for a meticulous analysis driven by the utilisation of SketchEngine. This analytical approach facilitates the generation of wordlists encompassing the ten (the number differs according to the objectives of the study) most prominent words according to the categories under investigation. The identification of keywords was also cross checked through a comparative analysis between the corpus and the reference corpus, namely the British National Corpus (BNC). The process involved in this study encompassed the identification and subsequent selection of comparable keywords from a pool of comparisons. These trigrams and qualgrams, which were derived from the keywords, were then utilised to generate concordances that are pertinent to the research at hand.

To evaluate sentiment, an examination was conducted on the noteworthy vocabulary and the top 5 ngrams that were identified. This analysis was carried out using Atlas.ti, a software tool that integrates automatic opinion-mining functionalities. This process will sort the words into their sentimentality categories. In order to uphold the utmost precision, a manual phase was undertaken, wherein three esteemed linguistic experts were engaged to authenticate the sentiment analysis. The ultimate outcomes were scrutinised in order to ascertain their congruity or disparity for the purpose of achieving the objectives of the study. The present comprehensive approach endeavours to establish a connection between the linguistic expressions found in media coverage and the official policy of China, thereby elucidating the sentiment conveyed in relation to the Rohingya crisis.

### 3. Method

#### 3.1 Research Design

The present study utilised a mixed-method research method, as outlined by Creswell & Creswell (2018). The quantitative phase was initiated through the utilisation of a corpus-driven research design and SketchEngine as the corpus tool. The forthcoming presentation shall encompass a comprehensive analysis of the ten most prominent nouns, verbs, and adjectives, accompanied by their respective frequency (f), frequency per million, Document Frequency (DOCF), and Relative Document Frequency (Relative DOCF). The subsequent step entails the utilisation of keywords to facilitate a comprehensive comparison with a reference corpus, specifically the British National Corpus (BNC), in order to conduct further analysis on ngrams. The qualitative analysis was performed by means of scrutinising concordances for each salient vocabulary and phrase that was chosen during the concluding stage. Ultimately, the entirety of the noteworthy lexicon and expressions have been subjected to sentiment analysis, specifically evaluating positive, negative, and neutral orientations, employing the software tool Atlas.ti. However, a subsequent iteration of manual analysis is undertaken in order to further validate the sentimentality of the data.

#### 3.2 News Sampling

The corpus pertaining to the Rohingya community was compiled by means of amassing a total of 78 newspaper articles, as reported by China Daily, spanning the period from 2012 to 2022. The search parameters utilised by the user pertained to the terms “Rohingya” and “China”. The Rohingya corpus comprised a total of 28,162 tokens, 24,769 words and 1,013 sentences, which were deemed suitable for subsequent analysis, as mentioned by Chen et al. (2023), the minimum words for a corpus was 8,000.

#### 3.3 Data Analysis Procedures

Firstly, the Sketchengine tool was executed. The aforementioned wordlist has yielded a selection of salient vocabulary. However, it is pertinent to note that solely the foremost decile of salient nouns, verbs, and adjectives have been taken into consideration. Frequency (f), frequency per million (henceforth, RF), document frequency (henceforth, DOCF) and relative DOCF were computed. The aforementioned vocabularies were subjected to a comparative analysis vis-à-vis the British National Corpus, with the aim of identifying salient keywords. The selection process would entail the identification of the keywords that exhibit similarity and have surfaced within the top 10 salient vocabulary.

Subsequent to the collection of salient vocabulary, an analysis of n-grams was conducted. The investigation focused on analysing the top five salient trigrams and qualgrams with regard to their expansion and resemblances related to the salient vocabulary identified. According to Leech et al. (2014), it was deemed advantageous to utilise a total of five words on both the left and right sides of a given sentence in order to enhance the overall presentation of language. The aforementioned vocabulary and phrases were subjected to a semantic analysis using Atlas.ti. The stance exhibited can be categorised as either positive, negative, or neutral. A visual representation is presented herewith to facilitate comprehension.

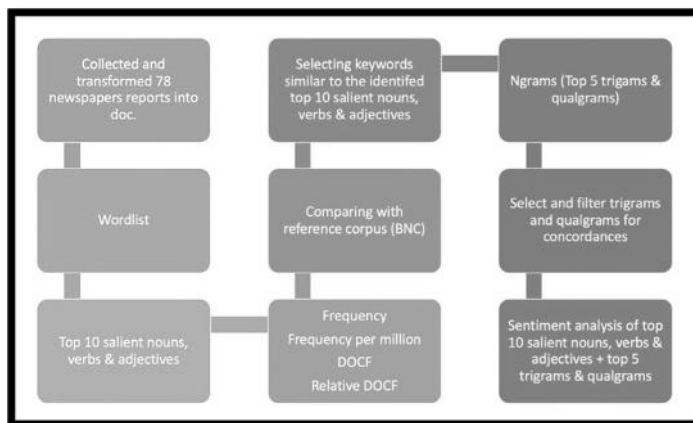


Figure 3. Data analysis procedures

### 3.4 Scopes of the Study

The scope of this study is confined solely to the examination of China Daily, thereby excluding any consideration of alternative newspapers or media sources. The objective at hand is to ascertain the extent to which linguistic alignment plays a role in shaping the newspaper's portrayal of China's position on the matter of Rohingya refugees. Table 3 below summarize the scope and parameter of this study.

Table 3. Scope and parameter of this study

Scope	Parameter
Years	2012-2022 (78 articles)
Materials	Newspapers only
Company	China Daily
Method	Corpus analysis only

## 4. Results

### 4.1 Research Question 1: What are the Top Ten Salient Nouns, Verbs and Adjectives Obtained from the Collected Corpus?

The top 10 salient nouns, verbs, and adjectives have been delineated as follows. The frequency of their occurrences within the Rohingya corpus was computed, as follows;

Table 4. Top ten salient nouns

Noun	Frequency (f)	Frequency (f) per million	DOCF	Relative DOCF (%)
Myanmar	235.00	8,344.58	56	71.79
people	139.00	4,935.73	56	71.79
Rohingya	135.00	4,793.69	58	74.36
country	121.00	4,296.57	52	66.67
migrants	99.00	3,515.38	29	37.18
Bangladesh	96.00	3,408.85	38	48.72
government	85.00	3,018.25	40	51.28
state	84.00	2,982.74	38	48.72
woman	81.00	2,876.22	20	25.64
year	80.00	2,840.71	43	55.13

\*DOCF = Document frequency

In accordance with the analysis conducted, the salient nouns that emerged are as follows: *Myanmar, people, Rohingya, country, migrants, Bangladesh, government, state, woman, and year*. Among the observed frequencies, *Myanmar* exhibits the greatest magnitude, with a value of 235.00. It is worth noting that the terms *people, Rohingya, and country* exhibit a slightly elevated frequency range of 121.00 to 139.00 in comparison to other nouns. The aforementioned nouns are arranged in accordance with their RF occurrences, with *Myanmar* exhibiting the highest frequency, trailed by *people, Rohingya, and country*. Whilst *Myanmar* is a salient noun, its DOCF and relative DOCF are comparatively lower than that of *Rohingya*. This is evidenced by the fact that *Rohingya* is extensively featured in 58 newspaper articles, whereas *Myanmar* is only identified in 56 articles, akin to individuals. The noun *migrants* has been observed to feature prominently in 29 newspapers, ranking among the top 5 salient nouns with a frequency of 99.00. The gender noun *woman* stands as the sole identifier,

however, a mere 20 news reports have been discovered utilising said noun. Despite the fact that the term *year* is utilised extensively in 43 newspapers, its frequency of 80.00 is surpassed by that of another term, *woman*, which boasts a frequency of 81.00.

Table 5. Top ten salient verbs

Verb	Frequency (f)	Frequency (f) per million	DOCF	Relative DOCF (%)
be	713.00	25,317.80	77	98.72
have	338.00	12,001.99	69	88.46
say	279.00	9,906.97	66	84.62
include	56.00	1,988.50	32	41.03
take	49.00	1,739.93	30	38.46
flee	36.00	1,278.32	24	30.77
make	35.00	1,242.81	19	24.36
come	35.00	1,242.81	29	37.18
do	32.00	1,136.28	23	29.49
displace	30.00	1,065.27	15	19.23

In response to this analysis, the most prominent verbs that have been identified are as follows: *be*, *have*, *say*, *include*, *take*, *flee*, *make*, *come*, *do*, and *displace*. The verb *be* exhibits the most frequent occurrence, with a frequency value of 713.00, and is observed in nearly all newspapers (77; 98.72%). The reason for its frequent usage lies in its ubiquitous nature as a verb within a wide range of sentences and contextual settings. The verbs *have* and *say* exhibit a notable frequency, with respective values of 338.00 and 279.00. Likewise, it is observed that they manifest in 69 (88.46%) and 66 (84.62%) newspapers. Upon scrutinising the remaining verbs, it is evident that the majority exhibit comparable frequency and DOCF. Within the context of the Rohingya, the verbs *flee* and *displace* hold significant relevance and are notably distinct as indicated by the table. The verbs *flee* and *displace* exhibit a frequency of 36.00 and 30.00, respectively. These frequencies were ascertained through an examination of 24 and 15 newspapers, respectively. The newspapers appear to centre their attention on the pertinent matters surrounding the fleeing and displacement of the Rohingya population on a global scale. Nevertheless, it is worth noting that certain verbs, namely *be*, *have*, and *do*, may not wield significant influence in the realm of sentiment analysis. Therefore, the aforementioned verbs are eliminated and substituted with the subsequent prominent verbs. The subsequent top ten salient verbs, following the application of a filtering process to the aforementioned three verbs, are presented herewith.

Table 6. Top ten salient verbs after filtration

Verb	Frequency (f)	Frequency (f) per million	DOCF	Relative DOCF (%)
say	279.00	9,906.97	66	84.62
include	56.00	1,988.50	32	41.03
take	49.00	1,739.93	30	38.46
flee	36.00	1,278.32	24	30.77
make	35.00	1,242.81	19	24.36
come	35.00	1,242.81	29	37.18
displace	30.00	1,065.27	15	19.23
see	29.00	1037.23	21	26.92
find	29.00	1037.23	19	24.36
help	29.00	1037.23	18	23.08

The three recently incorporated salient verbs (*see*, *find*, *help*) encompass the actions of perceiving, discovering, and aiding. The frequencies of 29.00 and 1037.23 are shared by these entities, with a frequency per million equivalent. The disparity in the distribution of their respective DOCF implies that *see* is employed more extensively in comparison to *find* and *help*.

Table 7. Top ten salient adjectives

Adjective	Frequency (f)	Frequency (f) per million	DOCF	Relative DOCF (%)
more	76	2,698.67	38	48.72
human	59	2,095.02	31	39.74
international	40	1,420.35	18	23.08
other	39	1,384.84	25	32.05
many	37	1,313.83	26	33.33
last	32	1,136.28	24	30.77
political	32	1,136.28	14	17.95
first	28	994.25	20	25.64
humanitarian	28	994.25	13	16.67
ethnic	26	923.23	15	19.23

In line with the analysis conducted, the ten salient adjectives are as follows: *more*, *human*, *international*, *other*, *many*, *last*, *political*, *first*, *humanitarian*, and *ethnic*. *More* appears most frequently (f=76.00). The present study reveals that *more* has been identified in a total of 38 newspapers. The findings indicate a notable prevalence of comparative sentences within the corpus of 38 newspapers analysed, which were utilised to explicate matters pertaining to the Rohingya community. Subsequent to this, it has been observed that the adjectives *human* and *international* hold the second and third positions in terms of salience, with a frequency of 59.00 and 40.00, respectively. These adjectives have been identified in 31 and 18 newspapers, respectively. Upon analysing the data pertaining to the DOCF, it has been observed that a

greater number of newspapers, precisely 31, exhibit a heightened level of interest in human-centric issues. This is in contrast to the 18 newspapers that demonstrate a relatively more globalised perspective, with a greater emphasis on international matters. The frequencies of *last* and *first* are comparable, with values of 32.00 and 28.00, respectively. Additionally, the DOCF for *last* and *first* are 24 and 20, respectively. It is postulated that the newspapers employed numerical adjectives to denote the Rohingya predicament. It is of note that the adjectives *humanitarian* and *ethnic* have been identified in a mere 13 to 15 newspapers. These adjectives hold the 9th and 10th positions, respectively, in terms of salience, with a frequency of 28.00 and 26.00. It has been deduced that the newspapers do not place significant emphasis on the utilisation of these two adjectives within the Rohingya corpus.

BNC was chosen as the reference corpus for the purpose of identifying keywords. Upon analysis of the 100 keywords presented, it has been determined that the salient vocabulary identified includes *Myanmar*, *Rohingya*, *migrant*, *Bangladesh*, and *humanitarian* are significant. The aforementioned quintet of nouns has been subjected to n-gram analysis, specifically targeting top five trigrams and qualgrams.

Table 8. Top five trigrams and qualgrams

Word	Trigram	Frequency	Qualgram	Frequency
Myanmar	Myanmar and Bangladesh	9.00	southeastern border with Myanmar	5.00
	in western Myanmar	7.00	from Myanmar and Bangladesh	3.00
	in Myanmar and	6.00	Rohingya Muslims from Myanmar	3.00
	from Myanmar and	6.00	persecution in Myanmar and	3.00
	Bangladesh and Myanmar	5.00	Myanmar's domestic peace process	2.00
Rohingya	Rohingya and Bangladeshi	6.00	Arakan Rohingya Salvation Army	5.00
	Rohingya Salvation Army	5.00	Rohingya and Bangladeshi migrants	5.00
	of the Rohingya	5.00	of Rohingya and Bangladeshi	3.00
	Arakan Rohingya Salvation	5.00	Rohingya Muslims from Myanmar	3.00
	the Rohingya refugee	4.00	A Rohingya Muslim illegal	3.00
migrants	thousands of migrants	5.00	Rohingya and Bangladeshi migrants	5.00
	migrants from Bangladesh	5.00	migrants stranded at sea	3.00
	migrants from Myanmar	5.00	illegal migrants from Bangladesh	3.00
	and Bangladeshi migrants	5.00	migrants on board capsized	3.00
	hundreds of migrants	4.00	of migrants from Myanmar	2.00
Bangladesh	Myanmar and Bangladesh	9.00	close to Bangladesh's southeastern	5.00
	migrants from Bangladesh	5.00	Bangladesh's southeastern border with	5.00
	Bangladesh and Myanmar	5.00	to Bangladesh's southeastern border	5.00
	Bangladesh's southeastern border	5.00	Bengal close to Bangladesh's	3.00
	close to Bangladesh's	5.00	from Myanmar and Bangladesh	3.00
humanitarian	top UN humanitarian	2.00		
	the humanitarian situation	2.00	n/a	
	international humanitarian law	2.00		

The table above has presented the top five trigrams and qualgrams. It is noteworthy that the noun *humanitarian* is comprised of three trigrams and does not contain any qualgrams. Upon conducting a thorough screening and filtering process, the resulting output showcases a collection of phrases that are suitable for concordance analysis: *Myanmar* (*in western Myanmar*, *Bangladesh and Myanmar*, *southeastern border with Myanmar*, *from Myanmar and Bangladesh*, *Rohingya Muslims from Myanmar*, *persecution in Myanmar and*, *Myanmar's domestic peace process*), *Rohingya* (*of the Rohingya*, *the Rohingya refugee*, *Arakan Rohingya Salvation Army*, *Rohingya and Bangladeshi migrants*, *Rohingya Muslims from Myanmar*, *A Rohingya Muslim illegal*), *migrants* (*thousands of migrants*, *hundreds of migrants*, *migrants stranded at sea*, *illegal migrants from Bangladesh*, *migrants on board capsized*, *of migrants from Myanmar*, *migrants from Bangladesh*), *Bangladesh* (*close to Bangladesh's southeastern*, *Bangladesh's southeastern border with*, *to Bangladesh's southeastern border*, *Bengal close to Bangladesh's*), and *humanitarian* (*top UN humanitarian*, *the humanitarian situation*, *international humanitarian law*).



UKTAW, Myanmar - The number of people killed in six days of unrest in western Myanmar reached at least 112 on Friday as security forces used deadly force to silence raged for a fifth day between Rohingya Muslims and Buddhists in western Myanmar on Thursday, pushing the death toll to more than 120 and testing the have been killed and thousands displaced in Buddhist-Muslim violence in western Myanmar, casting a shadow over the reformist government's attempts to remake es have been burnt down. Clashes with Buddhists in Rakhine in western Myanmar earlier this year also reportedly uprooted over 70,000 Rohingyas. He is going ahead with the trip despite recent sectarian violence in western Myanmar, which has drawn concern from the United States, the European Union ar's refusal to recognize the Rohingya, an ethnic minority group living in western Myanmar, as citizens. Most of Myanmar's 1.1 million Rohingya are state he left her home in a village in the violence-ridden district of Rakhine, in western Myanmar, which borders Bangladesh, on Sept 30, after her husband was killed lohingya issue. Wang said the issue should be resolved by Bangladesh and Myanmar through dialogue and consultation, and expected the two sides to t eign Minister Wang Yi said on Friday. China, friend of both Bangladesh and Myanmar, hopes to see the two countries properly settle the issue through t es and their defense partnership, Wu said. China supports Bangladesh and Myanmar in resolving the issue of Rohingya refugees through consultations i mp. Though a repatriation agreement was signed between Bangladesh and Myanmar in 2017, the plan has been delayed ever since. Mohammed rnement is seeking the help of other countries that are working with Bangladesh and Myanmar to arrange small repatriations. All dialogue that leads toward early Wednesday in the Bay of Bengal close to Bangladesh's southeastern border with Myanmar, leaving some 87 people missing, an official said. Lie, engal near Shah Porir Island in Teknaf close to Bangladesh's southeastern border with Myanmar, leaving an unknown number of people missing. Their n board capsized in the Bay of Bengal close to Bangladesh's southeastern border with Myanmar on Wednesday, leaving some 99 people missing, an official s engal near Shah Porir Island in Teknaf close to Bangladesh's southeastern border with Myanmar, leaving an unknown number of people still missing. T efugees capsized in the Bay of Bengal close to Bangladesh's southeastern border with Myanmar, said an official here on Tuesday. Hayat ibna Siddique y said it had turned away a boat carrying hundreds of migrants from Myanmar and Bangladesh to an uncertain fate. Thai crackdown leaves migrants str ists seeking to land in Malaysia, Indonesia and Thailand, mostly from Myanmar and Bangladesh, have been stranded at sea. Indonesia and Malaysia ha graves were believed to contain bodies of hundreds of migrants from Myanmar and Bangladesh. Grim secrets of trafficking camps found Malaysian polic i), told Xinhua "We've rescued 23 passengers, some of them Rohingya Muslims from Myanmar." On receipt of information, he said BGB, Bangladesh i alp them deal with the crisis. The migrants are mostly Rohingya Muslims from Myanmar and Bangladeshis - men, women and children who fled persec elieved to have been used by a network that smuggled ethnic Rohingya Muslims from Myanmar and migrants from Bangladesh and then abused them, held thi ng route used by criminals preying on Rohingya Muslims fleeing persecution in Myanmar and Bangladeshis seeking to escape poverty. Rohingya Muslim Aid groups estimate that thousands more migrants - who fled persecution in Myanmar and poverty in Bangladesh - are stranded in the Andaman Sea. of Indonesia, Malaysia and Thailand, as Rohingya Muslims flee persecution in Myanmar and poor Bangladeshis crowd onto boats in the hope of finding jobs in sought China would like to play a constructive role in Myanmar's domestic peace process and safeguard the safety and stability of the border region, sovereignty of Myanmar and pays strong attention to Myanmar's domestic peace process, Xi said. Noting that the traditional friendship betw

Figure 4. Concordances related to Myanmar

Upon examination of the aforementioned concordances, it becomes patently evident that the noun *Myanmar* is utilised as a noun to denote both a sovereign state and a geographic locale. One could posit that the utilisation of language in reference to the Rohingya population primarily serves to relay unfavourable occurrences, employing words such as *killed*, *death*, *shadow*, *violence*, *uncertain*, *flee*, *poverty*, and *poor*. It is noteworthy that a total of four phrases exhibit a frequency below 4.00, while three phrases demonstrate a frequency of 5.00 and above. The Rohingya Muslims hailing from Myanmar and Bangladesh have been a subject of extensive media coverage owing to the persecution they have faced in Myanmar. Various newspapers have reported on this issue. It is of academic interest to note that the domestic peace process of Myanmar has been identified twice in a certain newspaper. The recurrence of the issue in question may be attributed to the domestic tranquilly that was reported on the aforementioned day, thereby prompting the author to reiterate it.

at least frozen gestures toward Myanmar as a condition to stop abuse of the Rohingya. "President Obama is sending the message that crimes against akhine state, as well as measures to support the civil and political rights of the Rohingya people, the White House said. Violence erupted across Rakhine uman rights groups also remain deeply concerned about the oppression of the Rohingya Muslim minority, who remain stateless and largely confined to camps ne n Minister Aung San Suu Kyi came under fire for her country's handling of the Rohingya situation. In India, J. Jayalithaa, a popular leader and state chi ted the two sides to sign an agreement as scheduled on the repatriation of the Rohingya people who have crossed into Bangladesh. China is willing to cc health services and information. When floods badly affected the Rohingya refugee camps last year in Cox's Bazar, Bangladesh, midwife Shakila id. Bangladesh seeks humanitarian aid over refugees Life in the Rohingya refugee camp has never been easy for 26-year-old Senowara Begum. nd Relief, said Bangladesh needs more political support to deal with the Rohingya refugee issue. The international community can provide more dialogue that leads toward peaceful and durable solutions regarding the Rohingya refugee issue is encouraged so that they can return home, said Regin anmar authorities on Friday declared the armed attackers and Arakan Rohingya Salvation Army as extremist terrorist groups who have claimed the respon es and several civilians. The government declared the Arakan Rohingya Salvation Army, which claimed responsibility for the offensive, a terrorist c >The security forces have been entering villages to search for Arakan Rohingya Salvation Army extremist terrorists to maintain peace and stability in the ai d to elsewhere in Southeast Asia since August 2017 when the Arakan Rohingya Salvation Army extremists launched attacks on police outposts in Rakhine yanmar army only intervened after an attack by fighters of the Arakan Rohingya Salvation Army in August 2017. She pointed out that tensions have concerned about the urgent situation faced by thousands of Rohingya and Bangladeshi migrants at sea in Southeast Asia," State Department spokesman Jef hrough a regionally coordinated effort." Some 6,000 Rohingya and Bangladeshi migrants are believed to be still stranded at sea in precarious condit Indonesia, officials said Wednesday, the latest in a stream of Rohingya and Bangladeshi migrants to reach shore in a growing crisis confronting Southeast Asi country had "given more than it should" to help hundreds of Rohingya and Bangladeshi migrants stranded on boats by human traffickers. The foreign palace. Marsudi said Indonesia has sheltered 1,346 Rohingya and Bangladeshi migrants who washed onto Aceh and North Sumatra provinces last w B), told Xinhua "We've rescued 23 passengers, some of them Rohingya Muslims from Myanmar." On receipt of information, he said BGB, Bangladesh help them deal with the crisis. The migrants are mostly Rohingya Muslims from Myanmar and Bangladeshis - men, women and children who fled persec elieved to have been used by a network that smuggled ethnic Rohingya Muslims from Myanmar and migrants from Bangladesh and then abused them, held thi t tens of thousands of people displaced, according to UN. A Rohingya Muslim illegal immigrant cries as he prays at the Immigration Detention Centre n the Buddhist and Muslims in Myanmar's western Rakhine State A Rohingya Muslim illegal immigrant looks out from the Immigration Detention Centre durin Muslim holy fasting month of Ramadan in Kanchanaburi province A Rohingya Muslim illegal immigrant prepares plates of food at the Immigration Detention C

Figure 5. Concordances related to Rohingya

Upon analysing the concordances presented, it is evident that the noun *Rohingya* is frequently collocated with various nouns, including but not limited to *refugee*, *army*, *migrants*, and *Muslims*. Primarily, it is utilised to document the unfavourable occurrences encountered by individuals, as evidenced by the utilisation of vocabulary such as *abuse*, *oppression*, *repatriation*, *badly*, *attackers*, *terrorist* and *illegal* in various newspapers.

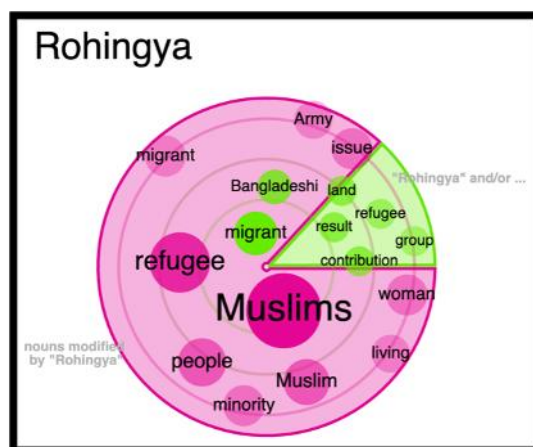


Figure 6. Visualisation of Word Sketch: Rohingya

With reference to the aforementioned figure, it is evident that the nouns that are modified by the vocabulary *Rohingya* bear resemblance to the concordances. The evidence suggests that the nouns *Muslims* and *refugee* are frequently employed in conjunction with *Rohingya*. Thus, it can be observed that China Daily predominantly enlisted individuals belonging to the Rohingya Muslim community, as well as those who have sought refuge as Rohingya refugees, to serve as nouns for their newspapers. The aforementioned analysis reveals that the words associated with the Rohingya issue encompasses a set of nouns, namely *migrant*, *Bangladeshi*, *land*, *result*, *refugee*, *group*, and *contribution*. The collocation of *migrant* and *Bangladeshi* is frequently observed in conjunction with *Rohingya*, as evidenced by the concordances.

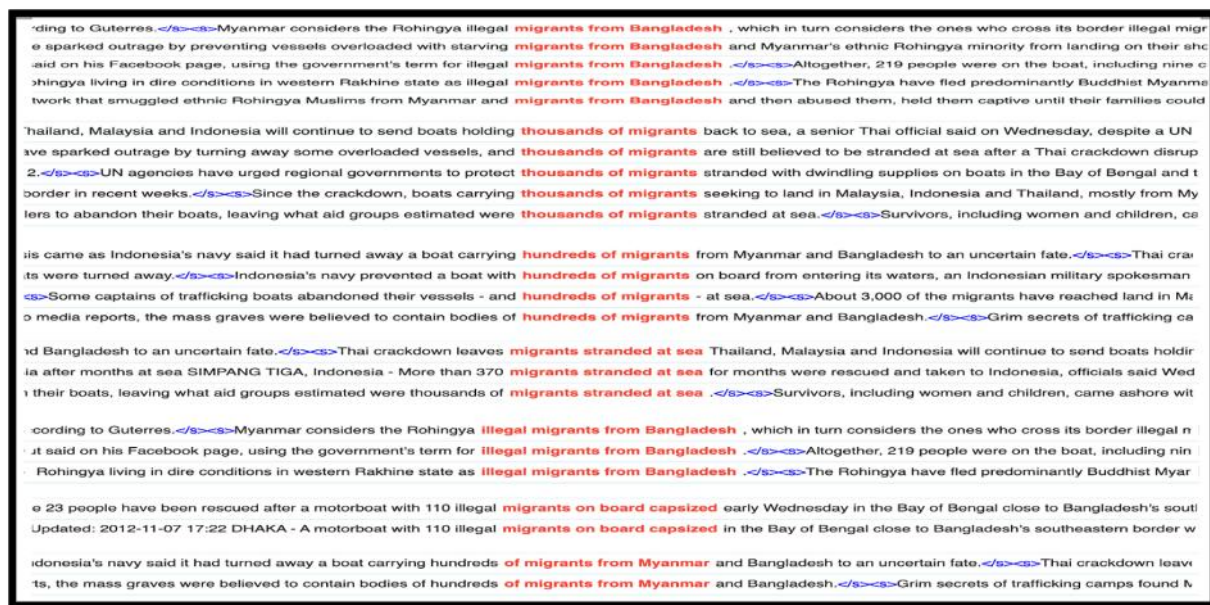


Figure 7. Concordances related to Migrants

Based on the concordances, one may infer that the migrants originate solely from Bangladesh and Myanmar. They are commonly known as the Rohingya. The word *illegal* connotes a lack of lawful entry into any given nation-state. The plight of those who find themselves famished and marooned in the vast expanse of the ocean has captured the attention of the global community. Each of the aforementioned phrases was discerned from distinct newspapers. The veracity of the reported figure of 110 undocumented migrants appears to be substantiated by the fact that two distinct newspapers have independently corroborated this numerical value. The newspapers primarily delineated the various challenges and hardships encountered by the Rohingya community.



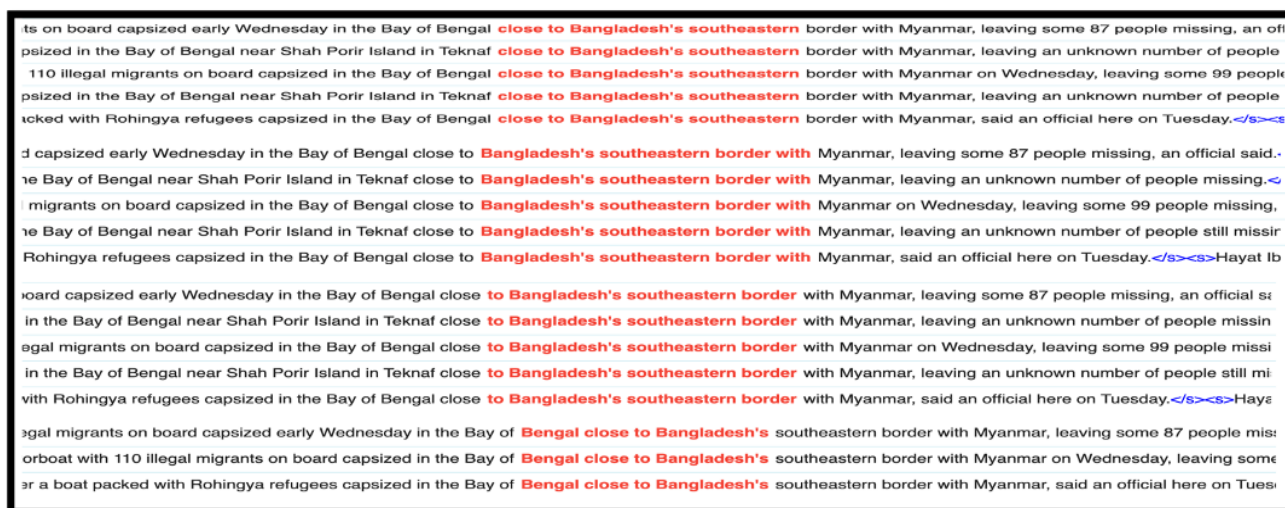


Figure 8. Concordances related to Bangladesh

Based on the aforementioned concordances, it is plausible to construct a lengthier phrase: *Bay of Bengal close to Bangladesh's southeastern border with Myanmar*. It has been observed that a limited number of newspapers, specifically newspapers numbered #2, #3, and #68, have reported the usage of the aforementioned phrase in their news coverage. The data presented indicates that during the specified timeframe, the aforementioned trio of newspapers covered incidents pertaining to the Rohingya community transpiring in the Bay of Bengal. The utilisation of the aforementioned phrase is indeed intriguing, particularly in light of the fact that no other newspapers had employed it. The data suggest that the occurrences that transpired within the Bay of Bengal hold considerable significance.

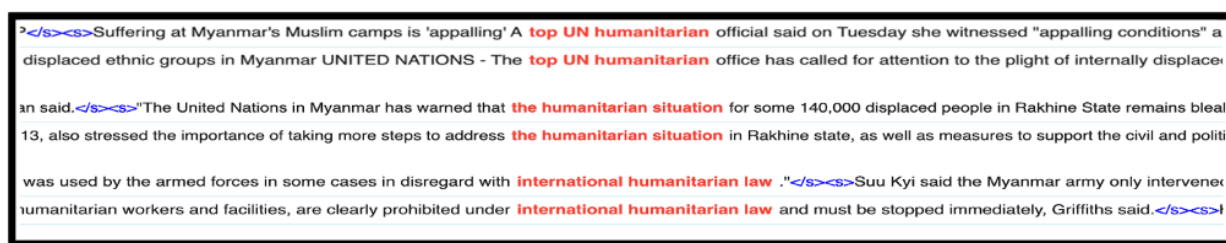


Figure 9. Concordances related to Humanitarian

It is noteworthy that the noun *humanitarian* lacks any qualgrams, as per the available information. The presentation encompassed concordances of three trigrams. The aforementioned incidents exhibited a parallelism in their manifestation, specifically a frequency of 2.00 in distinct newspapers. The acronym UN is commonly understood to refer to the esteemed international organisation known as the United Nations. The present concordances serve to indicate the interferences of the humanitarian office with regard to the Rohingya situation. Furthermore, it is noteworthy that the principles of humanitarian law were implemented in the context of the Rohingya situation. In a broader sense, it may be posited that the United Nations in Myanmar maintains a semblance of humanitarianism rather than exhibiting authoritarian tendencies.

#### 4.2 Research Question 2: What is the Sentimentality of the Aforementioned Salient Vocabulary and Phrases?

The table below presents the proposed sentimentality of salient nouns, verbs, adjectives, trigrams and qualgrams presented above by Atlas.ti.

Table 9. Words, phrases and their sentimentality

Words/Phrases	Sentimentality
<i>help, find, humanitarian, Myanmar's domestic peace process, Arakan Rohingya Salvation Army, top UN humanitarian, the humanitarian situation, international humanitarian law</i>	Positive (8)
<i>Myanmar, people, Rohingya, country, migrants, Bangladesh, government, state, woman, year, be, have, say, see, include, take, make, come, do, more, human, international, other, many, last, political, first, ethnic, in western Myanmar, Bangladesh and Myanmar, southeastern border with Myanmar, from Myanmar and Bangladesh, Rohingya Muslims from Myanmar, of the Rohingya, Rohingya and Bangladeshi migrants, Rohingya Muslims from Myanmar, thousands of migrants, hundreds of migrants, illegal migrants from Bangladesh, of migrants from Myanmar, migrants from Bangladesh, close to Bangladesh's southeastern , Bangladesh's southeastern border with, to Bangladesh's southeastern border, Bengal close to Bangladesh's</i>	Neutral (45)
<i>flee, displace, persecution in Myanmar and, the Rohingya refugee, A Rohingya Muslim illegal, migrants stranded at sea, migrants on board capsized</i>	Negative (7)



Figure10. Percentage of sentimentality

Upon completion of the analysis, the researchers determined that most of the vocabulary (75%) and phrases scrutinised exhibited a neutral affective valence. Merely 13% of the vocabulary and phrases were identified as possessing affirmative connotations, while 12% of the vocabulary and phrases were classified as negative. The analysis of the gathered texts indicates that China's stance on the Rohingya matter does not manifest any overtly favourable or unfavourable disposition. The statement connotes a dearth of robust affective or appraising stance on the part of China.

It is imperative to acknowledge that the present findings are circumscribed to the linguistic scrutiny undertaken in the research manuscript. The insights pertaining to the underlying political or diplomatic motives of China with respect to the Rohingya refugee crisis are not provided. The linguistic neutrality exhibited in the discourse emanating from Chinese sources may indicate a desire to uphold an equitable or tactful position. However, additional inquiry and scrutiny are imperative to gain a more profound comprehension of China's authentic policy or conduct concerning the Rohingya predicament.

## 5. Discussion

The Rohingya refugee crisis has resulted in the displacement of over one million individuals and drawn significant global media attention. While sentiment analysis of media coverage provides valuable insights into global perceptions of the crisis, understanding these perceptions necessitates a critical contextualisation within broader political, cultural and geopolitical frameworks. This discussion integrates these dimensions and critically assesses methodological variations across sentiment analyses.

Political and cultural contexts significantly influence sentiment patterns regarding the Rohingya crisis. Myanmar's domestic political landscape, particularly the deeply entrenched ethno-nationalism and anti-Muslim sentiments, has profoundly shaped public discourse. Scholars such as Lee (2019) underscore the role of Myanmar's state media in systematically dehumanising the Rohingya. It uses derogatory labels and framing them as threats to national sovereignty and security. This strategic framing normalised violence against the Rohingya. It effectively mobilises public support for oppressive state policies. Moreover, Kironksa and Peng (2021) highlight how influential Buddhist religious leaders reinforced these narratives. Thereby, they culturally legitimise exclusionary policies. These internal factors underscore the critical need to interpret sentiment analysis results within the broader context of state-driven ethno-political agendas.

Furthermore, geopolitical interests significantly shape international narratives around the crisis. China's substantial economic investments in Myanmar, notably the Kyaukphyu deep-sea port and various energy projects integral to the Belt and Road Initiative, necessitate a diplomatic approach prioritising regional stability over humanitarian accountability (Swadhin, 2024). Consequently, Chinese media narratives strategically avoid criticising Myanmar. It emphasises stability, sovereignty and anti-terrorism measures rather than human rights violations. Kironksa and Peng (2021) notes that China's official communications have actively reframed the Rohingya issue from a human rights crisis into a security matter. They downplay humanitarian abuses to protect bilateral economic and diplomatic interests. This calculated neutrality, consistently reflected in Chinese state media like China Daily, illustrates how geopolitical strategies can strongly influence media sentiment, shaping public perceptions to align with national interests.

The methodological frameworks applied in sentiment analysis studies further explain variations observed across different past studies. Studies employing automated sentiment analysis methods, such as Al-Zaman and Rashid (2024), provide extensive quantitative insights. They highlight how media in countries like the United States and the United Kingdom predominantly employ negative sentiment. They frame the Rohingya situation as a humanitarian and moral crisis. In contrast, the media in Asian host countries such as Bangladesh and Malaysia portray complex narratives that oscillate between compassion and burden-framing. They are driven largely by socio-economic strains and security fears in host communities (Das et al., 2022; Ehmer & Kothari, 2021). Such methodological divergences ranging from

corpus selection (news versus social media), sentiment categorisation approaches (manual coding vs. automated machine learning models) and analytical techniques (lexical versus algorithmic analyses) produce varied results that can complicate direct comparative interpretations. In examining specific comparisons with other scholarly studies, notable convergences and divergences emerge, Roe (2021), Nahar (2020), Samsi (2019) and Palakodety et al. (2020b) consistently identified certain salient terms such as Myanmar, Rohingya, and Bangladesh. They highlighted these as key lexicons essential for accurately framing the crisis. Such convergent findings underscore the linguistic necessity of including specific nouns that accurately describe the refugee crisis. Conversely, unique vocabulary such as "fleeing" and "government," highlighted explicitly by Samsi (2019), indicates how different corpora selection can yield distinct focal points within sentiment analyses. It emphasises varied narrative emphases reflective of particular events or temporal frames. Samsi's corpus, derived from BBC News and Jakarta Post articles, specifically foregrounded themes of escape and government response. It demonstrates how methodological differences in corpus selection influence narrative emphasis and sentiment outcomes.

Expanding on these methodological issues, comparative analysis with Hassan (2022) reveals further contrasts in sentiment findings across platforms. His analysis of Facebook comments yielded notably more polarised results. They are characterised by highly negative anti-Rohingya sentiments. It reflects how social media's anonymity fosters more extreme rhetoric compared to official news outlets like China Daily or BBC News, which must maintain diplomatic decorum and official neutrality. Similarly, Fournier-Tombs and Hendricks (2021) observed predominantly neutral Twitter sentiment towards Rohingya refugees. They suggested that different social media platforms host diverse user demographics whose expressions of sentiment vary in accordance with platform norms, accountability structures and user demographics. This divergence across platforms indicates the importance of clearly delineating corpus characteristics and considering the implications of platform selection on sentiment analysis outcomes.

Moreover, a critical observation from comparative studies is that ostensibly negative sentiments carry distinct meanings depending on the narrative context. Western media outlets, for example, often use negative sentiment to criticise Myanmar's government actions. They highlight humanitarian abuses against the Rohingya. Conversely, negative sentiment in Myanmar's domestic media targets the Rohingya community itself. It frames them as provocateurs or illegitimate outsiders. These nuanced differences underscore the limitations of purely computational sentiment analysis techniques, which may fail to capture underlying framing and contextual implications. As such, hybrid methodologies integrating qualitative content analysis with automated natural language processing techniques, such as those employed by Arifin et al. (2024) and Ehmer & Kothari (2021), may provide more accurate representations of the complex interplay of sentiment, framing, and political context.

Furthermore, international media narratives illustrate significant geopolitical biases shaped by diplomatic alliances and interests. Western media outlets, for instance, extensively report Rohingya victimhood and humanitarian abuses. They frequently invoke terms such as genocide to galvanize international condemnation and policy responses. This framing contrasts sharply with the narratives in regional media influenced by geopolitical realities. India and China, driven by strategic interests in Myanmar, typically adopt restrained media language that refrains from overt condemnation. They prioritise diplomatic engagement and economic cooperation over vocal humanitarian advocacy. This variation in international media responses emphasizes how geopolitical calculations shape narrative frames, subsequently influencing global public sentiment.

Finally, the broader literature indicates that media sentiment not only reflects prevailing political and cultural narratives but actively constructs public attitudes toward the Rohingya crisis. Studies such as those by Al-Zaman and Rashid (2024) and Irom et al. (2022) confirm the dominance of negative sentiment globally but also highlight critical divergences driven by national contexts, political agendas and methodological approaches. Sentiment analysis thus serves as both an analytical tool and a barometer reflecting broader cultural biases, political strategies and diplomatic imperatives.

In conclusion, the sentiment analysis of the Rohingya refugee crisis transcends simplistic polarity assessments. It demands critical engagement with political, cultural and diplomatic contexts that drive global perceptions. The convergence and divergence across studies underscore the complex interplay between geopolitical interests, national media practices, cultural narratives, and methodological choices. Future research would benefit from integrated methodologies combining quantitative and qualitative techniques and contextualised within robust theoretical frameworks, to better capture the nuanced dynamics shaping global sentiments toward humanitarian crises.

## 6. Conclusion

The present study investigated the alignment of newspaper language concerning China's stance on the Rohingya refugees, utilizing a corpus linguistics analysis of data from Chinadaily. It found that China's media portrayal is governed by principles of non-interference and sovereign equality, promoting a neutral stance that underscores the need for peaceful dialogue and consultation in resolving regional conflicts. These insights are expected to augment the existing literature by providing fresh perspectives on how China's approach to international humanitarian assistance might contribute to global peace and stability. Academically, the research enhances the field of media discourse analysis by deploying sophisticated tools like Sketchengine and Atlas.ti to scrutinize linguistic patterns in Chinese media's representation of the Rohingya issue. This methodological advancement offers a reliable and replicable framework for future investigations, which could be applied across various geopolitical contexts to explore the impact of media representation on national policy perceptions and international relations. For policymakers and international relations experts, the study provides nuanced insights into China's diplomatic posture as reflected in state-controlled media, which not only mirrors China's foreign policy strategies but also serves as a tool of soft power, influencing both domestic and international perceptions. Moreover, the study advocates for an empirical,

data-driven approach to the analysis of national stances on international issues, moving away from more subjective methodologies such as interviews. This shift is pivotal for achieving less biased and more substantiated interpretations, which are essential for informed policymaking and effective international dialogue. The societal implications of this research are profound, underscoring the media's role in shaping public perceptions, particularly in contexts involving humanitarian crises. It highlights the necessity of media literacy, enabling the public to critically assess news content and recognize biases, particularly in environments dominated by state-controlled media. However, the study recognizes its limitations, primarily the narrow scope of the corpus used, which may not completely encapsulate the complexity of China's public communications. Future research should expand to include a more diverse range of media sources and possibly multilingual materials. Longitudinal studies could further reveal shifts in China's diplomatic language over time, deepening the understanding of its evolving foreign policy and its ramifications for international humanitarian efforts. In conclusion, this study not only illuminates the strategic narratives China employs regarding the Rohingya crisis but also lays the groundwork for more extensive future research that could guide more effective international responses and interventions in global humanitarian issues.

### Acknowledgments

The authors wish to express deep appreciation to the two anonymous reviewers for their insightful suggestions and valuable comments which greatly enhanced this work.

### Authors' contributions

Dr. Chen Minjie and Dr. Wong Wei Lun were responsible for study design and revising. Dr. Warid Mihat was responsible for data collection. Liu Jing drafted the manuscript and Dr. Henri E. Lemana II revised it. All authors read and approved the final manuscript.

### Funding

This study is a preliminary result of the research project: "China's stance on the Rohingya refugee issue - based on corpus analysis of Chinese official media in the past ten years" [24ZBGJ-B010] from Research Center for China's Neighboring Countries. This study is also financially supported by Research Center for China's Neighboring Countries and Fudan University Press. (中国周边国家研究中心、复旦大学出版社资助, 项目号: 24ZBGJ-B010)

### Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Informed consent

Obtained

### Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

### Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

### Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

### Data sharing statement

No additional data are available.

### Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

### References

- Alsuhaibani, Z. (2022). Corpus-based instruction of gerunds and infinitives: The case of EFL learners. *3L: The Southeast Asian Journal of English Language Studies*, 4(28), 82-94. <https://doi.org/10.17576/3L-2022-2804-06>
- Al-Zaman, M. S., & Rashid, M. H. (2024). Global media sentiments on the Rohingya crisis: A comparative analysis of news articles from ten countries. *Journalism and Media*, 5(3), 1098-1111. <https://doi.org/10.3390/journalmedia5030070>
- Arifin, M. Z., Maulana, S. Y., Noertjahyana, A., & Mohamed Asghaiyer, A. (2024). Analyzing the Indonesian sentiment to Rohingya refugees using IndoBERT model. *Bulletin of Social Informatics Theory and Application*, 8(2), 180-191.

<https://doi.org/10.31763/businta.v8i2.749>

- Bennett, M. M., & Faxon, H. O. (2021). Uneven frontiers: Exposing the geopolitics of Myanmar's borderlands with critical remote sensing. *Remote Sensing*, 13(6), 1158. <https://doi.org/10.3390/rs13061158>
- Brandenburg-Weeks, T., & Abalkheel, A. (2021). The giver: A corpus-based analysis of word frequencies. *3L: The Southeast Asian Journal of English Language Studies*, 3(27), 215-227. <https://doi.org/10.17576/3l-2021-2703-14>
- Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). Affective computing and sentiment analysis. *A Practical Guide to Sentiment Analysis*, 1-10. [https://doi.org/10.1007/978-3-319-55394-8\\_1](https://doi.org/10.1007/978-3-319-55394-8_1)
- Casolari, M. (2018). Bangladesh 2018: Sheikh hasina's triumph. *Asia Maior*, 29, 247-264. <https://doi.org/10.52056/9788833132426/11>
- Chen, M., Wong, W. L., Yin, G., Swaran Singh, C. K., Mihat, W., & Yoong, S. M. (2023). Global intellectual trend of corpus linguistics studies among scholars in social sciences from September 2013 – September 2021. *Asian Journal of University Education*, 19(4), 613-631. <https://doi.org/10.24191/ajue.v19i4.24615>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design qualitative, and mix methods approaches*. SAGE.
- D'Aniello, G., Gaeta, M., & La Rocca, I. (2022). KnowMIS-ABSA: An overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis. *Artificial Intelligence Review*, 55(7), 5543-5574. <https://doi.org/10.1007/s10462-021-10134-9>
- Das, S., Anjum, A., Khushbu, S. A., & Rashed Haider Noori, S. (2022). Public sentiment analysis with opinion mining on Rohingya refugee crisis in Bangladesh. 2022 *13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. <https://doi.org/10.1109/icccnt54827.2022.9984563>
- Ehmer, E., & Kothari, A. (2021). Malaysia and the Rohingya: Media, migration, and politics. *Journal of Immigrant & Refugee Studies*, 19(4), 378-392. <https://doi.org/10.1080/15562948.2020.1821274>
- Field, J., Pandit, A., & Rajdev, M. (2021). Coping practices and gender relations: Rohingya refugee forced migrations from Myanmar to India. *Gender, Place & Culture*, 30(3), 329-349. <https://doi.org/10.1080/0966369x.2021.1997938>
- Fink, C. (2019). Myanmar in 2018: The Rohingya Crisis continues. *Asian Survey*, 59(1), 177-184. <https://doi.org/10.1525/AS.2019.59.1.177>
- Fournier-Tombs, E., & Hendricks, C. (2021). Safe deliberative spaces online between refugee-claimants and the international community: Rohingya refugee use of Twitter for advocacy and deliberation. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3954252>
- Fumagalli, M. (2018). Myanmar 2018: Botched transition and repatriation plan. *Asia Maior*, 29, 233-246. <https://doi.org/10.52056/9788833132426/10>
- Ghoshal, B. (2013). Democratic transition in Myanmar: Challenges ahead. *India Quarterly: A Journal of International Affairs*, 69(2), 117-131. <https://doi.org/10.1177/0974928413481882>
- Han, E. (2020). Myanmar's internal ethnic conflicts and their implications for China's regional grand strategy. *Asian Survey*, 60(3), 466-489. <https://doi.org/10.1525/as.2020.60.3.466>
- Han, E. (2021). Overconfidence, missteps, and tragedy: Dynamics of Myanmar's international relations and the genocide of the Rohingya. *The Pacific Review*, 36(3), 581-602. <https://doi.org/10.1080/09512748.2021.1996451>
- Hassan, M. K. (2022). Polarity of news headlines in social media and extent of anti-Rohingya rhetoric: Evidence from online news media in Bangladesh. In E. Gola, A. Volterraani, & F. Meloni (Eds.), *Communication, digital media and future: New scenarios and future changes* (pp. 27-40). <https://doi.org/10.13125/unicapress.978-88-3312-054-6>
- Hossain, A. N. (2022). Geopolitics of Rohingya refugee crisis and regional security. *International Journal of Safety and Security Engineering*, 12(2), 167-177. <https://doi.org/10.18280/ijssse.120204>
- Houston, K. (2020). Pseudo neutrality in intra-state conflict: Myanmar's official discourse on Rakhine. *Central European Journal of International and Security Studies*, 14(02), 25-52. <https://doi.org/10.51870/cejiss.a140202>
- Hu, G., Lin, T. E., Zhao, Y., Lu, G., Wu, Y., & Li, Y. (2022). UniMSE: Towards unified multimodal sentiment analysis and emotion recognition. arXiv preprint arXiv:2211.11256
- Hussain, I. (2019). Bangladesh and the changing global rivalry: An inside-out appraisal. *Global Political Transitions*, 1-30. [https://doi.org/10.1007/978-981-13-7240-7\\_1](https://doi.org/10.1007/978-981-13-7240-7_1)
- Irom, B., Borah, P., Vishnevskaya, A., & Gibbons, S. (2022). News framing of the Rohingya crisis: Content analysis of newspaper coverage from four countries. *Journal of Immigrant & Refugee Studies*, 20(1), 109-124. <https://doi.org/10.1080/15562948.2021.1906472>
- Islam, M. S., Ailian, H., & Jie, Z. (2018). Major challenges and remedies in building the Bangladesh-China-India-Myanmar economic corridor. *China Quarterly of International Strategic Studies*, 04(04), 613-629. <https://doi.org/10.1142/s2377740018500306>

- Kamal, M., & Kamal, M. (2022). Statelessness, marginality and influx of the Rohingyas: Impacts on environment and host communities of Bangladesh. *Internal Migration Within South Asia*, 91-110. [https://doi.org/10.1007/978-981-16-6144-0\\_5](https://doi.org/10.1007/978-981-16-6144-0_5)
- Karim, M. A. (2021). Ethnicity and geopolitics of Rohingya crisis. *Risks, Identity and Conflict*, 109-130. [https://doi.org/10.1007/978-981-16-1486-6\\_5](https://doi.org/10.1007/978-981-16-1486-6_5)
- Kelemen, B., & Turcsányi, R. Q. (2019). It's the politics, stupid: China's relations with Muslim countries on the background of Xinjiang crackdown. *Asian Ethnicity*, 21(2), 223-243. <https://doi.org/10.1080/14631369.2019.1677145>
- Kipgen, N. (2019). The Rohingya crisis: The centrality of identity and citizenship. *Journal of Muslim Minority Affairs*, 39(1), 61-74. <https://doi.org/10.1080/13602004.2019.1575019>
- Kironska, K., & Peng, N. (2021). How state-run media shape perceptions: An analysis of the projection of the Rohingya in the Global new light of Myanmar. *South East Asia Research*, 29(1), 16-31. <https://doi.org/10.1080/0967828x.2020.1850178>
- Lee, R. (2019). Extreme speech| extreme speech in Myanmar: The role of state media in the Rohingya forced migration crisis. *International Journal of Communication*, 13, 22-22.
- Leech, G., Rayson, P., & Wilson, A. (2014). *Word frequencies in written and spoken English: Based on the British national corpus*. Routledge. <https://doi.org/10.4324/9781315840161>
- Malji, A. (2021). People don't want a mosque here: Destruction of minority religious sites as a strategy of nationalism. *Journal of Religion and Violence*, 9(1), 50-69. <https://doi.org/10.5840/jrv202142086>
- Mallampalli, C. (2023). Minority vulnerability in South Asia and China. *International Journal of Asian Christianity*, 6(1), 29-50. <https://doi.org/10.1163/25424246-06010003>
- Mallick, A. H. (2020). Rohingya refugee repatriation from Bangladesh: A far cry from reality. *Journal of Asian Security and International Affairs*, 7(2), 202-226. <https://doi.org/10.1177/2347797020938983>
- Martel, S., & Glas, A. (2022). The contested meaning-making of diplomatic norms: Competence in practice in Southeast Asian multilateralism. *European Journal of International Relations*, 29(1), 227-252. <https://doi.org/10.1177/13540661221133194>
- Møller, H. K. (2021). Borderlines, livelihood, and ethnicity in the Yunnan–Myanmar borderlands. *Yunnan–Burma–Bengal Corridor Geographies*, 122-142. <https://doi.org/10.4324/9781003094364-9>
- Morada, N. M. (2021). ASEAN and the Rakhine crisis: Balancing non-interference, accountability, and strategic interests in responding to atrocities in Myanmar. *Global Responsibility to Protect*, 13(2-3), 131-157. <https://doi.org/10.1163/1875-984x-13020003>
- Mostafanezhad, M., Sebro, T., Prasse-Freeman, E., & Norum, R. (2022). Surplus precaritization: Supply chain capitalism and the geoeconomics of hope in Myanmar's borderlands. *Political Geography*, 95, 102561. <https://doi.org/10.1016/j.polgeo.2021.102561>
- Nahar, S. (2020). *A corpus-assisted discourse analysis of the representation of the 2017 Rohingya crisis in the Daily Star*. [Unpublished master's thesis]. University Malaya.
- Palakodety, S., KhudaBukhsh, A. R., & Carbonell, J. G. (2020a). Voice for the voiceless: Active sampling to detect comments supporting the Rohingyas. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(01), 454-462. <https://doi.org/10.1609/aaai.v34i01.5382>
- Palakodety, S., KhudaBukhsh, A. R., & Carbonell, J. G. (2020b). The refugee experience online: Surface positivity amidst hate. *24th European Conference on Artificial Intelligence - ECAI 2020*, 2925-2926. <https://doi.org/10.3233/FAIA200456>
- Park, J. (2020). Myanmar's foreign strategy toward China since Rohingya crisis: Changes, outlook and implications. *Journal of Liberty and International Affairs*, 6(1), 10-32.
- Rasul, A. (2022). ASEAN and the challenge of democracy. *Sustaining Peace in ASEAN and the Asia-Pacific*, 177-194. [https://doi.org/10.1142/9789811256547\\_0008](https://doi.org/10.1142/9789811256547_0008)
- Roe, J. (2021). Global responsibility: An exploratory corpus assisted discourse analysis of the Rohingya crisis in online media. *Journal of Modern Languages*, 31(1), 1-21. <https://doi.org/10.22452/jml.vol31no1.1>
- Samsi, Y. S. (2019). Corpus linguistic-semantic based approach: Revealing Rohingya's issue through nominalization as grammatical meaning in BBC news and the Jakarta post. *KREDO: Jurnal Ilmiah Bahasa dan Sastra*, 2(2), 338-350. <https://doi.org/10.24176/kredo.v2i2.2896>
- Shi, L., Xiong, L., & Jiang, M. (2022). The demographic characteristics of Rohingya refugees on the Myanmar-Bangladesh border and its impact on regional geopolitical relations—Based on analysis of refugee camps in Palong Khali district. *Tropical Geography*, 42(3), 499-508. <https://doi.org/10.13284/j.cnki.rddl.003449>
- Swadhin, S. (2024). China and India in the Rohingya crisis: Exploring a sustainable solutions in the geopolitical considerations and legal framework. *Malaysian Journal of International Relations*, 12(1), 35-55. <https://doi.org/10.22452/mjir.vol12no1.3>
- Tan, A. (2021). A critical assessment of human rights diplomacy by western states in Myanmar (Burma) from 2007 to 2020. *Asian Affairs*,



- 52(3), 655-687. <https://doi.org/10.1080/03068374.2021.1953783>
- Taufiq, H. A. (2019). China, India, and Myanmar: Playing Rohingya roulette? *Global Political Transitions*, 81-99. [https://doi.org/10.1007/978-981-13-7240-7\\_4](https://doi.org/10.1007/978-981-13-7240-7_4)
- Tieri, S. (2021a). Bangladesh 2021: The year of the golden jubilee and the second wave of pandemic. *Asia Maior*, 32, 277-295. <https://doi.org/10.52056/9791254690994/13>
- Tieri, S. (2021b). Bangladesh 2019-2020: Issues of democracy, disasters, development. *Asia Maior*, XXXI(2020), 275-304. <https://doi.org/10.52056/9788833138282/11>
- Tsai, Y., Chang, C., Chen, K., & Hwang, S. (2022). An integration of TextGCN and Autoencoder into aspect-based sentiment analysis. *Big Data Analytics and Knowledge Discovery*, 3-16. [https://doi.org/10.1007/978-3-031-12670-3\\_1](https://doi.org/10.1007/978-3-031-12670-3_1)
- Uesugi, Y., Wong, K. C., & Mahendra, L. (2020). The ASEAN regional forum's action-oriented preventive diplomacy mechanisms in the Rohingya crisis: Perspectives of ASEAN, China and Japan. *Series on Asian Regional Cooperation Studies*, 199-235. [https://doi.org/10.1142/9789811218576\\_0014](https://doi.org/10.1142/9789811218576_0014)
- Wang, T., Sun, B., & Tong, Y. (2022). Auto-ABSA: Automatic detection of aspects in aspect-based sentiment analysis. arXiv preprint arXiv:2202.00484
- Ware, A., Ware, V., & Kelly, L. M. (2022). Strengthening everyday peace formation after ethnic cleansing: Operationalising a framework in Myanmar's Rohingya conflict. *Third World Quarterly*, 43(2), 289-308. <https://doi.org/10.1080/01436597.2021.2022469>
- Wong, W. L., Muhammad, M. M., Mihat, W., Ya Shak, M. S., Rahman, M. A., & Prihantoro, P. (2023a). Vocabulary index as a sustainable resource for teaching extended writing in the post-pandemic era. *World Journal of English Language*, 13(3), 181. <https://doi.org/10.5430/wjel.v13n3p181>
- Wong, W. L., Mihat, W., & Rahmah, M. A. (2024) Incorporating corpus linguistics in the development of vocabulary and phraseological index. *Asian Journal of University Education*, 20(3), 920-940.
- Wong, W. L., Muhammad, M. M., Mihat, W., Rahman, M. A., Ya Shak, M. S., & Lee, M. C. (2023b). Using technologised computational corpus-driven linguistics study on the vocabulary uses among advanced Malaysian upper primary school English as a second language learners (ESL) in northern region. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 31(1), 298-314. <https://doi.org/10.37934/araset.31.1.298314>
- Xu, L., Pang, X., Wu, J., Cai, M., & Peng, J. (2023). Learn from structural scope: Improving aspect-level sentiment analysis with hybrid graph convolutional networks. *Neurocomputing*, 518, 373-383. <https://doi.org/10.1016/j.neucom.2022.10.071>
- Xue, G. (2021). Belt and Road initiative in Myanmar: Adapting to the slippery road. *Research Handbook on the Belt and Road Initiative*. <https://doi.org/10.4337/9781789908718.00025>
- Yoshikawa, S. (2022). China's policy towards Myanmar: Yunnan's commitment to Sino-Myanmar oil and gas pipelines and border economic cooperation zone. *Journal of Contemporary East Asia Studies*, 11(1), 143-161. <https://doi.org/10.1080/24761028.2022.2062857>
- Zahed, I. U. (2022). Myanmar and the Rohingya: In the shadow of Sino-Indian geopolitical competition. *Asian Journal of Political Science*, 30(1), 99-117. <https://doi.org/10.1080/02185377.2022.2048875>
- Zin, M. (2020). Myanmar in 2019: Deepening international pariah status and backsliding peace process at home. *Asian Survey*, 60(1), 140-145. <https://doi.org/10.1525/AS.2020.60.1.140>