

CHAPTER

5

CORPUS-BASED ANALYSIS OF DISASTER NEWS ARTICLES

Sarimah Shamsudin and Wan Najmiyyah Wan Md Adnan

5.1 INTRODUCTION

Analysing media representations of disasters is crucial in order to understand how events are portrayed and perceived by the public. Media coverage can significantly shape public opinion, influence policy decisions, and provide narrative strategies. By examining these representations, researchers can uncover biases, highlight the framing of events, and reveal the underlying socio-political dynamics at play.

Corpus-based analysis is an effective method for examining large datasets of textual material, allowing researchers to systematically identify and analyse linguistic patterns. By utilising frequency and concordance analyses, it can reveal the contextual usage of specific keywords, providing insights into the media's thematic focus and narrative strategies (Biber, 1993; Sinclair, 2004). This method is especially relevant for studying aviation disaster news, where understanding the representation of events, actors, and nations involved is critical (Wan Md Adnan & Shamsudin, 2019).

This chapter centres on the MH17 news disaster event, in which a passenger jet flying from Amsterdam to Kuala Lumpur was shot down over the Ukrainian-Russian border on 18 July 2014. Using a corpus compiled from an online English-language Dutch news portal, this

chapter introduces the international framings of the MH17 disaster between 2014 and 2016 and shows how the external narratives have shaped Malaysia's own crisis narrative and policy responses.

This chapter describes the use of user-friendly corpus analysis tools to identify and analyse 60 news articles from the *Dutchnews.nl* portal. *Dutchnews.nl* is a freely available news portal with a uniform HTML format, which makes it easy for users to download, clean, and run keyword searches. Hence, it is considered suitable for illustrating the corpus analysis method step by step. The findings aim to shed light on the media coverage of the disaster to reveal underlying narratives and biases. This chapter's objectives are to demonstrate the process of corpus construction, outline the methodologies of corpus analysis, and present findings and discussions on the portrayal of selected keywords in the media.

5.2 ANALYSING DISASTER NEWS DISCOURSE

Analysing news discourse is critical for understanding how media representations impact public opinion, shape policy decisions, and create narratives around specific events (van Dijk, 1988). For example, when it comes to disaster news coverage, the media's focus can be on particular aspects of a disaster, such as the human impact, recovery operations, or other underlying reasons. For example, Bromhead (2021) examines popular discourse about adverse weather conditions to evaluate the 2011 Queensland floods in Australia. It examined vocabulary and discourse trends through semantically enriched discourse studies based on natural semantic metalanguage (NSM). The extensive examination revealed cultural and historical influences in the terminology used to describe extreme weather events. Sarmiento and Poblete (2021) examined how people responded to both natural and man-made disaster news using social media communications from 41 real-world crisis occurrences, based on textual and linguistic variables. Mosurska et al., (2022) have examined disaster discourse, focusing on the portrayal of disasters and the involvement of indigenous

populations in news media. The study identified different disaster discourses of both natural disasters and human-made disasters, like systems of oppression, humanitarianism, and technocracy issues. These works indicate that media can prioritise certain elements in the public's mind, thereby influencing their overall perception of the disaster event.

When it comes to the methodology of analysing news, corpus analysis provides a structured approach to understand how language is used in specific contexts (Baker, 2006). Additionally, it also enables a systematic analysis of large datasets to reveal patterns and trends that are not immediately apparent through manual analysis (Biber, 1993). Methods like frequency and keyword analysis, concordance analysis, and machine learning classification are widely employed to effectively comprehend large amounts of textual data. Gopal et al., (2020) for one, proposed a machine learning method to categorise newsworthy content from disaster-related news articles. Their methodology achieved approximately 70% accuracy in filtering insightful information.

There were also approaches taken to use lexical, semantics, syntactic, and linguistics analysis when it comes to analysing disaster news to identify the representations of the disasters in the media (Chen, 2020; Skibina et al., 2022). Chen (2020) used reports from the BBC, New York Times, and China Daily in his study on the Boeing 737 MAX crash in 2019. It employs Halliday's systemic functional theory to examine the language aspects of news items and propose the function of language to represent the relationship between the speaker's internal and external world experience, primarily through transitivity, rhetoric, and mood structure. As a result, the study compared the varied language styles and characteristics of disaster news coverage in China and the West. Another study by Skibina et al. (2022) used keyword and collocation analysis to identify dominant issues in English online news discourse over a certain period of time. Meanwhile, Denti (2023) used diachronic analysis to examine how the Corona Virus Disease of 2019 (COVID-19) situation was portrayed in press articles from January 2020 to February 2022, utilising a corpus obtained from The Times. It investigates how language use has changed over time, as well as how media coverage affects public perception and sentiment during various stages of