False News Discourse Online: Corpus-Assisted Discourse Analyses

By Bashayer Baissa

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

Department of English Language and Linguistics
School of English, Drama, and Creative Studies
College of Arts and Law
University of Birmingham
March 2024

Abstract

This thesis examines the discourse characteristics of online false news articles using a mixed-method corpus-assisted discourse approach. The study employs three corpus-assisted analytical methods: move analysis (MA), key semantic domains analysis (KSDA), and multidimensional analysis (MDA) to analyse online false news articles in terms of discourse structure, discursive news values, and stylistic patterns respectively, with the goal of understanding how false news influences its audience. The corpus includes 137 verifiably online false news articles (totalling 106,673 words) on climate change, vaccination, and COVID-19. It also includes a comparative corpus consisting of 548 news articles (totalling 350,798 words) from reliable and reputable broadsheets, tabloids, web-based publications, and blogs covering the same topics.

The MA findings indicate that the primary goal of false news articles is challenging mainstream narratives. This aim is reflected in the discourse structure of false news articles, which combines elements of expository, narrative, and argumentative genres. The KSDA uncovers that false news discourse emphasises meanings constructing unique news values, such as *scepticism, corroboration, causality, mysticism,* and *historicity*. The MDA showed that false news shares stylistic similarities with traditional and reliable news sources, especially broadsheets. Both false news and broadsheets exhibit *informational, involved*, *narrative,* and *expository styles*. However, false news differs from all other news types in that it does not tend to use an *explicitly advocating style*. Overall, the study reveals that these similarities and distinctive characteristics contribute to the persuasive and viral impact of false news discourse.

The study highlights the significant threat posed by false news, noting that false news can gain credibility and traction by imitating credible journalism and that false news can be difficult to detect by detection algorithms based on surface-level stylistic features. Moreover, the study emphasises that the key factor for the linguistic variations between false news and true news is false news's primary goal of challenging mainstream narratives. Recognizing the contradictory nature of false news compared to established narratives and the potential for audience low credulity, false news writers strategically design their discourse to appeal to audiences in a manner distinct from true news. This finding underscores the pervasive issues of scepticism and information crisis in society. The thesis concludes with a reflection on the corpus and the mixed-method approach, as well as recommendations for future research and practical implications.

Acknowledgements

My sincere gratitude goes out to my supervisor, Prof. Matteo Fuoli, for his belief in my potential as a PhD candidate and for providing me with the opportunity to work under his guidance. His encouragement, meticulous review of my drafts, constructive criticism, and continuous support during challenging times have been pivotal in shaping this research. Prof. Matteo's emphasis on continuous exploration and learning has been a driving force in motivating me as a dedicated researcher.

I am also grateful for the guidance and motivation provided by my supervisor, Prof. Jack Grieve, who consistently pushed me to expand my research and enhance my academic skills. His valuable feedback has significantly improved the quality of this thesis.

I extend my heartfelt thanks to my husband, Meshal, for his love, patience, steadfast support, and constant encouragement during the difficult moments of this research journey. His unwavering belief in me has been a source of strength and motivation.

Deep appreciation is felt towards my family, especially my parents, who always believed I could do it, as well as my sisters - Malak, Razan, Wasan, and Jood, and my brother Muhammad. Their continuous support, understanding, and love are truly appreciated. I am thankful for their assistance with the children and for enabling me to allocate more time to my studies. Their presence, joy, and companionship have enriched and made this journey truly memorable and rewarding. My children, Nora and Sara, have been my greatest source of inspiration and determination. Their love and companionship have fuelled my perseverance in my research endeavours.

Lastly, I would like to express my gratitude to Ministry of Education in Saudi Arabia and Taif University for their generous financial support, which made this research possible. Their assistance has been invaluable in helping me achieve my academic goals and contribute to the field of knowledge.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CADA	Corpus-assisted Discourse Analysis
CDA	Critical Discourse Analysis
CL	Corpus Linguistics
DNVA	Discursive News Values Analysis
GA	Genre Analysis
KSDA	Key Semantic Domains Analysis
MA	Move Analysis
MDA	Multidimensional Analysis
NLP	Natural Language Processing

Chapter 1: Introduction

1.1. Examining the Discourse of Fake News: A Social Imperative

In 2019, the US experienced a notable measles outbreak with a total of 1,249 reported cases in the nine-month period from January to September, surpassing the 1992 record (Patel, 2019). The outbreak was primarily attributed to insufficient vaccination coverage (Burki, (2019). According to the Centres for Disease Control and Prevention's report, 89% of measles patients were unvaccinated or had an unknown vaccination status (Patel, 2019). Inadequate vaccination coverage can be caused either by vaccine shortage or vaccine hesitancy, which refers to reluctance to vaccination despite the availability of vaccination services (World Health Organization, 2017). Such reluctance is often fuelled by false information or fake news about vaccination (Burki, 2019). The World Health Organization (2017) noted that fake news concerning vaccination has led to public distrust, fear, and anger towards health organizations, contributing to vaccine hesitancy.

The 2019 measles outbreak is a clear example of how false information can have significant and widespread effects on society. The same holds true for the COVID-19 pandemic. One of the most notorious examples of such false information was an article published online that wrongly claimed that public health organizations had predicted up to 65 million deaths from coronaviruses prior to the outbreak in December 2019. In fact, The Johns Hopkins Centre for Health Security conducted a mock exercise known as Event 201 to evaluate emergency preparedness in the event of a severe pandemic. However, it is important to note that the exercise did not predict a scenario where the pandemic would result in the death of 65 million people (Palma, 2020). Believing such false narratives can erode trust in healthcare organizations and authorities, generating fear and anger among the public. Moreover, there have been several other false reports concerning the effectiveness of COVID-19 vaccination, making claims such as the death of initial vaccine trial volunteers, the vaccine causing more deaths, and linking the vaccine to female sterilization, miscarriages, population control, and immune system suppression[1].The

[1] These claims are derived from the corpus (Articles' ID: *C False 32, C False 71, C False 52, C False 57, C False 36, C False 69*).

dissemination of such false information has likely led to confusion and panic and could have had negative impacts on people's beliefs and decision-making, potentially affecting their health as a result (Hansson et al., 2021; Mian & Khan, 2020).

The negative impact of false information can be further exemplified by fake news related to climate change. Despite a strong scientific consensus attributing global warming to human activities, some individuals actively reject this consensus and perpetuate falsehoods that discredit climate science (Cook, 2017). Such falsehoods have a range of negative effects on the public. First, they evoke negative emotions such as uncertainty, fear, frustration, and anger towards environmental organizations and authorities. Additionally, they lead to political and international inaction when it comes to addressing the climate crisis. Ultimately, these falsehoods compromise society's environmental, intellectual, and social well-being (Treen et al., 2020).

The adverse effects of fake news are becoming increasingly evident, particularly with online news outlets and social media now serving as the main sources of information. Fake news can appear in multiple formats, including images, memes, social media posts, or entire news articles published by news outlets with a title and body texts that mimic the format of reliable news sources. Fake news appearing in the format of news articles, the focus of this study, is particularly problematic (Dentith, 2018), resulting in its wide dissemination and acceptance as truth (Rini, 2017).

There are several reasons why fake news is created and spread. These motives include supporting or stigmatizing political or religious figures or groups, promoting specific ideological beliefs, generating advertising revenue, gaining attention and status within a social community, and showcasing one's knowledge. Individuals might also disseminate fake news to conform to social norms or beliefs, to share emotionally evocative content, or because they genuinely believe in the false information they disseminate (Baptista & Gradim, 2020). Scholars, such as Lazer et al. (2018) and Wardle and Derakhshan (2017) argue that the intention to deceive, along with falsehood, distinguishes disinformation, from misinformation. Disinformation is false information that is deliberately created with the intent to mislead, while misinformation may be false or misleading but not necessarily intended to deceive (cf. Grieve & Woodfield's (2023) and Tandoc et al.'s (2018) typologies of fake news). In the literature, a distinction is made between

‘false’ versus ‘true’ news and ‘fake’ versus ‘real’ news. Determining whether news is 'fake' or 'real' often relies on the intention of the writer. On the other hand, differentiating between 'false' and ‘true' news is based on the factual accuracy of the content itself (Baptista & Gradim, 2020).

Regardless of the intentions of the creators, false information has detrimental effects on society. Exposure to fake news has led to the creation of alternative epistemologies that reject established scientific evidence (Lewandowsky et al., 2017; Mian & Khan, 2020). This has resulted in a significant portion of the public losing trust in reputable scientific institutions, including the World Health Organizations (WHO) (Bayram & Shields, 2021), as well as agencies such as the National Oceanic Atmospheric Administration (NOAA) or the National Aeronautics and Space Administration (NASA) (Motta, 2018). Instead, people are turning to unreliable sources for information. This phenomenon is commonly referred to as *post-tru*th or an *information crisis* (Hancock & Bailenson, 2021). Despite the efforts of fact-checkers to debunk false information, these alternative epistemologies continue to jeopardize public safety and well-being, while also posing a threat to social cohesion (Levy, 2017).

Fake news is a discursive phenomenon, as language plays a fundamental role in its creation, spread, and impact. Discourse, defined as language in use (Brown & Yule, 1983), constructs objects, interprets the world, and represents reality with power implications (Baker & Ellece, 2011). From this perspective, fake news can be seen as leveraging the characteristics of discourse to present a counter-representation of reality (Alba-Juez & Mackenzie, 2019). Therefore, analysing the discourse of fake news is crucial for understanding this phenomenon comprehensively and combating its spread.

In recent years, there has been a notable increase in linguistic research focusing on the language of fake news. This includes a significant number of computational linguistic studies (Ahmad et al., 2020; Gravanis et al., 2019a; Horne & Adalı, 2017; Rashkin et al., 2017; Rubin & Lukoianova, 2015; Volkova et al., 2017), alongside a few studies utilizing discourse analysis (e.g., Igwebuike & Chimuanya, 2021; Maci, 2019). The objective of computational linguistic research is to develop algorithms capable of automatically distinguishing between fake and real news based on linguistic features. While some studies have achieved high levels of accuracy, these studies have several limitations. One limitation is that these studies have primarily

concentrated on analysing superficial linguistic elements like syntax and vocabulary, neglecting the discursive, rhetorical, and ideological aspects of fake news. Furthermore, contextual factors that may impact the language of fake news are often disregarded. Moreover, these studies tend to highlight the disparities between fake and real news while failing to explore their similarities or offer explanations for the discrepancies. Hence, there is a need to conduct a discourse analysis of fake news that goes beyond surface-level linguistic examination and encompasses the contextual factors associated with the production of fake news. Such an analysis can foster a more comprehensive understanding of this phenomenon, including the rhetorical and ideological techniques employed to persuade and manipulate readers. This research is of great social importance and has the potential to generate insights that can help combat the issue of fake news more effectively.

1.2. Approaches to Analysing (Fake) News Discourse

Discourse analysis is an umbrella term which covers diverse theories, approaches, techniques, and objectives of examining naturally occurring language (Baker & Ellece, 2011; Biber et al., 2007). Tannen et al. (2015) identify three distinct approaches to discourse analysis: the study of linguistic structure (beyond the sentence), the study of language use, and the study of sociocultural practices associated with language use. The study of linguistic structure looks at sequences of utterances or sentences and the structure of texts (e.g., using genre analysis). The study of language use centres around core linguistic features such as words, phrases, and clauses, exploring why discourses can have different linguistic features (e.g., using register analysis). Lastly, the study of sociocultural practices associated with language use attempts to understand how discourse produces power and ideologies (e.g., using critical discourse analysis).

Similarly, Biber and Conrad (2019) and Biber et al. (2007) argue that discourse analysis can be approached from two distinct viewpoints: genre and register. The genre perspective focuses on the linguistic features that are conventionally associated with the structure of a particular type of text, such as the conventional structure of a letter. In contrast, the register perspective examines the linguistic features that are linked to a specific situational context, like academic writing or conversations. In essence, these different perspectives highlight the diverse ways in which the terms 'discourse' and 'discourse analysis' are used and underscore the fact that each approach offers a unique viewpoint on the discourse under study.

To fully analyse news discourse, as pointed out by Van Dijk (1983), we need to consider “its various levels, units, dimensions, modes, and social contexts” (p.25). Throughout history, the pursuit of comprehending news discourse has driven extensive research in discourse analysis, exploring numerous aspects such as its genre or structural elements (Bell, 2005; Van Dijk, 1988), discursive practices (i.e., production and comprehension) (Richardson, 2006; Van Dijk, 1988, 2004), style or language use (Biber, 1988; Biber & Conrad, 2019), images, evaluation (Bednarek & Caple, 2012) news values (Bednarek & Caple, 2017), and power and ideologies (Davies, 2014; Van Dijk, 2017) as well as representation of certain social groups (Baker et al., 2013).

Thus, to gain a deeper understanding of how fake news discourse works, it is necessary to engage in an analysis that considers different levels of fake news discourse and involves multiple approaches. This may entail examining the genre and structural elements of fake news, considering their communicative purposes. Additionally, analysing the stylistic patterns represented by co-occurring linguistic features and their communicative functions, while taking into account the situational context of the news texts, can provide further insight. Furthermore, studying the meanings mobilized in fake news texts to stimulate readers' engagement and influence their attitudes, known as discursive news values, is also essential.

Discourse analysis can be assisted by the techniques and tools developed within corpus linguistics (CL), which refers to computer-aided analysis of (typically) large amounts of naturally occurring language that represents a particular language variety (Baker, 2020). In Corpus-assisted Discourse Studies (CADS), researchers can systematically analyse larger and more representative samples of language, providing higher quality evidence compared to focusing on limited or specific texts (Sinclair, 1991). CL effectively addresses the criticism of cherry-picking in discourse studies by minimizing researchers' cognitive biases (Baker, 2006), resulting in a more transparent and replicable analysis process. Additionally, CL is a versatile approach capable of addressing a wide range of research questions. It has the potential to inspire promising research questions and influence the shaping of hypotheses that go beyond initial expectations (Gillings, 2023).

CADS are increasingly adopting the practice known as triangulation, which involves the use of multiple data sources or different methods for data collection and analysis. In the realm of

CADS, two approaches to triangulation have been employed: triangulation within corpus methods and triangulation between corpus methods and other linguistic methodologies such as psycholinguistics (Egbert & Baker, 2019). Triangulation in CADS has several advantages, including reducing the risk of bias, providing a comprehensive understanding of the discourse under investigation, facilitating the detection of contradictions and inconsistencies in data, verifying results and methods, and enabling a robust interpretation of the findings (Baker & Ellece, 2011; Egbert & Baker, 2019). Given the complex and multifaceted nature of fake news discourse, triangulating corpus methods in its analysis is invaluable in gaining a rigorous and comprehensive understanding of its different aspects and increasing confidence in interpreting the results.

1.3. Methodological Considerations in Analysing Fake News Discourse

There are several methodological considerations that should be taken into account when analysing fake news discourse. One important factor to consider is that there is no consensus on the definition of fake news, as well as on how to construct a representative corpus for analysis. Scholars adopt varying approaches, with some gathering data from sources known for their lack of reliability (e.g. Rashkin et al., 2017), while others focusing on verifiably false information (e.g., Asr & Taboada, 2019), and still others emphasising the intention to deceive as the defining factor of fake news data (e.g., Grieve & Woodfield, 2023). Consequently, defining fake news is a crucial factor in determining the data collection method employed for analysis (Asr & Taboada, 2019).

Grieve and Woodfield (2023) suggest that the main distinction between real and fake news lies in the intent to inform versus the intent to deceive. They argue that the writers' knowledge of falsehood or deliberate intent to deceive can have a direct impact on the linguistic style of fake news, distinguishing it from news written with the intention to inform. According to Grieve and Woodfield, falsehood alone may not necessarily influence the language used in news, particularly considering that some fake news writers genuinely believe in the information they propagate. They emphasise the need for research to prioritize the study of news with the intention to deceive, regardless of whether it contains falsehood or selectively omits certain information to construct a misleading narrative. Nevertheless, establishing the intent to deceive can be challenging (Dentith, 2018). As stated earlier, there are various reasons why people

produce fake news, including cases where writers genuinely believe in the veracity of their claims. In addition, solely focusing on cases of proven deceptive intent may limit the scope of research and may not effectively represent the prevalence of partisan and harmful fake news encountered online. Consequently, studying verifiably false news, which is the focus of this study, is of paramount importance. Moreover, although Grieve and Woodfield (2023) argue that systematic linguistic distinctions between real and fake news cannot stem solely from falsehoods, this is an empirical question that warrants further exploration, especially considering potential contextual disparities that could influence their language.

Another factor to consider is that the analysis of fake news should be performed in a comparative manner to understand the similarities and differences between fake news and real news discourse. However, finding a suitable comparative (or reference) corpus for a focus corpus of fake news can be challenging. For instance, comparing online fake news articles written by amateurs with printed articles by experienced journalists involves comparing two different writing styles that originate from different contexts. This can result in differences that are simply due to register variation, regardless of the veracity of the news or the communicative intent of the writer. To effectively compare fake and real news, it is essential to compare fake news articles across different types of reliable news sources, such as printed newspapers, online sources, and blogs. By conducting this multi-comparison, we can identify the distinctive characteristics that separate fake news discourse from other types of reliable news sources.

In light of the approaches and considerations discussed in Sections 1.2 and 1.3 regarding the analysis of fake news discourse, this thesis will employ a mixed-method approach. The approach will involve conducting three separate case studies, each focusing on a different aspect or level of false news discourse, in comparison to various types of reliable news sources. The aim of this multi-pronged approach is to gain a comprehensive understanding of the complex phenomenon of false news discourse. Moreover, conducting multiple case studies increases confidence in interpreting the findings.

1.4. Research Question and Overview of the Case Studies

This thesis seeks to answer the following research question:

What are the discourse characteristics of online false news articles?

To address this research question, I conducted three case studies, each of which utilizes a different corpus-assisted analytical method to analyse one aspect of false news discourse. The corpus used for this study consists of 137 online verifiably false news articles (a total of 106673 words). These articles covered three controversial topics with significant social ramifications: climate change, vaccinations, and COVID-19. A comparative corpus was also compiled consisting of four sub-corpora representing true news derived from reliable printed broadsheets and tabloids, web-based publications, and blogs on the three topics (a total of 548 articles and 350798 words).

The first case study is a *genre analysis* of online false news articles. This study aims to identify the constituent elements that comprise the structure of these articles. By examining these structural elements, it becomes possible to uncover their intended purpose(s), understand how the news genre is manipulated to serve these purposes, examine the strategic presentation of information used by false news writers to emphasise certain values and beliefs, and explore variations in discourse structures across articles. The study employs *move analysis* (MA), an analysis of a text's communicative purposes by segmenting it into units or rhetorical moves according to their specific communicative purposes (Swales, 1990). Since MA is qualitative and requires manual analysis, this case study only focuses on a small subset of the corpus, namely the false articles on climate change (total of 20 articles).

The second case study aims to identify the discursive news values in false news by examining its most prominent or 'key' meanings. *Discursive news values* are factors that shape the selection, production, and consumption of news and are usually constructed in news discourse. Since false news still adheres to the conventions of the journalism genre, analysing its discourse from a news values perspective is crucial. This analysis helps understand how news writers create articles that are relevant, engaging, and appealing to their audience while also highlighting underlying ideological influences. To achieve this, the study introduces a novel approach to news values

analysis. Unlike previous studies which relied on predetermined lists of news values and applied discourse analysis to find evidence of these news values (e.g., Bednarek & Caple, 2017), this study adopts an inductive approach. Using *key semantic domains analysis* (KSDA), this study first identifies the prevalent meanings in the false news corpus by comparing the whole corpus of false news to the entire reference corpus. Then, it analyses the concordances of these key semantic domains to identify reoccurring patterns of meanings. Finally, these patterns are analysed qualitatively to identify whether they represent potential news values that may be distinctive of false news.

The third case study focuses on identifying the *stylistic patterns* found in online false news discourse. A stylistic pattern is a consistent co-occurrence of lexico-grammatical features that collectively perform a communicative function. By examining the stylistic patterns employed in false news, we can understand the similarities and differences between false news and true news, as well as how these stylistic choices are manipulated to make false news more appealing. A subset of the corpus is used for this study, in accordance with methodological requirements explained in Chapter 6. The false news corpus (total of 101 articles) is compared to the four reference sub-corpora of printed broadsheets and tabloids, websites, and blogs separately (total of 504 articles). The study employs *multidimensional analysis* (MDA) (Biber, 1988), which identifies dimensions (i.e., patterns of co-occurring linguistic features which together serve a particular communicative function) in the news types under investigation, and then compares the news types along these dimensions.

The integration of these three methods in this thesis serves two primary purposes. Firstly, it aims to ensure a comprehensive and holistic approach to analysing false news by encompassing various levels of analysis. This methodology facilitates a thorough understanding of the complexities and nuances inherent in false news discourse. Secondly, integrating these methods enhances our confidence in interpreting the findings, ultimately providing a more robust understanding of the societal impact of false news.

1.5. Objectives of the Thesis

This thesis aims to achieve several objectives. Firstly, it seeks to enhance our understanding of online false news discourse by analysing three different aspects: discourse structure, discursive

news values, and stylistic patterns. This multi-pronged analysis will provide insights into the rhetoric of false news, uncover its underlying ideologies, and examine how it appeals to its readers. This knowledge is crucial in countering the impact and spread of false news.

In addition, this thesis aims to shed light on the psychological drivers behind the consumption and spread of false news. By examining the discourse structure, news values, and stylistic patterns specific to online false news discourse, we can gain a better understanding of its objectives and why people may engage with or potentially believe it. This understanding is vital in comprehending the reasons behind false news going viral.

Furthermore, the thesis aims to offer valuable linguistic insights to improve current methods for detecting false news. Unlike existing automatic detection research that relies on shallow linguistic sets, such as n-grams or parts of speech (Grieve & Woodfield, 2023), this research delves into deeper and more elusive aspects of false news discourse. As a result, the findings of this research can be utilized in the development of more effective detection approaches.

In addition, the thesis will contribute insights into the construction of a representative focus corpus of false news and a reference corpus for comparative analysis. This is particularly valuable for studies seeking to develop effective false news detection algorithms, as the lack of representative data remains a significant limitation in previous research (Asr & Taboada, 2019; Grieve & Woodfield, 2023).

Lastly, the thesis aims to test the benefits, practicality, and effectiveness of mixing methods in discourse analysis, specifically in problematic discourses like false news. By comparing the findings of the three case studies, we can explore the complementarity of the results, evaluate the corpus, and assess the balance between the utilized corpus methods and their ability to generate interpretable and generalizable results.

1.6. Overview of the Structure of the Thesis

The thesis is structured as follows: Chapter 2 offers an overview of the existing research on fake news. It reviews existing definitions of fake news, discusses the problem of defining it, and presents the working definition used in the thesis. It also discusses the linguistic studies on fake news and highlights the gaps this thesis aims to fill. Lastly, it surveys psychological and sociological literature exploring the contextual factors affecting fake news production and

perception, including its authors, audiences, processing circumstances, communicative purposes, and topics.

Chapter 3 introduces the approach to data collection adopted for this study and describes the corpus that was compiled. The chapter starts by contextualizing the three topics of the analysed articles: climate change denialism, anti-vaccination, and COVID-19 false information and explains the rationale for selecting these topics for analysis. Next, the chapter discusses the opportunities and challenges of compiling a representative corpus of false news and a reference corpus. It also describes the process used for compiling the corpus for this thesis, the rationale behind it, and the breakdown of both the focus and the reference corpora. The chapter concludes by acknowledging the limitations of the corpus.

Chapters 4, 5, and 6 present the three case studies in detail. Each chapter includes an overview of relevant theories and previous research, a description of the methodology employed, an explanation of the analytical procedure, and sections for results and discussion.

Chapter 7 offers a comprehensive discussion of the entire thesis. This chapter is divided into five main parts. The first part summarizes the findings of the three case studies. The second part examines the implications of these findings, considering how they collectively contribute to our understanding of false news discourse. The third part assesses the corpus utilized in the study, comparing it to corpora from previous research and discussing the broader implications of analysing this specific corpus. The fourth part evaluates the advantages and challenges of employing a mixed-method approach in the thesis. Lastly, the fifth part provides recommendations for future research on fake news.

Chapter 8 serves as the conclusion of the thesis, offering a summary of the key points discussed throughout the thesis. It also discusses the practical implications derived from the three case studies and underscores the contributions made by the thesis.

Chapter 2: Previous Work on Fake News

2.1. Introduction

Extensive research has been conducted in various disciplines, such as linguistics, sociology, psychology, and media literacy, to study fake news. This chapter aims to provide a comprehensive literature review, primarily focusing on linguistic research. The first section of this chapter reviews previous research definitions of fake news and establishes the adopted definition for this thesis. This definition is essential as it guides data collection and analysis, interpretation of results, and drawing conclusions aligned with the research objectives. The second section explores linguistic studies on fake news, categorizing them into computational linguistic analysis, forensic linguistic analysis, register analysis, and discourse analysis. Each approach offers a unique perspective to analyse fake news. By reviewing these approaches, I aim to identify the gaps that this thesis aims to fill in the existing body of literature. The final section of this chapter examines insights from sociology, psychology, and media literacy studies regarding the situational characteristics of fake news. These characteristics include author and audience attributes, processing circumstances, communicative purposes, and topics. Considering these situational aspects is crucial as they provide valuable insights into the factors that influence the adoption of discourse structure features, news values, and stylistic choices in fake news.

2.2. Defining Fake News

The phenomenon of fake news is not new. Barclay (2018) characterizes fake news as an age-old practice of deceit, stating that terms like propaganda, hoax, satire, rumours, and lies can all be considered synonymous. However, it was the 2016 US presidential election that brought significant attention to the issue of fake news (Allcott & Gentzkow, 2017; Asr & Taboada, 2019). Since then, researchers from various disciplines, including linguistics, psychology, sociology, and media have taken an interest in defining fake news, leading to the proposal of several definitions and typologies in the literature (Egelhofer & Lecheler, 2019; Tandoc et al., 2018; Zannettou et al., 2019). These definitions mainly approach fake news from four different perspectives, namely the (a) veracity of information, (b) intention, (c) design, and (d) target

audience. This section provides an overview of these definitions and presents the working definition of fake news adopted for this study.

Veracity of Information

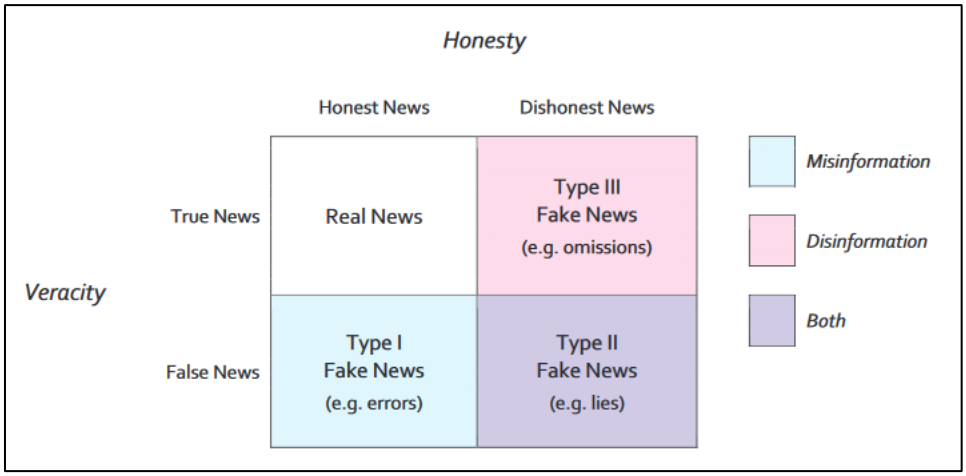
A simple definition of fake news is that fake news involves false information (Zhou & Zafarani, 2018). However, fake news can also be created by omitting certain important details that are necessary for a reader to fully understand the story. Dentith (2018) argues that significant omissions in reporting can distort a genuine story and make it misleading or fake. Galasiński, (2000) refers to this as *passive deception*, where truthful information is concealed from the audience, leading them to acquire or maintain false beliefs.

Intention

Fake news cannot be solely defined by the presence of falsehood; intentional deception also plays a crucial role (Dentith, 2018). Consequently, some definitions of fake news emphasise the intent to deceive along with the presence of verifiable falsehood. For instance, Allcott and Gentzkow (2017) define fake news as "news articles that are intentionally false and can potentially mislead readers" (p. 214). This definition excludes certain concepts often associated with fake news, such as reporting errors, non-news-based rumours, conspiracy theories generated by believers, satirical content unlikely to be interpreted as fact, false statements from politicians, and biased reports that may be misleading but not verifiably false.

The broader categories of *disinformation* and *misinformation* encompass the two perspectives of falsehood and intention to deceive (Bakir & McStay, 2018; Lazer et al., 2018; Wardle & Derakhshan, 2017). Disinformation refers to deliberately created misleading or incorrect information, while misinformation refers to unintentionally propagated misleading or incorrect information. The perspectives of falsehood and intention to deceive have been utilized by researchers to classify specific types of fake news, including propaganda, satire, hoaxes, conspiracy theories, and biased reports (cf. Tandoc Jr et al.'s, 2018; Zannettou et al.'s, 2019 typologies of fake news). Grieve and Woodfield (2023) have developed a categorization model that incorporates these two perspectives, referring to them as veracity and honesty. This model, depicted in Figure 2-1, outlines three types of fake news.

Figure 2-1 Grieve and Woodfield's Typology of Fake News, p. 13



- Type I Fake News: This category encompasses unintentional false news, where writers report information they believe to be true, but it is actually false. It often arises from mistakes, errors, or misunderstandings. Type I Fake News falls under the domain of misinformation rather than disinformation.
- Type II Fake News: This category involves intentionally false news, where writers purposefully report information they know to be false. Type II Fake News can encompass both disinformation and misinformation.
- Type III Fake News: This category comprises news that is true but is intended to deceive, such as fake news through omission and biased reporting. Type III Fake News falls under the domain of disinformation.

Grieve and Woodfield (2023) emphasise the need to prioritize research on disinformation due to its detrimental effects on society. However, determining the intention behind fake news creation is not always straightforward (Dentith, 2018). Fake news creators might have diverse motives including financial gain, ideological manipulation, or social and psychological factors (See Section 2.4.2). Besides, in some cases, it is difficult to determine whether the writer purposedly aims to mislead the reader or genuinely believe what they report. For instance, the Journal of American Physicians and Surgeons, established in 1943 to combat socialized medicine, publishes papers presenting medically discredited hypotheses, such as disputing the connection between HIV and AIDS or suggesting links between the MMR vaccine schedule and autism or

between abortions and breast cancer. However, it remains unclear whether the editors purposefully aim to mislead readers or genuinely believe the research they publish meets academic standards (Dentith, 2018). Moreover, fake news has far-reaching consequences that go beyond the intentions of its creators. Even if the writers of fake news sincerely believe in their content, it can still have detrimental effects on society. An evident case is the dissemination of misleading information about COVID-19, which can result in ill-advised choices regarding prevention methods or treatments, consequently putting public health at risk. In fact, such misinformation is what people continuously encounter online, making it impossible to undermine its harmful impact.

Design

Other definitions of fake news emphasise the design element of the source and media through which fake news is disseminated (Gelfert, 2018). Lazer et al. (2018) define fake news as "fabricated information that mimics news media content in form but not in organization process or intent" (p. 1094). Similarly, Levy (2017) states that "fake news is the presentation of false claims … in a format and with a content that resembles the format and content of legitimate media organisations" (p. 20). The real problem of fake news lies in the presentation of false information as *news* by *news outlets* (Dentith, 2018) as this results in it being widely disseminated (Rini, 2017). However, fake news is available on multiple online platforms, resulting in broad variations in the surface form and layout of fake news content. Parikh and Atrey (2018) classify fake news carrier platforms as follows:

a) Blog sites: These sites have more user-generated content than popular news websites and lack journalistic organisation, making them a good place for false information. While most blogs are personal and diary-like, blogs that are dedicated to specialized news are widespread nowadays. However, the distinction between the two is not clear cut; personal blogs sometimes present news items, while news blogs may present news with personal anecdotes or opinions (Trimarco, 2014).
b) Social media: Content on social media is circulated by sharing; platforms such as Facebook, Instagram, and Twitter allow users to create and share posts, which may include news stories, making it a good place for creating and sharing false information (Parikh & Atrey, 2018).

c) Emails: news can be created and sent via email. Like social media, it is difficult to determine the veracity of the news shared by emails (Parikh & Atrey, 2018).

Intended Audience

Beyond falsehood, intent, and design, certain fake news definitions underscore the significance of the intended audience. As Dentith (2018) notes, "fake news is a misleading story which is intended to deceive some target audience" (p. 24). Similarly, Baptista and Gradim (2020) define fake news as "a type of online disinformation, with totally or partially false content, created intentionally to deceive and/or manipulate a specific audience, through a format that imitates a news or report" (p. 5). Rini (2017) argues that although the ultimate goal of fake news is to be widely shared, it is intended to deceive at least some of its audience.

Identifying the target audience of a news story can sometimes be accomplished by considering the general readership of the news outlet that has published it or the types of stories it typically features. For example, The Guardian often attracts urban liberals, while The Daily Mail tends to have a conservative readership (Dentith, 2018). However, determining the intended audience of fake news presents challenges since information about readership is typically unavailable. In some cases, the intended readership can be tentatively inferred from certain characteristics of the fake news content. For instance, fake news that attempts to provide some evidence and minimizes obvious fabrications may appeal to a broader audience, while fake news that contains deliberate falsehoods may target a specific audience (Dentith, 2018). Nevertheless, such inferences are inevitably speculative.

The Definition Used in This Thesis

After examining various proposed definitions of fake news, it is clear that there is no universally accepted definition for this phenomenon. As a result, researchers often choose to emphasise specific aspects of fake news that align with their research objectives and enable efficient data collection (Asr & Taboada, 2019; Grieve & Woodfield, 2023). In light of this, it is important to establish a working definition of fake news that suits the purposes of this thesis. For the purposes of this thesis, *fake news is defined as false information that has been verified as false by fact-checkers, presented in the format of a news article, and published on a news website or blog*.

When referring to the data used in this thesis, the term ‘false news’ will be used instead of ‘fake news’ to reflect this more precise definition.

2.3. Linguistic Approaches to the Study of Fake News

In recent years, there has been an increasing amount of research that employs linguistic approaches to investigate fake news. These studies can be classified into four main categories based on their methodologies: computational linguistic analysis, forensic linguistic analysis, register analysis, and discourse analysis. The following sections provide a comprehensive review of studies conducted within these four approaches, emphasising their respective strengths and limitations.

2.3.1. Computational Linguistic Analysis

Researchers in the field of computational linguistics have conducted numerous studies to automate the detection of fake news. Most of these studies approached fake news detection as a classification task. They analysed fake news to identify its unique features and then developed computer algorithms that can determine whether texts are true or false based on these features (Asr et al., 2023; Bondielli & Marcelloni, 2019; Demestichas et al., 2020; Oshikawa et al., 2020; Zhou et al., 2020). This approach has been used to analyse different types of false information, such as propaganda, satire, and hoaxes, across various formats including tweets, Facebook posts, and news articles. The effectiveness of this approach is evident in its high accuracy rates. For example, the detection method proposed by Gravanis et al. (2019) achieved an accuracy rate of up to 95%. In addition, these studies have contributed to the development of user-friendly systems enabling users to assess the credibility of news items (Demestichas et al., 2020; Zannettou et al., 2019) like Fakebox[2] and Discourse Processing Lab[3].

Two types of features are commonly used for developing automatic fake news detection algorithms: style-based features and context-based features (Bondielli & Marcelloni, 2019; Demestichas et al., 2020; Shu, Wang, et al., 2019; Zhou et al., 2020). Style-based features (also called content-based features) primarily involve language features extracted from the text using natural language processing (NLP) tools. Alternatively, context-based features include

[2] https://machinebox.io/docs/fakebox
[3] http://fakenews.research.sfu.ca/

information about users' characteristics, social network propagation, and reactions of other users to the news item. While most fake news detection studies focus on style-based features, e.g., Gravanis et al. (2019), there are a few studies that specifically utilize context-based features to automate fake news detection, e.g., Castelo et al., (2019) and Vosoughi et al. (2018) (Bondielli & Marcelloni, 2019). However, some studies recognize the importance of considering both types of features to develop effective automatic fake news detection algorithms, as demonstrated by Shu et al. (2019). In this literature review, the focus will be on style-based NLP studies, as they are directly relevant to the thesis.

In NLP studies, three main steps are typically involved. Firstly, there is the text preprocessing or text representation, which includes tokenizing and stemming. In this step, textual data are transformed into numerical or vectorized representations that can be processed by machine learning algorithms (Ahmad et al., 2020; Bondielli & Marcelloni, 2019; Conroy et al., 2015; Demestichas et al., 2020; Oshikawa et al., 2020). The second step involves extracting linguistic features that are crucial in constructing a classifier algorithm capable of automatically detecting fake news (Ahmad et al., 2020; Bondielli & Marcelloni, 2019; Conroy et al., 2015; Demestichas et al., 2020; Oshikawa et al., 2020). Researchers have conducted multiple experiments to identify the best combination of features that could detect fake news with high accuracy (Gravanis et al., 2019).

Researchers have successfully identified multiple linguistic features that characterize fake news, which can be broadly grouped into seven categories: syntactic, lexical, semantic, discourse structure, complexity, sentiment, and subjectivity features. A concise overview of these linguistic features utilized in previous studies for automating fake news detection is presented in Table 2.1. According to the seven categories, I have classified the NLP tools and methods utilized for feature extraction and provided examples of features commonly associated with fake news in comparison to real news. Although some features may fall under multiple categories, I have roughly sorted them into the seven identified groups.

Table 2.1 Linguistic Features Utilized in Creating Fake News Detection Algorithms

Linguistic Features	**NLP Tools/Methods Used in Extracting the Features**	**Example Features that Typify Fake News in Comparison to Real News**
Syntactic Features: refer to sentence-level elements such as the frequency of content or function words, part-of-speech (POS) patterns, and punctuation	• POS Taggers and parsing trees, e.g., Natural Language Toolkit (NLTK) (Bird et al., 2009)	• Fake news uses more personal pronouns (Asr et al., 2023; Horne & Adalı, 2017), more adverbs (Asr et al., 2023; Horne & Adalı, 2017; Rashkin et al., 2017), especially modal and manner adverbs (Asr et al., 2023), more comparatives (Asr et al., 2023) and superlatives (Rashkin et al., 2017), and fewer punctuation marks (Asr et al., 2023; Horne & Adalı, 2017)
Lexical Features: include word-level aspects such as word frequency, characters per word, and unique words	• Term frequency • Term frequency-Inverse Document • N-gram models • Specialised lexicons • Bag of words	• Fake news uses more hedging language, while real news uses more assertive words (Rashkin et al., 2017; Volkova et al., 2017). • Fake news uses fewer nouns (Horne & Adalı, 2017)
Semantic Features: concern the meanings of words, phrases, and sentences	• Linguistic Inquiry and Word Count (LIWC) (Tausczik & Pennebaker, 2010) • Word embeddings, e.g., Word2Vec (Mikolov et al., 2013) • Bag of words • N-grams models	• Fake news appeals to moral foundations through the use of terms related to harm/care, loyalty/betrayal, and authority (Volkova et al., 2017) • Fake news uses more swearing, sexual language, perceptual words related to seeing, and negation (Rashkin et al., 2017) • Fake news is characterized by semantic unrelatedness, such as the introduction of new entities in the final sentence of an article or contradictions between the lead and final sentences (Rubin et al., 2016) • Words related to sex, death, anxiety, and biological and cognitive processes are used more frequently in fake news • Fake news's heavily weighted n-grams are related to divisive topics, e.g., 'Trump' and 'liberals', dramatic cues, e.g., 'breaking', abstract generalisations, e.g., 'truth' and 'freedom', vaguely facetious hearsay,

		e.g., ‘reportedly’ and ‘confirm’, or specific issues, e.g., ‘Syria’ and ‘vaccine’ (Rashkin et al., 2017)
Rhetorical Relations and Discourse Structure Features: concern features that relate and organize parts of a text	Rhetorical Structure Theory Parsers (e.g. the one by Mann & Thompson, 1988)	• Fake news has more isolated discourse units and a longer distance between units and their dependants, indicating that fake news is less coherent (Karimi & Tang, 2019) • Rhetorical relations, including conjunction (linked items), elaboration (additional information), evaluation, list, sequence, (next items), means (method) and solution-hood (a situation or method supporting satisfaction of the need) appear more in fake news, while summary (restatement of the content), preparation (making the text ready for reading), unconditional (independent), and disjunction (alternatives) appear only in fake news (Rubin & Lukoianova, 2015)
Complexity/ Readability: includes features like type-token ratio, depth of syntactic tree, average word length, and word fluency (i.e., how common a word is)	• Tools used to measure type/token ratio, average word length, depth of syntactic tree, and word fluency, e.g., Stanford Parser (Toutanova et al., 2003). • Grade-level readability indexes, e.g., Gunning Fog Index	• Fake news articles tend to be shorter (Horne & Adalı, 2017) • Fake news has more short words(Asr et al., 2023; Horne & Adalı, 2017; Rubin et al., 2016), less technical language (Horne & Adalı, 2017; Rubin et al., 2016) fewer quotes, and greater lexical redundancy (Horne & Adalı, 2017)
Sentiment: refers to the emotional tone or attitude expressed in a text	• Sentiment analysis conducted by tools like SentiStrength (Thelwall et al., 2010) or LIWC	• Fake news has been found to be more negative (Asr et al., 2023; Dey et al., 2018; Horne & Adalı, 2017; Rubin et al., 2016)
Subjectivity: refers to the extent to which a piece of text expresses opinions, beliefs, emotions, or personal perspectives	• Tools used for extracting lexical features were used to examine subjectivity • Subjectivity lexicons, e.g., the lexicon created by Recasens et al. (2013)	• Fake news contains more bias markers (Dey et al., 2018; Volkova et al., 2017) and more subjective words (Dey et al., 2018; Rashkin et al., 2017; Volkova et al., 2017)

The third step involves the development of a classifier algorithm that automatically detects fake news using machine-learning models (Ahmad et al., 2020; Bondielli & Marcelloni, 2019; Conroy et al., 2015; Demestichas et al., 2020; Oshikawa et al., 2020). Machine learning models are algorithms that learn from data to make predictions or decisions about new data without explicit programming. Most of the existing research on fake news detection uses supervised machine learning (Oshikawa et al., 2020). This entails training the algorithms on labelled news data, where the labels indicate whether the news is fake or real, to identify patterns and relationships between the extracted features and the corresponding labels. Examples of machine learning models used in fake news detection include Support Vector Machines, Naive Bayes Classifiers, Logistic Regression, and Decision Trees (Demestichas et al., 2020; Oshikawa et al., 2020).

Other computational linguistic research has used a different approach called deep learning, a more complex subfield of machine learning (Oshikawa et al., 2020). Deep learning trains deep neural networks with multiple layers, taking inspiration from how the human brain processes information. This allows the networks to automatically extract relevant features from news text and identify complex patterns. This approach is particularly effective for processing large amounts of data (Asr et al., 2023). Recurrent Neural Networks and Convolutional Neural Networks are commonly used deep learning models in fake news detection (Demestichas et al., 2020; Oshikawa et al., 2020).

It should be noted that some studies used a combination of machine learning models to achieve high accuracy of fake news detection (Ahmad et al., 2020; Gravanis et al., 2019a). For a comprehensive review of proposed fake news detection algorithms, Bondielli & Marcelloni (2019) and Zannettou et al., (2019) provide detailed analyses of these algorithms, highlighting their methodologies and datasets.

An example of NLP research is the notable study by Horne and Adalı (2017) who analysed fake news (i.e., stories known to be false and obtained from well-known fake news sources), in comparison to satire (i.e., stories obtained from sources that explicitly indicate they are satirical), and real news (i.e., stories known to be true and acquired from trusted sources). The researchers utilized various NLP tools, including NLTK, LIWC, Stanford Parser, Gunning Fog Index, and

SentiStrength, to analyse the data for stylistic features (such as nouns, pronouns, online slang terms), complexity features (such as type-token ratio, depth of syntactic tree, average word length), and psychological features (such as emotion words, risk words, strength of positive words).To develop the classifier algorithms, the authors initially employed ANOVA and Wilcoxon Rank Sum tests to identify the distinguishing features among the three news categories. Subsequently, they proposed a Support Vector Machine classifier based on the top four features of the article titles as well as the top four of the body texts. Their classifier achieved an accuracy of 67% in separating fake news from satirical news and 71% in classifying fake news from real news based on the top body texts' features.

Horne and Adalı's (2017) analysis showed that fake news exhibits certain features that differentiate it from real news. Firstly, the content of fake news exhibits similarities with satire news in various aspects. This includes the usage of shorter and less technical words, a reduced presence of analytical language, and a lower frequency of quotes, nouns, and punctuation marks. In terms of fake news titles, they are typically longer than titles of real news stories and use simpler, shorter, and less technical language. These titles also feature more all-capitalized words and proper nouns while containing fewer nouns and stop words. The authors suggested that this trend indicates that creators of fake news aim to pack as much information as possible into their titles.

Although Horne and Adalı's (2017) study aimed to identify linguistic features that could distinguish fake news from real news automatically, it also provided valuable insights into the discursive characteristics of fake news. One notable finding was that fake news adopts a heuristic approach to persuasion by presenting assumptions without logical reasoning or factual basis. This is evident in the abundance of information in fake news titles, which aim to provide as much content as possible, while the body texts of fake news often simply repeat the claims from the titles without adding substantial information. The authors suggested that this strategy might be intended to target readers who do not go beyond reading the titles.

The above review highlights the extensive research conducted on the linguistic features of fake news and their utilization in fake news detection. These studies clearly demonstrate the existence of linguistic variation between fake and real news. However, it is important to acknowledge the

limitations of these studies. One limitation is that many of them prioritize maximizing the accuracy of classification algorithms, rather than explicitly identifying the specific features that characterize or are more prevalent in fake news. As a result, these studies often do not offer an in-depth linguistic analysis or a comprehensive examination of the discursive and rhetorical aspects of fake news.

A second limitation is that the features utilized for automated fake news detection mainly reflect the linguistic content of the analysed articles (Grieve & Woodfield, 2023). Moreover, these features might be specifically related to the topics covered in the analysed news articles (Castelo et al., 2019). Therefore, it is crucial to approach the findings of these studies with caution for two reasons. First, some studies primarily concentrate on specific domains, particularly the realm of politics (e.g., Horne & Adalı, 2017). Thus, a proposed detection algorithm trained on features relevant to a specific domain may perform optimally within this domain while potentially being less effective when applied to articles from other domains (Ahmad et al., 2020). Second, these studies primarily delve into the linguistic content of fake news without considering how the broader social, cultural, political, and historical context in which fake news originates might affect its linguistic content. Fake news is a discursive phenomenon intertwined with the sociopolitical landscape, and such contextual factors are inevitably reflected in the language of fake news. Hence, there is a pressing need to take into account the contextual environment in which fake news is produced (Castelo et al., 2019; Shu et al., 2019).

Another limitation in the studies on detecting fake news is the limited utilization of linguistic, psychological, and social theories to inform the selection or extraction of linguistic features (Grieve & Woodfield, 2023; Zhou et al., 2020). These theories would provide a solid conceptual framework for understanding the factors that influence the creation and spread of fake news. The lack of these theories may impede the ability to fully capture the complex nature of fake news discourse. However, an exception is a recent research by Zhou et al. (2020) which addressed this limitation by incorporating theories from social and forensic psychology, such as *reality monitoring* and *four factor* theories, to develop a comprehensive list of syntactic, lexical, and discourse-level features for building a detection algorithm.

Another drawback is that although these studies focus on the linguistic differences between fake news and real news, they do not explain why these differences exist. Moreover, these studies do not acknowledge the similarities between fake and real news. As Bondielli and Marcelloni (2019) put it, in recent times, fake news has been employing a writing style that closely resembles genuine news (Bondielli & Marcelloni, 2019). Consequently, the proposed algorithms may have restricted applicability in real-world scenarios and may become less effective.

Lastly, a significant limitation in fake news detection studies is the use of unrepresentative data. Many studies have analysed data that do not accurately represent fake news. For instance, some studies analysed news texts from sources known for publishing fake news (e.g., Horne & Adalı, 2017; Rashkin et al., 2017; Rubin et al., 2016; Volkova et al., 2017). However, these studies failed to consider that these sources sometimes do publish real news, meaning the analysed news texts may not be necessarily false. The data collected in this manner are unsuitable for learning the linguistic patterns of fake news and instead capture the general writing style of specific news sources (Asr et al., 2023; Asr & Taboada, 2019). In addition, certain studies have employed artificial news articles that were either written by journalism students in training (e.g., Rubin & Lukoianova, 2015) or created by recruited participants (e.g., Pérez-Rosas et al., 2017). However, these artificial articles may not accurately capture the real-world context of fake news. Consequently, concerns arise regarding the reliability and applicability of the findings from these studies to real instances of fake news. Another issue is the failure to control for register variation. For instance, some studies have compared fake news sourced from social media platforms to real news from formal and traditional news outlets (e.g., Gravanis et al., 2019; Horne & Adalı, 2017) or to datasets that are not directly comparable to fake news (e.g., Rashkin et al., 2017, who compared fake news articles to a dataset derived from Gigaword, which contains facts, not news reports). Such comparisons could result in differences due to register variation rather than the text's truthfulness. Moreover, certain studies compared fake news articles to mainstream news articles obtained from a single news source (e.g., Ahmad et al., 2020), potentially resulting in a biased dataset due to each journalism organization's unique style guides (Gravanis et al., 2019).

In sum, NLP studies on fake news detection focusing on linguistic features have successfully demonstrated the existence of linguistic variation between fake and real news. However, as noted by Grieve and Woodfield (2023), these studies are limited in terms of feature selection,

interpretation of findings, and data collection methods. There remains a need for considerably more comprehensive linguistic and discourse research on fake news to address the limitations of NLP approaches to fake news detection.

2.3.2. Forensic Linguistics Analysis

Forensic linguistic analysis involves using linguistic theories and methods to investigate legal and criminal cases, considering language as evidence. To the best of my knowledge, there is one notable study conducted by Sousa-Silva (2022) who utilized forensic linguistic analysis to study fake news. The author posited that fake news could involve cybercriminal activities and emphasised the role of language in detecting and investigating fake news as well as in providing evidence in legal cases involving fake news. The author conducted an analysis on a small corpus of English and Portuguese articles that were sourced from outlets previously identified by fact-checkers as disseminators of hyper-partisan content, disinformation, and misinformation. A comparison was made with high-quality mainstream news articles. The analysis involved several stages. Initially, simple text statistics were calculated, including the average paragraph, sentence, and word lengths, as well as the usage of punctuation. Furthermore, idiosyncratic structural features at the typography, orthography, and morphosyntax levels, commonly observed in fake news texts, were identified. Lastly, discourse-level features like agency, theme, and rheme were examined using the principles of critical discourse analysis (CDA).

Although there were some differences between English and Portuguese fake news articles, the study revealed certain language-agnostic features (i.e., features that occur in both English and Portuguese fake articles). Firstly, paragraphs in mainstream news articles were of moderate length, whereas fake news articles exhibited a wide variation ranging from very short to very long paragraphs. Similarly, the length of sentences in mainstream articles fell in the middle range, while fake news sentences showed extreme variations at both ends, being both too short and too long. Another language-agnostic feature observed was that the use of punctuation in fake news articles was more associated with expressing emotions and opinions. This can be attributed to the high level of subjectivity present in fake news. Furthermore, fake news texts exhibited a notable number of grammar, spelling, and typographical errors. At the morphosyntax level, fake news articles utilized adverbs more extensively than mainstream articles, often reflecting negative evaluations. Regarding agency, theme, and rheme, fake news tended to foreground

positive actions of the interested party while portraying opposing parties in a negative and marginalized light.

While the study has the advantage of identifying language-agnostic features of fake news, it is important to acknowledge its shared limitations with other computational linguistic studies. One limitation is that the analysed articles were sourced from outlets known to publish fake news; this does not guarantee that all the analysed articles were fake. Furthermore, the analysis primarily focused on the linguistic content of fake news articles, examining superficial formal features such as word, sentence, and paragraph length, spelling, and grammar. However, the study did recognize the significance of going beyond linguistic indicators and incorporated CDA to explore agency, theme, and rheme.

2.3.3. Register Analysis

A notable recent study on fake news is the research conducted by Grieve and Woodfield (2023). They approached their analysis from the perspective of register theory, which suggests that differences in language use reflect distinct communicative purposes and contexts. The researchers argued that since the aim of fake news is to deceive, while genuine news aims to inform, there should be differences in their language usage. They also emphasised the importance of considering differences in communicative contexts, purposes, and attributes of the authors and audience when comparing fake news to real news, as these factors can contribute to register variation.

To address this issue, Grieve and Woodfield (2023) focused their analysis on the writings of a single author, Jayson Blair. Blair was a former employee of The New York Times newspaper in the early 2000s and gained notoriety for fabricating and plagiarizing content for personal reasons. The researchers utilized the linguistic features identified by Biber (1988) and examined the linguistic variations between Blair's fake and real news articles. Using Cliff's delta, they identified 28 substantial differences in the use of linguistic features between Blair's fake and real news.

Through an examination of the communicative functions and actual usage of these features within Blair's articles, Grieve and Woodfield (2023) identified two main patterns of variation: *information density* and *conviction*. When Blair told the truth, the authors argued, his writing

exhibited a denser style with greater conviction compared to when he was fabricating information.

The informationally dense style present in Blair's real news articles was characterized by features such as *long words, nouns, nominalizations, gerunds*, and *participial post-nominals*, as well as idea expansion markers like *that-relatives, infinitives*, and *time adverbials*. In contrast, Blair's fake news displayed a less dense style and was associated with features that: (a) replaced nominal density with various verb forms, such as *present tense, copula be*, and *perfect aspect*, and *predicative* and *attributive adjectives*, (b) emphasised informality using different types of *adverbs*, and (c) connected ideas using *subordinators, the pronoun 'it', third-person pronouns*, and *demonstrative pronouns*.

The study also found that Blair exhibited greater confidence and certainty in his truthful articles compared to his fabricated ones. *Suasive verbs, prediction modals, by-passives*, and *public verbs* were more frequent in his truthful writing, indicating the expression of stance and evidentiality and reflecting confidence in the information reported and his sources. On the other hand, Blair's fake news demonstrated a less confident and uncertain style, characterized by features such as *agentless passives*, *downtoners*, and *WH-relative clauses*.

Grieve and Woodfield's (2023) study is advantageous in two ways: First, it examined fake news articles whose writers' intentions are known, and second, it compared them to genuine articles written by the same author, effectively controlling for register variation. However, it is important to acknowledge that the study had some limitations. Firstly, the analysis was confined to the style of only one journalist, which greatly limits the generalizability of this research. Additionally, the study focused exclusively on articles from The New York Times, a mainstream news source, whereas fake news proliferates on various platforms such as tabloids, news websites, and blogs. Consequently, the fake news articles written by Blair may not accurately represent the type of fake news commonly encountered online, specifically the more polarizing and misleading kind, such as anti-vaccination news, which can have a harmful impact. Therefore, an analysis of articles written by various journalists and published on multiple news platforms is necessary to test the generalisability of this finding.

2.3.4. Discourse Analysis

To date, there have been limited studies applying the discourse analysis approach to the study of fake news. These studies have primarily analysed fake news discourse in accordance with established linguistic and sociological theories, including argumentation theory, Martin and White's (2005) appraisal theory, Goffman's theory of frame (1975) and theory of footing and alignments (1981), Walton's (2007) framework for identifying propaganda, and Van Leeuwen's (2007) discourse model of legitimation. Unlike computational linguistic studies, discourse studies offer valuable insights into the rhetorical strategies employed by fake news writers and how they construct arguments to persuade their readers.

For example, Alba-Juez and Mackenzie (2019) examined scientific and political news articles that were verifiably false, as well as articles obtained from known fake news publishers such as WhatDoesItMean.Com. They also investigated fake news shared on social media. The focus of their qualitative analysis was on how truth is manipulated and distorted using socio-pragmatic resources. To accomplish this, the authors employed argumentation theory, appraisal theory, and Goffman's (1975, 1981) theories of frame and footing and alignments, with specific attention to the expression of stance, emotions, and judgments.

Alba-Juez and Mackenzie's (2019) study revealed several rhetorical strategies employed by fake news writers. First, fake news on scientific topics often adopt a popular scientific writing style that involves nominalization and abstraction. Furthermore, quotes, cited references, and photos were fabricated to make the readers reject evidence presented by conventional science. Fake news also employed various argumentation fallacies, such as *ad hominem* and *ad populum*. In addition, fake news writers used heteroglossic engagement by attributing claims to others, thereby distancing themselves from personal responsibility and allowing readers to form negative evaluations by themselves. The study emphasised that the persuasive intent of fake news relied heavily on manipulating readers' emotions, particularly by evoking fear, envy, hatred, and negative judgments towards opposing groups.

Vamanu's (2019) study on fake stories with propagandistic impact highlighted similar findings. The study took an interdisciplinary approach, drawing insights from communication studies, argumentation theory, library and information science, and discourse research. To identify

instances of propaganda, the author utilized CDA along with Walton's (2007) framework for identifying propaganda, which is derived from argumentation theory.

Vamanu's (2019) study revealed the following results: First, fake stories were goal-oriented, conveying an implicit message while narrating an event. Second, these stories often relied on unfounded assumptions, lacked logical reasoning, and employed fallacious arguments like *false dilemma*. Third, they frequently presented a one-sided argument, selectively emphasising a single aspect while disregarding alternative perspectives. Fourth, fake stories sought to trigger emotional responses from readers by negatively evaluating opposing groups. Fifth, propagandists positioned themselves as credible spokespeople for specific groups while simultaneously discrediting speakers from opposing groups. They skilfully employed powerful dichotomies, such as ‘good versus evil’ or ‘patriot versus traitor,’ to marginalize certain individuals or groups. The author concluded that CDA can serve as a valuable tool for detecting and debunking fake news.

While Alba-Juez and Mackenzie's (2019) and Vamanu's (2019) studies provided valuable conclusions about argument and persuasive strategies in fake news discourse, they share the limitation that they did not provide a clear definition of fake news. A clear definition enables effective data collection (Asr & Taboada, 2019), sharp analysis, and generalizability of the results. Additionally, both studies lacked clarity regarding their data collection procedures, such as the number of texts analysed, the genres represented, the topics discussed, the publication time frame, and the background information of the publishers, writers, or speakers. This weakens the reliability, replicability, and generalizability of these findings. Moreover, the texts used for analysis in Alba-Juez and Mackenzie's study were unbalanced, as they included texts that were both verifiably false and obtained from sources known for publishing fake news. It is possible that the data contain truthful information and real news discourse, as these sources occasionally publish real news. Consequently, the analysis of this type of data may focus on stylistic or register features specific to certain sources (e.g., Twitter or news blogs) (Asr et al., 2023; Asr & Taboada, 2019), rather than on the distinctive discourse of fake news. Furthermore, both studies lacked sufficient detail in describing their qualitative analysis methodologies. While examples were provided to support their findings, the extent to which the theoretical frameworks were

effectively applied remains unclear. This suggests a potential lack of systematic and comprehensive data analysis.

Another discourse analysis of fake news is the study conducted by Igwebuike and Chimuanya (2021). This study stands out from the previously two mentioned studies by providing a definition of fake news that allowed the collection of balanced data and by offering both quantitative and qualitative analyses, resulting in a comprehensive and systematic examination of the data. In their research, Igwebuike and Chimuanya defined fake news as “news comprises fabricated information geared towards legitimating falsehood to achieve political gains” (p.45). They focused on verifiably false news that circulated on social media during the 2019 Nigerian election, utilizing Van Leeuwen's (2007) discourse model of legitimization. Their study identified three primary legitimation strategies used to achieve persuasion in fake news, ranked by frequency: (a) authorization (referring to the authority of tradition, custom, law, or public figures), (b) moralization (referring to norms, moral values, and ethical systems), and (c) rationalization (referring to commonly accepted notions of what is right and rational). In addition, Igwebuike and Chimuanya performed a qualitative analysis to delve into the form and functions of these strategies. They argued that these strategies served as effective persuasive tools, conveyed through manipulated images, emotive language, appeals to emotions, logical conclusions, derogatory comments, verbal indictments, and coercive verbs.

The three studies mentioned thus far relied on pre-existing theories and conducted manual searches for evidence within their datasets. However, by utilizing this approach, there is a risk of overlooking certain patterns or discourse trends, as the analysis may be influenced by the researcher's cognitive biases (Baker, 2006) or the constraints of the selected theoretical frameworks. Consequently, only patterns that are readily apparent, directly relevant to the chosen theoretical framework, or personally interesting to the researcher may be considered. Furthermore, analysing fake news discourse solely based on pre-existing theories may not capture the distinct discursive characteristics of fake news. It is plausible that real news sources employ similar legitimization strategies or emotional manipulation techniques. Thus, it is imperative to analyse fake news discourse in comparison to real news to accurately identify the genuine differences between the two and the truly distinctive features of fake news.

The study conducted by Maci (2019) aimed to address this issue by adopting a multidisciplinary approach that combined corpus-based CDA and socio-semiotics to examine the discourse strategies used in tweets containing false information about vaccination. In the linguistic analysis, the author conducted a sentiment analysis and extracted key POS in 16,782 anti-vax tweets by comparing them to the spoken British National Corpus (BNC). In the visual analysis, pictures in the tweets were categorized based on their content. Overall, the linguistic and visual patterns identified in the corpus indicated that the discourse surrounding anti-vax sentiments was emotionally charged and emphasised the negative side effects of vaccination; approximately 35% of the tweets reflected negative emotions, while about 15% of the visuals portrayed an emotive response and another 15% highlighted the negative effects of vaccination. These patterns also revealed that the anti-vax discourse was grounded in scientific information and presented in a way that was easily understandable. Around 45% of the tweets conveyed scientific information without judgment, while about 52% of the pictures emphasised the importance of being informed. In addition, figures, statistics, and visual metaphors were used to establish credibility; for example, about 10% of the pictures highlighted the profits made by the pharmaceutical industry through the vaccination process.

However, it is important to note that while Maci's (2019) study utilized corpus-based CDA, the reference corpus used for the analysis, which was the spoken BNC, was not directly comparable to the corpus of anti-vax tweets. This is because the conversational language on various topics differs significantly from the language of microblogs specifically discussing vaccination. Consequently, the analysis might reveal differences in register and theme rather than genuinely distinguishing features between the language of fake news and real news. This is evident in the key POS analysis. For example, the most frequent words in the key POS analysis were items with hashtags. However, this prominence of hashtags is a characteristic unique to Twitter and cannot be found in spoken language. Another example is that the most frequent nouns included words like ‘children’, ‘autism’, ‘risks’, and ‘big pharma’, which are clearly associated with (anti)-vax discourse. To accurately identify key linguistic features in anti-vax tweets, the comparable corpus should be as balanced and representative as the focus corpus (cf. McEnery, 2005). In other words, the reference corpus should consist of tweets about vaccinations that are verifiably true, allowing for an accurate extraction of key POS found in anti-vax tweets.

Another corpus-based approach was employed by Dance (2019) who used the method of key semantic domains analysis (KSDA) (Rayson, 2008) to investigate the motivations behind the online sharing of disinformation. He compared a corpus of shared links containing disinformation with a reference corpus containing non-disinformation shared links and conducted a KSDA using Wmatrix (Rayson, 2009). The analysis identified the significantly most frequent semantic domains in the disinformation sharing corpus, including (a) *Closed; Hiding/Hidden* (containing words such as 'cover up', 'hidden', 'secret'), (b) *Open; Finding; Showing* (e.g., 'show', 'find', 'reveal'), (c) *Evaluation: False – words conveying falsehood* (e.g., 'lies', 'disingenuous', 'fabricated'), (d) *Negative Judgment of Appearance* (e.g., 'nasty', 'horrible, 'soul-less'), and (e) *Degree Booster* (e.g., 'really', 'very', 'extremely'). Dance argued that these domains suggest that people engage with and share fake news because they might believe it uncovers hidden truths, and that emotional involvement also plays a role in reading and sharing fake news. Dance asserted that the emotional element could impede critical thinking and lead to the acceptance of fake news.

Dance's (2019) study has several advantages over previous studies. Firstly, it focuses specifically on the analysis of *shared* disinformation, highlighting its viral nature and impact on individuals. Secondly, it is a corpus-based discourse analysis, which allows for a systematic and replicable analysis of the prevalent meanings in fake news. Most importantly, Dance's research emphasises the importance of analysing the semantic resources used in fake news as they might provide insights into the reasons behind its creation, consumption, sharing, and belief (cf. news values in Bednarek & Caple, 2017).

After reviewing discourse studies on fake news, several noteworthy conclusions can be drawn. Firstly, establishing a clear definition of fake news is imperative, as it enables focused research, the compilation of a representative corpus, and rigorous interpretation. Secondly, a comparative analysis approach is essential for a comprehensive understanding of fake news discourse (Jankowski, 2018). This can be accomplished through corpus-assisted analysis to identify key characteristics. Furthermore, it is crucial to account for register variation when conducting a comparative analysis, ensuring a comparison between a representative focus corpus of fake news and a comparable corpus, rather than one from a different register. Thirdly, although social media platforms have received significant attention in discourse studies on fake news, there is a need

for more targeted investigations into full-length fake news articles published on news websites and blogs. These articles can be more hazardous due to their close resemblance to real sources in formatting.

2.3.5. Related Study

In a related study by Mourão and Robertson (2019), a linguistic analysis was not conducted. Instead, they conducted a content analysis of news stories published by prominent fake news sites during the 2016 US election. The researchers categorized the content of these sites based on the degree of four factors: misinformation (factual errors), bias (political or ideological leaning), clickbait (headline manipulation to entice readers to click), and sensationalism (emphasis on emotional and dramatic elements). They also measured the engagement of users on Facebook and Twitter for each category. The findings indicated that the content of these news sites predominantly consisted of biased content rather than outright fabrications or sensationalism. In addition, this biased content garnered the highest levels of social media engagement and sharing. The study further highlighted that fake news content often blended genres, integrating elements of news, entertainment, and opinion.

However, it is important to note some limitations of this study, which also apply to some of the previously reviewed studies. Firstly, the analysis was conducted on articles sourced from sites known to publish fake news; however, these sites may also publish real news. This means that some of the analysed content may contain accurate information. Additionally, the study focused specifically on political content, which may limit its generalizability to other types of fake news discourse. Despite the absence of linguistics or discourse analysis in this study, it underscores the significance of examining fake news discourse from a genre perspective (cf. Biber et al., 2007). Genre analysis can provide insights into the structure of fake news, identifying its constituent elements (i.e., moves) and their communicative functions.

2.3.6. Bridging the Gaps in Fake News Studies

The above literature review demonstrated that various linguistic approaches have been employed to analyse the language of fake news, including computational linguistic analysis, forensic linguistic analysis, register analysis, and discourse analysis. These approaches offer distinct insights into the linguistic features and discursive strategies utilized in fake news articles and

social media posts. However, there are still areas that require consideration and further exploration in understanding fake news discourse, which are the focus of this thesis.

First of all, it is crucial to develop a representative corpus of fake news for analysis, conduct a comparative analysis, and consider register variation in the process. The common research practice of comparing online fake news articles to traditional printed broadsheet articles can highlight differences due to register variations rather than to the veracity of the analysed articles. Therefore, comparing fake news to various sub-registers of news can ensure a balanced comparison and enable the identification of distinctive features in fake news discourse.

Secondly, a research gap exists regarding the discourse structure of fake news articles. Analysing the structure of fake news articles from a genre perspective can aid in identifying their communicative purposes, understanding how the news genre is manipulated to achieve these purposes, and examining how information is strategically presented to emphasise specific values and beliefs.

Thirdly, while discourse studies have shed light on the rhetorical strategies that make fake news persuasive, it is crucial to investigate the rhetoric of fake news from a news values analysis perspective (cf. Bednarek & Caple, 2017). Since fake news is bound by the genre conventions of journalism, studying the news values of fake news is essential. This analysis can help unravel its distinctive discourse, comprehend its appeal, and uncover the ideological underpinnings within its discourse.

Fourthly, although numerous studies have attempted to automatically detect fake news by analysing its style and comparing it to real news, it is crucial to delve deeper into the stylistic choices employed in fake news. This involves examining a comprehensive range of lexico-grammatical features, their communicative functions, and how they are manipulated to make fake news more appealing.

Lastly, it is crucial to recognize the impact of situational characteristics on the production of fake news, as highlighted by Jankowski (2018). These contextual factors can greatly influence the discourse structure, news values, and stylistic choices employed in fake news. The subsequent

section reviews these situational characteristics, which have been the focus of sociology, psychology, and media literacy research on fake news.

2.4. Situational Characteristics of Fake News

This section examines the situational characteristics of fake news which encompass various aspects related to the circumstances in which fake news is produced and disseminated. Key situational characteristics include the participants involved, communication channels utilized, processing circumstances, setting, communicative purposes, and topics covered. Analysing these situational characteristics is essential for comprehending why fake news differs or resembles real news and how fake news appeals to its audience.

To examine the situational characteristics, I reviewed previous surveys of fake news studies, covering both theoretical and experimental research (e.g., Baptista & Gradim, 2020; Parikh & Atrey, 2018; Shu et al., 2017; Zhang & Ghorbani, 2020) along with studies from the fields of sociology, psychology, and media literacy. In this review, I focused on the studies' findings rather than methodologies as they are not based on linguistics.

To summarize the accumulated research on the situational characteristics of fake news, I adopted the framework proposed by Biber and Conrad (2019). This framework was specifically developed to analyse the situational characteristics of a register. By applying this framework, I compiled and presented the findings related to the situational features of fake news in Table 2.2. Each of these characteristics is discussed in detail in the rest of this section.

Table 2.2 Situational Characteristics of Fake News, as Identified in Previous Studies

Biber and Conrad's (2019) Framework Classifications		**Findings from Previous Studies**
Participants	Addressor: Creators and Propagators	• Malicious authors who create and spread fake news intentionally • Benign authors and users who create or spread fake news unintentionally
	Addressee	They could be: • Students • Voters • Parents • Senior people

		• People who belong to a certain political, religious, or ideological group, etc.
	Consumers of fake news	They could be: • Intuitive people • People with psychotic thoughts • Less literate and less well-educated people • Very young and very old users • Female users • Users who spend a significant amount of time on social media • Users with highly active political, religious, or ideological identities or participation
Relationships amongst Participants		Fake news is likely to be consumed by readers who share the same ideological, political, or religious beliefs as those of the writers of fake news
Channel	Medium	News websites and blogs and social media platforms
	Mode	Text, audio, or video
Processing Circumstances	Production	Fake news originates as planned writing and goes through three phases: • Creation • Production • Distribution and reproduction
	Comprehension	Fake news can be: • Fully accepted • Partially accepted • Rejected
Setting		The production of fake news and its perception may not always coincide in time or location The popularity of fake news depends on its social community and context
Communicative Purposes	Fake News Creation Purposes	• Ideological: imposing beliefs • Financial: generating advertising profits • Social or psychological: gaining status or a sense of control

	Fake News Sharing Purposes	**Intention-to-deceive motivations:** • Ideological • Financial • Social or psychological **General news-sharing motivations:** • Attaining social approval • Adherence to norms • Showing off knowledge • Sharing emotional or controversial content
Topics		• Covers a wide range of topics • Focuses on differences and divisions amongst political parties, religious groups, nationalities, races, and socioeconomic classes • Sheds light on controversial, surprising, bizarre, sensational, exaggerated, and emotional topics

2.4.1. Fake News Creators, Propagators, and Consumers

Fake news stories might be created and spread either intentionally by malicious authors or unintentionally by benign authors or users (X. Zhang & Ghorbani, 2020). People who intentionally create fake news can include journalists, criminals, terrorist organizations, activists, governments, or individuals who have something to gain from spreading false information. On the other hand, people who unintentionally spread fake news may include true believers, conspiracy theorists (Zannettou et al., 2019), or people who are mistaken or lack journalistic professionalism in verifying information (Baptista & Gradim, 2020). It should be noted that fake news can be spread by bots, which are intentionally programmed to diffuse fake news (Zannettou et al., 2019). Based on the goals of the authors or spreaders, the addressees can be students, voters, parents, senior people, or certain religious or political groups (Rini, 2017; X. Zhang & Ghorbani, 2020).

Psychological theory suggests that there are two psychological factors inherent in human nature that contribute to people's susceptibility to fake news (Shu et al., 2017). The first factor, known

as *naïve realism*, reveals that individuals tend to believe that their own perception of reality is true and regard differing opinions as irrational, biased, or uninformed. The second factor, termed *confirmation bias*, refers to the tendency for individuals to actively seek information that supports their existing beliefs. Consequently, people are more prone to accepting and believing fake news that aligns with their preconceived notions.

Experimental psychological research (e.g., Bronstein et al., 2019; Martel et al., 2020; Pennycook & Rand, 2020, 2021) revealed that belief in fake news is associated with various factors. These include a tendency to accept statements without concern for the truth (referred to as *bullshit receptivity*), delusionality, dogmatism or overclaiming, emotionalism, religious fundamentalism, and political partisanship. Research reveals that individuals who rely less on analytical and critical thinking skills, prefer simplistic reasoning, and seek easy access to news are more likely to consume and believe fake news (Bronstein et al., 2019; Pennycook & Rand, 2020, 2021). In addition, users who hold psychotic thoughts, believing in conspiracy theories and paranormal phenomena, are also more susceptible to fake news (Bronstein et al., 2019; Pennycook & Rand, 2021). Furthermore, research indicates that less literate and less educated individuals (Baptista & Gradim, 2020), very young and elderly users (Shu et al., 2018), and female users (Shu et al., 2018; Zannettou et al., 2019) are more susceptible to fake news. Further, individuals who claim to have higher levels of knowledge are also more prone to believing fake news (Pennycook & Rand, 2020, 2021). Moreover, emotional reactions play a role, as individuals who experience strong emotions, particularly negative emotions, when exposed to fake news are more likely to believe it (Fernández-López & Perea, 2020; Martel et al., 2020).

Furthermore, spending significant time on social media can increase susceptibility to fake news due to repeated exposure to false content and beliefs, known as the *illusory truth effect* (Baptista & Gradim, 2020; Pennycook et al., 2018; X. Zhang & Ghorbani, 2020). This can lead to the formation of *echo chambers*, where individuals primarily encounter false information, ideas, and opinions that align with their existing beliefs (Baptista & Gradim, 2020; X. Zhang & Ghorbani, 2020).

Political, religious, and ideological identities also influence susceptibility to fake news, with active participants in these domains often falling victim to fake news that supports their existing

beliefs (Bronstein et al., 2019; Pennycook & Rand, 2021). For instance, conservatives, republicans, and right-wingers tend to be more prone to believe fake news (Baptista & Gradim, 2020). This highlights a relationship between the writers and readers of fake news, as readers who share ideological, political, or religious beliefs similar to the fake news writers are more likely to consume and believe such news (Pennycook & Rand, 2021; Rini, 2017).

2.4.2. Motivations for Creating and Sharing Fake News

The motivations behind fake news can be broadly classified into two categories: (a) creating motivations and (b) sharing motivations. The existing literature shows that there are three primary motivations for creating fake news (Wardle & Derakhshan, 2017; Zannettou et al., 2019):

a) Ideological motivation: This includes supporting or denigrating political or religious figures or groups and imposing specific beliefs.
b) Financial motivation: This involves generating advertising opportunities through clickbait.
c) Social or psychological motivation: This motivation can be driven by the desire to gain attention, status, and acceptance within online communities. It can also involve gaining control by creating false content that causes chaos or entertains others.

The motivations for sharing fake news can be categorised as: (a) intention-to-deceive motivations and (b) general news-sharing motivations. Like the motivations of creating fake news, intention-to- deceive sharing motivations can be ideological, financial, and social or psychological while general news-sharing purposes include (Baptista & Gradim, 2020):

a) Attaining social approval: Making friends, gaining followers online, and reinforcing popularity
b) Adherence to norms: Disseminating information that aligns with the norms of a particular community (cf. *social identity theory* and *normative influence theory* in Shu et al., 2017)
c) Showing off knowledge and expertise
d) Sharing emotional or controversial content
e) Sharing partisan or polarized political content that supports one's political beliefs (Osmundsen et al., 2021a)

2.4.3. Fake News Production and Comprehension Circumstances

The creation of fake news follows a similar process as real news production, involving three stages: conception by the author, publication on the web, and subsequent sharing by interested individuals leading to reproduction (Wardle & Derakhshan, 2017). Unlike spoken registers, fake news begins as planned writing and is not limited by time constraints. However, the production of news articles, both fake and real, often necessitates a swift response to time-sensitive events. It is important to note that fake news typically lacks the editorial processes that ensure the reliability of information, in contrast to trustworthy news published by reputable sources (Lazer et al., 2018). Like real news, fake news is disseminated through a variety of channels, including news websites, blogs, and popular social media platforms like Facebook, Twitter, WhatsApp, and Instagram (Parikh & Atrey, 2018). It can take the form of written texts in the form of articles or memes, as well as audio or video content (Wardle & Derakhshan, 2017).

In terms of comprehending fake news, individuals tend to respond to fake news stories in one of three ways: (a) accepting the story entirely (hegemonic), (b) accepting certain aspects while questioning others (negotiated), or (c) outright rejecting the story (oppositional) (Wardle & Derakhshan, 2017). As fake news is predominantly spread online, its setting, including the specific time and place of communication, is not commonly shared by the authors and recipients. In addition, fake news can be archived and accessed online for extended periods of time, contributing to its potential longevity and impact.

2.4.4. Fake News Viral Dissemination

Vosoughi et al. 's (2018) study on fake news spreading on Twitter revealed that fake stories propagate more quickly and broadly than real stories. In fact, the popularity and rapid dissemination of fake news depend on the social context or community (Baptista & Gradim, 2020; X. Zhang & Ghorbani, 2020). Online users often share news within specific social groups, such as friends and followers. When a large group of individuals with similar opinions share and promote a particular news item, it has the potential to go viral and amplify the influence of its message. This can contribute to the creation of an *illusionary truth effect* and *echo chambers,* as previously mentioned (Baptista & Gradim, 2020; X. Zhang & Ghorbani, 2020). Rini (2017) argued that social media platforms act as a form of testimony, where users accept fake news

based on the transmission and acceptance of others who share and believe it to be true. This type of biased testimony, Rini contends, is often perpetuated through social media when the testifier and audience share a partisan orientation.

The literature has identified multiple factors that contribute to the virality and believability of fake news:

a) Emotional content: Fake news often utilizes striking headlines and exaggerates stories to evoke negative emotions such as shock, fear, anger, and moral outrage. This type of content attracts public attention (Baptista & Gradim, 2020; Fernández-López & Perea, 2020; Pennycook & Rand, 2021).
b) Political content: It was found that political fake stories, especially those with partisan polarization spread widely and rapidly (Vosoughi et al., 2018), exhibiting particularly more intensified effects when compared to other fake stories related to terrorism, science, or urban legends (Vosoughi et al., 2018).
c) Heuristic persuasion: Unlike real news, fake news persuades through heuristic means, using informal and simple language without logical or well-supported arguments. This approach appeals to readers who rely on cognitive shortcuts and trust news sources easily, often not engaging in critical thinking or reading beyond the headlines (Bronstein et al., 2019; Horne & Adalı, 2017; Martel et al., 2020; Pennycook & Rand, 2020, 2021; Rini, 2017).
d) Imitation of journalistic format: Fake news attempts to mimic real news in terms of website design, article structure, and photo attributions. This imitation makes it challenging for readers to distinguish between real and fake news (Dentith, 2016; Rini, 2017).
e) Clickbait: Fake news articles often employ long, content-rich headlines with hyperbolic words, enticing readers to click on them (Baptista & Gradim, 2020; Horne & Adalı, 2017).

2.4.5. Fake News Topics

Fake news can cover a wide range of financial, social, political, and technology-related topics (X. Zhang & Ghorbani, 2020). However, fake news tends to focus on topics that accentuate

differences and divisions amongst political parties, religious groups, ethnicities, nationalities, races, and socioeconomic classes (Wardle & Derakhshan, 2017). Moreover, fake news usually sheds light on controversial, surprising, bizarre, sensational, exaggerated, and emotional topics. Examples of these topics are money laundering, crimes, particularly sexual offences, and fraudulent, imaginary, and politically motivated inventions (Baptista & Gradim, 2020). This thesis specifically examines false news articles related to the controversial topics of vaccinations, climate change, and COVID-19, which will be explored in detail in the subsequent chapter.

To conclude, while studies in the fields of psychology, sociology, and media literacy have provided valuable insights into the situational factors influencing the production and consumption of fake news, they do have certain limitations. Primarily, some of the studies are more theoretical or observational in nature rather than empirical (e.g., Dentith, 2016; Lazer et al., 2018; Rini, 2017). The empirical research that does exist is often limited in scope and focused on specific populations. For example, Martel et al. (2020) and Pennycook & Rand (2020) recruited participants from Amazon Mechanical Turk. Such participants may not be representative of the general population as they are typically young and skilled in using technology and usually engage in studies for payment. These two limitations raise questions about the generalizability of the findings. Moreover, these studies do not delve into how fake news resonates with its target audience or why certain groups, such as conservatives or religious individuals, are more susceptible to it. Moving forward, discourse analysis can enrich our comprehension by scrutinizing the language of fake news while considering these situational factors. This analysis has the potential to reveal more about the motivations driving the creation and dissemination of fake news. Furthermore, it can offer a better understanding of why certain audiences are attracted to fake news, how writers tailor their content to appeal to their audience, why particular fake news topics gain popularity, and what discursive elements contribute to the viral spread of fake news.

2.5. Conclusion

In conclusion, the extensive literature review conducted across various disciplines, including linguistics, psychology, sociology, and media literacy, has revealed that fake news is a timely and extensively researched topic. The review has emphasised the ongoing debate surrounding the

definition of fake news and the need for a clear and specific definition that aligns with research objectives and facilitates systematic data collection, focused analysis, and rigorous interpretation.

Furthermore, the review has identified distinct linguistic variations between fake news and real news, as well as the utilization of rhetorical strategies to enhance the credibility of fake news. However, further investigation is required, specifically in the examination of a balanced and representative corpus of fake news, comparing it to comparable real news, and exploring its discourse at different levels.

Moreover, the review has highlighted the substantial amount of research, both theoretical and experimental, on the situational factors related to the creation, distribution, and consumption of fake news. These research findings are crucial for understanding the phenomenon of fake news and should inform any analysis of its discourse. Indeed, conducting a discourse analysis on fake news can not only leverage the insights from these studies but also contribute valuable perspectives regarding the situational characteristics of fake news.

This comprehensive literature review from an array of disciplines not only delineates existing gaps and unanswered questions in our comprehension of fake news but also furnishes valuable insights for the multi-level analysis undertaken in this study. A thorough understanding of the definitions, linguistic aspects, rhetorical strategies, and situational characteristics of fake news is crucial for a more nuanced analysis and interpretation of this complex discursive phenomenon. The subsequent chapter will focus on the corpus compilation for this thesis.

Chapter 3: Compiling the Focus and the Reference Corpora

3.1. Introduction

This chapter focuses on the data collection process for the thesis, which analyses false news discourse related to climate change, vaccination, and COVID-19. These topics are highly controversial and have resulted in the spread of significant amounts of false information in recent years, with important societal implications. To effectively analyse false news discourse on these topics, two key aspects must be addressed: the socio-historical context surrounding them and the criteria for creating a representative corpus. Therefore, the chapter begins by providing the socio-historical context for the three topics and justifying their selection for analysis. It then discusses general guidelines for designing a representative corpus, as well as the challenges and opportunities encountered during the compilation of a focus corpus of false news and a reference corpus for comparison. The chapter proceeds with a step-by-step procedure for compiling these two corpora, including a breakdown of their contents. Finally, the chapter discusses the limitations of the corpus used in this study and stresses the importance of considering these limitations when interpreting the findings.

3.2. Contextualizing Climate Change Denialism, Anti-vaccination Movement, and COVID-19 False Information

To properly interpret the analysis results of false news discourse on climate change, vaccination, and COVID-19, it is important to consider the socio-historical context of these topics. This section offers a concise overview of the socio-historical background and contextual factors associated with these three issues.

3.2.1. Climate Change Denialism

In June 1988, James Hansen, a NASA scientist, testified before the United States Senate Committee on Energy and Natural Resources regarding anthropogenic global warming. Since then, there have been persistent organized efforts, known as the *denial campaign*, to refute the existence of anthropogenic global warming (Dunlap & McCright, 2015). The major claims of the denial campaign are threefold: (a) denying the existence of global warming, (b) disputing its

human-made nature, and (c) downplaying its harmful impacts (Coan et al., 2021; Dunlap & McCright, 2011). Other claims include casting doubt on the reliability of climate science and dismissing the effectiveness of climate solutions (Coan et al., 2021). This campaign has continued and grown for decades, often peaking during climate change policy debates.

The denial campaign is led by industrial groups, particularly those associated with fossil fuels, as well as conservative foundations and think tanks. The campaign is supported by contrarian scientists, conservative media, politicians, and more recently, sceptical bloggers, forming a worldwide network of advocates (Dunlap & McCright, 2011). Through this global advocacy network, the denial of climate change has spread to other countries, resulting in an increase in public scepticism towards climate change. This trend was particularly noticeable in 2009 and 2010 following the fabricated Climategate scandal and some relatively minor errors found in the Intergovernmental Panel on Climate Change Fourth Assessment Report (Dunlap & McCright, 2011). False information about climate change has significantly decreased in mainstream print and online media globally (Lynas et al., 2023). Nevertheless, there are still instances of recent false information circulating in unreliable news sources, as the corpus for this research includes a couple of recently published articles denying the connection between climate change and wildfires (Articles' ID: *A False 1* and *A False 20)*.

Psychological factors play a significant role in contributing to climate change disbelief. Political affiliation and ideology strongly influence attitudes towards climate change, as proposed solutions are often associated with specific politicians (Cook, 2017; Treen et al., 2020). Furthermore, social and cultural influences, including remarks from influential figures (e.g., elites or conservative leaders), contribute to the scepticism surrounding climate change and create obstacles in addressing this issue (Cook, 2017).

According to Treen et al. (2020), the dissemination of false information regarding climate change has various detrimental effects on the public. Firstly, it can trigger negative emotions such as panic, fear, worry, and anger, which in turn can influence individuals' decision-making and behaviour. Secondly, it can result in political inaction or the rejection of efforts to mitigate climate change. Lastly, it can generate confusion and erode trust in scientific organizations and factual information, thus undermining societal and intellectual well-being (Treen et al., 2020).

3.2.2. Anti-Vaccination Movement

Opposition to vaccination has long been a concern, influenced by various factors including religious beliefs, scepticism towards medical institutions and pharmaceutical companies, and objections to vaccine regulations and laws that mandate immunization for children (Hussain et al., 2018). However, in recent years, there has been a significant rise in anti-vaccination sentiments, asserting that vaccines pose more harm than benefit to children's health. These sentiments are primarily fuelled by the negative portrayal of vaccines in news, entertainment outlets, and social media platforms. Consequently, vaccination rates have declined in certain Western countries, resulting in the re-emergence of diseases that were previously nearing eradication (Hussain et al., 2018).

The flawed study conducted by former British doctor and researcher Andrew Wakefield had a significant impact on the resurgence of the anti-vaccination movement (Hussain et al., 2018; Royal Society for Public Health, 2019). Wakefield suggested a link between the measles, mumps, and rubella (MMR) vaccine and the development of autism in young children. However, his claim was thoroughly debunked, and his research was deemed unethical, leading to his removal from the UK Medical Registry (Hussain et al., 2018).

Despite the discrediting of Wakefield's study, its influence extended to various parts of the world, particularly Western Europe and North America (Hussain et al., 2018; Royal Society for Public Health, 2019). This false information and distrust surrounding vaccines have resulted in public fear, anger, and decreased trust in health organizations, contributing to vaccine hesitancy (World Health Organization, 2017). The consequences of the anti-vaccination movement have become evident, with persistent measles outbreaks in the USA and numerous countries in 2019. These outbreaks have reversed decades of progress made in achieving herd immunity, with a global count of over 400,000 confirmed cases (Gardner et al., 2020).

3.2.3. COVID-19 False Information

The COVID-19 pandemic refers to a worldwide outbreak of an infectious disease known as novel coronavirus. It is caused by the severe acute respiratory syndrome coronavirus (SARS-CoV-2). Initially, the virus was detected in China in December 2019, and it quickly spread to other countries. In response to the growing concern, the WHO declared a Public Health

Emergency of International Concern on 30 January 2020. Later, on 11 March 2020, the WHO officially classified the outbreak as a pandemic (World Health Organization, 2023).

The spread of the pandemic was accompanied by an *infodemic*, which refers to the rapid global dissemination of false information, rumours, and conspiracy theories (Pian et al., 2021). This false information exacerbated the already challenging situation. Hansson et al. (2021) classify false information surrounding COVID-19 as follows: (a) misrepresenting the origin of the virus, such as claiming that the pandemic originated from cities with 5G networks in widespread activation, (b) misrepresenting the transmission of the virus, such as the notion that individuals with blood group A are more susceptible to the virus, (c) discouraging protective measures against contracting or spreading the virus, such as spreading rumours about death in COVID-19 vaccine trials, (d) promoting false or harmful treatments, like using colloidal silver to prevent or treat the coronavirus, and (e) downplaying the risks related to the pandemic, with claims suggesting that the virus would be eliminated in warm weather.

According to Pian et al. (2021), the infodemic was brought about by several causes, including extensive use of social media, low levels of health literacy, the rapid publication process, and the availability of preprint services. The consequences of the infodemic, as emphasised by Pian et al. (2021), were far-reaching. They included various psychological issues such as feelings of loneliness, post-traumatic stress disorder, depression, obsession, compulsions, and even psychotic symptoms. Healthcare providers also experienced adverse effects, such as a sense of worthlessness and guilt. The infodemic also resulted in a significant lack of trust in science and health institutions. In addition, some individuals resorted to improper protective measures or completely refrained from using any protective measures. Panic buying was another negative impact of the infodemic, with people rushing to purchase unproven remedies and stockpile groceries in their homes.

3.3. Rationale for Investigating False News Discourse on Climate Change, Vaccinations, and COVID-19

Understanding false news discourse surrounding climate change, vaccinations, and COVID-19 is vital for several reasons. Firstly, these topics are known to be highly controversial and divisive, leading to intense public debates and increased polarization of opinions. Additionally, they

provide fertile grounds for emotional manipulation. Secondly, given their global impact, these issues generate a substantial amount of false information that can be disseminated rapidly and widely, making them ideal for creating a sizable corpus for study. Thirdly, false information about climate change, vaccinations, and COVID-19 has the potential to significantly impact global health and the environment. Decision-making processes influenced by false information can result in significant health risks, such as vaccine hesitancy or ineffective measures to combat the spread of COVID-19. Similarly, false information concerning climate change can impede efforts to address environmental challenges and worsen ecological degradation.

3.4. General Guidelines for Designing a Representative Corpus

A corpus is a collection of texts or speech transcriptions that can be analysed using computers. To ensure its usefulness and accuracy in answering a research question, a corpus needs to be well-constructed and representative of the language variety under investigation (Brezina, 2018). According to Biber (1993), representativeness can be evaluated in two ways: situational representativeness and linguistic representativeness. Situational representativeness (also called target domain or target population representativeness) refers to the extent to which a corpus represents the range of text types in the target population, while linguistic representativeness refers to the extent to which a corpus represents the range of the linguistic features in the target population. While situational representativeness determines whether a corpus can be generalised to a larger population of interest, linguistic representativeness determines whether a corpus is suitable for answering linguistic research questions.

Ensuring situational representativeness requires defining the population that the corpus aims to represent and random sampling techniques (Biber, 1993; Brezina, 2018). The definition should consider the boundaries of the population as well as the hierarchical organization within the population. This can be done by establishing a sampling frame that specifies which texts are to be included or excluded, as well as their defining categories and respective criteria (Biber, 1993; Brezina, 2018). Random sampling means that each produced text has an equal chance to be included in the corpus (Brezina, 2018).

Linguistic representativeness refers to the ability of a corpus to facilitate the analysis of the distribution of linguistic features in the target population since different linguistic features

usually have different distributions within texts or across texts and text types (Biber, 1993). Notably, a corpus can be situationally but not linguistically representative, but the opposite is not true since linguistic representativeness ensures situational representativeness (Egbert, 2019). For example, a well-sampled corpus that represents the range of text types in university writing might not be large enough to investigate a question related to a low-frequency linguistic feature. This means that the size of the corpus, such as the number of words per text and the number of texts per text type, determines whether a corpus is linguistically representative (Biber, 1993). However, small corpora can be adequate if they represent a narrow or specialized target domain and the research questions pertain to highly frequent linguistic features such as *articles*, *propositions*, or *nouns* (Egbert, 2019). To design a representative corpus, Egbert (2019) suggests the following steps:

1) The research objectives and design are established. This is important because no single corpus serves every research objective.
2) Extensive research is conducted to define the target domain of interest.
3) A corpus design plan is developed to ensure that the corpus is representative of the target domain. This plan involves the sampling frame, sampling unit, sampling method, sample size (text lengths and count), cost, timeline, and storage format.
4) The plan is implemented to collect the corpus.
5) The corpus is annotated for external characteristics (e.g., source, date, speaker demographics, and register) and internal characteristics (e.g., POS tagging).
6) The corpus is evaluated in terms of situational representativeness.
7) The corpus is evaluated in terms of linguistic representativeness.
8) Steps 3-5 are repeated if the corpus lacks situational or linguistic representativeness.
9) The methods used to conduct all the steps are documented in detail.

The guidelines mentioned above are generally applicable for compiling a corpus that focuses on a well-defined and easily accessible language variety. However, they can be challenging to implement meticulously when compiling a corpus representing a target domain that is under-researched, complex to define, and difficult to collect, as is the case with fake news. In the following sections, I will discuss the challenges and opportunities encountered in compiling a focus corpus of false news and a reference corpus for comparison.

It is important to acknowledge that the main challenge I faced was ensuring situational representativeness, i.e., the extent to which the collected corpus of false news can be considered representative of false news. On the other hand, linguistic representativeness was not a major concern because the study focused on analysing three discourse aspects: discourse structure, news values, and stylistic patterns which do not require investigating specific or rarely used linguistic features. The procedure used to compile both corpora will be detailed after discussing the challenges.

3.5. Challenges of Compiling the Focus Corpus

The prevalence of false information online, particularly on social media platforms, is widely recognized. However, the primary obstacle faced by linguists in analysing fake news lies in collecting reliable data (Asr et al., 2023; Asr & Taboada, 2019). When compiling a corpus of fake news, researchers encounter five primary challenges. Firstly, fake news is disseminated across various mediums and formats, including websites, blogs, and social media posts, and secondly, false information encompasses a diverse range of types, such as rumours, clickbait, and fake reviews (Bondielli & Marcelloni, 2019). However, these two challenges can be overcome by establishing a clear definition of fake news, as the data collection strategy relies on the adopted definition (Asr & Taboada, 2019). Thirdly, the volume of false information available online is limited compared to the vast amount of published content (Bondielli & Marcelloni, 2019). Fourthly, analysing fake news on social media is complicated due to constraints imposed by social media companies, which hinder the analysis of published posts (Bondielli & Marcelloni, 2019). Last but not least, distinguishing between truth and falsehood presents a challenge, and news outlets tend to accuse each other of publishing fake news with biased viewpoints (Asr et al., 2023; Asr & Taboada, 2019).

There are several approaches that can be adopted for collecting fake news items. The first approach involves collecting data from news sources known for publishing false information, specifically conspiracy-pseudoscience websites, such as Natural News or The Alternative News websites. This method has been employed in many NLP studies to gather data for developing fake news detection algorithms (e.g., Horne & Adalı, 2017; Rashkin et al., 2017; Volkova et al., 2017). However, this approach has a limitation in that some well-known fake news outlets also publish true news, leading to the inclusion of noise in the collected data. Consequently, this data

is not suitable for studying the distinctive features of fake news discourse, but it may be more suitable for stylistic analysis of specific news sources (Asr et al., 2023; Asr & Taboada, 2019). Moreover, determining the factual nature of a source necessitates objective measures and cannot be solely determined by the linguist.

The second approach involves using existing datasets, such as the Buzzfeed News collections of false news (2016; 2017; 2018) or corpora created for other studies like Asr et al.'s (2023) corpus and FakeNewsNet by Shu et al. (2019). However, it is important to note that these datasets may have limitations. Some datasets may be unbalanced in terms of genres and topics. For example, FakeNewsNet contains both tweets and news articles, and Horne & Adalı's (2017) dataset contains both Facebook posts and website articles. This restricts their usefulness for analysing specific genres or topics of fake news. In addition, certain datasets may have been collected from sources known to publish fake news but also occasionally publish true news (e.g., the one by Rashkin et al., 2017). As stated earlier, this can result in noisy data that is not suitable for studying the distinctive features of fake news discourse (Asr & Taboada, 2019). Lastly, certain datasets have been collected for specific purposes, making them less applicable to other research objectives. For example, the FEVER dataset by Thorne et al. (2018) was compiled for stance detection. It consists of claims generated from information extracted from Wikipedia, along with supporting or refuting Wikipedia texts for these claims. Table 3.1 provides an overview of existing fake news datasets. The table illustrates the genre and topic of the datasets, indicating whether they consist of articles based on content veracity (verifiably false news items) or source reputation (collected from sources known to publish fake news). In addition, it indicates whether these datasets include a comparable corpus of real news

Table 3.1 Review of Fake News Datasets Available Online

Corpus/Dataset	**Genre**	**Topic**	**Veracity Method**		**Comparative Datasets**	**Notes**	**Links**
			content-veracity	**source-reputation**			
Põldvere et al.'s (2023) PolitiFact Oslo Corpus	News articles and social media posts	various	✓		✓	The corpus is available upon request. The corpus has metadata information.	https://www.mdpi.com/2078-2489/14/12/627
Asr et al.'s (2023) MisInfoText	news articles and Facebook posts	various	✓		✓		https://github.com/sfu-discourse-lab/MisInfoText
Shu et al.'s (2019) FakeNewsNet	news articles and tweets	various	✓		✓	It includes information about the social context of the collected fake news and about the users who propagated the news.	https://github.com/KaiDMML/FakeNewsNet
Zhang et al. (2018) Credibility Coalition	news articles	health and climate science				The data contains highly shared articles, each of which has a	https://data.world/credibilitycoalition/webconf-2018

						credibility score based on a set of credibility indicators	
Thorne et al.'s (2018) FEVER	claims extracted from Wikipedia	various	✓		✓	The claims were extracted by annotators from Wikipedia articles.	https://fever.ai/
Allcott & Gentzkow (2017)	news articles	2016 US presidential election	✓				https://www.openicpsr.org/openicpsr/project/101153/version/V1/view
Rashkin et al. (2017)	news articles	various		✓	✓	The articles are classified into satire, hoax, propaganda, and real.	https://hrashkin.github.io/factcheck.html
Horne and Adalı (2017)	news articles and Facebook posts	mainly political (US 2016 election)	✓	✓	✓	The texts are classified into satire, fake, and real.	https://github.com/BenjaminDHorne/fakenewsdata1
Wang's (2017) LIAR Dataset	short statements	various	✓		✓	The statements are derived from news releases, TV/radio	https://paperswithcode.com/dataset/liar

						interviews, campaign speeches, TV ads, tweets, debates, Facebook posts, etc.	
Tacchini et al.'s (2017) FacebookHoax	Facebook posts	conspiracy, pseudo-science		✓	✓		https://github.com/gabll/some-like-it-hoax
McIntire (2017)	news articles	political		✓	✓	The articles are categorized according to their topics and political leaning.	https://shorturl.at/blCKS
Rubin et al. (2016)	news articles	various		✓	✓	The articles are categorized into legitimate and satire	https://shorturl.at/wDHV8
Zubiaga et al.'s (2016) PHEME	tweets	specific events	✓		✓	The tweets are classified into rumours and non-rumours.	https://dx.doi.org/10.6084/m9.figshare.2068650
Ferreira and Vlachos' (2016) Emergent Dataset	claims and their associated news articles	various	✓		✓		https://github.com/willferreira/mscproject

Mitra and Gilbert's (2015) CREDBANK	tweets	various topics/ events	✓		✓	The tweets are categorised according to their topics.	http://compsocial.github.io/CREDBANK-data/
BuzzFeedNews (2016, 2017, 2018)	Facebook posts	mainly political	✓		×		https://shorturl.at/gFMZ7 https://github.com/BuzzFeedNews/2017-12-fake-news-top-50 https://github.com/BuzzFeedNews/2018-12-fake-news-top-50
Kaggle.com datasets	news articles	various		✓	✓	This website includes data collected by different users.	https://www.kaggle.com/competitions/fake-news/overview https://www.kaggle.com/datasets/jruvika/fake-news-detection

							https://www.kaggle.com/datasets/mrisdal/fake-news

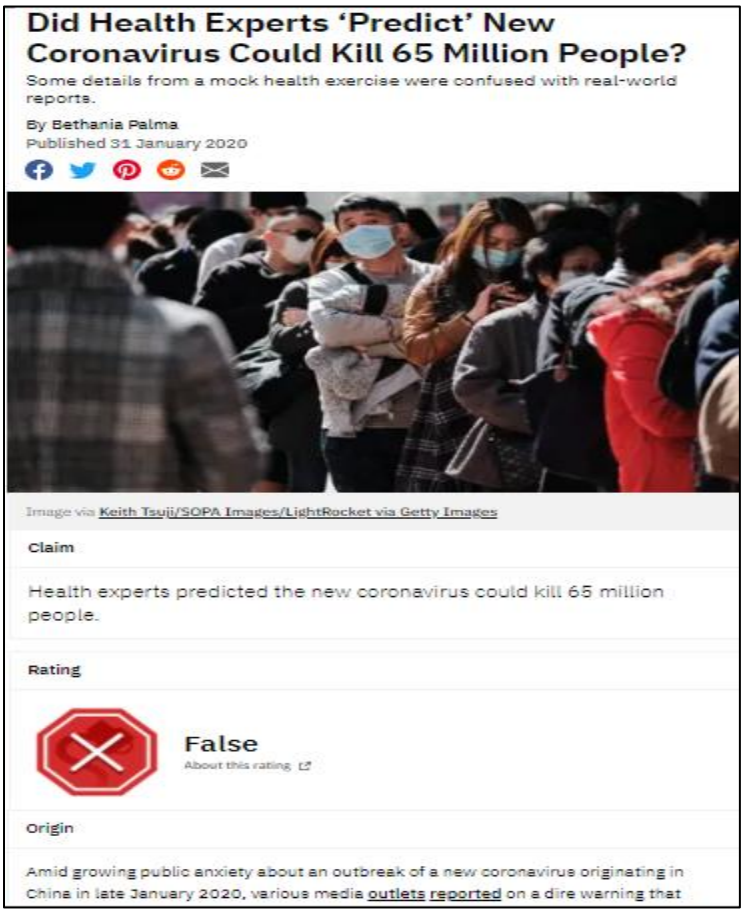
The third approach to collecting fake news data involves manually verifying each article by the researcher. However, this method is time-consuming, labour-intensive, and likely beyond the expertise and capacity of the researcher, as it requires conducting fact-checking research for each individual article. Furthermore, there is a potential for bias to be introduced during the data collection process due to the subjective assessment of the researcher.

The fourth approach to obtaining fake news data is via crowdsourcing. This involves soliciting contributions from a large group of people to gather information. Some studies have used crowdsourcing to ask people to rate the accuracy of social media posts (e.g., Mitra & Gilbert, 2015) or to create fake news texts artificially (e.g., Pérez-Rosas et al., 2017). This method allows for the collection of a large quantity of fake news data and provides some objectivity for researchers. However, artificially created fake news items lack the contextual factors, such as socio-political influences, that are inherent in real-world instances of fake news.

The fifth approach involves using online algorithms specifically designed to fact-check the content of news items (e.g., Fakebox and Discourse Processing Lab). This automated approach can be helpful in quickly fact-checking large amounts of articles. However, it may have limitations in terms of accuracy, as it relies solely on algorithms without human judgment and may struggle with nuanced or complex claims that require contextual understanding.

A final approach is collecting news items that have been identified as false by experts who are widely recognized and trusted. Numerous websites now offer fact-checking services for various types of news, including both political and non-political stories. Examples of such websites are *Snopes, PolitiFact, FactCheck.org, Reuters Fact Check,* and *AFP Fact Check.* Figure 3-1 illustrates what a fact-checking webpage typically contains. These websites have become valuable resources for studies investigating fake news. A notable example of research using data from these fact-checking websites is the study conducted by Asr et al. (2023). In this study, the researchers collected extensive datasets of false and true news articles, totalling 1,649 false articles and 1,349 true articles, by utilizing Snopes and BuzzFeed. This corpus was then used for conducting classification experiments.

Figure 3-1 A Screenshot of a Fact-checking Article from Snopes



However, there are several disadvantages to the fact-checking approach. One major concern is that it can potentially introduce bias in data collection if the fact-checking organization has certain political leanings (Graves, 2017; Grieve & Woodfield, 2023). Nonetheless, many instances of fake news have been repeatedly identified as false by multiple fact-checking sites; this provides some assurance of the inaccuracy of these false news items. Furthermore, fact-checking websites often outline their procedures and claim to employ objective strategies. Graves (2017), in his study on fact-checking practices through interviews and observations, identified a five-step process commonly employed in fact-checking. This process is designed to ensure objectivity in fact-checking efforts:

1) Choosing claims to check: Fact-checkers select statements that are newsworthy, useful and interesting to the public, fair and balanced (i.e., checking sources from opposing views), and verifiable.

2) Contacting the speaker: This step emphasises fairness and investigation, as fact-checkers give speakers the opportunity to explain their statements before making a judgment.
3) Tracing false claims: Fact-checkers trace the origin of a claim and reconstruct its dissemination. This step provides important contextual information for assessing the claim's truthfulness.
4) Contacting experts: Fact-checkers consult primary sources of information, such as government agencies, academic experts, and think-tank organizations, for additional insights and verification.
5) Showing fact-checking work: As a matter of transparency, fact-checkers disclose all the sources used to analyse a claim in their reports. This step serves as the basis for objectivity, providing evidence to support the verdict and addressing any sceptics.

The second issue with using fact-checking websites for data collection concerns the veracity labels used to rate the truthfulness of news items. News articles often contain a combination of true and false information, and thus truth cannot be measured as a gradable or quantitative variable (Grieve & Woodfield, 2023). In addition, different fact-checking websites employ varying labelling systems. Some websites, like Reuters Fact Check, AFP Fact Check, and FactCheck.org, debunk false stories without utilizing labels, while others such as Snopes and PolitiFact provide graded labels such as pants on fire, false, mostly false, mixed, and unproven. This poses a challenge when selecting data, particularly in situations where researchers aim to create two comparable sets of true and false news articles with an equal number of articles and matching veracity labels. Furthermore, if a researcher intends to expand data collection by utilizing multiple fact-checking websites, where one employs graded labels and another does not, the presence of labels creates difficulties in collecting balanced data as determining the degree of truthfulness for unlabelled news items becomes complex.

Another limitation of using fact-checking websites is the time-consuming nature of collecting data from them. Automatic scraping of data from fact-checking websites does not provide a readily usable corpus of false news items for linguistic analysis (Allcott & Gentzkow, 2017; Asr et al., 2023). In fact, manual effort is required to gather news texts from these websites. Fact-checking websites typically present their work on webpages, each discussing a specific claim, citing the sources where the claim originated, providing a veracity label for the claim, and

offering the reasons and evidence behind the assigned label, often with additional supporting sources (see Figure 3-1). Unfortunately, most fact-checking websites lack advanced search tools and comprehensive categorizations based on topics, veracity labels, genres, or authors. Therefore, researchers must manually go through all fact-checking articles related to a topic of interest, locate hyperlinks to the original news items, assess their suitability, and then archive the text, all of which is very time-consuming.

The final drawback is that fact-checking websites rely on archive websites to store the news item's original webpages, but sometimes these archives are not indexed properly, making it difficult to retrieve the original webpages. Obtaining the original text online is especially difficult when dealing with websites that regularly publish fake news, as they tend to have short lifespans (Allcott & Gentzkow, 2017).

Despite its drawbacks, using fact-checking websites for collecting fake news data offers several advantages. Firstly, the veracity of news items is established by experts through ad-hoc investigations, making this method more reliable than alternative approaches. Secondly, this method can increase the objectivity of data collection since the researcher is not personally involved in determining the truthfulness of the news items. Thirdly, it enables the collection of only those articles that have been verified to be false, which may not be possible when collecting data from sources known to publish fake news in addition to real news. Lastly, using fact-checking websites allows researchers to have control over the genre and topic of the verified news items, which may not be possible when using existing datasets. Therefore, the current study adopted this method to build its focus corpus.

3.6. Challenges of Compiling the Reference Corpus

To fully understand fake news discourse, a comparative analysis is necessary. However, it is challenging to design a suitable corpus for comparison because fake news items found online have a distinct register. This means that the communicative contexts, purposes, and attributes of the authors and audience of fake news published online differ from those of reliable printed news (Grieve & Woodfield, 2023). One notable difference is the genre or discourse structure of fake news online, which differs from that of printed news articles due to online formatting (Alba-Juez & Mackenzie, 2019). Moreover, fake news is typically written by amateur bloggers, whose

writing style is likely to differ from that of professional journalists. Additionally, fake news originates from individuals with diverse backgrounds, encompassing differences in nationality, social, religious, ethnic, and educational profiles. Consequently, writing styles can vary and be influenced by authors' social backgrounds (Grieve & Woodfield, 2023), while content can be shaped by a range of authors' interests and reader preferences. Therefore, constructing a reference corpus that closely aligns with the focus corpus presents a significant challenge.

The aim of this thesis is to facilitate the study of the distinct characteristics of false news discourse. To achieve this, two critical variables in the reference corpus need to be controlled: the veracity of the content and the register of the collected news articles. Determining the factual accuracy of an article requires extensive investigation, making it challenging for researchers to determine on their own. As a result, one approach to collecting verifiably true news articles is to rely on texts that have been checked by fact-checkers and marked as true. However, it can be difficult to find verifiably true news articles through fact-checking websites, as they tend to examine fake news stories more extensively than real ones.

Another approach is to collect news items published by sources known for their factual reporting, such as broadsheets or sources rated highly factual by fact-checking websites like *The* Media Bias/Fact Check website. Comparing fake news items to broadsheet articles has been commonly used in previous research (e.g., Ahmad et al., 2020; Gravanis et al., 2019; Horne & Adalı, 2017; Rashkin et al., 2017; Rubin et al., 2016). However, comparing fake news articles posted online to broadsheet articles is likely to involve analysing two different registers, involving different communicative contexts and purposes and writing styles, regardless of the veracity of the news items. This register variation could be a confounding variable, potentially obscuring the true differences between fake and real news discourses (Grieve & Woodfield, 2023).

An approach worth considering is to collect both fake and real news articles written by the same person for the same type of news outlet (Grieve & Woodfield, 2023). This approach offers the advantage that it minimizes the risk of register variation being a confounding variable, ensuring consistent contextual circumstances and writing styles. In addition, it can allow researchers to identify the author's intentions behind creating fake news. By comparing an individual's real and

fake news, any differences can be confidently attributed to their intentions of writing fake news. Nevertheless, applying this method becomes challenging when studying fake news that circulates online because there is a vast amount of fake news authored by various individuals, published on multiple outlets, and motivated by diverse purposes. Moreover, it is crucial to investigate widely circulated online fake news, as it can have detrimental effects that should not be ignored.

Considering the limitations of the previous three approaches, this thesis adopts a more systematic and comprehensive method. It compares a corpus of false news to a corpus of news articles obtained from reliable sources, encompassing various types of news outlets like broadsheets, tabloids, websites, and blogs from different countries. This comprehensive analysis considers not only the accuracy of the news content but also takes into account register variation that could potentially impact the comparison between false and true news. Furthermore, this approach facilitates an exploration of the distinctions not only between false and true news but also among false news and variations across different types of true news.

3.7. The Procedure for Compiling the Focus Corpus

In this thesis, the focus is on analysing verifiably false news articles published by news websites or blogs on three topics: climate change, vaccinations, and COVID-19. The term *false news* is used to refer to this data. To select articles for the false news corpus, five fact-checking websites were utilized: Snopes, PolitiFact, FactCheck.org, AFP Fact Check, and Reuters. Several methods were employed to obtain false news articles from these websites:

a) Collections and archives: PolitiFact and FactCheck.org maintain collections of false news related to vaccination, climate change, and COVID-19. These collections served as the primary sources of false news articles for this study.
b) Keyword search: Searches using keywords such as ‘vaccine(s),’ ‘vaccination,’ ‘climate change,’ ‘global warming,’ ‘covid-19,’ ‘covid,’ and ‘coronavirus’ were conducted on all the five fact-checking websites. The search results were then filtered using dates and veracity labels.
c) Source search: This method involved searching for false articles specifically written by bloggers (i.e., news published on blogs rather than social media platforms). This

technique was implemented on PolitiFact, as it allows the filtering of false news items by sources or authors.

The selection of false news articles was based on two main criteria:

a) Genre: The articles had to be in the format of news articles published by news websites or blogs, excluding social media posts, WhatsApp messages, memes, or other formats.
b) Veracity labels: Only articles that were rated as Pants on Fire, False, or Mostly False were considered for inclusion. This decision was made considering that truthfulness is not a measurable attribute. Additionally, since false news articles often contain a combination of both true and false information, articles rated as mostly false were included in the corpus.

Using these search methods and criteria, I carefully examined all the fact-checking articles related to the three topics. I located hyperlinks to the fact-checked news items, assessed their suitability, and, if appropriate, archived the text. The collected texts were saved as separate text files, including only the title and body text of the articles. Other elements such as the authors' names, publication dates, figures and their captions, advertisements, notes, comments, authors' contact information, tags, or other links (such as links to social media, subscriptions, related articles, or sources) were excluded from the saved text files. A total of 137 articles were included in the focus corpus. A metadata table was created, recording the claim, fact-checking webpage URL, original false article webpage URL, veracity label, publication date, topic, and source type. Table 3.2. shows the breakdown of the corpus and its counts.

Table 3.2 Breakdown of the Focus Corpus

False News Topics	**No. of Articles**	**Average No. of Words**	**Standard Deviation of No. of Words**	**Total No. of Words**	**Sources**	**Publication Dates**
Climate Change	20	887	708	17,736	Online tabloids, news websites, blogs	2011–2020
Vaccination	27	903	568	24,372	News websites, blogs	2013–2019
COVID-19	90	717	724	64,565	Online tabloids, news websites, blogs	January 2020–December 2020
The Whole Focus Corpus	**137**	**779**	**693**	**106,673**	---	---

3.8. The Procedure for Compiling the Reference Corpus

The comparative corpus comprised four sub-corpora, each representing a news type (broadsheets, tabloids, web-based publications, and news blogs) and including 137 articles. The articles were collected using the news dataset Nexis.uk and the keywords used were: 'vaccine(s)', 'vaccination', 'climate change', 'global warming', 'coronavirus', 'covid-19', and 'covid'. An advanced search was conducted to retrieve articles that contained these keywords in the headlines and leading section. For the broadsheets and tabloids, only the printed versions were used, not the online versions. As for the web-based publications, formal news sources that are published solely online and do not have printed versions were included. For these news types, a selection of five to seven reputable news sources representing three English-speaking countries (the US, the UK, and Australia) was made. The news blog corpus consisted of news articles collected by Newstex LLC. This company scans full-text content from high-quality blogs

covering various topics including art, career, economics, environment, finance, food, health, law, marketing, and medical technology. These articles were delivered on demand through Nexis.uk. (*About Newstex | We Distribute Authoritative Content*, 2022; Newman, 2008). The news sources used for compiling the reference corpus are listed in Table 3.3.

Table 3.3 News Sources Used to Compile the Reference Sub-corpora

Broadsheet Sources	1. New York Times 2. USA Today 3. The Times (London) 4. The Independent (United Kingdom) 5. The Australian 6. Sydney Morning Herald (Australia)
Tabloid Sources	1. Daily News (New York) 2. The Sun (England) 3. The Daily Mail and Mail on Sunday (London) 4. The Daily Telegraph (Australia) 5. The Age (Melbourne, Australia)
Web-based Publication Sources	1. CNN.com 2. Business Insider US 3. Politics Home (Online) 4. Progressive Media – Company News[4] 5. TVEyes-BBC World 6. WA Today 7. ABC Premium News (Australia)
Blog Sources	Newstex blogs[5]

For each news type, texts were randomly selected from Nexis.uk with the aid of a random number generator[6]. The number of the collected articles and their publication dates were matched with those of false news articles. In addition, the minimum and maximum word limits per article were set between 300 and 4000 words, in order to align with the range of word counts in the false news articles. Furthermore, the comparative corpus includes both news and opinion articles, mirroring the composition of the false news articles. For example, the comparative corpus

[4] This news source does not have articles on COVID-19
[5] The blogs delivered by Newstex LLC are listed at: https://shorturl.at/uyzHP
[6] The random number generator used in selecting the articles is found at: https://rb.gy/93yspt

contains 108 articles on vaccination in total; 27 articles were collected from each news type. The collected articles were between 300 and 2305 words long, and they were published between 2013 and 2019. Table 3.4 shows the breakdown of the reference corpus and its counts.

Table 3.4 Breakdown of the Reference Corpus

Reference Sub-corpora	**No. of Articles on Climate Change**	**No. of Articles on Vaccination**	**No. of Articles on COVID-19**	**Total No. of Articles**	**Average No. of Words**	**Standard Deviation of the No. of Words**	**Total No. of Words**
Broadsheets	20 articles for each sub-corpus	27 articles for each sub-corpus	90 articles for each sub-corpus	137	683	273	93,533
Tabloids				137	600	260	82,183
Web-based Publications				137	700	282	95,884
Blogs				137	578	203	79,198
The Whole Reference Corpus	**80**	**108**	**360**	**548**	**640**	**261**	**350,798**

The articles downloaded from Nexis.uk were checked for their relatedness, converted to text files, and cleaned. Only the titles and body texts were included. The cleaning process was performed by SarAnt (Anthony, 2016). Using search and replace tools, the metadata generated by Nexis.uk (including news source names, author names, publication dates, load-dates, and marks signifying the body and end of the article) were deleted. Any duplicated articles were removed from the corpus. The articles were also checked manually to delete texts such as authors' contact information, notes, tags, or other links (e.g., social media, signing up or subscribing, related articles, or sources). This meticulous cleaning process aimed to align the comparative articles with the false news articles, by including only the title and body text in the final corpus.

3.9. Limitations of the Corpus

It is important to acknowledge that the corpus used in this thesis has certain limitations. Firstly, the focus corpus is relatively small, consisting of only 106,673 words. However, as indicated earlier, even small corpora can yield valid results if they are carefully designed and focus on a

specific target domain, and if the research question does not focus on rarely occurring linguistic features (Egbert, 2019). The focus of this research is the discourse characteristics of online false news articles, including their discourse structure, news values, and stylistic patterns, which can effectively be investigated in small corpora. Therefore, the study is positioned to generate significant findings despite the limitations imposed by the corpus size.

Secondly, while the collected articles were rated as false at the time of data collection, it should be recognized that the truthfulness of a claim can fluctuate over time. What is considered false today might be proven true after further investigation and verification. For example, the assertion that prolonged use of face masks can lead to health issues due to excessive inhalation of carbon dioxide was initially debunked by fact-checkers (Ibrahim, 2020), but later it was proven to be true through research (Kisielinski et al., 2023). In addition, a similar limitation to acknowledge lies in the reference corpus. The reference corpus was compiled using well-reputed news sources which might publish false information. This decision was made because it was challenging to collect a sufficient number of verifiably true articles that represent different news types. However, it can be generally assumed that the majority of the articles obtained from these sources are true, while most false news articles are inaccurate. Hence, the corpus remains valid for comparing the discourse of false news with that of true news, despite this limitation.

Thirdly, the focus corpus contains both disinformation and misinformation, as it was not feasible to determine the intentions of the writers of online false news articles. Most research studies on fake news also focus on the veracity of the claims, without delving into the intentions of the writers (Zafarani et al., 2019). Moreover, restricting research to cases of disinformation may overlook extensively circulated partisan and detrimental false news.

Fourthly, the focus corpus is limited in that it covers only health and environmental topics which are highly controversial. This may restrict the generalizability of the findings to other domains. However, the corpus has the advantage that it encompasses articles from diverse authors and various news outlets across a span of almost a decade (2011-2020), thereby enhancing the applicability and relevance of the study's results.

Overall, the focus corpus was meticulously constructed, consisting solely of online news articles without a mix of news articles and social media posts. In addition, the articles were verified as

false and were not randomly sourced from outlets known for disseminating false information. The reference corpus encompasses a diverse range of reputable and reliable news sources, closely matching the false news corpus in terms of topics, publication dates, length, and number of articles. Such a comparable corpus can facilitate the identification of similarities and differences between false and true news while recognizing the variability within true news. This methodological approach enhances the comprehension of the implications of the identified distinctions and similarities and bolsters confidence in interpreting the study results. Consequently, despite these limitations, it is expected that this study will yield valuable insights into the analysis of false news discourse.

3.10. Conclusion

In summary, this chapter provides an overview of the false news articles utilized in this research. It examines the socio-historical background and contextual factors surrounding the three topics under investigation: climate change denialism, the anti-vaccination movement, and false information related to COVID-19. The chapter also explores the challenges and opportunities encountered during the compilation of the corpus, elucidating the rationale behind the chosen data collection procedure for obtaining a representative corpus of false news and a comparable corpus. However, it is essential to acknowledge that the corpus still has certain limitations that should be taken into account when interpreting the analysis results. The next three chapters will present the three analyses conducted on the corpus.

Chapter 4: The Discourse Structure of Online False News Articles

4.1. Introduction

This chapter introduces the first case study, which specifically looks at the discourse structure of online false news articles. It examines how information is grouped, organized linearly, and strategically presented. While there has been a growing interest in fake news research, there is a notable lack of in-depth studies focusing on its discourse structure.

Some previous studies have concentrated on developing algorithms to detect fake news by analysing differences in coherence and discourse structure between real and fake news articles (e.g., Karimi & Tang, 2019; Rubin & Lukoianova, 2015; Uppal et al., 2020). These studies have highlighted the importance of discourse structure analysis in automated fake news detection, emphasising its complementarity to lexicogrammatical analysis. However, they have not extensively explored the constituent elements that serve rhetorical or communicative functions and shape the structure of fake news articles. Furthermore, contextual factors influencing the construction of fake news articles have not been adequately considered in these studies.

This study aims to address these gaps by analysing the discourse structure of online false news articles using genre analysis (GA). GA is a methodology established in the 1970s and 1980s that concentrates on the conventional structures employed to construct a comprehensive text within a language variety (Biber et al., 2007). Specifically, the study utilizes GA using rhetorical moves, commonly known as move structure analysis or move analysis (MA). This method was developed by Swales (1990) to describe the rhetorical organizational patterns of academic research articles by segmenting texts into rhetorical moves according to their communicative functions.

Given that GA involves qualitative and manual analysis, this case study narrows its focus to a specific subset of the corpus, which consists of 20 articles on climate change. The study aims to: (a) delve into the internal structure of these articles, (b) elucidate the motives behind false news,

(c) examine the rhetorical strategies used to present information, (d) and explore variations in discourse structure among the false news articles.

The chapter is structured as follows: The first section serves as a methodological background, offering an overview of the methodological principles of GA and MA. The second section includes a review of previous studies on news genres. This review is crucial for understanding how the structure of false news differs from the structural conventions observed in traditional hard news reports. The third section describes the analytical procedure for analysing the data. Finally, the analysis results are presented and discussed.

4.2. Methodological Principles of Genre Analysis and Move Analysis

This section starts by exploring the concept of genre and genre analysis (GA) method, followed by an overview of the primary schools of GA. It then discusses commonly used analytical methods in GA, with a focus on move analysis (MA), a well-established method specifically applied in this study. Finally, the section concludes by justifying the employment of GA using rhetorical moves for studying the discourse structure of false news articles.

4.2.1. Genre and Genre Analysis

A ‘genre’ is commonly defined as a categorization of a particular type of text or social practice. Such categorizations can further be divided into sub-genres (Baker & Ellece, 2011). For example, the genre of fiction can be divided into historical, adventure, romance, spy, and so on. Swales (1990) defines genre as:

> a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert members of the parent discourse community and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style. (p. 58)

According to this definition, genres are more than just discourse categories; they are social and communicative events. The crucial element that defines a genre is the communicative purpose, which determines the structure, style, content, and intended audience of texts within the genre. Moreover, experts within the genre's community can recognize the rationale behind a genre. For

example, scientists can identify the purposes of research papers, while journalists can determine the purposes of news articles (Askehave & Swales, 2001).

In line with Biber et al. (2007) and Biber and Conrad (2019), the term ‘genre’ is a perspective for describing and analysing texts. It has been used to refer to “culturally recognized ‘message type’ with a conventional internal structure, such as an affidavit, a biology research article, or a business memo” (Biber et al., 2007, p. 8). Consequently, genre is often associated with a discourse analysis that focuses on “the conventional structures used to construct a complete text within the variety” (Biber & Conrad, 2019, p. 2).

GA is often conducted within three schools: New Rhetoric, English for Specific Purposes (ESP), and the Sydney School or Systemic Functional Linguistics (SFL) (Fakhruddin & Hassan, 2015; Hyon, 1996). The New Rhetoric focuses on the cultural contexts of genres and particularly emphasises the actions genres perform (Fakhruddin & Hassan, 2015). Miller (1984), a key figure in this school, argues that a rhetorically sound definition of genre should be centred on the actions it accomplishes, rather than on its substance or form. Moreover, Miller argues that a genre cannot be studied without understanding the community that utilizes it, aiming to uncover the attitudes, values, and beliefs reflected in the genre (Fakhruddin & Hassan, 2015). Unlike other schools, New Rhetoric pays little attention to the formal linguistic characteristics of genres (Hyon, 1996). Researchers in this school adopt ethnographic methods, such as observations and interviews, to describe the cultural contexts and actions associated with genres (Hyon, 1996). The typical procedure of GA within this school involves (a) collecting genre samples, (b) identifying the cultural contexts, (c) describing recurrent patterns and features, and (d) reflecting on their implications for the genre and cultural setting (Fakhruddin & Hassan, 2015).

According to ESP, a genre is a class of communicative events recognized by the genre community for their communicative purposes, schematic structure, content, style, and target audience (Swales, 1990). ESP focuses on analysing the communicative purpose, schematic structure, and formal linguistic features of genres within their specific contexts. In practical terms, ESP aims to equip non-native English speakers with the language resources and skills necessary for academic and professional settings, such as research articles, literature reviews, conference abstracts, job application letters, and legislative documents (Bawarshi & Reiff,

2010; Fakhruddin & Hassan, 2015; Hyon, 1996). Much of the research in this school is influenced by Swales (1990), who introduced MA. This approach examines a text's constituent moves which perform specific rhetorical or communicative functions. The common procedure for GA within ESP involves: (a) identifying a genre within a discourse community, (b) determining its communicative purpose, (c) analysing its structure using MA, and (d) examining the textual and linguistic features of these rhetorical moves (Fakhruddin & Hassan, 2015).

According to the Sydney School or SFL, a genre refers to a purposeful social activity that individuals participate in as members of their culture (Martin, 1984). This school originated within the framework of SFL, developed by Michael Halliday (1978) which focuses on the relationship between language, its function, and the situational context. In SFL, situational context is seen as being made up of three components: field (the activity taking place), tenor (the relationship between participants), and mode (the communication channel). These correspond to Halliday's metafunctions of language: ideational (action), interpersonal (participant interaction), and textual (cohesion and coherence within and between texts). GA in this approach aims to examine the stages or moves within a genre, as well as their lexical, grammatical, and cohesive patterns. The typical procedure for GA in the Sydney School involves four steps: (a) identifying the social purposes as structural elements and staging possibilities within a particular genre, (b) analysing the text's register in terms of field, tenor, and mode, and (d) analysing the microfeatures of the genre, which include semantic, lexicogrammatical, and phonological/graphological features (Fakhruddin & Hassan, 2015).

In general, the different schools of GA offer diverse viewpoints on the study of genres, focusing on aspects such as communicative purposes, context and users, internal discourse structure, and lexico-grammatical features shaping genres. Analytical methods for GA are varied, with genre analysts often combining approaches to gain a deep understanding of a genre's goals, setting, users or community (Tardy & Swales, 2014), or to achieve other purposes like teaching writing or exploring how language structure serves users' purposes in specific contexts. Historically, genre studies have employed various analytical methods (Tardy & Swales, 2014), including:

a) Text analysis: which identifies the grammatical or phraseological patterns specific to a genre.
b) Comparative GA: which examines a genre produced by multiple communities.
c) Diachronic GA: which analyses changes in a genre over time.
d) Genre system analysis: which investigates genre clusters that collaborate to achieve a shared objective.
e) Critical GA: which combines tools and methods from critical discourse analysis and GA.
f) MA: which identifies a text's communicative purposes by segmenting it into units or moves according to their rhetorical or communicative purposes.

4.2.2. Move Analysis

As mentioned previously, MA was originally developed by Swales (1990) to address the needs of non-native English speakers in reading, writing, and publishing research articles. It has since been successfully applied to a wide range of academic genres, argumentative essays (Hyland, 1990; Kanestion et al., 2017), and professional genres (Bhatia, 2014; Upton & Cohen, 2009), as well as news reports (Jabbari & Farokhipour, 2014) and press releases (Catenaccio, 2008; Chavanachid & Sripicharn, 2019). This type of analysis is usually qualitative and often used to study a limited number of texts (Upton & Cohen, 2009).

Moves are characterized by recognizable linguistic boundaries, such as a transition word or a switch in tense. It should be noted that they are rhetorical rather than grammatical categories (Biber et al., 2007; Tardy & Swales, 2014) and, as such, vary in length. A move can be a single word, clause, sentence, paragraph, or multiple paragraphs. It can also be realised through images or other audio-visual materials.

In a particular genre, some moves occur more frequently than others. Such moves can be described as *obligatory* or *conventional*. In contrast, moves that occur less frequently are *optional* (Biber et al., 2007; Tardy & Swales, 2014). Some researchers argue that a move must appear in each text under investigation to be obligatory, while others consider a move obligatory if it appears in half of the texts (Samraj, 2014). Samraj (2014) contends that a move's importance is indicated by its textual size and frequency of repetition.

MA is also concerned with the sequential order of moves, known as *move structure*, because genres are defined not only by the occurrence of certain moves but also by the organisational patterns in which they appear (Samraj, 2014). Move structure is usually hierarchical in that moves may have multiple propositions (i.e., the smallest and independent meaning units of discourse) that together fulfil the purpose. Such propositions are called *strategies* or *steps* and primarily serve the purpose of the move to which they belong (Biber et al., 2007; Samraj, 2014; Tardy & Swales, 2014).

To illustrate the process of MA, Hyland's (1990) analysis of the move structure in argumentative essays is a notable example. His data consists of the top 10% of argumentative essays (65 papers) submitted for the Papua New Guinea High School admission in English. Table 4.1 demonstrates the moves found in the three stages of the argumentative essay: thesis, argument, and conclusion. Moves written between brackets are optional.

Table 4.1 Hyland's (1990) Move Structure of Argumentative Essays

Stage	**Move**
1) Thesis Introduces the position to be argued	**(Gambit)** Attention grabber - controversial statement of dramatic illusion. **(Information)** Presents background material for topic contextualization. **Proposition** Furnishes a specific statement of position **(Evaluation)** Positive gloss - brief support of the proposition. **(Marker)** Introduces and/or identifies a list
2) Argument Discusses grounds for thesis. (Four-move argument sequence can be repeated indefinitely)	**Marker** Signals the introduction of a claim and relates it to the text **(Restatement)** Rephrasing or repetition of proposition. **Claim** States reason for acceptance of the proposition **Support** States the grounds which underpin the claim

3) Conclusion Synthesizes discussion and affirms the validity of the thesis	**(Marker)** Signals conclusion boundary. **Consolidation** Presents the significance of the argument stage to the proposition. **(Affirmation)** Restates proposition. **(Close)** Widens context or perspective of proposition.

4.2.3. Rationale for Using Genre Analysis to Analyse the Structure of False News Articles

By applying GA using rhetorical moves on false news articles, we can gain valuable insights into this phenomenon in multiple ways. Firstly, GA enables an exploration of the underlying purpose(s) of the texts under examination. GA's exploratory nature, as highlighted by Bhatia (2017), allows for a deeper understanding of genres by examining their constituent parts to infer communicative intentions (Swales, 1990). This approach moves beyond mere description, offering explanatory insights into why a genre is structured in a specific manner (Bhatia, 2017; Biber et al., 2007). Furthermore, according to Bhatia (2017), communicative purposes can be analysed at both general and specific levels, indicating that a genre may serve single or multiple purposes. Therefore, by analysing online false news articles from a genre perspective, this case study can uncover the broader and specific communicative goals behind the dissemination of false information.

Secondly, GA plays a crucial role in exploring the relatedness and variations within genres (Samraj, 2014). As Bhatia (2017) and Biber et al. (2007) suggest, GA extends further than merely describing language usage by providing a dynamic explanation of how various users manipulate a shared genre to accomplish their communication objectives. Consequently, this analysis can uncover how false news writers manipulate the news genre to strategically present information in a way that serves their objectives, influencing the perceived credibility and widespread acceptance of false news content.

Finally, GA assists in revealing the cultural assumptions and values emphasised in false news articles. Scholars, such as White (1997), emphasise how analysing the discourse structure of

news articles exposes the structured delivery of informational and interpersonal meanings. Similarly, Richardson (2006) argues that the discourse structure of news articles establishes links between events, situations, and states, offering insights into how values, beliefs, and knowledge are intertwined and communicated in false news articles. Given that news articles reflect significant events, their discourse structure plays a pivotal role in shaping perceptions of what is considered "important, trivial, fortunate, tragic, good, and evil" (Richardson, 2006, p 74).

In conclusion, GA is instrumental in delving into the discourse structure of online false news articles and in facilitating the examination of multiple critical aspects. This includes uncovering the motivations behind the creation of false news, investigating variations in discourse structure among false news articles, comparing these variations to structural norms in typical hard news reports, and highlighting the values and assumptions embedded in false news discourse.

4.3. Previous Research on News Genres

The section starts with a review of theories related to the genre of printed hard news reports. It then explores studies on the genre of online news reports (published on websites and blogs), examining their historical evolution and distinct characteristics. This review is crucial for establishing connections between the findings of the move analysis conducted in this study on online false news articles and the existing literature on the structural conventions of printed hard news reports, as well as their online counterparts.

4.3.1. The Genre of Printed Hard News Reports

The news genre has been extensively studied in discourse research, leading to the development of various theories regarding the structure of printed hard news reports. While these theories were not specifically developed within the scope of MA, they provide insights into the fundamental elements that constitute the structure of news articles.

An early example of this research is Van Dijk's (1983, 1988) proposal that the superstructure of news discourse consists of three categories: *summary/introduction, episode(s),* and *comments*. The summary/introduction includes the headline, whether it is super-, main-, or sub-headlines, captions, and lead. It expresses the semantic macrostructure of the text and introduces the main events, participants, location, and time. The episode consists of the following elements: a review

of information given previously; a report of the facts preceding the event; an explanation of the actual event, including its context and background; and a description of the event's consequences and verbal reactions (reporting statements of political figures or observers). The comments are the journalist's or newspaper's beliefs and evaluation. This hierarchical schemata describes the order of the news texts from left to right or from top to bottom. However, Van Dijk emphasises that this structure is not fixed. For example, describing consequences or reporting verbal reactions can be located before reporting previous information or preceding events.

White (1997) proposes that news reports are divided into two phases: an *opening nucleus*, which contains the main informational and interpersonal meanings (headline and lead), and a *satellite structure*, which develops the meanings presented in the first phase through *elaboration, contextualization, explanation, evaluation,* and *justification*. Similarly, Bednarek and Caple (2012) contend that the structure of most printed news genres consists of three parts: *headline, intro/lead,* and *body/lead development*. The headline frames the event, summarizes the story, and attracts the reader. The intro/lead performs these same functions but also construes newsworthiness. The body/lead development adds details and attributes some information to external sources. Bednarek and Caple (2012) also emphasise that images constitute an important part of printed news structure, but most theories on news structure are text-oriented and have paid little attention to images.

The genre of news reports is often categorized as expository due to its informative and analytical nature (Ricketson, 2014). However, some scholars classify news reports as narrative. For example, Richardson (2006) argues that the news article genre is similar to that of storytelling or narrative and that the study of news narrative involves analysing both the content and presentation of news stories—in other words, the narrative content and the narrative form. The narrative content is the structure of events, i.e., the sequence of events as they actually occurred. This structure consists of a three-part plot: the setting, the event, and the outcome. However, this structure does not always offer a final resolution. Understanding the connections between these three parts requires applying knowledge and creating a clear and logical link. The narrative form, on the other hand, refers to the sequence of events as they are presented. News usually follows an *inverted pyramid* or *climax-first* structure, in which the most important information is placed at the beginning. The following paragraphs identify the setting, answering the five questions of

who, what, why, where, and when, and finally, the outcome is presented. In this structure, news hardly follows a chronological order. Another narrative form is the *dropped introduction*, where the news writer starts with the setting, sometimes using cataphoric references to create a sense of anticipation, and then provides more details as they progress with the narrative, sometimes referring to their sources. This type of narrative form, Richardson (2006) argues, *blurs genres,* incorporating elements of both news reporting and entertainment.

Similarly, Bell (2005) suggests that news writers do not write articles; instead, they write stories. To support his argument, he compares Labov's (1999) model of face-to-face personal narrative structure with news stories. Labov's model divides the personal narrative structure into six elements:

1) The *abstract:* which summarises the main event and presents the main purpose of the narrative.
2) The *orientation:* which presents the setting of the events, i.e., the time, place, and people involved.
3) The *complicating action:* which presents the events in chronological order.
4) The *evaluation*: which justifies the significance of the narrative, i.e., why the story is narrated.
5) The *resolution:* which presents the final event and concludes the complicating action.
6) The *coda*: which marks the end of the narrative and returns it from the time of past events to the present time of narration.

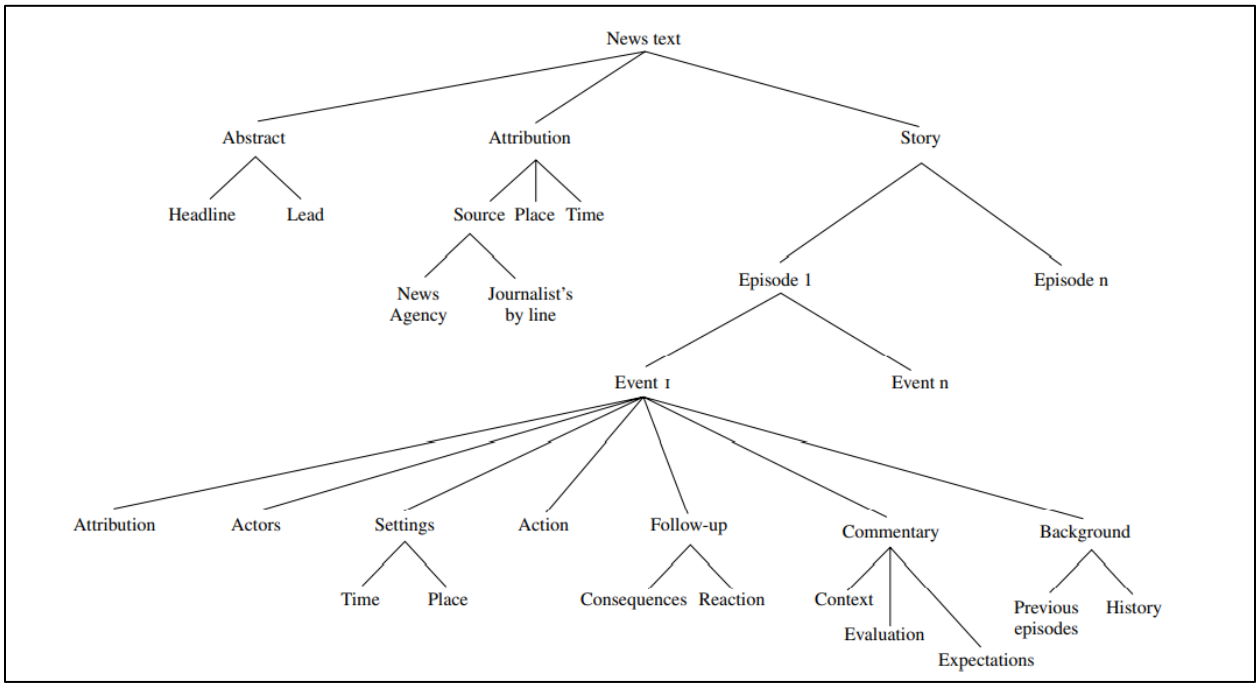
Except for evaluation, which can occur in different parts of the narrative, these six elements occur in a fixed order.

Bell (2005) argues that news stories have some narrative elements but differ from personal narratives in terms of the elements' importance and order. First, the *lead* in news stories serves the same function as the *abstract* in personal narratives. It summarises the main event and establishes the story's newsworthiness. However, while the abstract in personal narratives is optional, the lead in news stories is obligatory. Second, the *orientation* also occurs in news stories, setting the scene and answering the questions of who, what, when, and where. The orientation occurs at the beginning of personal narratives, while in news stories it can be detailed

further down. Third, like personal narratives, news stories have an *evaluation* element, which establishes the story's significance and justifies claiming readers' attention. While the evaluation in personal narratives occurs throughout the story, in news stories it is usually concentrated in the lead. Fourth, the *complicating action* is the heart of both personal narratives and news articles. Personal narratives are characterised by the temporal sequence of their sentences, i.e., actions are narrated chronologically. However, in news stories, actions are seldom reported in chronological order; instead, they are presented in an order of decreasing importance. Finally, unlike personal narratives, news stories lack a *resolution* and *coda*. They do not present the results or end of the complicating action, nor do they mark the end of the story. Instead, "they finish in mid-air" (Bell, 2005, p. 402).

Based on the comparison with Labov's (1999) model of personal narrative and theories on news schemata (Van Dijk, 1983, 1985, 1988), Bell (2005) proposes that the structure of news stories consists of three main elements: the *abstract*, *attribution*, and *story proper*, as illustrated in Figure 4-1.

Figure 4-1 Bell's (2005) Model of the Structure of News Texts



The abstract mainly comprises the headline and lead. The lead reports the main event of the story (and sometimes a second event) and provides information about the actors and setting.

Occasionally, the lead includes supplementary elements, such as evaluation. The attribution clarifies the story's origins. This can be a place, time, or news source, like a news agency or the journalist's byline. The attribution may be incorporated into the lead or the story proper.

The story proper consists of one or more episodes, each of which containing at least one event. Events, in turn, consist primarily of *actors*, *action*, and *setting*. In addition to these central elements, events include three additional elements*: follow-up, commentary*, and *background*. The follow-up reports any action that occurred after the main action of an event. It can be a description of a verbal reaction or other subsequent actions. Commentary involves evaluation, which can be explicit or implicit, and expectations of how the event could develop further. Commentary can be provided by either the writer or an actor, embedded in quotations. Finally, the background reports any events that occurred before the main action. These events can be classified into two main categories. If they are recent events, they are classified as previous episodes; otherwise, they are classified as history.

To sum up, several theories have been developed regarding the structure of printed hard news reports. According to these theories, hard news reports follow an inverted pyramid structure in which information is reported in an order of decreasing importance. That is why news reports must have a lead, abstract, introduction, summary, or nucleus in which the most important information is reported. Further, while hard news reports are generally categorized as expository, some theories suggest that the structure of hard news reports is similar to that of a narrative; it establishes the newsworthiness, gives the setting of the story, includes an evaluation element at the beginning, and then details the actions of the main story event. These theories also state that hard news reports include other elements that support the narrative of the main story event, like giving background information, describing consequences and expectations, and reporting comments and reactions of both journalists and news actors.

4.3.2. The Genre of Online News Articles

Technological innovations have played a pivotal role in the evolution of news. Today, news appears on websites and blogs as well as microblogs and social media. As a result, numerous traditional newspapers now have online versions and social media accounts. This section

discusses the genre of online news articles, published by websites and blogs, as they are the focus of this study.

News articles published online differ from printed ones in terms of their content and format. First, online news articles are characterized by hypermedia containing audio-visual materials (Toolan et al., 2001). These are used to emphasise the writer's viewpoint, promote authority, or add visual or aesthetic effects (Myers, 2010). As Pavlik (2000) puts it, the traditional inverted pyramid structure of news stories is being replaced by a more immersive and interactive multimedia structure. This new structure aims to create an experience for readers that makes them feel as though they are personally involved in the events being reported. Additionally, online news articles usually contain advertisement panels and sidebars (Toolan et al., 2001). Online news articles are also distinguished by hypertextuality; that is, the news text is embedded with links to other related stories or social media (Toolan et al., 2001). Thus, the structure of online news reports is less linear and orderly than printed news. This means that their narrative structure is distorted (Toolan et al., 2001), fragmented, and interdependent with other websites (Wall, 2005).

After the September 11, 2001 attacks, news blogs gained popularity as a result of the public's need for immediate and accurate information (Allan, 2011). This form of journalism has been referred to by various names, such as *personal journalism*, *do-it-yourself journalism*, *black-market journalism*, *we media*, and *postmodern journalism* (Wall, 2005). One notable aspect of blogs is their ability to foster a sense of community among users who share similar social identities (Myers, 2010). Wall (2005) emphasises that news blogs embrace a journalistic approach that combines personalization with audience participation in content creation. Bloggers, whether amateur or professional, report on events, offer personal opinions, and provide platforms for audiences to comment or engage in discussion (Allan, 2011).

To summarize, the internet has caused a great change in the genre of hard news reports. Online news articles have a deeply engaging structure due to their hypermedia, and they have a fragmented structure due to their hypertextuality. News blogs, in particular, have contributed to the transformation of news in that personalization, readers' engagement in content creation, and the writer-reader relationship are more intense than in other kinds of news.

4.4. Corpus

This case study focuses on a specific subset of the corpus, specifically the 20 false news articles related to climate change outlined in Chapter 3. The decision to focus solely on this subset of the corpus was based on two main factors. Firstly, genre analysis especially move analysis, is primarily qualitative and demands meticulous manual analysis, making it more feasible to analyse a smaller subset thoroughly. Secondly, the objective of the study includes identifying a shared move structure (a sequence of moves) i.e., within online false news articles, a task that may be more effectively accomplished by examining a more condensed set of articles on the same topic.

Moreover, it is essential to highlight that this case study did not involve a comparative analysis between the false news corpus and the reference corpus outlined in Chapter 3. This decision was influenced by the significant time and effort required for the qualitative and manual GA, which spanned approximately eight months, including the subsequent semantic analysis discussed in Section 4.5.3. The overall goal of this PhD project is to triangulate corpus-assisted analytical methods to gain a comprehensive understanding of false news discourse. Therefore, due to time constraints and to allow for the completion of the other planned case studies, a comparative MA was not conducted in this particular case study.

It is noteworthy that that the analysis was conducted on the original versions of the false news articles as they appear on the webpage, rather than the text versions compiled for this thesis. This choice was made to ensure that the analysis encompasses all relevant structural elements, such as advertisements, comments, and sources, which are not preserved in the text versions. Nvivo, an analytical software utilized for both statistical and qualitative data analysis (QSR International Pty Ltd, 2020), was utilized in this study. The webpages hosting the articles were captured utilizing NCapture, a web browser add-on, in conjunction with Nvivo.

4.5. Analytical Procedure

As Biber et al. (2007) indicate, move analysis has no strict rules. However, researchers are advised to develop a flexible analytical procedure that could be adapted to the specific structural features of the genre being studied. Biber et al. propose two main approaches to analyse discourse structure: the top-down approach and the bottom-up approach.

In the top-down approach, the first step is to create an analytical framework by identifying possible discourse unit types (or moves) in the genre being studied based on prior knowledge of the major communicative functions that the discourse units can fulfil. This framework is then applied to analyse all texts in a corpus. On the other hand, the bottom-up approach starts by automatically segmenting all texts in the corpus into discourse units using linguistic criteria. Next, the linguistic features of these discourse units are analysed, and the units are grouped into clusters based on their linguistic similarities. These groupings can be identified as discourse unit types (or moves) by determining their typical functions in texts.

To analyse move types and their functions in this study, I employed Biber et al.'s (2007) bottom-up corpus-based approach. While this approach is commonly applied with the assistance of advanced computational tools to handle large text datasets, I found it more appropriate to manually apply this approach in analysing the genre of online news reports for two reasons. Firstly, there is limited prior knowledge available regarding the discourse structure of false news articles. Consequently, it is challenging to identify the move types without first segmenting articles into units and analysing the linguistic features and their communicative functions. Secondly, the corpus used in this study is relatively small, comprising only 20 articles, making it feasible to manually employ the bottom-up approach.

Table 4.2 provides an overview of the seven stages of Biber et al.'s (2007) bottom-up approach. However, I did not follow all these stages in their sequential order. Instead, I adjusted the order to better facilitate a more effective analysis of the online false news genre. In the subsequent sections, each stage will be thoroughly discussed in the order I followed in my analysis.

Table 4.2 Stages of the Bottom-up Corpus-based Approach to Discourse Organization

Stages	
1) Segmentation	Segmenting each text into discourse units according to shifts in vocabulary or other linguistic features
2) Linguistic analysis of each unit	Analysing the lexico-grammatical features of each discourse unit in each text of the corpus
3) Classification	Identifying a set of discourse unit types or categories that are linguistically identifiable

4) Linguistic description of discourse categories	Describing the typical linguistic characteristics of each discourse unit type, based on the linguistic analysis of each discourse unit belonging to a particular type
5) Communicative functional categories	Describing the typical communicative functions of each discourse unit type, based on an analysis of the communicative functions of each discourse unit belonging to a particular type
6) Text structure	Analysing the sequence of discourse unit types in each text
7) Discourse organizational tendencies	Describing the whole corpus in terms of typical discourse structure

4.5.1. Segmentation and Identification of Communicative/Functional Categories

The initial stage involved selecting a subset of the corpus and segmenting the articles into rhetorical moves. This segmentation process was achieved by analysing shifts in vocabulary, discourse markers, and meanings, which helped in understanding the communicative and functional purposes of the corpus as a whole and of each segment.

Following this, an analytical framework was developed inductively, guided by the segmentation process, to identify a set of communicative and functional categories (i.e., moves) within online false news articles. Additionally, a set of sub-functional or semantic categories (i.e., strategies) that supported the broader moves was also identified. Multiple readings and reflections were necessary to clearly define these move types. A coding manual was then created with explicit definitions and examples of move types and their corresponding strategies detailed in Appendix 4.1.

Several challenges arose during this stage. One challenge was determining the overall purpose of a text, a challenge acknowledged by previous move analysts like Samraj (2014). As suggested by Askehave and Swales (2001), not all genres have a single identifiable purpose. This complexity is particularly evident in false news, given the diverse motives driving its creation, encompassing ideological, financial, social, and psychological factors.

To address this challenge, it was important to consider the situational and social context, as noted by Bhatia (2017) and Biber et al. (2007). As recommended by Askehave and Swales (2001), GA

should involve a comprehensive examination of both the text and its context. Van Dijk (1998) defines context as the structured set of all relevant properties of a social situation that impact the production, structure, interpretation, and function of text and discourse. In order to determine the overall communicative purpose of the false news articles in the study, a review of the social and political context of climate change false news was conducted, as detailed in Chapter 3. In addition, the fact-checking articles that verify the claims made in the analysed article were read as they provided background information on the claims' origin and context, as well as the reasons why they were false.

Another challenge encountered was determining the boundaries between moves. Samraj (2014) states that identifying moves within a genre can be difficult, and even experienced readers may disagree on where one move begins and another ends. This is particularly amplified in genres with a high level of hierarchical complexity, where one move may be embedded within another. Additionally, certain segments can serve multiple purposes and be interpreted as different moves.

In MA studies, analysts employ various criteria, both linguistic and non-linguistic, to determine move boundaries. Some analysts rely on switches in tense or transitions, while others focus more on the semantic functions of texts (Tardy & Swales, 2014). Furthermore, analysts may also draw on their intuition and knowledge of text structure to identify moves and their boundaries, as noted by Swales (2004). To address this challenge, a comprehensive approach was taken in this study, utilizing a range of linguistic and non-linguistic criteria, and considering all available evidence when making decisions about move definitions and boundaries. Furthermore, the coding was conducted multiple times to ensure consistency in identifying move boundaries.

4.5.2. Moves Classification

The articles were coded using Nvivo (QSR International Pty Ltd, 2020) where nodes were used to code moves and child nodes to code strategies. Each article was saved as a separate file (i.e., source) and segmented according to the rhetorical moves it contained. In this stage, additional move types or strategies were identified, leading to refinements in the coding manual. The segmentation, identification of functional categories, and classification processes were conducted multiple times, allowing for the resolution of any inconsistencies and the recoding of problematic texts.

4.5.3. Linguistic Analysis and Linguistic Description of Moves

Given that move types are defined according to their functions, they can vary in their linguistic realization, drawing on an open-ended set of linguistic features such as grammar, word choice, and phraseology. Upton and Cohen (2009) state that analysing the linguistic realisations of each move type plays a pivotal role in understanding how linguistic features are used and interact with each other for specific communicative purposes.

At this stage, all segments belonging to each move type were analysed linguistically using the corpus method of frequency lists of semantic domains (Archer et al., 2003) (see Appendix 4.2 for details). This method was selected for its comprehensive nature, as it allows for the categorization of words into semantic groups, providing a more insightful understanding of how each move is meaningfully conveyed compared to a frequency list solely based on parts of speech. By conducting a concordance analysis of the most frequently occurring semantic domains within each move type, interesting findings about news values were revealed. These findings proved to be influential and provided inspiration for the second case study, which specifically investigates the news values intrinsic to false news (refer to Chapter 5).

4.5.4. Identification of Move Structure and Discourse Organisational Tendencies

After the MA, two further levels of analysis were conducted. The first level was analysing the move structure or move sequence of each individual text in the corpus. This facilitates the description of the internal discourse organisation in terms of the move types used. The second level was identifying the typical and/or alternate move structure of the corpus as a whole. It should be noted that move types may not necessarily occur in a fixed order (Biber et al., 2007), a point also emphasised by theories on the discourse structure of hard news reports (refer to Section 4.3). Hence, in this level of analysis, the objective was to identify the most frequent pattern occurring in the corpus. The analysis of move structure might be useful for determining whether a news article is false. As stated previously, automatic false news detection based on discourse-level structure has proven useful, so a corpus-assisted analysis of move structure could be equally useful in distinguishing the discourse structure of false news.

Identifying the move structure of texts has been recognized as challenging in previous MA studies (Samraj, 2014). This is particularly evident in analysing online news articles as they are

characterized by hypertextuality and a non-linear presentation of text and other materials, including audio-visual content, pull quotes, boxed news stories, ads, social media links, and other links. To address this issue, a unique letter code was assigned to each move type, a method suggested by Groom and Grieve (2019). The vertical move sequence of each article was represented using a series of letters to indicate the order of moves, with horizontally adjacent or successive moves identified by a hyphen.

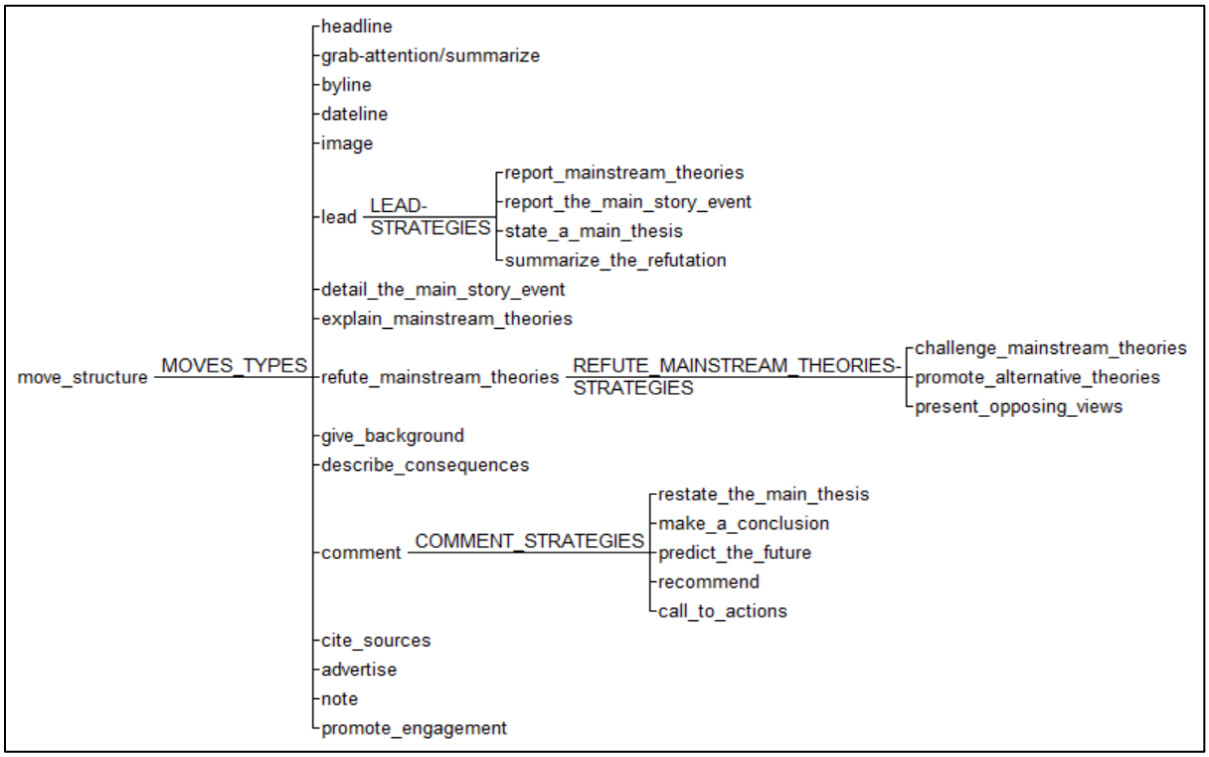
4.6. Analysis

This section presents the results of the move analysis. First, it presents the move types used in the collected false news articles, along with their functions, typical locations, and illustrative examples. Following this, the section reports the frequency of moves' occurrences and the proportion of text dedicated to each move. Subsequently, it discusses the variations among articles according to their move types and move structures.

4.6.1. Move Types and Their Functions and Typical Locations

Informed by Biber et al.'s (2007) approach to the analysis of discourse organization, sixteen distinct moves were identified that appear to be functional components of false news articles. Each move was assigned a letter code for reference. Figure 4-2 visualises these moves and their constituent strategies.

Figure 4-2 Move Structure of Online False News Articles on Climate Change



The following outline discusses these moves, their functions, strategies, and typical positions in the news articles. The moves and their strategies are illustrated through examples extracted from the corpus.

Move 1: Headline (H)

False news articles always begin with a headline, which contains the main story event and sometimes the main thesis adopted by the news writer or outlet. In Example 1, the title suggests that Google and mainstream media utilize fact-checking services to distort our perception of truth.

(1) How Google and MSM Use "Fact Checkers" to Flood Us with Fake Claims *(A False 11)*

Move 2: Grab Attention/Summarise (G)

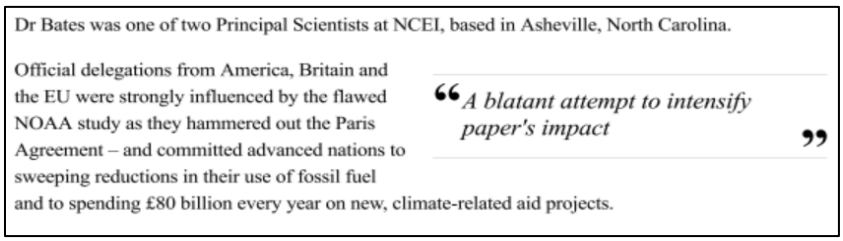
This move functions as an explanation of the main title or as a summary of the article and captures readers' attention. It can appear in various forms as subtitles or tags after the headline; bullet summaries, subheadings, and pull quotes within the text; or captions after images.

Example 2 is a bullet summary provided after the headline of one false news article. Figure 4-3. presents an example of a pull quote taken from the main body of an article and displayed in a larger font size to draw the reader's attention.

(2)

- 533,000 more square miles of ocean covered with ice than in 2012
- BBC reported in 2007 global warming would leave Arctic ice-free in summer by 2013
- Publication of UN climate change report suggesting global warming caused by humans pushed back to later this month *(A False 14*)

Figure 4-3 An Example of the Move 'Grab Attention/summarize' (A False 13)


Dr Bates was one of two Principal Scientists at NCEI, based in Asheville, North Carolina.

Official delegations from America, Britain and the EU were strongly influenced by the flawed NOAA study as they hammered out the Paris Agreement – and committed advanced nations to sweeping reductions in their use of fossil fuel and to spending £80 billion every year on new, climate-related aid projects.

"A blatant attempt to intensify paper's impact"

Move 3: Byline (W)

This move indicates the author's name, sometimes with a short description, as seen in Example 3. It usually appears after the headline.

(3) By DAVID ROSE FOR THE MAIL ON SUNDAY *(A False 13)*

Move 4: Dateline (D)

This move indicates the publication date (see Example 4). It usually appears after the headline along with the byline.

(4) 15 hours ago October 30, 2017 *(A False 11)*

Move 5: Image (I)

This move consists of an image, usually located before the lead of the news article (see Figure 4-4). It is sometimes followed by a caption, which as indicated previously, constitutes an instance of the move *grab attention/summarize*.

Figure 4-4 An Example of the Move 'Image' (A False 7)

Move 6: Lead (L)

This move constitutes the beginning of the news article itself; it usually follows the image. The main function of this move is to summarise the article and establish the newsworthiness or significance of the story. Sometimes, this move begins with the name of the news outlet written between brackets. This move can be accomplished through one or more strategies. One strategy is reporting the main story event, which is usually used as a starting point to delegitimise mainstream theories. Another strategy is reporting mainstream theories, in which the writer reports what has already been established by mainstream media or governmental organisations about the topic of the news article. Another strategy is expressing the main thesis adopted by the news writer or outlet towards the topic. This strategy clearly states that the mainstream theories are incorrect and/or proposes alternative theories. A final strategy is summarising the refutation of the mainstream theories, which is usually developed in the main body text.

All these strategies are illustrated in Example 5, a lead extracted from a false news article. Initially, the author presents the main story event detailing the catastrophic impact of the inferno in Australia. Subsequently, the author reports the mainstream narrative, notably promoted by radical environmental extremists and celebrities, attributing the fires to climate change. After that, the author states the main thesis of the article emphasising that climate change is not the

primary cause of the crisis in Australia. Finally, he summarized his refutation of the mainstream narrative, highlighting that many of the fires were deliberately set by arsonists. In the example, the strategies are denoted by subscripts.

> (5) The raging inferno that has engulfed a large swath of Australia is horrifying. More than 1,500 homes have been destroyed, 25 people have died and millions of kangaroos and koala bears and other animals may have been lost. [report main story event] Radical environmental extremists and Hollywood celebrities have blamed the fires on climate change. [report mainstream theories] But it turns out — climate change has nothing to do with the humanitarian crisis unfolding Down Under. [state a main thesis] It turns out — many of the fires were set by arsonists. [summarize the refutation] *(A False 1)*

Move 7: Detail the Main Story Event (M)

This move provides more details about the main story event reported in the lead. It may contain information about the place, time, and people involved. This move is either located at the beginning of the body text after the lead or occurs multiple times in the middle of the body text. Example 6 illustrates this move by offering specific information about John Bryson, the former US secretary of commerce. In this example, Bryson is reported to endorse world governance:

> (6) In the video – given to The Daily Caller by sources on Capitol Hill – Bryson refers to the 2009 United Nations climate negotiations in Copenhagen as "the closest thing we have to a world governance organization," implying it provided the best model for imposing climate regulations on countries around the globe. Bryson also boasts about his role as an adviser to the U.N. secretary general on climate change in the video. *(A False 10)*

Move 8: Explain Mainstream Theories (E)

This move elaborates on what has already been established by mainstream media or governmental organisations about the topic of the news article. The explanation can be textual or visual. It tends to be located at the beginning of the body text, after the lead. However, in many articles, it occurs in multiple places in the middle of the body text. Examples 7 and 8 provide an elaboration on the mainstream narrative that a pause or slowdown in global warming did not

occur. They detail the sources of this narrative, dates, and specific information related to the assertion.

> (7) The report claimed that the 'pause' or 'slowdown' in global warming in the period since 1998 – revealed by UN scientists in 2013 – never existed, and that world temperatures had been rising faster than scientists expected. *(A False 13)*

> (8) The paper was published in June 2015 by the journal Science. Entitled 'Possible artifacts of data biases in the recent global surface warming hiatus', the document said the widely reported 'pause' or 'slowdown' was a myth. *(A False 13)*

Move 9: Refute Mainstream Theories (R)

This move refutes mainstream theories and is located either at the beginning of the body text after the lead or occurs multiple times throughout the body text. This move is established through one or more of the following strategies. The first strategy is disputing mainstream theories by citing authoritative sources (textual or audio-visual), using logical reasoning, reporting unattributed facts, or using evaluation. Examples 9, 10, and 11 demonstrate the first strategy of disputing mainstream narratives. In Example 9, the author refers to a NASA chart in his article as evidence to substantiate his claim that the degree of ocean level increase has been exaggerated. The author encourages the readers to examine the evidence themselves.

> (9) However, these people wildly exaggerate the degree of ocean level increases to the point of absurdity. If you zoom out on the NASA chart, you'll see a long-term trend of sea levels rising 3.4mm per year on average, according to NASA's own analysis. This means that over an entire century, the oceans would rise 340mm, or 13.4 inches … a little over a foot. *(A False 8)*

Example 10 showcases the author employing logical reasoning to argue that climate change is advantageous for the planet. The author contends that climate change, with increased CO2 levels and warmer temperatures, would lead to a greener Earth with augmented rainfall, potentially restoring deserts to lush plains or forests. The argument challenges the prevailing belief that climate change might result in ecological degradation.

> (10) Planet Earth would actually be far more green with higher CO2 levels and warmer temperatures that promote more continental rainfall. Deserts would be restored to plains or forests. Sparse areas, now devoid of plant life, could become food production areas. *(A False 8)*

In Example 11, the author highlights Dr Bates' distinguished reputation as a top scientist at NOAA. The author then reports that Dr Bates provided compelling evidence to The Mail on Sunday, indicating that the referenced paper relied on misleading and unverified data.

> (11) But the whistleblower, Dr John Bates, a top NOAA scientist with an impeccable reputation, has shown The Mail on Sunday irrefutable evidence that the paper was based on misleading, 'unverified' data. *(A False 13)*

The second strategy is promoting alternative theories by citing authoritative sources that evaluate them positively or describe positive reactions to them. For instance, in Example 12, the author presents the work of Milankovitch as a counterpoint to the mainstream narrative, focusing on the collective effects of changes in Earth's movements on its climate over long periods of time. The author discusses how Milankovitch's theory was initially overlooked but was later recognized, validated, and adopted.

> (12) When Milankovitch first put forward his model, it went ignored for nearly half a century. Then, in 1976, a study published in the journal Science confirmed that Milankovitch's theory is, in fact, accurate, and that it does correspond to various periods of climate change that have occurred throughout history.
> In 1982, six years after this study was published, the National Research Council of the U.S. National Academy of Sciences adopted Milankovitch's theory as truth, declaring that "… orbital variations remain the most thoroughly examined mechanism of climatic change on time scales of tens of thousands of years and are by far the clearest case of a direct effect of changing insolation on the lower atmosphere of Earth." *(A False 2)*

The third strategy is presenting views that neither support the theories adopted by the news writer or outlet nor accord with mainstream theories. By employing this strategy, objectivity can be enhanced, and some support can be provided to counter mainstream theories. Example 13 demonstrates this strategy. The author reports on Dr Ed Hawkins acknowledging that natural factors contribute to global warming but asserts that they are not solely responsible.

(13) Others are more cautious. Dr Ed Hawkins, of Reading University, drew the graph published by The Mail on Sunday in March showing how far world temperatures have diverged from computer predictions. He admitted the cycles may have caused some of the recorded warming, but insisted that natural variability alone could not explain all of the temperature rise over the past 150 years. *(A False 14)*

Move 10: Give Background (B)

This move provides background information or reports prior events that are necessary to understand the current news story or to support the main thesis. This move is located in the middle of the body text, and it can occur once or multiple times. For instance, in Example 14, the author references a 2015 report by BBC News about the significant growth of Arctic ice following an atypically cool summer.

(14) Back in 2015, BBC News reported that Arctic ice had grown by a staggering 30 percent after what they called an "unusually cool summer" – unusual indeed, if the global warming narrative is to be believed. *(A False 7)*

Move 11: Describe Consequences (K)

This move describes the implications or consequences of the main story event. It further involves the reporting of insights from authoritative sources that underscore the implications or consequences of such events or theories. It can be found at various points within the news article and may appear multiple times. Example 15 demonstrates this move, with Colin Hanna, a president of a public policy think tank, asserting that John Bryson's, the former United States secretary of commerce, endorsement of global carbon emissions standards would yield adverse effects such as decreased exports, escalated unemployment, and diminished global competitiveness.

(15) But Bryson's advocacy of global carbon emissions standards would accomplish the exact opposite, according to Hanna.
"If we followed the standards of Copenhagen, we would have decreased exports, increased unemployment and a loss of global competitiveness," Hanna said. *(A False 10)*

Move 12: Comment (C)

This move establishes the newsworthiness or the significance of the news story and expresses the writer's position through several strategies. It is usually located at the end of the news article; however, there are many instances of comments spread throughout the body text. One of the strategies of these moves is to restate the main thesis adopted by the news writer or outlet towards the topic of the news article by stating that the mainstream theories are wrong and/or that the alternative theories are true. For instance, in Example 16, the author asserts that NASA manipulation of climate data is a deliberate scientific fraud to deceive the global population about global warming.

> (16) This is a clear case of scientific fraud being carried out on a grand scale in order to deceive the entire world about global warming. *(A False 15)*

Another strategy is providing a conclusion in which the writer interprets all the information previously provided in the article and makes a judgment about the mainstream theories (e.g., inferring the motivation behind governmental behaviours or declarations). In Example 17, the author concludes that the fact-checking conducted by Google and certain social media platforms is deemed to be a form of political censorship.

> (17) Of course, this fact-checking by Google and some "social media" websites is nothing but a political censorship. Google is a monopoly created by the Obama administration through failure to enforce anti-trust laws, and then through subjecting the Internet use of 300 million citizens to the Title II regulation for the benefit of Google and other content providers. I dare to say that through Google, Facebook, and some other hand-picked corporate favorites, the Obama administration attempts to control our speech from the grave. *(A False 11)*

Another strategy is making predictions about what is likely to happen, or even stating what is going to happen, in the future as a result of adhering to the mainstream theories. As shown in Example 18, the author predicts that certain disclosures will likely strengthen President Trump's determination to fulfil his pledges of reversing his predecessor's environmental policies and exiting the Paris Agreement, potentially leading to a contentious political dispute.

(18) His disclosures are likely to stiffen President Trump's determination to enact his pledges to reverse his predecessor's 'green' policies, and to withdraw from the Paris deal – so triggering an intense political row. *(A False 13)*

The comment may also involve the strategy of providing recommendations of certain behaviours, beliefs, or reading material to the reader, e.g., suggesting other related or non-related news reports, as in Example 19.

(19) Read More Stories About: Big Government, Environment, Aden, Albert Parker, Clifford Ollier, Climate Change, Global Warming, Karachi, Mumbai, NASA, NOAA, PSMSL, sea level, sea-level rise. *(A False 17)*

The last strategy is a call to action, in which the writer encourages people in power to take some desired action. Example 20 showcases this strategy, where the writer stresses the importance of scientists following evidence-based practices rather than manipulating data to fit predetermined outcomes.

(20) It's understandable that a trend of continuous heating for 30 years would raise concerns, but by the same token, a continuous cooling trend over the last five years must also be taken into consideration. Scientists are supposed to go where the evidence takes them, not try to keep editing the evidence to fit a predetermined outcome. *(A False 7)*

Move 13: Cite Sources (S)

This move provides a list of the sources of the information in the news article, usually links to other websites, as shown in Example 21. It usually occurs at the end of the news article.

(21) Sources for this article include: HalTurnerRadioShow.com NaturalNews.com BBC.com *(A False 2)*

Move 14: Advertise (A)

This move involves advertisements which can be textual or audio-visual, as seen in Example 22. This move can appear in different locations throughout the news article.

(22) LEARN HOW TO SAVE OUR NATION FROM THE SOCIALISTS. CLICK HERE TO ORDER TODD'S LATEST BEST-SELLING BOOK, "CULTURE JIHAD: HOW TO STOP THE LEFT FROM KILLING A NATION." *(A False 1)*

Move 15: Note (N)

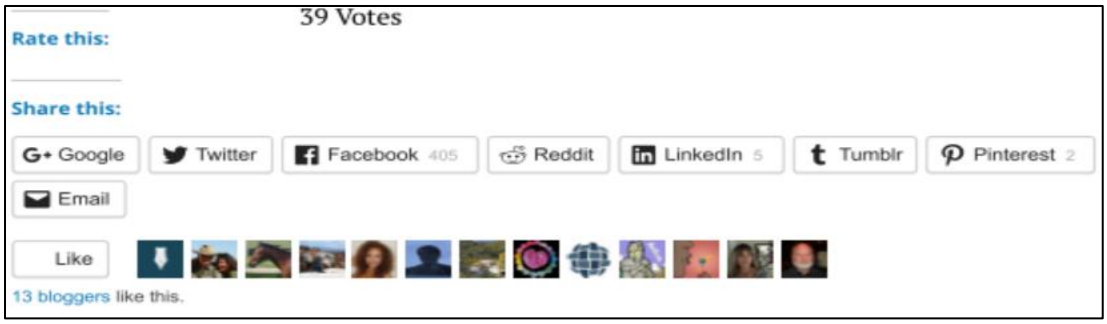
This move contains notes that comment on something outside the news text, for example, notes about the author, website/blog, or the comments on the news report itself. Notes usually appear at the beginning or end of the news article. In Example 23, a disclaimer note clarifies that the views presented do not necessarily align with MailOnline's perspective.

(23) The views expressed in the contents above are those of our users and do not necessarily reflect the views of MailOnline. *(A False 13)*

Move 16: Promote Engagement (P)

This move promotes readers' engagement with the news article's content by presenting links that enable readers to post comments or share the article on social media platforms, as illustrated in Figure 4-5. This move is usually located at the beginning or end of the news article and sometimes in the middle.

Figure 4-5 An Example of the Move 'Promote Engagement' (A False 11)



Figures 4-6 and 4-7 illustrate the move structure of two articles *(A False 7* and *A False 20)*, with move types indicated next to their respective segment.

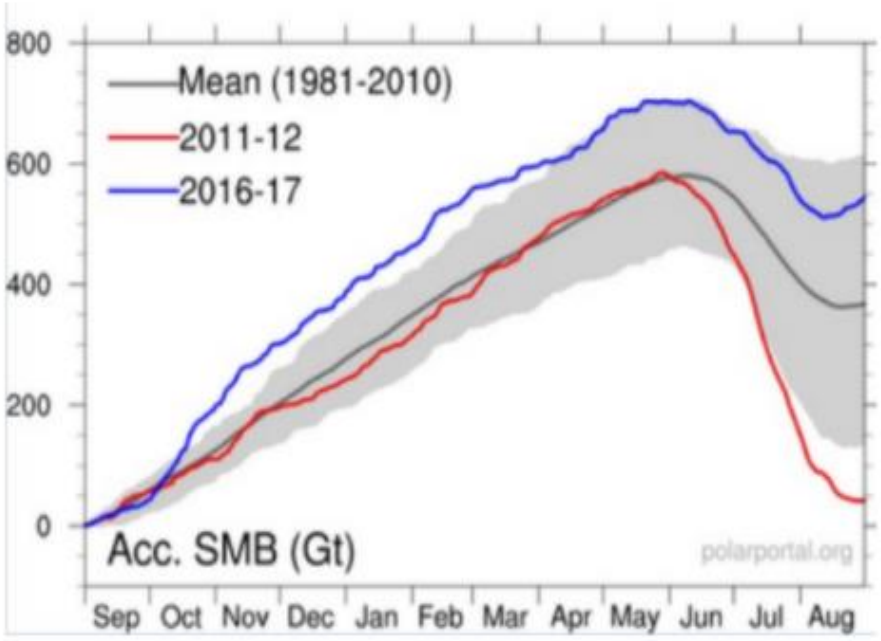
Figure 4-6 An Example of the Moves Identified in a False News Article (A False 7)

There has been quite some discussion about Greenland in the climate blogosphere this year. Heavy snow and rain in winter with a relatively short and intermittent summer melt season have ***left the Greenland ice sheet with more ice than has been usual over the last twenty years*** *– in fact we have to go back to the 1980s and 90s to see a year similar to this one in terms of snow fall and ice melt, though perhaps not for iceberg calving. …*

If we rank the annual surface mass balance since 1981 from low to high, the lowest on record was 2011-2012 (38 Gt) and ***this year is the 5th*** *highest out of the 37 year record. The highest on record 1995-1996 had an end of year SMB of 619 Gt in our records.*
[Emphasis added]

In fact, Greenland experienced a 10 times higher level of surface ice than it did five years ago. And confirming that this is not a fluke occurrence taking place in only one year, Greenland's most well-known glacier – the Petermann Glacier – has been growing slowly and steadily for the past five years.

As you can see from the chart below, **the sea ice growth for 2016 – 2017 is much higher than the mean from 1981 – 2010**:



refute mainstream theories

This has been the pattern in the Arctic over the last few years. Back in 2015, *BBC News* reported that Arctic ice had grown by a staggering 30 percent after what they called an "unusually cool summer" – unusual indeed, if the global warming narrative is to be believed. That trajectory continued into 2014, and the increases in ice for those two years exceeded all recorded losses in the preceding three years. (Related: Mercury pollution contaminating Arctic wildlife, causing death and reproductive problems among bird populations, studies show.) comment

give background

That 30 percent constituted a massive amount of physical land area – the *Daily Mail* reported at the time that a cooler Arctic summer had left over 530,000 additional square miles of ice than the previous year. refute mainstream theories

Astoundingly, the mainstream media, in spite of having this information at their fingertips, continues to spout the same old global warming nonsense. explain mainstream theories

In the very same *BBC* article cited earlier, for example, the writer immediately insisted that "2013 was a one-off" and went on to stress that the Arctic region had warmed more than most other places on Earth over the past three decades. (Related: Discover the truth at ClimateScienceNews.com) give background

It's understandable that a trend of continuous heating for 30 years would raise concerns, but by the same token, a continuous cooling trend over the last five years must also be taken into consideration. Scientists are supposed to go where the evidence takes them, not try to keep editing the evidence to fit a predetermined outcome.

It's like scientists, governments and the mainstream media have all decided that the Earth is warming, and now there's no turning back. No matter what the physical evidence shows, they're all just going to save face by insisting that they were right all along. comment

Sources include:

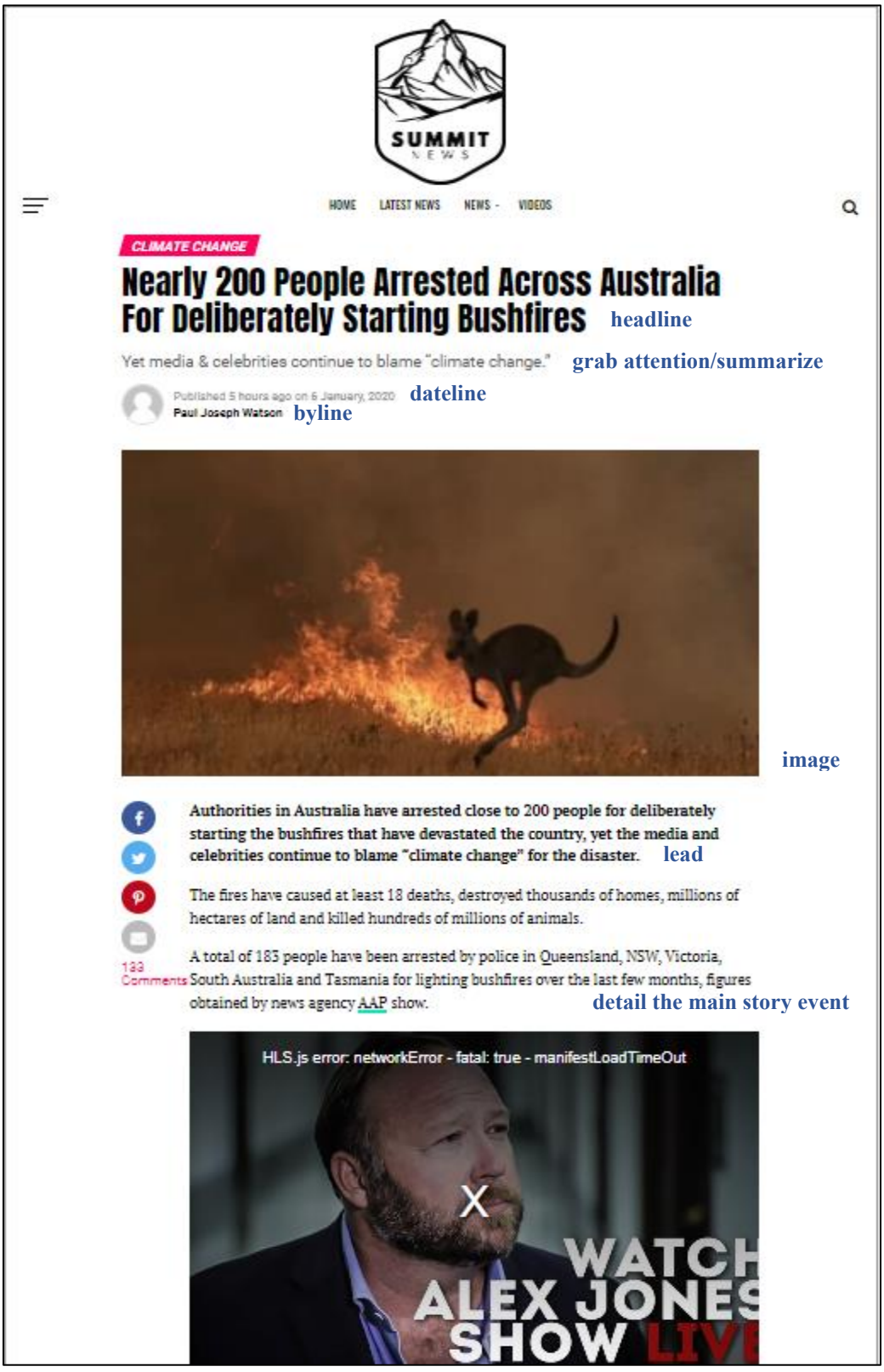
ClimateDepot.com

PolarPortal.dk

BBC.com

DailyMail.co.uk cite sources

Figure 4-7 An Example of the Moves Identified in a False News Article (A False 20)



Alex Jones is live on air right now breaking down how the carbon tax cult is seizing this tragedy to push for planetary government and taxation. He's also taking calls from Aussies!

refute mainstream theories

In New South Wales, 24 people were arrested for arson, risking prison sentences of up to 25 years.

In Queensland, police concluded that 103 of the fires had been deliberately lit, with 98 people, 67 of them juveniles, having been identified as the culprits.

detail the main story event

"The link between arsonists and the deadly fires that devastate Australia every summer is well known and well documented, with the rate of deliberately lit fires escalating rapidly during the school holiday period," reports Breitbart's Simon Kent.

Around 85 per cent of bushfires are caused by humans either deliberately or accidentally starting them, according to Dr Paul Read, co-director of the National Centre for Research in Bushfire and Arson.

"About 85 per cent are related to human activity, 13 per cent confirmed arson and 37 per cent suspected arson," he said. "The remainder are usually due to reckless fire lighting or even just children playing with fire."

Read also highlighted the link between school holidays and kids starting fires, commenting, "School holidays are a prime time for fire bugs, but especially over summer." The kids have got time to get out there and light, and the most dangerous adults choose hot days."

refute mainstream theories

"Police are now working on the premise arson is to blame for much of the devastation caused this bushfire season," reports 7 News Sydney.

New environmental policies that restrict "prescribed burning," where land owners burn off flammable ground cover in cooler months in a controlled manner so it doesn't contribute to bushfires, have also exacerbated the problem.

detail the main story event

The fact that the bushfires were deliberately started and have nothing to do with man-made climate change hasn't prevented that being the dominant narrative.

comment

The fact that the bushfires were deliberately started and have nothing to do with man-made climate change hasn't prevented that being the dominant narrative.

Last week, Bernie Sanders blamed those who were "delaying action on climate change" for "the blood-red sky and unbreathable air in Australia because of raging forest fires."

Virtue signaling celebrities have also pounced on the issue to push their dogma, including at the Golden Globes last night when Australian-born Cate Blanchett asserted, "When one country faces a climate disaster, we all face a climate disaster."

explain mainstream theories

SUBSCRIBE on YouTube:

Paul Joseph Watson
YouTube 1M
Follow on Twitter: Follow @PrisonPlanet

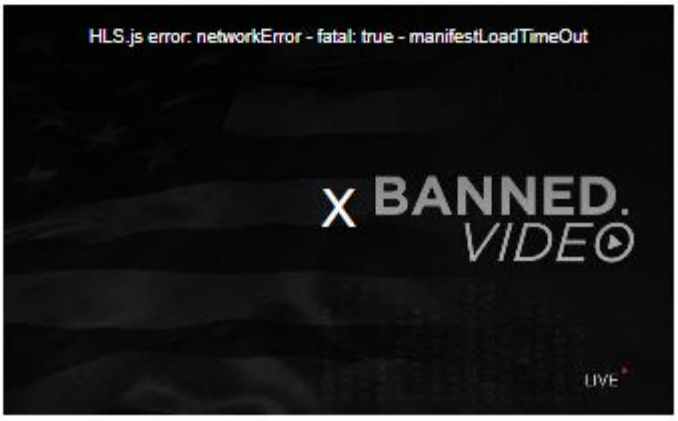
My voice is being silenced by free speech-hating Silicon Valley behemoths who want me disappeared forever. It is **CRUCIAL** that you support me.

Please sign up for the free newsletter here. Donate to me on SubscribeStar here.

Support my sponsor – Turbo Force – a supercharged boost of clean energy without the comedown.

promote engagement

RELATED TOPICS comment



advertise

4.6.2. Frequency of Moves' Occurrence

This section and subsequent sections present four analyses aimed at providing a comprehensive overview of the moves within the entire false news corpus. The focus will be on (a) move frequencies, (b) text sizes coded for each move, and (c) variations in move types as well as (d) variations in move structures across false news articles.

The debate regarding whether a move type should be classified as obligatory or optional varies among researchers, with some suggesting a 100% occurrence rate for obligatory status while others propose a threshold of 50%. In this case study, a threshold of 60% was set to identify obligatory moves. This percentage was selected because it represents the median value among the percentage frequencies of the moves' occurrence.

Figure 4-8 shows the percentage frequency of each move. As the bar graph shows, *headline* and *lead* occurred in all the articles in the corpus, and *byline*, *image*, *comment*, *dateline*, *refute mainstream theories*, and *promote engagement* occurred in most of the articles (ranging from 90% to 80%). Furthermore, the moves *advertise* (65%) and *grab attention/summarise* (60%) were notably frequent. Given their high prevalence, the above-mentioned moves can be considered obligatory, while the remaining moves are considered optional. The optional moves include *explain mainstream theories* (50%); *detail the main story event* and *give background* (35%); *note* (30%), *cite sources* (25%), and *describe consequences* (20%).

Figure 4-8 Frequency of Moves' Occurrence

4.6.3. Amount of Coded Text for Each Move and Strategy

Figure 4-9 shows how much text was coded at each move and its strategies. The size of the rectangles in the hierarchy chart indicates the amount of coded text for each move and strategy in relation to others. Larger areas are displayed towards the top left of the chart, and smaller rectangles are displayed towards the bottom right.

Figure 4-9 Amount of Coded Text for Each Move

As the chart shows, the greatest amount of text was dedicated to *refute mainstream theories* (24%), followed by *comment* (16%). The amount of text coded for the strategies of these two moves differed substantially, with *challenge mainstream theories* and *make a conclusion* having the largest amount of text. Other very frequent moves include *grab attention/summarise* (9%), *lead* (7%), *promote engagement* (7%), and *explain mainstream theories* (6%). A large proportion of the *lead* text was allotted to *summarise the refutation* and *state the main thesis*. This may be because the largest proportion of text as a whole was coded at *refute mainstream theories*.

One important observation from the hierarchy chart is that moves that are typical elements of hard news reports (*give background, detail the main story event, describe consequences*) constituted a very small portion of text (being 4%, 3%, and 2% respectively). The amount of text coded as *detail the main story event* and *describe consequences* is smaller than that coded for typically short elements like *headline* and *byline* (4%). Another important result is that *promote*

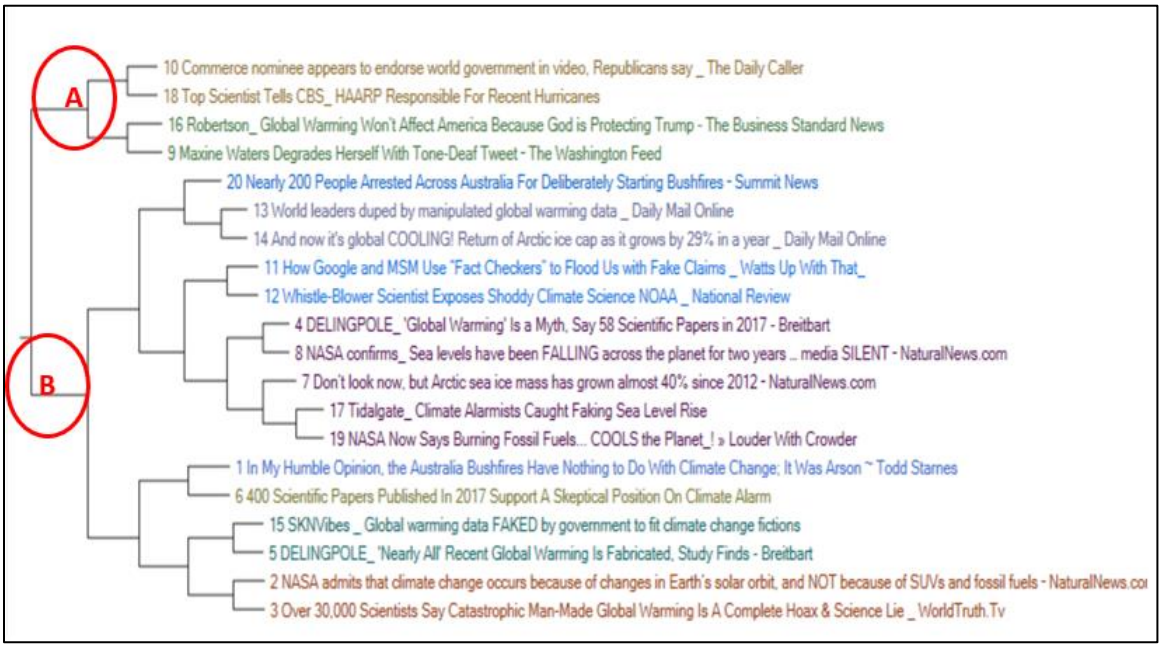
engagement (7%) and *advertise* (5%) constitute a larger amount of text than the aforementioned typical elements of hard news reports. This might be due to the online nature of news articles or the writers' high interest in disseminating their articles.

4.6.4. Variations in Move Types Across False News Articles

A cluster analysis was performed using Nvivo to group articles based on their move types. This analysis involved comparing each pair of articles using a similarity metric, specifically Jaccard's coefficient, which measures the percentage of shared moves between articles, using the complete linkage (farthest neighbour) hierarchical clustering algorithm. A value of 0 indicates the least similarity, while 1 indicates the highest similarity. The calculated similarity index was then used to cluster the articles into distinct groups (*NVivo 11 for Windows Help - How Are Cluster Analysis Diagrams Generated?*, n.d.). Figure 4-10 is a dendrogram presenting the cluster analysis results which offer insights into the similarities and differences among the analysed articles in terms of their move types. In this dendrogram, the articles were given a serial number and labelled by their titles.

As illustrated in Figure 4-10, the analysis identified two main groups: Cluster A and Cluster B. What sets Cluster B apart from Cluster A is the absence of *refute mainstream theories* in the latter, while it is present in the former. As moves refer to the communicative purposes of text segments, the absence of *refute mainstream theories* in some articles suggests that false news articles on climate change exhibit variations in their specific intentions. Cluster A articles (without *refute mainstream theories*) seem to serve other purposes, such as delegitimizing a political or religious figure.

Figure 4-10 A Cluster Analysis of the Articles Based on Their Move Types



The following are example extracts from articles from Cluster A. Example 24 is an extract from *article 10*, suggesting that Obama's commerce secretary nominee, John Bryson, supports the world government of climate change. The author of the article bases such speculation on the fact that Bryson endorses climate change mitigation actions. Therefore, this article aims to delegitimize a political figure.

> (24) Bryson's nomination has already run into problems stemming from his involvement in the founding of the Natural Resources Defense Council, a liberal environmental group, during the 1970s. Critics say Bryson's advocacy of cap and trade, among other environmental regulations, is at odds with the pro-business mission of the Commerce Department. *(A False 10).*

Example 25 is an extract from *article 16*, claiming that Pat Robertson, a climate change denier and a Trump supporter, says that global warming will not affect America because God is protecting Trump. This story seems to be satirical, criticizing Robertson and Evangelicals.

> (25) Televangelist Pat Robertson has released some predictions for the New Year. And he claims that the United States will be protected from the ravages of climate change because of President Donald Trump.

...

> Like Robertson, the majority of Evangelicals are still solidly behind the bumbling Trump, who was recently accused of running national security based on his gut. Eighty-one percent of Evangelicals backed Trump. *(A False 16)*

Example 26 is an extract from *article 18*. While this article mainly *details a main story event* and lacks *refute mainstream theories*, it does refute mainstream theories indirectly. It attributes natural disasters to government research programs.

> (26) World renowned physicist Dr. Michio Kaku made a shocking confession on live TV when he admitted that HAARP is responsible for the recent spate of hurricanes.
>
> In an interview aired by CBS, Dr. Kaku admitted that recent 'man-made' hurricanes have been the result of a government weather modification program in which the skies were sprayed with nano particles and storms then "activated" through the use of "lasers". *(A False 18)*

4.6.5. Variation in Move Structures Across False News Articles

To analyse variation among the false news articles in terms of move structure, each move type was assigned a letter code to analyse the move sequence of each news article. Additionally, the sequences of moves were categorized into introductory moves (including headline, byline, dateline, grab attention/summarize, image, and lead) and body moves (the remaining moves) to simplify the tracking process. In cases where two moves were placed horizontally adjacent to each other, they were connected with a dash to indicate their adjacent position. This helps reduce each news article into a code string. For example, the move sequence of the news article illustrated in Figure 4-6 is as follows.

- Sequence of Introductory Moves: HDWGIL:

Headline, Dateline, Byline, Gran attention/summarize, Image, and Lead

- Sequence of Body Moves: RBCREBCS:

Refute mainstream theories, Give background, Comment, Refute mainstream theories, Explain Mainstream theories, Give background, Comment, and Cite sources

Table 4.3 presents the move sequences of the 20 false news articles. The table reveals considerable variation in the move structure of the false news articles. This indicates that there is no specific or typical move sequence that can be regarded as representative of online false news articles. However, two identical patterns of introductory move sequences can be identified, as highlighted in the table. Interestingly, the articles sharing this same pattern originated from the same news source. This suggests that certain news sources may have a preference for starting their articles in a specific manner.

Table 4.3 Move Sequences in False News Articles on Climate Change

Article No.	**Coding Strings**	
	Introductory Moves	**Body Moves**
1	HDIAL	RARAPA
2	HDWGPIL_A	RCRARCGRCSP
3	HIAL	RARGRGCSA
4	HIWDL_A	ERCN
5	HIWDL	RCRSN
6	HWDGIL	RSP
7	HDWGPIL_A	RBCREBCS
8	HDWGIL_A	RCGERCEGRCRP
9	HIL	MCAMCWPA
10	HIPWDL	MBCMCACKRBCP
11	HWDL	ECERGRCGCRGCGCPC
12	HWD_PIGLA	ERAECKRCGRCA
13	HGWDPL_IG	REKRCP_CEGEGRK_GBR_GEKRERGCEKMRKR_GEGEGRKCRGE GCKCRCR_GRGB_EGCRCP_CPA
14	HGWDPL	RM_GRBKCRGBMGRC_PERGCRCRPA
15	DHIWL	RGRCRGRCGRGRGCGCN
16	DHPIGWL	MCMP
17	HIWDPL	RCRCRCBERCN
18	HDWPIPL	MGMGAGCN_PA

19	HPWDIL	ERERCBCP
20	HGDWIL	MRMRMCEPCA

In addition, as highlighted in Table 4.3, it is apparent that *refute mainstream theories* is often followed by *comment*, indicating a common sequence. Example 27 presents a conclusion made by the author that asserts the presence of widespread scientific fraud aimed at manipulating public understanding of global warming. This conclusion is derived from the preceding evidence, which includes instances of deliberate manipulation of weather data to amplify the perception of global warming. The two moves in this example are denoted by subscripts.

> (27) Using the exact same data found in the chart shown above (with a few years of additional data after 2000), NASA managed to misleadingly distort the chart to depict the appearance of global warming:
> …
> This new, altered chart shows that historical data -- especially the severe heat and droughts experienced in the 1930's -- are now systematically suppressed to make them appear cooler than they really were. At the same time, temperature data from the 1970's to 2010 are strongly exaggerated to make them appear warmer than they really were. [refute mainstream theories]
> This is a clear case of scientific fraud being carried out on a grand scale in order to deceive the entire world about global warming. [comment] *(A False 15)*

4.7. Discussion

This section discusses the findings of the move analysis conducted on the online false news articles on climate change and examines these results in relation to previous discourse theories regarding the structure of conventional hard news reports. The section concludes by addressing the limitations of this case study and offering suggestions for future research directions.

4.7.1. The Discourse Structural Characteristics of Online False News Articles

Based on the previous theories on the genre of printed hard news reports and the results of this case study, false news reports share some structural elements with typical hard news reports but, at the same time, have additional distinct elements. Like typical hard news reports, false news

reports use *headline*, *grab attention/summarise*, *byline*, *image*, *lead*, *detail the main story event*, *give background*, *describe consequences*, and *comment*. This finding is consistent with previous research indicating that fake news often mimics the format of legitimate news (Lazer et al., 2018; Tandoc et al., 2018). This imitation makes it difficult for readers to recognize that the news is fake, leading to its wide dissemination (Dentith, 2016; Rini, 2017). However, some specific strategies for achieving the above-mentioned moves have not been attested in hard news reports, e.g., *state a main thesis*, *report mainstream theories*, and *summarise the refutation* in *lead,* and *restate the main thesis*, *make a conclusion*, *recommend*, and *call to action* in *comment*. Moreover, there are five distinct moves: *dateline*, *cite sources*, *advertise*, *note*, and *promote engagement* which can be attributed to the analysed articles' online publication.

Prior research has indicated that traditional hard news reports typically serve an expository function, aiming to inform, describe, or explain events (Ricketson, 2014). Additionally, they often incorporate narrative elements, presenting a sequence of events (Bell, 2005; Richardson, 2006). However, the MA conducted here reveals that false news articles, while containing expository and narrative components, lean more towards being argumentative in nature. False news articles are characterized by two primary moves: explaining mainstream theories and refuting them.

In this respect, false news articles share similarities with argumentative essays, as they express a central thesis, present opposing perspectives, refute them, draw conclusions, and provide references (refer to Table 4.1 displaying the MA of argumentative essays by Hyland, 1990). The analysis of move frequencies and text allocations to each move indicates that false news reports primarily focus on debunking mainstream theories, as opposed to elaborating on the main story event. The analysis showed that *refute mainstream theories* is an obligatory element of false news articles and *detail the main story event* is optional. This result contradicts previous theories about the structural elements of typical hard news reports which mainly focus on elaborating on an episode or episodes.

The combination of narrative, expository, and argumentative genres likely contributes to the appeal of false news, as it achieves its argumentative purpose while adhering to journalistic conventions. This blending of genre has also been highlighted by the content analysis conducted

by Mourão & Robertson (2019) who found that fake news incorporates elements from various genres, including news, opinion, and entertainment.

In addition to the purpose of refuting mainstream theories, the cluster analysis shows that some of the articles were characterized by focusing on delegitimizing political or religious figures. This finding indicates the political context and partisan nature of false news articles on climate change, as politicians play a significant role in the climate change debate (Cook et al., 2017). This political partisanship not only motivates the writing of these articles but also might influence belief in them, as indicated by psychology research conducted by Bronstein et al. (2019) and Pennycook and Rand (2021). In addition, this partisan nature of false news might contribute to its viral dissemination, as noted by Osmundsen et al. (2021) and Vosoughi et al. (2018).

The move sequence analysis of the false news articles suggests that false news reports do not follow a fixed structure. This result contrasts with previous research on the genre of typical hard news articles, which has found that such articles tend to follow a pyramid structure, first reporting the main story in broad terms, highlighting the most important information, before moving on to the details (Bednarek & Caple, 2012; Richardson, 2006). In the case of false news articles, this pattern is not consistently observed. Instead, false news articles exhibit a greater degree of variability in their move structure.

This finding may indicate that online false news articles are not only distinct from the conventional hard news genre but do not even constitute a coherent genre themselves. This finding may also indicate that the motivation for writing false news is relatively inconsistent, leading to greater variability in how these articles are written. In other words, the communicative motivations for false news creation may encourage greater flexibility in how author shape their narratives, and how certain elements of false news are emphasised to achieve the desired impact.

In addition, the results revealed that some moves were used repeatedly and occurred in different places within one article. This suggests a lower level of coherence in false news articles compared to traditional hard news reports, which aligns with Karimi and Tang's (2019) findings regarding the structure of false news articles. The observed incoherence in false news articles can therefore also be attributed to the lack of factual accuracy, reflecting the writers' deliberate intent

to manipulate readers by employing a range of strategies (e.g., promoting alternative theories, providing background information, describing consequences, issuing calls to action, etc.) repeatedly within a single article to support their main thesis. Additionally, this incoherence can be attributed to the rapid production of false news articles, as writers strive to meet the demands of the fast-paced and viral nature of online media platforms. Furthermore, the fact that false news writers are generally not professional journalists may also contribute to the lack of coherence in their articles.

4.7.2. Limitations and Future Research Directions

While the MA has provided valuable insights into online false news articles, it is important to recognize that this study has certain limitations. First. this case study did not involve a comparison between the move structure of false news articles and that of true news. To fully understand the distinctive discourse structure of false news, further research is needed to conduct a parallel comparative analysis between the two. Additionally, it is important to note that this case study specifically focused on articles related to climate change. As a result, the generalizability of the findings may be limited. To address this limitation, future investigations should encompass articles covering a range of diverse topics to enhance our understanding of false news discourse.

4.8. Conclusion

This study employed genre analysis using rhetorical moves to investigate the discourse structure of online false news articles on climate change. The research focused on determining the internal structure of these articles, exploring the motives behind creating false news, analysing the rhetorical strategies employed to convey information, and examining the variations in discourse structure across the false news articles.

The analysis yielded three significant findings. Firstly, false news articles on climate change primarily serve two purposes: challenging mainstream theories and delegitimizing political and religious figures. Secondly, while the structure of false news articles shares some common elements with typical hard news reports, it aligns more closely with argumentative essays. In other words, they display elements of narrative and expository genres by detailing main story events, giving background, and describing consequences. Simultaneously, they embrace an

argumentative genre by expressing and restating a main thesis, reporting and refuting mainstream theories, drawing conclusions, offering recommendations, and citing references. Lastly, the examined articles displayed a lack of fixed move structure, with many moves occurring in various positions within an article and occasionally appearing more than once. The MA not only revealed the structural features of false news articles but also provided insights into their underlying rationale, political contexts, and the strategies employed by false news writers to serve their objectives.

Chapter 5: News Values of Online False News Articles

5.1. Introduction

This chapter presents the second case study which aims to identify the news values constructed in online false news articles, by comparing them with four types of news articles: broadsheet, tabloid, web-based publication, and blog obtained from reputable sources. Research in communication and journalism (e.g., Harcup & O'Neill, 2001) and linguistics (e.g., Bell, 1991; van Dijk, 1988) has identified a variety of factors, known as news values, which are thought to guide the selection, production, and consumption of news. These values are believed to be shaped by material, cognitive, and social factors, and are reflected in news discourse (Bednarek & Caple, 2017). News values vary according to the news outlet; however, they typically include factors such as *timeliness* (the news event is new, recent, ongoing, about to happen, or otherwise relevant to the immediate situation or time), *impact* (the news event is having significant effects or consequences), *proximity* (the news event is geographically or culturally close to the target audience), and *unexpectedness* (the news event is unexpected, unusual, or strange) (Bednarek & Caple, 2017).

The news values underpinning false news stories have not so far been explored. However, previous work on false news discourse has unveiled a number of recurring themes that may point to salient news values. For example, Golbeck et al. (2018) identified seven major themes in fake news: (a) a hyperbolic position against one person or group (for example, Trump, Islam or refugees), (b) a hyperbolic position in favour of one person or group, (c) discrediting a normally credible source, (d) sensationalist crimes and violence, (e) racist messaging, (f) paranormal theories (such as aliens or Flat Earth theories), and (g) conspiracy theories. These themes indicate the stories that are usually deemed newsworthy for false news, but further research is needed to understand why they are newsworthy, as well as the values that make them so.

There are several reasons why studying the news values in online false news discourse is important. Firstly, it allows for the investigation of whether false news constructs news values that are distinct from or similar to those found in reliable news sources. This exploration helps in

understanding the similarities or disparities between false news discourse and true news. Secondly, such an investigation can help explain why false news becomes appealing and goes viral. Richardson (2006) indicates that news values represent the "imagined preferences of readers" (p. 94), so examining these values can provide a better understanding of the factors that drive engagement with false news. Thirdly, the examination of news values offers insights into the ideologies conveyed within the discourse of false news. Ideologies are sets of beliefs, ideas, values, and principles that shape how individuals perceive reality. According to Bednarek and Caple (2017) repeated associations of news stories with specific news values can potentially have ideological implications. In fact, news values can reflect a wide range of social and political ideologies, going beyond the conventional notions of newsworthiness attributed to events (Bednarek & Caple, 2017; Bell, 1991; Cotter, 2010; Fowler, 1991; Montgomery, 2007).

Previous discourse research on news values (e.g., Bednarek & Caple, 2017; van Dijk, 1988) has theorised about what makes an event newsworthy and has provided lists of defined news values. In some discourse research, corpus linguistics tools have been utilized (e.g., Bednarek & Caple, 2017; Chen & Liu, 2023; Fruttaldo & Venuti, 2017) to ascertain how these predefined news values are discursively constructed. However, this study takes a different approach. Instead of simply confirming the presence or absence of predetermined news values, it will utilize the corpus method of key semantic domains analysis (KSDA) to identify potential news values. By using KSDA, this case study will identify dominant meanings and patterns of language usage to determine whether such patterns can construct news values similar to or different from those identified by earlier research on news discourse.

The chapter is organized as follows. Firstly, it provides an exploration of the concept of news values by incorporating insights from journalism and linguistics studies. Next, it outlines the corpus utilized in this study, and then it introduces the novel approach employed to analyse the news values in online false news articles. Following this, the chapter details the analytical procedure that was implemented. Moving forward, the chapter presents the findings obtained from the analysis. Finally, the chapter concludes by offering a discussion on the results, addressing the limitations encountered, and providing recommendations for future research.

5.2. Previous Work on News Values

This section reviews previous work on news values. The first subsection provides a concise overview of journalism studies that have addressed news values. This is followed by a review of linguistic studies in the second subsection, which have examined this topic from a linguistic standpoint. The subsequent section offers a review of Bednarek and Caple's (2017) framework of Discursive News Values Analysis (DNVA). The DNVA framework is considered the most recently published and influential work on the discourse of news values and will be referenced in the analysis. Finally, the section concludes by discussing the application of corpus methods in facilitating the discursive news values analysis.

5.2.1. Conceptualisation of News Values in Journalism Studies

Galtung and Ruge's (1965) work is widely acknowledged as being the starting point of research on news values in journalism studies. Their conceptualisation of news values, referred to as 'news factors' in their paper, has become a major theory in journalism and communication studies. Through a content analysis of the coverage of three international crises (Congo, Cuba, and Cyprus) in selected Norwegian newspapers, the authors were able to support their hypotheses regarding news values via common sense psychology, that is, based on observations and experience. They identified twelve key values that made an event worthy of being reported. Examples of these values are *relevance, unexpectedness, reference to elite people,* and *reference to something negative*.

A critical perspective on Galtung and Ruge's (1965) work has emerged within communication and journalism studies, with some researchers using it as a foundation to create contemporary lists of news values (e.g. Harcup & O'Neill, 2001). These lists often share similarities (Caple & Bednarek, 2013), with a common perspective that news values are inherent qualities of news events themselves, as well as strategies used in news processing and production. In other words, news values serve as criteria for journalists to determine what should be investigated or reported and how it should be presented (Bednarek & Caple, 2014, 2017). According to researchers in journalism, news values have the power to determine whether a story is pursued, published, and consumed (Conley & Lamble, 2006). However, researchers in journalism perceive news values as "pre- and a-textual," meaning that they influence the news content but are not explicitly

negotiated within (Bednarek & Caple, 2017, p. 35). This view contradicts the perspective of some linguists who study news discourse.

5.2.2. Conceptualisation of News Values in Linguistic Studies

Early linguistic studies have focused on conceptualising news values from a linguistic perspective; however, these studies do not provide in-depth analyses of how these news values are constructed in news discourse. A brief review of linguists' conceptualisations of news values is presented below. The complete list of news values and their definitions proposed by linguists are far too numerous to be discussed here; however, they are covered by Bednarek and Caple's (2017) framework, which is presented in the next subsection (see Caple & Bednarek, 2013, for a comprehensive review of definitions of news values that have been discussed in both journalism and linguistics studies).

Van Dijk (1988), Fowler (1991), Bell (1991), Richardson (2006), Montgomery (2007) and Cotter (2010) have made significant contributions to the definition of news values. These definitions generally agree that news values are criteria that journalists use consciously or subconsciously to determine the newsworthiness of an event, to select and represent news input, and to prioritise the collection and production of news. As Montgomery (2007) explained, news values are seen as "principles of selection, paradigms of relevance, and frames of including and excluding materials" (p. 5).

Linguists have proposed various lists of news values, most of which overlap. For example, Van Dijk (1988) identified three different types of news values: The first type is related to economic production, such as *budgets* and *sales/subscriptions*. The second type pertains to the gathering of news, such as *deadlines, sections*, and *accessibility of sources*. The third type is related to cognitive constraints, such as *novelty, recency, presupposition, consonance, relevance, deviance and negativity,* and *proximity*.

Similarly, Bell (1991) classified news values according to three groups. The first is related to news content, including *negativity, recency, proximity, consonance, unexpectedness, superlativeness, relevance, personalisation, eliteness, the quality of attribution*, and *facticity*. The second group pertains to news gathering and processing, including *continuity, co-option,*

competition, composition, predictability, and *prefabrication*. The third group was related to the quality or style of the news text, including *clarity, brevity*, and *colour*.

Early linguists have suggested that news values have a significant effect on news processing, production, and perception. Van Dijk (1988) argued that news values shaped the selection, attention, understanding, representation, recall and use of news information, and reflected the economic, social, and ideological values of a news community. Bell (1991) contended that news values affect the way in which news stories are gathered, structured, and presented, thus leading to a particular framing of events. He further argued that news values are not impartial but represent the ideologies and priorities of the news community. Montgomery (2007) highlighted that news values are ideologically laden, which contributed to certain versions of reality being included while others are excluded. Consequently, as Montgomery (2007) stated, news values must be viewed as an inescapable part of both the news production process and the wider social processes. Cotter (2010) asserted that news values act as guidelines "to establish a hierarchy of importance" in most of the steps in news processing and production (p. 85). These values guide every step, from story conceptualization and construction to positioning in terms of time and space, as well as evaluation. He further suggested that news values are reflected in news texts, thereby emphasising the ideology behind news selection and shaping the structure of the news text itself. Although linguists have highlighted the influence of news values on the structure and presentation of news texts, they have yet to provide clear explanations and examples illustrating how news values accomplish this.

The epistemological status of news values, as Montgomery (2007) claimed, is somewhat unknown. In their seminal work, Galtung and Ruge (1965) analysed the content of Norwegian newspapers' coverage of three international crises to derive their understanding of news values and find evidence supporting their hypotheses. Consequently, it is plausible that the news values they identified may be influenced by the specific type of coverage they examined. In fact, news values are shaped by social (Fowler, 1991; Van Dijk, 1988) and journalistic conventions (Fowler, 1991). In addition, news values are constrained by the imagined audience, as journalists often need to guess what the audience will prefer and select news stories accordingly (Richardson, 2006). Moreover, news values can evolve according to the changing preference of readers (Richardson, 2006). Thus, news values are situationally sensitive and indexical; that is, they must

be interpreted according to the context in which they are applied (Montgomery, 2007). Cotter (2010) argued that news values can be interpreted differently across publications and editors, over time, and depending on whether they are viewed from the inside (by a reporter) or from the outside (by a reader or analyst).

5.2.3. The Discursive News Values Analysis Framework

A series of studies led by Bednarek and Caple (Bednarek, 2016; Bednarek & Caple, 2014, 2017; Caple & Bednarek, 2013; Potts et al., 2015) examined news values from a discourse analysis perspective and provided empirical analyses of news texts. According to Bednarek and Caple, (2017), news values refer to "the newsworthiness of events – their potential newsworthiness in a given community, as evaluated and determined by news workers, or as constructed through discourse" (p. 42). Bednarek and Caple differentiated news values from news writing objectives (such as clarity of expression, brevity, colour, and accuracy) and news selection factors (such as commercial pressures, availability of reporters, deadlines, and audience analytics). In addition, Bednarek and Caple regarded moral-ethical values (such as truth and fairness) and commercial values (such as speed and access via multiple platforms) as not being within the scope of news values.

Bednarek and Caple (2017) argued that news values have four dimensions:

a) Material: an event has potential news value in its material reality,
b) Cognitive: news workers and audience members have beliefs about news values and newsworthiness,
c) Social: news values are applied as selection and production criteria in journalistic routines and practices, and
d) Discursive: news values can be established or negotiated through discourse.

Bednarek and Caple argue that these various aspects of news values interact in different ways. For instance, events with potential news value occur (material), and news workers use their beliefs about news values, as well as those of their target audience, to determine which events to include in the news, how to order them, and how to present them (cognitive, social). Crucially, news workers utilize language to construct these news values within the actual news texts (discursive). Bednarek and Caple's focus was on how an event is constructed discursively as

news rather than on how and why it is selected as news. To explore the latter issue, Bednarek and Caple argued that an ethnographically informed study of the social, cognitive, and material aspects of news values should be conducted.

The work of Bednarek and Caple (2017) has been significant in creating the Discursive News Values Analysis (DNVA) framework. This framework includes a concise list of news values, refined definitions of news values that have been proposed in the literature, and an inventory of linguistic resources for systematic analyses. To reduce the amount of overlap between the different lists of news values that have been proposed in the literature, Bednarek and Caple combined related concepts under one value of news. For example, the concepts of *newness, recency, immediacy, imminence, currency*, and *seasonality* were included in the news value of *timeliness*.

Table 5.1 (adapted from Bednarek and Caple, 2017, and from Caple and Bednarek, 2013) illustrates the DNVA framework, including news values, their definitions, linguistic resources, and the related concepts that have been discussed in previous studies. As Bednarek and Caple (2017) explained, the linguistic resources have been allocated to a specific news value based on their meaning potential, typical usage, and function. Note that the term 'event' in the provided definitions of news values was used by Bednarek and Caple as a cover term for events, issues, and happenings, including their elements or aspects, as well as locations and news actors.

Table 5.1 Bednarek and Caple's DNVA Framework

News Values[7]	**Definition**	**Linguistic Resources**	**Related Concepts**
Consonance	The event is discursively constructed as (stereo)typical (limited here to news actors, social groups, organizations, or countries/nations)	References to stereotypical attributes or preconceptions, assessments of expectedness/typicality, similarity with past, explicit references to general knowledge/traditions, and so on	*stereotypes, expectation*
Eliteness	The event is discursively constructed as of high	Various status markers, including role labels, status indicating adjectives,	*elite-centred, elites, attribution, the*

[7] The news value of aesthetical appeal is not included in the review here because it refers to how an event is discursively constructed as being visually beautiful through images, and this is not the focus of the present study.

	status or fame (including but not limited to the people, countries, or institutions involved)	recognized names, descriptions of achievement/fame, use by news actors/sources of specialized/technical terminology, high-status accent, or sociolect (esp. in broadcast news)	*power elite, celebrity, prominence, worth, power*
Impact	The event is discursively constructed as having significant effects or consequences (not necessarily limited to impact on the target audience)	Assessments of significance, representation of actual or nonactual significant/relevant consequences, including abstract, material, or mental effects	*importance, relevance, social impact, consequence, significance, interest*
Negativity	The event is discursively constructed as negative, for example, as a disaster, conflict, controversy, or criminal act	References to negative emotion and attitude, negative evaluative language, negative lexis, descriptions of negative behaviour	*conflict, deviance and negativity, drama, bad news,*
Personalization	The event is discursively constructed as having a personal or 'human' face (involving non-elite actors, including eyewitnesses)	References to ordinary people, their emotions, experiences; use by news actors/sources of everyday spoken language, accent, sociolect (especially in broadcast news)	*personification, personalities, familiarity, human interest*
Positivity	The event is discursively constructed as positive, for example, as a scientific breakthrough or heroic act	References to positive emotion and attitude, positive evaluative language, positive lexis, descriptions of positive behaviour	*valence: success, good news*
Proximity	The event is discursively constructed as geographically or culturally near in relation to the publication location/target audience	Explicit references to place or nationality near the target community, references to the nation/community via deictics, generic place references, inclusive first-person plural pronouns, use by news actors/sources of (geographical) accent/dialect, cultural references	*identification, meaningfulness, ethnocentrism, cultural relevance*

Superlativeness	The event is discursively constructed as being of high intensity or large scope/scale	Intensifiers, quantifiers, intensified lexis, metaphor and simile, comparison, repetition, lexis of growth, only/just/alone/already + time/distance or related lexis	*threshold, size, magnitude, scale/scope, intensity, discontinuity*
Timeliness	The event is discursively constructed as timely in relation to the publication date: as new, recent, ongoing, about to happen, or otherwise relevant to the immediate situation/time (i.e., current or seasonal)	Temporal references, present and present perfect, implicit time references through lexis, reference to current trends, seasonality, change/newness	*recency, currency, topicality*
Unexpectedness	The event is discursively constructed as unexpected, for example, as unusual, strange, rare	Evaluations of unexpectedness, references to surprise/expectations, comparisons that indicate unusuality, references to unusual happenings	*novelty, dynamics, deviance, surprise, novelty/rarity/oddity, unusualness*

Bednarek and Caple (2017) emphasised that their inventory of linguistic resources should not be considered a closed and automatic checklist because there is no one-to-one relationship between language and news values. This is primarily due to two reasons. Firstly, the same linguistic resource can be used to construe different news values. Secondly, the same linguistic resource can construct several news values simultaneously. This means that it is necessary to pay close attention to the meaning potential of the linguistic resource as it is used in the news story, as well as for the intended audience. Bednarek and Caple concluded that this is not an exact science, as multiple interpretations are possible.

Bednarek and Caple (2017) also emphasised that news values are contingent upon the *target audience* and the *preferred reading*. The concept of the *target audience* refers to the role of the audience. Bell (1991) outlined the distinctions among the target or the addressed audience: (a) the target audience (who is directly addressed), (b) the auditors (expected but not targeted), (c) the overhearers (not expected to be present in the audience), and (d) the eavesdroppers (expected to be absent from the audience). News writers can identify their target audience by making assumptions about imaginary readers or by using market analytics. The concept of *preferred*

reading pertains to how readers are positioned in relation to the content of the news and is influenced by underlying ideologies. It indicates the attitude or perspective that a text encourages the reader to adopt (Martin & White, 2005). Therefore, in order to analyse news values, one must make an educated guess about the *target audience* and identify the *preferred reading* (Bednarek & Caple, 2017).

5.2.4. Utilizing Corpus Methods in Discursive News Values Analysis

Bednarek and Caple (2017) and Potts et al. (2015) explored the role of corpus linguistics (CL) in DNVA. Their research aims to understand how language use patterns contribute to the representation of news values in news discourse. They explore various corpus methods such as frequency lists, keywords, ranges, collocations, semantic tags, word sketches, concordances, and search terms, to analyse these language patterns. According to the authors, these corpus methods are valuable tools that can aid researchers in examining the construction of news values through figurative devices, rhetorical strategies, and phraseologies.

Previous corpus-assisted discourse studies on news values (e.g., Bednarek & Caple, 2017; Chen & Liu, 2023; Fruttaldo & Venuti, 2017) have followed a deductive approach whereby they used CL to identify how a preset list of news values is constructed in news texts. These studies aimed to investigate the discursive construction of established news values and how they align with specific ideologies. However, these studies did not aim to identify news values beyond those attested in previous work. For example, Chen and Liu (2023) used the corpus method of keywords and the DNVA framework to compare the discursive construction of news values in news reporting in *China Daily* and CNN regarding the COVID-19 vaccine with the aim of exploring how the two media presented the vaccine and improved the international community's acceptance of the vaccine. One of the major findings was that most of the keywords in *China Daily* focused on international concepts, such as 'cooperation', 'foreign', 'continent', 'international' and 'Cambodia'. The collocations of these keywords demonstrated China's international contribution to providing COVID-19 vaccines by highlighting the professionalism of China in COVID-19 vaccine research, thus also highlighting the news value of 'eliteness'.

The literature review demonstrates the complexity of news values, which are influenced by social, cognitive, and material factors, and are reflected in news discourse. Notably, the reviewed

journalism and linguistic studies often provide fixed lists of news values that are presumed to be universally used in news production. Even researchers like Bednarek and Caple (2017), who argue for the discursive construction of news values, suggested utilizing CL methods to identify prominent linguistic features in a corpus and align them with a predefined list of news values. Nevertheless, as Bell (1991) stated, "news values can be identified through a variety of means such as analysing textbooks which teach news skills or by deduction from *what actually appears in the media* [emphasis added]" (p. 155). Therefore, this case study aims to identify the potential news values in online false news discourse by adopting an inductive approach, as explained in more detail below.

5.3. Corpus

In this case study, the analysis was conducted using the entire focus corpus, consisting of 137 false news articles, as well as the entire reference corpus. The reference corpus comprised articles sourced from reputable outlets, including broadsheets, tabloids, web-based publications, and blogs, and encompassed a total of 548 articles. For more details regarding the two corpora, please refer to Chapter 3.

5.4. Method

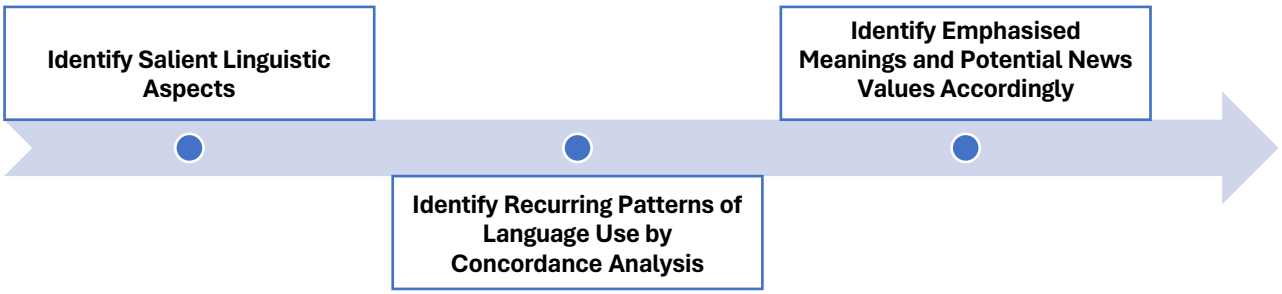
This section provides a detailed explanation of the novel methodological approach employed in this study. It begins by outlining the major steps of this approach, followed by a justification for why this approach is necessary when analysing news values in online false news articles. The section then discusses the pilot corpus analyses that were conducted and explores their limitations. Finally, the chosen corpus method, key semantic domain analysis, is described, along with an explanation of the rationale behind its selection.

5.4.1. An Inductive Corpus-assisted Approach to News Values Analysis

In this study, I argue that *potential* news values can be derived *inductively* through CL analysis. Instead of relying solely on established theories or predefined lists of news values, CL tools are utilized to identify potential news values. The main objective is to uncover values that may have not been attested in previous theories.

To achieve this objective, I propose the following steps. Firstly, a comparison is made between the corpus of false news and the reference corpus, utilizing corpus methods. The purpose of this comparison is to identify significant linguistic aspects that are specifically characteristic of false news. Subsequently, a concordance analysis is conducted to explore these noteworthy aspects and identify recurring patterns of language use of these aspects. These patterns are then qualitatively examined to ascertain if they highlight specific meanings that have the potential to represent news values. A visual representation of this analysis process is provided in Figure 5-1.

Figure 5-1 An Inductive Approach to News Values Analysis



5.4.2. The Rationale for the Inductive Approach

The inductive approach to identifying false news's potential news values is necessary for several reasons. Firstly, as stated previously, news values have four dimensions, which are material, cognitive, social, and discursive (Bednarek & Caple, 2017). Thus, news values are affected by the news makers, the audiences, the news event itself, the time and place of news publications, and the purpose of publishing news. Since these dimensions are not static, it is possible that news values are in a state of flux, with new values emerging while existing values may evolve or become less important. As Richardson (2006) pointed out, news values can evolve and change over time to satisfy the evolving preferences of the anticipated audience.

Secondly, false news has situational characteristics, including purposes, writers, audiences, modes, and media used for communication, which are different from those of true news, as discussed in Chapter 2. In fact, research has demonstrated that fake news articles tend to resonate with their target audience by aligning with their emotions, biases, beliefs, and interests (Baptista & Gradim, 2020). Moreover, Bednarek and Caple (2017) stressed the importance of considering the *target audience* and the *preferred reading* when analysing news values. Accordingly, it is

possible that false news has distinct news values that make an event newsworthy for its specific target audience.

Thirdly, Bednarek and Caple (2017) warned against treating their inventory of linguistic resources for news values as an automatic checklist. Additionally, they emphasised the need for careful consideration of the context, usage, and potential meanings of linguistic resources when analysing the discursive construction of news values. Consequently, it is possible that linguistic resources can be used to highlight novel meanings, and thus, novel news values that have not been identified in prior research.

5.4.3. Exploring News Values Using Keyness Methods: Pilot Analyses

As mentioned earlier, comparative corpus analysis is needed to identify linguistic aspects or patterns that are characteristic of false news compared to other news types. The corpus-assisted method known as keyness can be particularly valuable for this task for two reasons. Firstly, keyness allows for a comparison between two corpora using statistical metrics that measure the size and/or significance of the difference between the frequency of a particular linguistic feature or aspect in a focus corpus and that in a reference corpus (Gabrielatos, 2018). Secondly, keyness is mainly used to explore *aboutness* (Baker, 2004; Gabrielatos, 2018), which encompasses not only the topics or themes being discussed but also the various aspects of negotiated or emphasised meanings. In fact, these aspects can be associated with news values.

Different methods of keyness can be used to compare two corpora on various levels, including words (keywords, Scott, 1997; key keywords, Scott, 1998; and text-dispersion keyness, Egbert & Biber, 2019), lexico-grammatical features (key features analysis, Biber & Egbert, 2018) and meanings (KSDA, Rayson, 2008). In the pilot analyses conducted for this study, a combination of keyness methods, including keywords, text-dispersion keyness, and KSDA, were utilized to determine the most suitable approach for conducting a comprehensive analysis of news values. However, it is important to acknowledge that keywords and text-dispersion keyness methods demonstrated some limitations, which are discussed below.

A keyword list was generated using AntConc (Anthony, 2022) by comparing the focus corpus to the reference corpus. The log-likelihood (4-term) was used as a significant measure of keyness,

and the dice coefficient, an effect-size measure, was used to rank the generated list. The generated list contained 259 words with $p < 0.05$ (3.48 with Bonferroni; see Appendix 5.1).

Considering the dispersion of these keywords, it was found that many keywords had text-dispersion of fewer than five texts. It seems that they were extracted as keywords because they had a high frequency in a limited number of texts. For example, the word 'HCQ' were repeated 25 times in one article. This limitation was also highlighted by Baker (2004) and Egbert and Biber (2019). Words that have low text-dispersion could be due to small corpus size. They could also be due to the author's style or the topic under discussion. Such words cannot be considered distinctive of or generalisable to the content of a corpus (Baker, 2004; Egbert & Biber, 2019). To address this issue, it has been suggested that a threshold for the number of texts should be established in order for a word to be considered a keyword. Nevertheless, there is no consensus regarding the optimal threshold for text-dispersion that is appropriate for all corpora of different sizes (Baker, 2004; Egbert & Biber, 2019). In addition, this is not practical when the majority of keywords occur in a small number of texts. In the pilot analysis, 40% of the generated keywords occurred in fewer than five texts; in fact, 40% of these words occurred in only one text.

Examining the concordances of some keywords revealed another limitation: some keywords often co-occurred, forming one term. This reduced the number of keywords, as they were not key by themselves. For instance, the words 'biological' and 'warfare', 'Bill' and 'Gates', 'dr' and 'Bates', and 'Qui' and 'Xiangguo' were all seen to occur together.

Egbert and Biber (2019) proposed a method to measure keyness by comparing word use between two corpora based on the number of texts in which a word occurs rather than its frequency. They used log-likelihood to measure text-dispersion keyness, contending that words that appear in a large number of texts are likely to have at least a moderate frequency. Additionally, the authors argued that texts are self-contained units of discourse and thus are linguistically valid to be considered as a basis for extracting keywords. Furthermore, Egbert and Biber claimed that their method has the advantage over the traditional method of keywords in that it excludes words that lack distinctiveness and generalisability to the content of the texts in the focus corpus.

AntConc (Anthony, 2022) was used to generate a keyword list based on Egbert and Biber's measure of text-dispersion keyness. However, I found the generated list (refer to Appendix 5.2)

to be surprisingly short, consisting of only 17 words with text dispersions ranging from 48 to 11 texts. This concise list may be due to the small size of the corpus. However, the list may fail to capture important patterns and may not fully reveal the aspects of meanings that are being negotiated in the focus corpus. Therefore, this list was considered inadequate for addressing the question regarding the news values constructed in false news discourse.

Finally, it is important to note that individual keywords do not hold much significance in isolation. According to Baker (2004), analysing keywords becomes more meaningful when grouping them into semantic domains based on discursive similarities. This highlights the significance of meaning in assessing keywords. Specifically, when evaluating news values, it is crucial to examine the meaning of keywords, as news values are conveyed through semantic resources (Bednarek & Caple, 2017).

5.4.4. The Advantages of Key Semantic Domains Analysis for Exploring News Values

Rayson (2008) proposed the method of KSDA which is based on the UCREL Semantic Analysis System (USAS), an automatic analytical framework which looks beyond the lexical level and groups words according to their meaning or semantic domain (Archer et al., 2003). To apply this method, Wmatrix5 (Rayson, 2008) is used to assign semantic domains (or tags) to each word or multiword expression in the focus and reference corpora. The semantic domains contain 21 major discourse fields and 232 subfields, as listed in Archer et al. (2002). Some tags can be assigned plus or minus signs to indicate where the meaning is located on a binary or linear level. For example, the tag T3 refers to ***time: old, new and young; age***. Thus, the word 'kid' is assigned T3-, whereas the word 'pensioner' is tagged as T3 +. Following the semantic field assignments, a keyness calculation (e.g., log-likelihood) is performed to identify the semantic domains that occur significantly more frequently in the focus corpus.

In the pilot analysis, the method of KSDA was proven to be a valuable tool for analysing the news values of false news for two reasons. Firstly, news values are constructed through semantic resources or "meaning-making devices" (Bednarek & Caple, 2017, p. 3), which encompass a wide range of lexical and grammatical resources. For example, the news value of *superlativeness* can be expressed by intensifiers (e.g., 'severe' or 'dramatically'), quantifiers (e.g., 'thousands' or 'huge'), metaphor and simile (e.g., 'a tsunami of crime', 'like a World War II battle'), or

comparison (e.g., ‘the largest drug ring in Detroit history’) (Bednarek & Caple, 2017). By analysing key semantic domains, we can address the open-ended nature of the semantic resources expressing a news value. This is because key semantic domains can represent general patterns of meaning; each of which can in turn represent specific news values.

Secondly, the method of KSDA can improve the interpretability of these semantic resources because the semantic tags themselves provide a better understanding of the functions of such semantic resources and, by extension, the news values potentially expressed by them. For example, if the semantic tag of ***comparing: unusual*** is a key domain in a news corpus, this may indicate that the news value of *unexpectedness* is frequently constructed.

Furthermore, this method has multiple advantages over other keyness methods. Firstly, it allows researchers to detect patterns that cannot be recognised at the word or part-of-speech levels (Rayson, 2008). In particular, it can group words that might not be extracted as keywords individually due to their low frequency and could thus be overlooked (Culpeper & Demmen, 2015). Moreover, it allows researchers to focus on fewer domains, which can serve as clues to patterns that exist in larger keyword lists (Culpeper & Demmen, 2015). In addition, compared to other keyness measures, this method is useful for small corpora, such as the one used in this study, as small corpora often yield short lists of keywords or keywords with low dispersion that cannot be sufficient to identify meaningful patterns. In summary, the method of KSDA can be regarded as a reliable and comprehensive tool for exploring the news values in false news discourse.

5.5. Analytical Procedure

Using Wmatrix5, the key semantic domains were extracted by comparing the false news corpus to the reference corpus. Statistical significance and effect size measures were combined, and only those domains with a log-likelihood value of 7.00 or higher ($p < 0.01$) were considered. Subsequently, the domains were ranked by Log Ratio, and those with a Log Ratio of less than 0.50 were removed, leaving the top 59 semantic domains (See Appendix 5.3). These domains were further categorized into larger semantic categories based on their shared meanings, ultimately leading to the identification of seven major semantic categories, as illustrated in Table 5.2.

A qualitative concordance analysis of the key semantic domains was conducted using the following steps, adapted from Baker's (2006) step-by-step guide:

1. Obtain the concordances of the semantic domain under question.
2. Clean the concordances by removing irrelevant lines (e.g., concordances of words that are not accurately tagged).
3. Look for concordances that share discourse patterns.
4. Look for other discourse patterns in the remaining concordance lines.
5. Note rare or unavailable cases of discourse patterns based on intuition.
6. Hypothesise the rationale behind the existence of certain discourse patterns and relate this to the context of the text.

The purpose of the concordance analysis was to gain insight into the functions conveyed by the semantic domains, categorize them into major semantic categories, identify any recurring usage patterns, hypothesize the reasons for the repetition of these patterns, and understand how they contributed to shaping the news narrative to make it appealing to readers. Through careful examination of the contextual use of these semantic domains, it became possible to discern their role in constructing news values.

5.6. Analysis

The analysis found there are seven major semantic categories that are distinctive of the false news corpus, namely *connecting ideas*, *evidentiality*, *evaluation*, *religion and supernatural*, *science*, *time*, and *news topics*. It should be noted that the major as well as the specific semantic categories created by Wmatrix5 are not mutually exclusive. Some words may fall into multiple categories due to the focus of the analysis being on meanings rather than specific parts of speech. As a result, defining clear-cut semantic categories may prove challenging. The subsequent sections present the qualitative analysis of these semantic categories and discuss how these groups could potentially embody news values.

5.6.1. Key Semantic Domains in Online False News Articles

Table 5.2 illustrates the major (and sub-) semantic categories, along with examples for each semantic domain generated by Wmatrix5. Subsequently, in the sections following the table, the

key semantic domains under each semantic category are discussed, and examples that represent the patterns identified in the concordance analysis of the semantic domains are analysed.

Table 5.2 Key Semantic Domains in the Focus Corpus

Major Semantic Groups		**Specific Semantic Domains Generated by Wmatrix5**	**Example Lexis**
I. Connecting Ideas	1. Comparative Relation	+Comparing: Unusual	novel, unprecedented, mysterious, unusual, exotic, abnormal
		Comparing: Usual	natural, common, usually, normal, commonly, generally
		+Comparing: Similar	same, replicated, identical, equal
		Unexpected	Surprise, wonder, unexpected, astonishing, surprised, astoundingly
		Degree: non-specific	degree, even, so-much-as
	2. Cause-effect Relation	Cause & effect/connection	cause, why, caused, because of, effects, due to
II. Evidentiality	1. Spoken Evidence	Speech: Not communicating	kept quiet, speechless
		Sound: Quiet	quietly, silent, quiet, silently, silence
	2. Written Evidence	The Media	published, media, editor-in-chief, edition, censorship
		The Media: Newspapers etc.	article, journal, magazine, newspaper, press release, press conference
		The Media: Books	book, authors, literature, readers, co-authors, public library
		Knowledge	data, database, databases, informative
		Information technology and computing	program, website, computer, dataset, internet, online, datasets
	3. Sensory Evidence	Seen	notice, look outside, noticed, noticeably
		Inattentive	ignored, ignoring, ignore, in dream, disturbed, distraction
		Open; finding; showing	found, discover, revealed, showing, findings, exposed,
	4. Deductive Evidence	Seem	appear, seems, apparently, looks, reportedly
		Probability	perhaps, maybe, odds, probability, likelihood
III. Evaluation	1. Evaluation of Credibility	Evaluation: authentic	actual, effectively, authenticity, proper, really, authentic
		Evaluation: Unauthentic	fake, forged, forgery, copy, pretended, unconfirmed
		Evaluation: True	evidence, fact, in fact, truth, facts, TRUE
		Evaluation: False	hoax, lie, FALSE, misleading, lies, fabricated
	2. Evaluation of Rationality	Foolish	stupid, gullible, mindless, absurdity, lunacy, baloney
	3. Evaluation of Morality	Unethical	scandal, scam, tricks, corruption, trick, rogue
		Bravery	courage, dare, courageous, courageously
		Crime	fraud, crime, terrorism, criminal, illegally, murderous

	4. Evaluation of Emotion	Fear/shock	alarmists, shock, panic, scare, fear, scary
IV. Religion and Supernatural		Religion and supernatural	churches, God, pray, prophet, catholic, pastors
V. Science		Science and technology in general	scientists, science, scientific, laboratory, technology
		Substances and materials generally	protein, Toner, fuels, substance, formaldehyde, hydrochloride
		Substances and materials: Solid	ferrous, aluminium, sodium, plastic, zinc, soils
		Substance and materials: Liquid	oil, water, calcium chloride, mercury, moisture, saline
VI. Time		Time	time, never, times, watch, timeline, never ever
		Frequent	always
VII. News Topics		Temperature	temperatures, temperature, melted
		+Temperature: Hot/ on fire	warmer, blistering, warming, fires, burning, heat
		+Temperature: Cold	colder, cooler, frozen, freeze, icy, chilly
		Life and living things	biological, wildest, lifesaving, life threatening, mortal, wild
		Food	peanuts, food, eating, sugar, seafood market, salt
		Plants	trees, seeds, plant, palm trees, pollen, plant life
		Geographical terms	global, earth, sea, sea level, earths, fossil
		Weather	climate, weather, flooding, hurricanes, cloud, drought
		Sailing, swimming, etc.	launched, buoys, ships, launch, yachts, sail
		Damaging and destroying	harm, bushfires, destroy, broke, damage, victim
		Warfare, defence and the army; weapons	shot, military, biological warfare, soldiers, raid, bombs
VIII. Miscellaneous		No people	depopulate
		No constraint	release, leaked, freedom, leak, loose, liberal
		Weight: Heavy	heavy
		Participating	interference, interfere
		Non existing	missing, non-existent
		Being	phenomenon, phenomena, availability
		Using	use, used, using, uses, function, utilizing
		Games	trumps, checkers, volleyball, checker, softball
		Discourse bin	however, of course, right, in other words, on the other hand, etc., once again, basically

Note. The semantic tags marked with plus signs were repeated in the top 59 tags but with varying degrees.

5.6.1.1. Connecting Ideas

The first group of prominent meanings in the false news corpus involves connecting different ideas or concepts by using words that indicate cause-effect and comparative relations. The semantic domain ***cause & effect/connection*** is concerned with the identification of reasons and the description of consequences, i.e., one event leads to another. Cause-effect meanings can appeal to the readers' sense of logic, as they are used to construct persuasive arguments (Fahnestock, 2011; Van Dijk, 1988) by providing logical reasoning. For instance, in Example 1, the writer uses a cause/effect connection to explain how a report's claims about global warming were based on unverified data, therefore leading to the conclusion that the pause in global warming never existed. In Example 2, the writer uses the cause/effect connection to describe how Bill Gates has been spending a lot of money on contraception, abortion, and sterilization in Third World countries, and this has led to the conclusion that his vaccines inadvertently containing sterilization medication. Thus, the frequent usage of cause-effect meanings in false news serves to build persuasive and logical arguments for the readers.

> (1) The report claimed the pause in global warming never existed, but it was **based on** [cause & effect/connection] misleading, unverified data. *(A False 13)*

> (2) How would vaccines **lead to** [cause & effect/ connection] depopulation, when they're ostensibly designed to make people healthy? Well, Gates is spending a whopping sum on contraception, abortion, and sterilization in Third World countries, but there's a little secret about his vaccines as well. Bill Gates vaccines have made news reports in recent days for (supposedly inadvertently) containing sterilization meds. *(C False 38)*

The semantic domains of ***comparing: similar, comparing: usual, comparing: unusual, unexpected,*** and ***degree: non-specific*** all involve comparative meanings. Comparison is associated with examining similarities, differences, expectations, and peculiarities, and comparing and contrasting existing knowledge with new knowledge (cf. mirativity in Aikhenvald, 2012; Bednarek, 2006). Comparison is used for building persuasive arguments and providing classification (Fahnestock, 2011). Through comparison, patterns, trends, and relationships can be recognized, and conclusions or inferences can be drawn. Thus, by using comparative meanings, writers can appeal to readers' logic. For instance, the word 'same', in Example 3, emphasises the shared characteristics of three organizations: NASA, NOAA, and the

UK Met Office. By pointing out that they all use identical, corrupted data, the author conveys the idea that the conclusions drawn by these organizations cannot be trusted.

> (3) What this means, the report concludes, is that claims by NASA, NOAA, and the UK Met Office that the world is experiencing unprecedented and dramatic warming should be taken with a huge pinch of salt: they all use the **same** [comparing: similar] corrupted global average temperature (GAST) data. *(A False 5)*

In Example 4, the use of 'natural' and 'normal' emphasises that the changes occurring to planetary weather patterns are not unusual or extraordinary. By using these comparative meanings, the writer downplays the idea of man-made global warming, suggesting that it is a hoax and that the space agency has chosen to let this false narrative continue without correcting it.

> (4) For more than 60 years, the National Aeronautics and Space Administration (NASA) has known that the changes occurring to planetary weather patterns are completely **natural** [comparing: usual] and **normal** [comparing: usual]. But the space agency, for whatever reason, has chosen to let the man-made global warming hoax persist and spread, to the detriment of human freedom. *(A False 2)*

By contrast, the use of the word 'mysteriously', in Example 5, emphasises the unexpected and unusual nature of the coronavirus and implies that its source is unknown or uncertain. This helps to create an atmosphere of suspense and conspiracy. Similarly, in Example 6, the adverb 'amazingly' is used to describe how surprising it is that a child was given more inoculations after showing a negative reaction to the first vaccine. The use of this word intensifies the emotional engagement of the reader by emphasising the shock that readers should feel when reading the text.

> (5) Coronavirus **Mysteriously** [comparing: unusual] has HIV Glycoprotein Inserts not Present in any other known Coronavirus JUST PUBLISHED Folks, this coronavirus epidemic is getting more interesting and possibly scary. *(C False 46)*

> (6) Reports indicate that after the initial inoculation Marcella had, lost her appetite, and became feverish and fatigued. **Amazingly** [unexpected], more inoculations were given despite her negative reaction from the first jab. *(B False 24)*

These examples show that false news writers use comparative meanings to challenge conventional assumptions about normalcy, create new expectations of normalcy, and highlight events that deviate from them. This allows them to create a narrative of the expected and unexpected, where the expected is more likely to be accepted, while the unexpected is more likely to be challenged. In doing so, false news writers can manipulate readers into accepting their claims by describing them as normal, usual, and expected, while discrediting their opponents' claims by describing them as unusual and shocking.

The use of the word 'incredible', in Example 7, emphasises the surprising and unexpected nature of the claim that there was a patron saint of pandemics called Corona; however, it is followed by the clause 'but it's seemingly true'. Therefore, 'incredible' serves to sensationalize the story and capture the reader's attention, prompting them to consider the implications of this statement.

> (7) It's **incredible** [comparing: unusual] but it's seemingly true there is a Saint Corona and she is one of the patrons of pandemics. *(C False 24)*

To sum up, as Bednarek and Caple (2012) state, comparative meanings are utilized in news in order to make persuasive arguments and capture readers' attention. The examples discussed above showed that the high frequency of comparative meanings in false news serves to emphasise similarities and differences, as well as to emphasise expectations and peculiarities encouraging readers to consider the implications of arguments. Comparative meanings were also used to sensationalize stories, as they imply evaluation and evoke emotion.

5.6.1.2. Evidentiality

The second group of prevalent meanings in the false news corpus pertain to evidentiality, and it can be broken down into spoken, written, sensory, and deductive. Spoken evidence encompasses the semantic domains of ***speech: not communicating*** and ***sound: quiet***. Examining the context of these meanings shows that they are used to emphasise the intentional concealing or ignorance of spoken evidence. In Example 8, the words 'quiet' and 'quietly' highlight the intentional concealment of the correct data by the National Snow and Ice Data Centre (NSIDC). The two words indicate that NSIDC was not forthcoming or transparent in their correction of the figures. In Example 9, the use of the word 'quietly' emphasises the sudden and unnoticed disappearance

of the cottonseed oil allergies that allegedly were caused by a vaccine. It implies that the change in the vaccine formula happened without much attention or publicity.

> (8) HOW NSIDC GOT ITS FIGURES WRONG AND THEN **KEPT QUIET** [speech: not communicating]
>
> Since publication of the original version of this article, the US source of the figures the NASA-funded National Snow and Ice Data Centre (NSIDC) - was discovered to have made a huge error and then **quietly** [sound: quiet] corrected the figure without mentioning it. *(A False 14)*

> (9) In the 1930s there was cottonseed oil in vaccines, followed by a short-lived spate of cottonseed oil allergies of about a decade that **quietly** [sound: quiet] went away with a change in formula. *(B False 21)*

Written evidence encompasses the semantic domains of ***the media, the media: newspapers etc., the media: books, knowledge, inattentive,*** and ***information technology and computing***. These meanings were used in the false news articles to emphasise either the intentional disregard and suppression of written evidence or the availability of written evidence that supports theories alternative to mainstream ones. For instance, in Example 10, the word 'published' accentuates the presence of written evidence, i.e., a chart published on the EPA.gov website, indicating that this chart was not simply created but was made available to the public. Likewise, the word 'website', in Example 11, emphasises the source of the information and that it was not reported by conventional news outlets but by an independent news outlet. Additionally, in Example 12, the word 'press' emphasises the source of the reported statement by Bill Gates, indicating that it was not just hearsay but was reported by the press. The word 'literature', in Example 13, emphasises that the theories of peanut allergies were not just supported by the cited study but also by the scientific literature that existed before this study.

> (10) The following chart, **published** [the media] on the EPA.gov website (4), clearly shows modern-day heat waves are far smaller and less severe than those of the 1930 's. *(A False 14)*

> (11) The news went mostly unreported, but independent outlets like The Liberty Beacon finally began publishing the ground breaking news. The **website** [information technology and

computing] wrote last month, In a recently published December 13, 2012 vaccine court ruling, hundreds of thousands of dollars were awarded to Ryan Mojabi, whose parents described how MMR vaccinations caused a severe and debilitating injury to his brain, diagnosed as Autism Spectrum Disorder (ASD). *(B False 8)*

(12) And Gates now argues that mass gatherings (churches) will not meet again until a vaccine is found, and everyone is vaccinated. **The press** [the media: newspapers, etc.] reports Gates said, What does opening up look like? Which activities, like schools, can be done in a way that the risk of transmission is very low? And which activities, like mass gatherings, maybe in a certain sense more optional? *(C False 38)*

(13) It turns out that Israeli children ate massive quantities of peanuts, and British children almost none, so the eating peanuts theory was blown out of the water by not only this study, but by existing scientific **literature** [the media: books] which has concluded that you cannot develop a peanut allergy from eating peanuts. What the study didn't mention however is that vaccines produced for Israeli children predominantly use sesame seed oil instead of peanut oil as an adjuvant. *(B False 20)*

By contrast, in Example 14, the word 'data' highlights the intentional concealing of evidence, i.e., the data were corrected not to make sure the results of the study were accurate but to hide evidence. Similarly, the word 'ignored', in Example 15, implies the intentional neglect of written evidence, Milankovitch's model, for nearly half a century although its accuracy was confirmed later by a scientific study.

(14) A Top Climate Scientist Blows the Whistle on Shoddy Climate Science.

The NOAA corrected **data** [knowledge] they didn't like and surprise didn't archive the evidence. *(A False 12)*

(15) When Milankovitch first put forward his model, it went **ignored** [inattentive] for nearly half a century. Then, in 1976, a study published in the journal Science confirmed that Milankovitch's theory is, in fact, accurate, and that it does correspond to various periods of climate change that have occurred throughout history. *(A False12)*

Sensory evidence is manifested in words having the semantic domains of ***seen*** and ***open; finding; showing***. These semantic domains are used to either stress the existence of sensory

evidence or emphasise the deliberate concealing or ignorance of sensory evidence that supports the false claims. The word ‘notice’, in Example 16, emphasises that doctors had observed and seen the similarities between the food reactions and the serum sickness caused by vaccines. In Example 17, the word ‘find’ indicates that the researchers had actively searched and found a study that supports the author’s claim, and the word ‘found’ indicates that the study actually provided evidence of a link between flu shots and coronavirus. The three words ‘notice, ‘find’, and ‘found’ imply that the evidence does exist and is seen rather than transmitted orally.

> (16) Soon other doctors began to **notice** [seen] striking similarities between food reactions and the serum sickness that was associated with vaccines. *(B False 22)*

> (17) In searching the literature, the only study we have been able to **find** [open; finding; showing] assessing flu shots and coronavirus is a 2020 US Pentagon study that **found** [open; finding; showing] that the flu shot INCREASES the risks from coronavirus by 36%. *(C False 47)*

The last type of evidentiality in false news is deductive evidence; it is reflected in the semantic domain of ***seem*** which includes words like ‘seem’, ‘seemingly’, and ‘appear’ and ***probability*** which includes words like ‘perhaps’, ‘maybe’, ‘odds’, ‘probability’, and ‘likelihood’. These words are among evidentiality markers and are used to express reasoning or predict future consequences, indicating low reliability (cf. Chafe, 1986). In Example 18, the word 'seem' suggests that there is a possibility that medical organizations use flu shots with the intention of suppressing immune systems. This low level of reliability is not enough to assert with certainty that it is indeed the case; however, it suggests that it could be true. Similarly, the word 'perhaps', in Example 19, indicates that there is a possibility that the current emissions-mitigation policies are not just ineffective but could even be harmful to the environment. This low level of reliability is not enough to assert with certainty that the policies are harmful; however, it indicates that it could be the case. The use of hedging in these examples could be a rhetorical device to avoid arousing suspicion by expressing uncertainty or lack of absolute confidence in these outlandish claims.

> (18) Flu shots are being hyped to suppress immune systems and create an explosion in coronavirus infections. Flu shots, it **seems** [seem], are one more way the medical establishment can literally murder children and adults, creating widespread vulnerability

> to the coronavirus pandemic which will no doubt surge back this Fall as flu shots work to suppress the immune systems of millions. *(C False 44)*

> (19) And current emissions-mitigation policies, especially related to the advocacy for renewables, are often costly, ineffective, and **perhaps** [probability] even harmful to the environment. On the other hand, elevated CO2 and a warmer climate provide unheralded benefits to the biosphere (i.e., a greener planet and enhanced crop yields). *(A False 6)*

In summary, the examples discussed highlight that false news stories make more frequent use of four types of evidentiality – spoken, written, sensory, and deductive – compared to other types of news. Evidentiality in news is utilized to support the subjective elements of a story (Bednarek & Caple, 2012) while maintaining an appearance of objectivity and evaluating the reliability of the reported information (Bednarek & Caple, 2012; Van Dijk, 1988). Thus, evidentiality creates an illusion of truth and enhances the perceived objectivity and reliability of the news (Van Dijk, 1988). In the context of the analysed false news articles, evidentiality serves to establish credibility for the false news' version of reality and at the same time to challenge the credibility of mainstream theories.

5.6.1.3. Evaluation

The third group of meanings prevalent in the false news corpus is associated with evaluation. Four types of evaluation can be identified: credibility, rationality, morality, and emotions. Evaluation of credibility involves the semantic domains of ***evaluation: authentic, evaluation: unauthentic, evaluation: true,*** and ***evaluation: false***. These meanings are used to present mainstream theories and their evidence as false and inauthentic and alternative theories and their evidence as true and authentic. For instance, in Example 20, the word 'proved' emphasises that the identities of the relics were verified, confirming that the two religious figures, Cedar and Saint Corona, really existed. 'Authenticity', in Example 21, suggests the unquestionability and reliability of the evidence that has been presented in the article, indicating that there have been cooling trends after the 1930s. In Example 22, 'lie' is used to describe the mainstream theories and their evidence regarding the global warming phenomena as false and inauthentic. In Example 23, 'forged' emphasises Time magazine's act of fabricating or altering evidence in order to provide a different narrative.

(20) Two investigations in 1943 and 1981, **proved** [evaluation: true] that the relics belong indeed to a man and a woman. Cedar pollen was found confirming an original burial in Syria and then in Cyprus. Saint Corona is especially venerated in Austria and Bavaria as the patron-saint of treasure hunters and against epidemics. *(C False 23)*

(21) It shows the highest temperatures actually occurred in the 1930s, followed by a cooling trend ramping downward to the year 2000: The **authenticity** [evaluation: authentic] of this chart is not in question. It is published by James Hansen on NASA 's website. *(A False 15)*

(22) In other words, the so-called Consensus on global warming is a massive **lie** [evaluation: false]. And Donald Trump was quite right to quit the Paris agreement which pretended that the massive **lie** [evaluation: false] was true. *(A False 4)*

(23) That cover is a seriously inconvenient truth for climate alarmists and their media accessories. So, Time attempted to re-write a history. It published a **forged** [evaluation: unauthentic] version of its own cover. *(A False 11)*

The second type of evaluation is concerned with rationality; it is manifested in the semantic domain of ***foolish***. Concordances of this domain show that mainstream theories and people who believe in them are described as irrational. For instance, in Example 24, 'mindless' is used to refer to the superficiality of the media's coverage of the measles resurgence. It suggests that the media are not taking the time to investigate the facts of the situation and are instead simply relying on sensationalism for attention.

(24) Once again, the media is discarding factual reporting in favor of **mindless** [foolish] sensationalism, attributing an alleged measles resurgence -- even this claim is specious -- to the unvaccinated. *(B False 7)*

The third type of evaluation is concerned with morality which encompasses the semantic domains of ***unethical***, ***bravery,*** *and* ***crime***. The concordances of these semantic domains show that false news typically portrays governments' and health organizations' practices as unethical, illegal, and criminal, while simultaneously characterizing those who speak out against mainstream theories as brave. In Example 25, the word 'corruption' hints at the dishonest adjustments that have been made to exaggerate the appearance of global warming, suggesting

that those in charge are behaving in an unethical and possibly illegal manner. In Example 26, the word ‘criminal’ emphasises the magnitude of the unethical and immoral deeds of drug companies. In contrast, the word 'bravery', in Example 27, describes how Dr Bates stepped forward to tell the truth about NOAA’s practices, suggesting that he acted in a courageous and admirable manner.

> (25) When alarmists in charge of surface temperature datasets make dishonest adjustments to exaggerate the appearance of global warming, it looks like **corruption** [unethical]. *(A False 14)*

> (26) The drug companies will cash into the tune of hundreds of billions of dollars while innocent children and adults all across America needlessly suffer and die. As usual, it’s all about the money, and the vaccine/drug cartels are murderous, **criminal** [crime] regimes that twisted science (Dr. Fauci ?) to maximize profits at any cost. *(C False 44)*

> (27) Last night Mr Smith thanked Dr Bates for **courageously** [bravery] stepping forward to tell the truth about NOAAs senior officials playing fast and loose with the data in order to meet a politically predetermined conclusion. *(A False 13)*

The fourth type of evaluation is concerned with emotion, and security in particular (cf. Martin & White, 2005). This is reflected in the semantic domain of ***fear/shock***. The analysis of this domain shows that false news portrays governmental organizations’ practices as frightening and shocking. In this way, it appeals to the reader’s emotions by evoking their feeling of insecurity and anxiety. Such is the case with Example 28 in which the word ‘terrifying’ is used to compare Bill Gates’ vaccine trials in India to the horrific excesses of 20th century evil regimes. Similarly, Example 29 describes the failure to archive and make available fully documented data as a ‘bombshell’ that goes against NOAA and Science rules.

> (28) Mainstream investigators have found that these vaccine trials in India led to thousands of injuries and deaths of young women. The philanthropist has also funded secretive sterilization programs, and says we should form death panels to differentiate between those who are worthy of life and those who have no benefit whatsoever to society. This cold, skewed psychopathic logic is **terrifying** [fear/shock] and brings to mind the worst excesses of the 20th century’s most evil regimes. Now we are starting to see the horrific results of Gates depopulation agenda playing out around the world. *(B False 9)*

> (29) Dr Bates revealed that the failure to archive and make available fully documented data not only violated NOAA rule, but also those set down by Science … Then came the final **bombshell** [fear/shock]. Dr Bates said: I learned that the computer used to process the software had suffered a complete failure. *(A False 13)*

To sum up, evaluation plays an important role in false news. It is used to discredit mainstream narratives and their evidence as false and inauthentic, while simultaneously praising alternative narratives and their evidence as true and authentic. Evaluation is also used to describe mainstream theories and their supporters as irrational, unethical, and frightening. These techniques are used to appeal to readers' emotions and evoke feelings of fear and insecurity.

5.6.1.4. Religion and Supernatural

The fourth dominant group of meanings in the false news corpus is linked to the semantic domain of ***religion and supernatural*** which involves references to mystical elements that are religious, spiritual, magical, and supernatural. In comparison to other news types, the false news corpus contains a higher frequency of references to religious figures, places, practices, and legends in relation to topics of climate change and COVID-19. For instance, in Example 30, it is stated that a self-proclaimed prophetess claimed to possess accurate prophecies and divine authority to stop natural disasters such as Hurricane Florence and the California wildfires. In Example 31, it is reported that the Democratic mayor of Kansas City had ordered pastors to turn over membership lists and details of anyone who entered church houses during the COVID-19 pandemic, which implies a conspiracy theory involving surveillance, tracking, and monitoring of religious individuals. Example 32 features a quote from the Pope saying he was close to those infected with COVID-19 and expressing his gratitude towards healthcare workers. Example 33 reports a religious legend which claims that there was Saint Corona who is said to have been martyred for her religious beliefs and is recognized as a patron saint against epidemics.

> (30) Similarly, self-proclaimed **prophetess** [religion and supernatural] Kat Kerr declared that her guarantee of a red tsunami had also been accurate … Kerr also claimed she had **divine** [religion and supernatural] authority to stop natural disasters such as Hurricane Florence and the California wildfires. *(A False 16)*

(31) Kansas City Mayor Quinton Lucas, a Democrat, has ordered **pastors** [religion and supernatural] to turn over the names, address and phone numbers of anyone who enters **church** [religion and supernatural] houses.

…

"Never in our wildest dreams could we have imagined Nazi-like measures designed to surveil, track and spy upon what was once a FREE American people," Liberty Counsel founder Mat Staver said. *(C False 15)*

(32) The **pope** [religion and supernatural] was seen coughing and blowing his nose during the Asha Wednesday Mass. Yesterday, he told well-wishers: I wish, again, to express my closeness to those who are ill with coronavirus and to health-care workers who are caring for them. *(C False 43)*

(33) According to **legend** [religion and supernatural], **Saint** [religion and supernatural] Corona was arrested and tied by her feet to the tops of two palm trees which were bent to the ground. *(C False 24)*

The extensive use of religious and supernatural themes in the false news corpus reveals a pattern where false news creators exploit beliefs, spiritual elements, and mystical narratives to add credibility to their stories. The incorporation of religious and supernatural elements in false news can also fuel conspiracy theories. For instance, in Example 31, there is a suggestion of a clandestine agenda to track individuals visiting churches, contributing to a sense of mystery and foreboding. Thus, the inclusion of religious and supernatural elements in false news can create a compelling blend of emotional triggers and supposed truths within the false content.

5.6.1.5. Science

The fifth group of meanings prevalent in the false news corpus is science as seen in the semantic domains of ***science and technology in general, substances and materials generally, substances and materials: solid,*** and ***substances and materials: liquid***. For instance, the word 'scientific', in Example 34, emphasises the evidence-based foundations for climate change sceptics. In Example 35, the word 'glucose' denotes a type of substance that can break down when exposed to low oxygen levels, supposedly caused by wearing masks. In Example 36, the word 'aluminium' refers to a type of solid, metallic element found in some vaccines that had an adverse effect on an

infant's brain and led to his eventual death. In Example 37, the word 'mercury' denotes a type of liquid element, found in the flu vaccine, which, the false claim alleges, can weaken the immune system and increase one's risk of catching the flu.

> (34) Climate change proponents remain undeterred in their mission, ignoring numerous recent **scientific** [science and technology in general] findings indicating that there has been no warming trend at all for nearly two decades. *(A False 14)*

> (35) In addition, oxygen deficiency causes **glucose** [substances and materials generally] breakdown and endangered lactic acid rise. Some people drive their car with the mask on, that is very dangerous, because, the stale air can make the driver lose consciousness. *(C False 30)*

> (36) A breakthrough study suggests that **aluminium** [substances and materials: solid] from vaccines was likely present in baby Alfie Evan's brain, which resulted in his death. *(B False 2)*

> (37) For most people, says Dr. Blaylock, flu vaccines don't prevent the flu but actually increase the odds of getting it. The **mercury** [substances and materials: liquid] contained in vaccines is such a strong immune depressant that a flu shot suppresses immunity for several weeks. *(B False 23)*

Overall, false news, compared to other news types, more frequently refers to technology, science, scientists, scientific substances, and lab work. This suggests that false news creators make significant efforts to present their claims as credible and trustworthy by incorporating scientific language and technical terminology and by offering in-depth medical, physiological, and biological explanations. These strategies may be employed to provide seemingly factual and logical justifications for false assertions.

5.6.1.6. Time

The sixth group of dominant meanings in the false news corpus pertains to time which is reflected in the semantic domains of ***time*** and ***frequent***. Examples 38, 39, and 40, use 'times', 'never', and 'always' respectively to emphasise that a certain event has been or has not been continuous or repeated and will or will not likely continue to occur. 'Times', in Example 38, highlights the multiple times the 'trick' of using false climate data has been repeated, making it clear that the event is a continuous occurrence. 'Never', in Example 39, indicates that no bat was

ever traded in the Wuhan market, signalling a conspiracy theory that coronavirus has leaked from a lab. Finally, 'always', in Example 40, implies that climate has constantly been changing, and it is not the result of human actions. In Example 41, the word ‘time’ denotes when the flu-related deaths occurred compared to the same time period in the previous year.

> (38) Then, climate alarmists had a joyful and profitable time debunking the forgery they and/or their accomplices had created. This trick has been repeated multiple **times** [time] with other arguments, convincing the majority of onlookers that climate realists are liars or simply ignore these uncertainties on board. *(A False 11)*

> (39) According to municipal reports and the testimonies of 31 residents and 28 visitors, the bat was **never** [time] a food source in the city, and no bat was traded in the market. There was possible natural recombination or intermediate host of the coronavirus, yet little proof has been reported. The researchers screened the area around the seafood market and identified two laboratories conducting research on bat coronavirus. *(C False 1)*

> (40) Based on these different variables, Milankovitch was able to come up with a comprehensive mathematical model that is able to compute surface temperatures on earth going way back in time, and the conclusion is simple: Earth’s climate has **always** [frequent] been changing, and is in a constant state of flux due to no fault of our own as human beings. *(A False 2)*

> (41) The Los Angeles Daily News reported there were 36 flu-related deaths this season in California for adults under age 65 compared to 13 for the same **time** [time] last year. (*B False 18)*

In light of these examples, the more frequent use of time words in false news articles, compared to other news types, serves multiple purposes. Firstly, time words provide a context for events and help readers establish a comprehensive timeline. Additionally, they highlight the relevance of past and present events, as well as their potential impact on the future. By referencing the past, false news articles create an illusion of historical accuracy and consistency with records and facts. Indeed, this is a common characteristic of news discourse in general. As Van Dijk (1988) argues, news discourse often relies on describing past events as causes or conditions and future events as possible or real consequences in order “to promote a persuasive process for assertion”

(p. 84). However, the analysis of key semantic domains shows that the use of time words is more pronounced in false news.

5.6.1.7. Topic

The last group of dominant meanings is topic-specific, comprising semantic domains that involve lexis related to the topics discussed in the false news articles. Examples of these semantic domains are ***weather, temperature, plants, geographical terms,*** and ***sailing, swimming, etc.,*** which pertain to the topic of climate change, ***warfare, defence and the army; weapon,*** which pertains to the topic of COVID-19, and ***damaging and destroying*** which is related to the three topics of climate change, vaccinations, and COVID-19**.** Despite the reference corpus being larger than the false news corpus and both focusing on the same topics, the analysis reveals that vocabulary denoting the topics of the analysed articles, particularly climate change, stands as key in the false news corpus.

The concordance analysis of the key semantic domains related to the topics discussed in the analysed articles revealed two distinct patterns. Firstly, these topic-specific words predominantly convey negative connotations. For instance, in Example 42, the word 'harm' is used to emphasise the adverse effects of vaccination. In Example 43, the mention of China's 'biological warfare' program hints at a conspiracy theory that suggests the coronavirus was created in a lab as part of this program. This narrative could provoke negative sentiments towards China.

> (42) a recent study puts to light a not-so-new problem with vaccines. And it is mainly because its adjuvant Aluminium may be doing more **harm** [damaging and destroying] than good. *(B False 16)*

> (43) China 's **Biological Warfare** [warfare, defence and the army; weapon] Program is believed to include full range of traditional chemical &; biological agents with a wide variety of delivery systems including artillery rockets, aerial bombs, sprayers, and short-range ballistic missiles. *(C False 34)*

The second pattern observed is the prevalence of repetitive content in false news articles. In other words, the high frequency of topic-related words is due to the repetitive content in false news articles. This pattern aligns with previous findings by Horne and Adalı

(2017). For instance, in Example 44, topic-specific words such as ‘bushfires’, ‘fire’, and ‘climate’, along with mentions of ‘Australia’ and ‘Australians’ (not belonging to the selected key semantic domains), are recurrent throughout the text. The repetition of these words reflects the focus on the causes and consequences of the Australian bushfires, as well as the debate surrounding whether climate change or arson is to blame for the fires.

> (44) The raging inferno that has engulfed a large swath of **Australia** is horrifying. More than 1,500 homes have been destroyed, 25 people have died and millions of kangaroos and koala bears and other animals may have been lost. Radical environmental extremists and Hollywood celebrities have blamed the **fires** [temperature: hot/on fire] on **climate** [weather] change. But it turns out **climate** [weather] change has nothing to do with the humanitarian crisis unfolding Down Under. It turns out many of the **fires** [temperature: hot/on fire] were set by arsonists. The **Australian** reports more than 180 people have been arrested since the start of the **bushfire** [temperature: hot/on fire] season. Swinburne University professor James Ogloff told the newspaper about 50 percent of **bushfires** [temperature: hot/on fire] were lit by firebugs and impending **fire** [temperature: hot/on fire] seasons excited them. They’re interested in seeing **fire** [temperature: hot/on fire], interested in setting **fire** [temperature: hot/on fire] and quite often the information around how **fires** [temperature: hot/on fire] burn and accelerate excites them, Ogloff told the newspaper. Note that the professor mentioned the impending **fire** [temperature: hot/on fire] seasons. The **bushfires** [temperature: hot/on fire] are a regular occurence in **Australia**. Deputy Prime Minister Michael McCormack dismissed concerns that **climate** [weather] change had anything to do with the massive **fires** [temperature: hot/on fire]. He blamed the outcry on raving inner-city lefties who were ignoring the needs of rural **Australians**. Weve had **fires** [temperature: hot/on fire] in **Australia** since time began, he told BBC.NASA admits that **climate** [weather] change occurs because of changes in Earth’s solar orbit, and NOT because of SUVs and fossil fuels. *(A False 1)*

5.6.2. News Values of Online False News Articles

This section discusses the major semantic groups identified in the false news corpus and their connection to news values. It presents the news values covered by the DNVA framework and those that may be unique to false news.

Summarizing the analysis of key semantic domains in the false news corpus, the following major groups of meanings were identified:

1) Connecting ideas: involves words that indicate cause-effect and comparative relations.
2) Evidentiality: involves meanings related to the availability, collection, and dismissal of evidence.
3) Evaluation: involves meanings related to the assessment of credibility, rationality, morality, and emotion.
4) Religion and supernatural: involves religious and mythical characters, practices, and places.
5) Science: encompasses technical and scientific jargon as well as detailed medical, physiological, and biological explanations.
6) Time: involves words that refer to the continuity or non-existence of events.
7) News topics: involves words that are related to the topics discussed in the false news reports.

Notably, these meanings have the potential of indicating news values of false news for three reasons. First, these meanings were found to occur more frequently in the false news corpus compared to the reference corpus. Second, this observation aligns with previous research which has also identified these meanings as commonly employed in fake news (as discussed below). Moreover, consistent usage patterns were observed in the concordances of these meanings, indicating that false news writers emphasise these meanings to enhance the newsworthiness of their articles, make them more compelling and impactful, and promote their dissemination.

The analysis suggests that false news, compared to true news, emphasises several news values that are covered by the DNVA framework, namely *negativity, unexpectedness*, *consonance*, and *superlativeness*. *Negativity* is realized by lexis describing damaging, destroying, fires, war, and crime, and by negative evaluation of credibility, rationality, morality, and emotion. This finding confirms previous research that has identified negativity as a defining characteristic of fake news (Alba-Juez & Mackenzie, 2019; Asr et al., 2023, 2023; Dance, 2019; Dey et al., 2018; Horne & Adalı, 2017; Maci, 2019; Rubin et al., 2016; Vamanu, 2019). *Unexpectedness* is established by

lexis with comparative meanings such as unusual and unexpected, and by evaluation of emotions denoting shocking. *Negativity* and *Unexpectedness* might make false news viral as well as memorable. According to Van Dijk (1988), deviant and negative events are more likely to be remembered due to their atypical or extraordinary nature. Indeed, sociological and psychological research on fake news has also shown that experiencing negative emotions while reading news stories can lead to a belief in them (Fernández-López & Perea, 2020; Martel et al., 2020; Pennycook & Rand, 2021).

Consonance is reflected in lexis with comparative meanings such as similar and usual. Bednarek and Caple (2017) limited consonance to news actors, social groups, organizations, or countries being described as stereotypical. However, as shown in the analysis, many news *events* were described as natural, normal, typical, or expected, and such lexis was used to support false news arguments and oppose mainstream theories. *Consonance*, especially when emphasising expectations, can enhance the credibility of the news report, as Galtung and Ruge (1965) states:

> A person predicts that something will happen and this creates a mental matrix for easy reception and registration of the event if it does finally take place. Or he wants it to happen and the matrix is even more prepared, so much so that he may distort perceptions he receives and provide himself with images consonant with what he has wanted. (p. 67)

Similarly, Douglas (2018) states humans tend to put faith in what they wish to be factual rather than to acknowledge the difficult truths they do not wish to be true (cf. the concepts of *naïve realism* and *confirmation bias* in Shu et al., 2017).

The semantic domain of ***weight: heavy*** (not belonging to the major seven semantic groups presented in the analysis), suggests that *superlativeness* might be a news value of false news. However, the evidence for this is limited, as there was only one semantic domain related to *superlativeness,* and it contained only one word ‘heavy’ that was repeated six times in the corpus.

The analysis also suggests that there are potential news values that might be distinctive of false news. Although these values are not covered by the DNVA framework, some of them are related to concepts discussed by other researchers. One possible news value is ***scepticism*****.** This is

reflected in lexical items that denote an evaluation of credibility, morality, and rationality. The analysis showed that false news writers portrayed established theories or mainstream knowledge as untrue or inauthentic while presenting alternative theories as true and authentic. In addition, the analysis showed that those who believe in mainstream theories were negatively evaluated while those who support alternative theories were positively viewed (e.g., described as brave). The concept of truth has also been recognized as a recurring theme in other linguistic studies on fake news (Dance, 2019; Rashkin et al., 2017). Previous studies have indicated that discrediting normally credible sources and conspiracy theories are predominant themes of fake news (Golbeck et al., 2018), suggesting that these sceptic themes make a story newsworthy according to false news writers. *Scepticism* provides false news writers with a means to present themselves as uncovering hidden truths and possessing knowledge that is beyond the reach of their readers. Moreover, *scepticism* can be appealing to readers who seek to discover the hidden truth. As Masterton (2005) argues, *secrecy* increases the newsworthiness of a story, noting that "when one person or group wants kept secret some information [,] the general public will find important or interesting. Such information is almost always newsworthy" (p. 47).

Moreover, the analysis showed that false news emphasised evidentiality. This relates to Bell's (1991) concept of *facticity*, which suggests that news stories that include facts and figures are newsworthy. However, the analysis showed that false news not only provided facts and figures but provided multiple types of evidence (written, spoken, sensory, deductive) to ensure that the facts presented were accurate and reliable. This finding is consistent with other discourse and linguistic research on fake news (Alba-Juez & Mackenzie, 2019; Dance, 2019; Maci, 2019; Rashkin et al., 2017). Consequently, this finding suggests that ***corroboration*** could be a news value of false news. It can foster scepticism in readers, providing evidence to accept alternative theories and refute mainstream ones. By *corroboration*, false news writers appeal to readers by enabling them to discern the truthfulness of the claims being made. This emphasis on *corroboration* further suggests that false news writers are aware that the information they present is unconventional and potentially controversial. Hence, they invest extra effort in providing ample evidence for their claims.

In addition, the analysis revealed that false news articles placed significant emphasis on providing logical and fact-based reasoning. This is evident in the use of words that convey cause-

effect relationships and those related to science, technology, substances, and materials. While some previous research has suggested that fake news tends to use less technical language (Horne & Adalı, 2017; Vamanu, 2019), other studies align with this case study by suggesting that fake news does incorporate technical language and scientific terms (Alba-Juez & Mackenzie, 2019; Maci, 2019). This suggests that a potential news value of false news is ***causality***, which involves constructing a news event as the result or the cause of another event. This news value can satisfy curious readers who are looking to understand the causes of current and past events or the rationale behind the claims proposed in the false news articles.

Furthermore, as shown in the analysis, false news included many mystical elements, such as religious figures and places, spiritual beliefs, supernatural occurrences, or mythical legends. This is in line with Douglas' (2018) finding that religion is often the subject of fake news and that religious believers are among the targeted audiences of fake news. ***Mysticism*** can thus be a news value of false news, as it appeals to readers' religious beliefs and increases the perceived credibility of the false news report (Bronstein et al., 2019; Pennycook & Rand, 2021). It can provide readers with a sense of comfort and security, especially those with a strong connection to faith and religious practice.

Lastly, the analysis revealed that false news placed emphasis on the aspect of time, specifically focusing on past events and ongoing events that originated in the past. In this way, false news writers claim their possession of genuine knowledge of history, which seems inaccessible to their readers. This suggests that false news utilizes historical narratives, incorporating historical facts, figures, and records, to underscore their authenticity and establish ***historicity*** as a news value. This news value appeals to readers as referencing historical events enables them to gain a comprehensive understanding of the timeline of an event and gives them the perception that they are being provided with historical evidence.

5.7. Discussion

The section commences with a reflection of the implications that stem from the distinct news values associated with false news. It then delves into the limitations of utilizing key semantic domain analysis and provides recommendations for overcoming these challenges. Additionally, the section highlights directions for future research that arise from the findings.

5.7.1. News Values, Rhetorical Strategies, and Ideological Implications

The prevalence of distinct news values in false news articles can be ascribed to their primary objective of questioning conventional narratives. These news values play a crucial role in achieving this objective. By adopting a sceptical perspective (*scepticism*), these articles utilize a range of persuasive techniques, such as presenting various forms of evidence (*corroboration*), appealing to logical reasoning (*causality*), invoking religious beliefs (mysticism), and showcasing historical insights (*historicity*). The qualitative analysis showed that these values serve as effective rhetorical tools, enhancing the allure, persuasiveness, and credibility of false news articles.

The news values of false news also reflect the potential ideologies that false news writers may attempt to construct. The news value of *scepticism*, supported by the news values of *causality*, *corroboration, historicity, negativity, consonance*, and *unexpectedness*, can contribute to fostering a mindset of distrust towards established knowledge and a preference for alternative theories. This aligns with the notions of *post-truth era* and *information crisis*, where conventional wisdom is undermined (Hancock & Bailenson, 2021). Furthermore, the news value of *mysticism* can reinforce religious ideologies by validating or reinforcing existing faith or emotional responses rooted in religion.

Furthermore, identifying the news values of false news helps to recognize certain characteristics of false news consumers. For instance, religious individuals may be attracted to the news value of *mysticism*, while sceptics may be intrigued by the news values of *scepticism, corroboration, causality*, and *historicity*. In addition, those who are emotionally responsive may be more susceptible to the news value of *negativity* and *unexpectedness*, which can greatly impact their emotions and reactions.

5.7.2. Limitations and Future Research Directions

Although the method of KSDA was helpful in identifying potential news values of false news compared to other news types, it does have its limitations. Firstly, KSDA may not account for words that are highly frequent in isolation but do not belong to any of the top semantic domains. These words can still be explicit indicators of news values. For example, ‘Dr.’, extracted as a keyword in the false news corpus (refer to Appendix 5.1), can establish the news value of

eliteness; however, its meaning was not among the top key semantic domains. Additionally, the analysis showed that many semantic domains cover lexis that is topic-specific but does not explicitly represent news values. Therefore, it may be beneficial to consider collocations of these words to identify news values that might be established in the immediate co-text (cf. Potts et al., 2015). Moreover, the analysis was limited in that it focused only on individual words under each semantic domain and the news values explicitly expressed by these words. However, the analysis does not consider news values that are implicitly expressed by longer chunks of language, such as phrases, clauses, or sentences.

Furthermore, through concordance analysis, I was able to identify frequent patterns of usage among lexis belonging to a certain semantic domain; such patterns can be related to news values. However, it should be noted that not all lexis under a semantic domain represent one particular news value. Some words may be multi-functional and can have different meanings depending on the context. For instance, not all instances of the semantic domain ***time*** establish the news value of *historicity*. The word 'time', in Example 45, refers to the summer holidays when kids have more leisure time to start fires.

> (45) School holidays are a prime **time** [time] for fire bugs, but especially over summer. The kids have got time to get out there and light, and the most dangerous adults choose hot days. *(A False 20)*

Finally, it is important to acknowledge that the accuracy of the semantic tagger used in the analysis may have limitations. For example, the word 'shot' was tagged as ***warfare, defence; weapons*** whereas most instances of this word refer to the flu shot (see Figure 5-2). Therefore, it is important to consider the concordances and collocations of the semantic tags.

Figure 5-2 Sample Concordances of the Semantic Domain 'Warfare, defence; weapons'

ng children who received FLU	SHOTS		while health authorities madl	115	More
authorities madly push more	shots		for the coming flu season A 2	116	More
e : The study finds that flu	shots		quite literally inactivate pa	117	More
n to avoid administering flu	shots		this year and therefore reduc	118	More
] . Those who push flu	shots		may be worsening the coronavi	119	More
at those who are pushing flu	shots		for children may in fact be w	120	More
accessories to murder . Flu	shots		are being hyped to suppress i	121	More
n coronavirus infections Flu	shots		, it seems , are one more way	122	More
surge back this Fall as flu	shots		work to suppress the immune s	123	More
that the people pushing flu	shots		are medical murderers , and m	124	More
nstream Pentagon Study : Flu	Shot		Raises Risk of Coronavirus by	125	More
flu shotyou should get a flu	shot		. Setting safety and efficacy	126	More
Andersons claim that the flu	shot		will help people fight COVID-	127	More
n able to find assessing flu	shots		and coronavirus is a 2020 US	128	More
tudy that found that the flu	shot		INCREASES the risks from coro	129	More
tory infections from the flu	shot		: 2018 CDC Study : Flu shots	130	More
shot : 2018 CDC Study : Flu	shots		increase risk of non-flu acut	131	More

Despite these limitations, the implications of this case study are crucial for future research. Not only does it provide valuable insights into the news values of false news, but it also sheds light on the discourse of news values in general. This study highlights the potential for using an inductive corpus-assisted approach to examine the news values of both false and true news discourse. Therefore, further studies on using corpus methods such as frequency lists and collocations to identify potential news values are suggested.

5.8. Conclusion

This case study presented a unique methodology for investigating the news values associated with false news. Rather than employing corpus analysis methods to validate predetermined news values, this study utilized the corpus method of key semantic domains to uncover potential news values. The analysis of the key semantic domains in the false news corpus revealed seven major semantic groups: connecting ideas, evidentiality, evaluation, religion and supernatural, science, time, and news topics. Examining the concordances of the lexis under these semantic groups showed that false news emphasises news values that have already been discussed in previous literature, such as *negativity, unexpectedness*, and *consonance*. The analysis also revealed that *scepticism* (i.e., discrediting credible sources), *corroboration* (i.e., providing various types of evidence: sensory, spoken, written, or deductive), *causality* (i.e., explaining the cause and effect of an event and connecting it with other events), *mysticism* (i.e., reference to religious, spiritual, mysterious, or mythical elements), and *historicity* (emphasising historical relevance of current events and claiming historical authenticity) could be news values that are distinct to false news.

The analysis highlighted how news values function as rhetorical tools in false news narratives. By accentuating these values, false news creators can enhance the appeal, engagement, memorability, and persuasiveness of their stories. Additionally, placing emphasis on news values enables false news creators to bolster their credibility. As Bednarek and Caple (2017) state, resources that construct newsworthiness "enhance truthfulness, credibility, and objectivity" (p. 104). Furthermore, the analysis illustrated that these news values play a crucial role in shaping and reinforcing ideologies, such as promoting scepticism towards established knowledge and influencing religious beliefs among audiences exposed to false news. Additionally, the analysis allowed for speculation on the potential target audiences of false news, including sceptics,

religious individuals, and those who react emotionally. Ultimately, these news values are likely contributing factors to the widespread dissemination of false news in the digital age.

Chapter 6: Stylistic Patterns in Online False News Articles

6.1. Introduction

This chapter introduces the third case study, which examines the stylistic patterns found in online false news articles. Stylistic patterns, in this context, refer to the consistent co-occurrence of specific lexico-grammatical features that collectively contribute to distinct communicative functions. The stylistic variation between false news and true news is viewed as a form of register variation, assuming differences in the communicative context of production between false and true news. Therefore, understanding these stylistic patterns requires the use of techniques from register analysis to describe and ultimately decipher such variation.

Analysing false news from a register perspective also helps address a limitation in prior research that has attempted to automatically differentiate online fake news from true news based on style and lexico-grammatical features. Past studies have compared online fake news to mainstream news articles like broadsheets (e.g., Horne & Adalı, 2017; Rashkin et al., 2017; Volkova et al., 2017); however, the potential influence of other sources of stylistic variation, acting as confounding variables, has not been considered. Online false news is situated in a distinct communicative context with unique formatting, authors, audiences, and objectives compared to broadsheets. As a result, the identified differences may be attributed to contextual factors rather than the factual accuracy of the text or the author's communicative intent.

To lessen the impact of the confounding contextual factors, this study compares false news articles circulated online with four different types of news: broadsheets, tabloids, web-based publications, and news blogs, all sourced from reputable sources. The main objective is to investigate whether there are any similarities or differences in the stylistic patterns employed by false news and these four news types. To achieve this goal, the study utilizes multidimensional analysis (MDA), an empirical approach developed by Biber (1988) that allows for the analysis of systematic register variation. Additionally, this study aims to shed light on how the stylistic patterns used by false news writers contribute to the appeal of their articles.

This chapter is structured as follows: First, it provides an overview of the methodological principles of MDA. Following this, a review of previous studies on the linguistic features and stylistic patterns found in different types of news reports is presented. This review serves to provide a solid foundation for interpreting the results of the analysis. Next, the chapter outlines the corpus used in this study and details the analytical procedure that was employed to examine it. Lastly, the chapter presents and discusses the analysis results.

6.2. Methodological Principles of Multidimensional Analysis

This section discusses the methodological principles of multidimensional analysis. It starts by presenting the theoretical foundations of MDA. Then, it outlines the methodological procedure of MDA, specifically referring to Biber's (1988) analysis. Finally, it explores the various approaches and current applications of MDA to address different research questions.

6.2.1. Theoretical Foundation of Multidimensional Analysis

The MDA approach was developed by Biber (1988) to analyse linguistic variation among different registers of spoken and written English. In other words, Biber explored how different registers (e.g., press reportage, academic prose, telephone conversations, etc.) differ in their linguistic features (e.g., tense and aspect markers, pronouns, passives, etc.). According to Biber and Conrad (2019), a register is described in terms of three components: (a) situational characteristics, such as the communicative purpose, action performed, channel of communication, and relationship between participants, (b) the linguistic (lexical and grammatical) features, and (c) the functional relationship between the linguistic features and situational characteristics.

Based on register theory, MDA assumes that the co-occurrence of certain linguistic features in a text reflects a specific communicative function. Thus, different texts with different communicative functions and situational characteristics tend to differ in terms of their co-occurring linguistic features (Biber, 1988; Biber et al., 2007; Biber & Conrad, 2019). For example, *contractions*, *false starts*, and *generalised content nouns* (e.g., ‘thing’) highly co-occur in real-time conversations while *past tense*, *perfect aspect verbs*, *third-person pronouns*, and *public verbs* frequently co-occur in fiction.

MDA has three major goals: (a) to empirically identify salient co-occurring patterns of linguistic features through multivariate stylistic analysis, (b) to compare different registers in terms of these co-occurring patterns, which are understood as dimensions of variation, and (c) to analyse what situational, social, and cognitive functions these dimensions represent. Accordingly, there are three main characteristics of MDA. First, it is inherently comparative as it assumes that different registers differ linguistically and functionally. Second, it is empirical and quantitative as it focuses on how frequently certain linguistic features are used in different registers. Third, it is also qualitative as it aims to explain the communicative functions shared by co-occurring patterns of linguistic features (Biber et al., 2007).

6.2.2. Methodological Procedure of Multidimensional Analysis

Table 6.1, adapted from Conrad and Biber (2001), provides an overview of the eight steps involved in MDA. Below, these steps are explained in more detail with reference to Biber's (1988) analysis.

Table 6.1 Methodological Steps of MDA

1	A representative corpus is designed based on previous research and analysis. The situational characteristics of the corpus/register are recorded, such as purpose, participants, and mode.
2	Research is conducted to determine the linguistic features that will be included in the analysis and to identify their functions.
3	Part of Speech (POS) tagging programs are developed to identify the linguistic features.
4	The corpus is automatically tagged. The accuracy of the tagging is reviewed, and systematic errors are analysed and remediated.
5	The frequency counts of each linguistic feature in each corpus text are computed.
6	A factor analysis of each frequency count is conducted to identify the co-occurrence patterns (i.e., dimensions) among the linguistic features.
7	An analysis of the functional representation of the factors (i.e., dimensions) from the factor analysis is conducted.
8	A dimension score for each text with respect to each dimension is calculated. The mean dimension score for each register is then calculated to analyse the variation among the registers under investigation.

1) Designing a Representative Corpus

The first task of MDA is designing a corpus that is representative of the registers under study (See Chapter 3, Section 3.4). For factor analysis, the minimum corpus sample size should be a function of the number of variables (i.e., linguistic features) included in the factor analysis. In

other words, a large sample size is needed for a large set of features. Egbert (2019) suggests a rule of thumb: a minimum ratio of five texts per linguistic feature, but not fewer than 100 texts in total. If the size of the corpus is fixed, the number of features included in the analysis must be reconsidered.

2) Identifying Linguistic Features

The second task of MDA is identifying the linguistic features that will be used in the analysis. Identifying the features requires extensive research or a review of studies on linguistic features and their functions in the registers under analysis. However, in general, the aim is that the list of features should be as inclusive as possible, including lexical classes, grammatical categories, and syntactic constructions that might have communicative functions or be used to different degrees in different registers (Biber & Conrad, 2019). The choice of the linguistic features is important because they influence the success of the factor analysis as well as the interpretability of the extracted dimensions (Egbert & Staples, 2019). Biber's (1988) list includes 67 linguistic features categorised into 16 major grammatical and functional groups. Throughout the history of MDA, researchers have used Biber's list but also added certain features related to the domains under investigation (Egbert & Staples, 2019).

3-5) Grammatical Tagging of Features and Frequency Counts

The third task is tagging words in the texts according to the identified linguistic features. Throughout the history of MDA, multiple taggers have been used, including Biber's (1988) tagger, the Constituent Likelihood Automatic Word-tagging System (CLAWS) (Garside, 1987), the Stanford Tagger (Toutanova et al., 2003), and the Multidimensional Analysis Tagger (MAT) (Nini, 2019). After tagging, the frequency of each linguistic feature is automatically counted in each text. These frequency counts are then normalised to a common basis. This is important to enable comparison among texts of unequal length.

MAT, the tagging program used in this case study, was developed by Nini (2015–2019). It is a program that replicates Biber's (1988–1989) research. The MAT first tags texts using the Stanford Tagger (Toutanova et al., 2003) and then applies the algorithms provided in Biber (1988) to compute the frequencies of the 67 linguistic features. The program then plots the

analysed texts along Biber's six dimensions and labels them according to Biber's text types (see Section 6.2.3).

6) Identifying Factors

The most crucial step of MDA is identifying the dimensions, which Biber (1988) defines as "bundles of linguistic features which co-occur in texts because they work together to perform some underlying function" (p. 55). The statistical analysis used to identify the dimensions is called factor analysis, and each pattern of co-occurring features is called a factor (i.e., a dimension). Factor analysis, known as a dimension reduction technique, reduces a large number of variables (i.e., linguistic features) measured over a set of observations (i.e., texts) to a small set of derived and underlying variables (i.e., dimensions) (Biber & Conrad, 2019). In other words, factor analysis identifies sets of variables that are correlated with each other, but that are distinct from other sets of variables (Egbert & Staples, 2019). Identifying factors involves analysing the extent to which certain features tend to vary in similar ways (Conrad & Biber, 2001). For instance, if certain features are frequently found together in some texts but are rarely seen in others, they are said to have a high level of shared variance.

When extracting factors, the first steps involve: (a) examining the factorability of the variables, (b) evaluating the adequacy of the initial extracted factors, and (c) determining the number of factors to be extracted for the analysis (Egbert & Staples, 2019). Factorability is determined by evaluating the strength of correlations between variables and utilizing the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (Egbert & Staples, 2019). Kaiser (1974) provided criteria to assess the factorability of a sample based on the KMO score as follows: .90 or above is considered marvellous, .80 is meritorious, .70 is middling, .60 is mediocre, .50 is miserable, and below .50 is deemed unacceptable. Factor adequacy is typically assessed by measuring variable communalities. Communalities represent the variance in a variable explained by the factors. Higher communalities indicate greater variance explained (Egbert & Staples, 2019). To decide on the number of factors to be extracted and analysed, eigenvalues and the scree plot are examined. Eigenvalues indicate the proportion of variance explained by each factor. The scree plot visually displays the eigenvalues, showing a break (an elbow) in the dots followed by a flattening. Factors before the elbow on the scree plot are usually considered for further analysis (Egbert & Staples, 2019).

After identifying potential numbers of factors (or factor solutions), the researcher should consider each solution. Two main considerations help determine the best solution. Firstly, the variable communalities in each factor and the total variance explained should be examined. A solution with higher variable communalities and total explained variance is preferred. Secondly, the interpretability of the functions related to each factor should be considered. If one factor solution is more easily interpreted, it should be chosen as the preferable solution (Egbert & Staples, 2019).

Each linguistic feature has some relation (or weight) to each factor, and the importance of this relation is represented by factor loadings. Conrad and Biber (2001) define factor loading as "the amount of variance that a feature has in common with the total pool of shared variance accounted for by a factor" (p. 264). Factor loadings can range from 0.0 (no relationship) to 1.0 (a perfect correlation). The size of the loading shows the importance of the relationship between a feature and the group of co-occurring features under one factor. In other words, it shows the extent to which a feature is representative of a dimension (Conrad & Biber, 2001).

Although each linguistic feature has a loading on each factor, only features associated with substantial loadings are generally considered when interpreting a factor. Notably, loadings can be positive or negative, indicating that two sets of features occur in a complementary pattern. That is, when the positive features occur frequently together, the negative features likely occur less frequently (Biber & Conrad, 2019). In Biber's (1988) analysis, features with loadings of more than the absolute value of .30 were taken into consideration during interpretation.

7-8) Interpreting Factors as Dimensions of Variation and Calculating Dimension Scores

Interpreting the extracted factors as functional dimensions involves two steps: (1) analysing the communicative functions that the co-occurring features of a factor share based on the factor loadings and (2) analysing the similarities and differences among the registers under investigation in terms of the extracted factors based on the factor scores (Biber & Conrad, 2019). The first step can be illustrated by Biber's (1988) Dimension 2: *narrative vs. non-narrative.* Dimension 2 includes the following positively weighted features: *past tense verbs, third-person pronouns, perfect aspect verbs, public verbs*, and *present participial clauses* among others. These features are typically used in past-time narration. *Past tense* and *perfect aspect verbs* describe

past events, *third-person pronouns* refer to actors in the events, *public verbs* (e.g., ‘say’, ‘tell’) are used as indirect speech acts, *present participial clauses* provide descriptions and imagery. The absence of these features indicates that a text is non-narrative.

The second step of the interpretation involves considering the similarities and differences among registers with respect to a dimension and determining whether the similarities and differences are consistent with the previously suggested interpretation of the dimension (e.g., narrative vs. non-narrative). This step is accomplished by calculating a dimension score for each text and then calculating the mean dimension score for each register.

There are two main types of factor score calculation methods: refined and non-refined methods (DiStefano et al., 2009). Non-refined methods involve simple procedures to determine the placement of each individual item on the factor distribution. The most frequently used non-refined methods are summing scores by factors, summing scores using a cut-off value, or using standardised scores. Non-refined methods have the advantage that they are easy to produce and interpret. Refined methods, on the other hand, involve more sophisticated and technical procedures to provide standardised scores for each item. Examples of refined methods are regression scores and Bartlett scores. Refined methods aim to ensure accuracy by generating factor scores that strongly align with a specific factor and by obtaining unbiased estimates of the true factor scores. Moreover, these methods seek to preserve the relationships between factors, thereby maintaining their interconnectedness (DiStefano et al., 2009).

Biber (1988) used a non-refined method by first standardizing the frequencies of salient features, which can be accomplished by calculating z-scores, and then by calculating the dimension score for each text by subtracting the sum of the standardized negative scores from the sum of the standardized positive scores. The dimension score for each register can be obtained by calculating the mean of that register’s texts’ scores.

By plotting the average dimension score for each register across all dimensions, it becomes possible to compare the different registers, identify their distinct linguistic characteristics, and gain a more comprehensive understanding of each dimension. For instance, in Biber's (1988) Dimension 2: *‘narrative vs. non-narrative’*, fiction and biographies tend to have high positive scores, while academic prose and official documents exhibit notably negative scores. Utilizing

statistical techniques like ANOVA, it is feasible to assess whether the variations observed in dimension scores among registers are statistically significant (Biber & Conrad, 2019).

6.2.3. Biber's (1988) Dimensions of English Variation and Text Types

In his analysis, Biber (1988) identified six major dimensions of English register variation. Table 6.2 shows the label of each dimension and provides a description and examples of registers and features associated with it.

Table 6.2 Biber's (1988) Dimensions of Variation in English

Dimension	Description
Dimension 1: Involved vs. Informational Production	High scores on this dimension are associated with affective, interactional, and generalised content, while low scores mark high information density and precise informational content. Involved texts include many verbs and pronouns (among other features), as, for example, in casual conversation. Informationally dense texts include many nouns, longer words, and adjectives (among other features), as, for example, in academic prose.
Dimension 2: Narrative vs. Non-Narrative Discourse	High scores on this dimension represent past-time narration, while low scores are associated with non-narrative content. Narrative texts such as novels include many past tense verbs and third-person pronouns.
Dimension 3: Elaborated vs. Situation-Dependent Reference	High scores on this dimension mean that a text is context-independent, where referents are identified and elaborated. Such referentiality indicates that the text is densely integrated and informational. Low scores indicate that the text is situation-dependent, where references to places and times outside the text are frequently used. Elaborated texts, such as academic prose, have many nominalisations whereas situation-dependent texts have many adverbs.
Dimension 4: Overt Expression of Argumentation	High scores on this dimension are associated with texts with linguistically overt attempts of persuasion, while low scores indicate a lack of such attempts. Texts with high scores, such as editorials, include many modal verbs.
Dimension 5: Abstract vs. Non-Abstract Style	High scores on this dimension mean that a text provides information with technical, abstract, impersonal, and formal language. Low scores are associated with texts that present information in a more personal and non-abstract style. Abstract texts, such as scientific discourse, have many passive clauses and conjuncts.
Dimension 6: On-Line Informational Elaboration Marking Stance	High scores on this dimension mean that a text is informational but produced under time constraints, as, for example, in prepared and spontaneous speech. Texts with high scores involve many post-modifications of noun phrases.

Biber (1989) conducted a cluster analysis to group the texts in his 1988 corpus according to their dimension scores. He defined each cluster of texts as a *text type* with common linguistic features and similar dimension scores. According to Biber, there are eight English text types: *intimate interpersonal interaction, informational interaction, scientific exposition, learned exposition, imaginative narrative, general narrative exposition, situated reportage*, and *involved persuasion*.

For example, *general narrative exposition* texts have high scores on Dimension 1 and low scores on Dimension 2, as in editorials and biographies, whereas *involved persuasion* texts, such as interviews and spontaneous speeches, have high scores on Dimension 4. However, Biber (2013) stresses the need to replicate this research using larger corpora because the corpora he analysed in 1988–1989 was the most inclusive corpus he could compile, and there were no other corpora at the time.

6.2.4. Approaches to Multidimensional Analysis and its Current Applications

There are two approaches to conducting MDA: (a) applying dimensions identified in previous work, such as Biber (1988), to analyse new texts and (b) conducting a new and complete MDA (Conrad & Biber, 2001). Methodologically, the former omits the methodological steps of 2, 6, and 7 while the latter includes all eight steps. Originally, MDA was developed to analyse linguistic variation in English written and spoken registers. However, both approaches of MDA are currently utilized to explore linguistic variation within sub-registers or to identify the stylistic patterns in specific text varieties.

Deciding whether to conduct a new MDA or apply Biber's dimensions, for example, depends on the nature of the research. The two approaches provide valuable but different perspectives. On the one hand, applying the established dimensions to new texts allows researchers to compare the texts in question to other registers in English (Conrad & Biber, 2001). For instance, Nini (2017) used Biber's dimensions to analyse threatening, abusive, and defaming texts that were treated as evidence in forensic cases. He found that such malicious texts belong to the *involved persuasion* text type, overtly expressing modality.

On the other hand, conducting a new MDA allows researchers to investigate a specialised discourse domain and explore the dimensions of (stylistic) variation within it (Conrad & Biber, 2001), especially since Biber's (1988) dimensions were derived from a relatively small corpus of limited British registers. For example, Gray (2013) conducted a new MDA to analyse variation among three sub-registers of research articles: theoretical, qualitative, and quantitative, in six disciplines: philosophy, history, applied linguistics, political science, biology, and physics. The analysis revealed four dimensions: *academic involvement and elaboration vs. informational density*, *contextualized narration vs. procedural discourse*, *human vs. non-human focus*, and

academese. Gray concluded that discipline is only one of the parameters that correspond to linguistic variation. Other parameters included the nature of evidence, the presence or absence of data, qualitative vs. quantitative vs. theoretical research paradigms, and the object of study.

6.3. Previous Work on the Linguistic Features and Stylistic Patterns in News Reports

Before reviewing the linguistic features and stylistic patterns of news reports, it is crucial to discuss the situational characteristics that typically surround the production of news reports. The situational characteristics related to fake news were previously discussed in Chapter 2. Both discussions are essential as they enable us to comprehend the reasoning behind the utilization of specific linguistic features and stylistic patterns in news reports, as well as the communicative functions these features and patterns serve. Therefore, the initial section of this literature review will examine these characteristics, while the subsequent sections will review previous research studies that have investigated the linguistic features and stylistic patterns employed in various types of news reports, including printed news reports, web-based news reports, and news blog posts. Notably, these studies have taken a comparative approach, which is emphasised as necessary by Biber and Conrad (2019) in order to identify the distinctive linguistic features of a particular register. While these studies have compared news articles with registers other than the main focus of this study (e.g., academic articles), they have provided valuable insights into the essential linguistic features found in news articles.

6.3.1 Situational Characteristics of News Reports

News reports are typically written by journalists and are aimed at readers interested in reading about recent newsworthy events. Their primary communicative purpose is to inform and report events. Professional revision and editing are commonly carried out on news reports. News reports are time-sensitive in that they must be published and updated as quickly as possible (Biber & Conrad, 2019).

News articles can differ significantly depending on their medium (e.g., print newspapers or online publications), target audience (e.g., elites, well-educated individuals, the general public, or specialists), topics (e.g., political news, sports news, entertainment/lifestyle, etc.), and communicative purposes (reporting an event or providing commentary and analysis, such as in

editorials, readers' comments, and letters to the editor). Consequently, it is expected that their linguistic features will also vary. The following paragraphs will provide a brief overview of the situational characteristics of the news types being examined: broadsheets, tabloids, web-based publications, and news blogs.

In the context of printed newspapers, Breuer and Napthine (2008) differentiate between broadsheets and tabloids as follows: Tabloids cater to a wide popular audience. They are usually easy to read, are characterised by accessible and direct language, and are mainly descriptive with some analyses. Articles in tabloids tend to be short and make extensive use of photographs and illustrations. Tabloid news articles are primarily crafted with a dual objective: to captivate and enlighten the audience. Consequently, they frequently adopt a sensationalist approach, embellishing and amplifying stories to stir up emotions like fear and anxiety. This tactic is often employed to align with and influence public opinion. Tabloids cover news stories about celebrities, local news, sports, and personal stories. In addition, they focus on stories that incite public sentiments, such as AIDS, asylum seekers, paedophilia, and homosexuality.

By contrast, broadsheets are written for professionals, businesspeople, and well-educated readers. Thus, the language in broadsheets is more sophisticated than that in tabloids. Unlike tabloids, which tend to be descriptive, broadsheets focus on analysis and attempt to cover a variety of viewpoints, comments, and perspectives in their news reports. Articles in broadsheets are usually longer and contain fewer photographs compared to tabloids. Broadsheets cover a wide range of issues, including politics, business, and social concerns, as well as both national and international news (Breuer & Napthine, 2008).

The advancement of technology has played a pivotal role in transforming the landscape of news dissemination. In today's digital age, news has found its way onto websites, blogs, microblogs, and social media platforms. Many traditional newspapers as well as TV news channels have adapted to this shift by establishing online versions and maintaining a presence on social media. As a result, a considerable number of individuals now access their preferred news sources online.

The rise of news blogs can be traced back to the aftermath of September 11, 2001, when there arose a pressing need for rapid and accurate accounts of the events that unfolded (Allan, 2011). This style of journalism has been labelled with various terms, such as personal journalism, do-it-

yourself journalism, black-market journalism, we media, and postmodern journalism (Wall, 2005).

One distinctive characteristic of blogs is their ability to foster a sense of community among users who share specific social identities (Myers, 2010). According to Wall (2005), news blogs encompass a form of journalism that emphasises both personalization and audience participation in content creation. While traditional news outlets regard their audience as passive customers receiving news as a product, news blogs see their audience as supporters or even companions. Wall (2005) notes that personalisation and the writer-audience relationship build the credibility of the news stories; the more personal and open the writer is, the more trustworthy they become.

Blogs have also challenged the credibility standards established by traditional media, as bloggers are not bound by the same fact-checking principles (Myers, 2010). This has resulted in the emergence of black-market journalism, wherein unsanctioned voices can be heard (Wall, 2005). In addition, bloggers prioritize attention over accuracy, as evidenced by their focus on links, hits, track-backs, and mentions in popular blogs, rather than the credibility of their evidence and sources (Myers, 2010). Furthermore, bloggers tend to regard facts as provisional, time-bound, and limited, and thus consider them as factoids (Myers, 2010).

Another characteristic of blogs is that they are placeless but time-bound. Myers (2010) found that there were fewer references in blogs to the place of writing, contrary to those found in newspaper articles. He argued that blogs are placeless as bloggers do not locate themselves, and they assume that they have the global vision of the blogsphere and see everywhere at once. However, Myers argued that blogs are uniquely time-bound as they are automatically stamped with a date and usually a time of publication.

In recent years, there has been a substantial increase in news premier blogs that are specialized in specific topics. These blogs often provide a more in-depth analysis of particular subjects, at times surpassing traditional news outlets in providing a comprehensive understanding of various issues. Specialized news blogs can cover anything from politics and current events to sports and entertainment. They can also provide a platform for independent journalists to share their work and opinions.

6.3.2. Printed News Reports

Biber and Conrad (2019) conducted a study comparing the register of news reports to academic prose as well as editorials. The authors found that there are notable linguistic differences between news reports and academic writing. These differences may be due to variations in their purposes and topics. One major difference is the use of tense. Academic writing tends to favour the *past tense*, whereas news reports employ both *past* and *present tenses* equally. News reports utilize the *past tense* to narrate recent events and the *present tense* to describe current happenings. Additionally, news reports make frequent use of *place* and *time adverbials* to clarify the sequence of events. This choice of linguistic features aligns with the primary objective of news reports, which is to report on newsworthy events. Another distinction between the two registers is the prevalence of *nominalizations* in academic writing. This can be attributed to the fact that academic writing often discusses abstract concepts and patterns. Furthermore, the way sources of information are attributed differs between news reports and academic writing. In academic writing, precise details such as names and dates are provided, whereas news reports tend to attribute information to vague sources such as documents, organizations, or unnamed individuals. Lastly, news reports use the *passive voice* less frequently compared to academic writing. This can be attributed to the fact that news reports typically focus on agents and their actions as the main points of interest.

According to Biber and Conrad (2019), news reports typically do not contain an overt expression of stance, but rather just report what happened. However, editorials are meant to express an opinion and argue in support of it; thus, they are characterised by an overt expression of stance. In their study, Biber and Conrad found that *modals* and *conditionals* are more common in editorials than in news reports. *Modals* are used to direct readers on how to act and to predict future events, while *conditionals* are used to discuss hypotheticals and to predict future consequences depending on certain actions. To summarize, the key features of news articles include the utilization of both *past* and *present tenses*, the inclusion of *place* and *time adverbials*, less specific attributions of sources, a reduced frequency of *passive voice* usage, and a tendency to avoid overt expressions of stance.

6.3.3. Web-based News Reports

Biber and Egbert (2018) studied the linguistic variation across six online registers: narrative, information description/explanation, opinion, information persuasion, how-to instructions, and lyrical texts. News reports were a sub-genre of narrative registers. The authors conducted an MDA and key features and keywords analyses.

The study findings shed light on the distinctive style of online news reports relative to other online registers. Firstly, news reports place a greater emphasis on providing factual information and reporting events rather than delving into individuals' thoughts and words. Secondly, in comparison to other narrative registers, news reports exhibit the highest level of informational density, conveying a significant amount of information within their content. Thirdly, news reports stand out as the register that employs reported communication the most. This can be attributed to their frequent use of quoted materials to recount events, as well as people's reactions to them. Fourthly, news reports tend to lack audience involvement and the characteristics commonly associated with oral discourse. Fifthly, news reports typically do not establish a particular stance or take a clear position on a specific subject. Finally, news reports were found to have an inclination towards describing humans, as they occasionally portray news characters.

Similar findings were highlighted by Biber and Egbert's (2018) key features analysis, which identifies the linguistic features that occur more frequently in a corpus compared to another corpus. The analysis revealed that *communication verbs* were more prevalent in news reports than in other online registers. Another key feature of news reports is the inclusion of *proper nouns* to refer to the names of people, places, and organisations, which are prevalent in news reports. The analysis also showed that *long words* tended to occur more frequently in news reports, which might be due to the use of *nominalisations* and other special terms that tend to be long. Finally, the analysis showed that *first-* and *second-person pronouns* occurred significantly less frequently in news reports than in other online registers, which might be due to their lack of interaction and involvement.

A similar study is that of Titak and Roberson (2013), who applied MDA to investigate linguistic variation in six online registers: blogs, microblogs, workplace emails, discussion posts and readers' comments, online news articles, and opinion columns. The online news articles collected for this study were from the web-based *New York Times*. Like, Biber and Egbert's (2018) study,

this study found that the online news articles displayed a high level of descriptiveness and informativeness but were less interactive compared to other online registers. Furthermore, online news articles exhibit some features of what the authors called *complex statement of opinion* which involves expressing one's stance while simultaneously reporting the speech of others. In addition, online news articles were found to have both narrative and analytical orientations.

To conclude, the studies discussed above indicated that online news reports resembled their printed counterparts, as they were descriptive and informationally dense registers that relied heavily on reporting facts, public information, and experts' and public figures' speech; in addition, they are not interactive and do not focus on people's opinions and attitudes.

6.3.4. News Blog Posts

There is a dearth of studies specifically examining the language used in news blogs; existing studies primarily explore language in various types of blogs that cover a range of topics, including news. However, these studies have managed to identify some features specific to blogs reporting or commenting on news. For instance, Grieve et al. (2010) employed MDA and cluster analysis to investigate the linguistic variation among American blogs, aiming to classify them into different types. Their analysis revealed three types: personal diary blogs, expert blogs, and commentary blogs. Expert blogs, although less prevalent in the data, were characterized as impersonal, informative, and focused on specific subjects. They bore a resemblance to newspapers and academic articles. Similarly, commentary blogs were also informative but tended to be more personal and directed towards a specific audience. Authors of commentary blogs expressed their opinions on various topics, including current news events. In a similar vein, Daems et al. (2013) examined the linguistic variation among blogs to distinguish between two contrasting occupational sectors: the humanities and sciences. Utilizing MDA, they observed that blogs discussing diverse subjects, including news events, tended to have a more formal tone than a casual one.

Sardinha (2018) conducted an MDA to compare online registers, including Twitter and Facebook posts, emails, webpages, and blogs. The blog corpus was sampled from the most frequently visited blogs, which were news and corporate blogs. The findings revealed that these blogs exhibited a lower degree of overt expression of stance in comparison to other online registers, and they were also identified as the most informational type of online register. This contrasts

with the results of Titak and Roberson's (2013) study, which suggested that blogs were characterized by high levels of personal engagement, interactivity, and narrative. However, it is important to note that Titak and Roberson's study did not specify the sampling method used for their blog corpus, which may have included predominantly personal blogs. Therefore, this discrepancy in findings could be attributed to the composition of their blog corpus. In summary, while the mentioned studies do not solely focus on news blogs, they provide insights into the characteristics of blog posts discussing news events, which tend to be formal, impersonal, informational, and express stance less explicitly.

6.4. Corpus

As discussed in Chapter 3, the complete corpus collected for this study consists of 137 false news articles. Multidimensional analysis, however, generally requires texts that are long enough to measure the relative frequencies of a large set of grammatical features. To ensure the validity of the analysis, texts must be sufficiently long to accurately estimate feature usage rates. Thus, articles containing fewer than 300 words were omitted from the corpus utilized in this case study, leading to the exclusion of 36 articles from the original focus corpus. Table 6.3 shows the breakdown of the focus corpus.

Table 6.3 Breakdown of the MDA Focus Corpus

False News Topics	**No. of Articles**	**Average No. of Words**	**Standard Deviation of No. of Words**	**Total No. of Words**	**Sources**	**Publication Dates**
Climate Change	18	917	713	16500	Online tabloids, news websites, blogs	2011–2020
Vaccination	22	986	554	21695	News websites, blogs	2013–2019
COVID-19	61	931	770	56788	Online tabloids, news websites, blogs	January 2020–December 2020

The Whole Focus Corpus	**101**	**940**	**713**	**94983**	**---**	**---**

The original reference corpus was also modified accordingly, so each reference sub-corpus included 101 articles, none shorter than 300 words. Table 6.4 shows the breakdown of the reference corpus. The finalized corpus utilized for the MDA analysis comprises a total of 548 articles. It is important to note that in the MDA process, the register or text type of the articles is not taken into consideration, as each article is treated as a single observation.

Table 6.4 Breakdown of the MDA Reference Corpus

Reference Sub-corpora	**No. of Articles on Climate Change**	**No. of Articles on Vaccination**	**No. of Articles on Covid-19**	**Total No. of Articles**	**Average No. of Words**	**Standard Deviation of No. of Words**	**Total No. of Words**
Broadsheets	18 articles for each sub-corpus	22 articles for each sub-corpus	61 articles for each sub-corpus	101	739	266	74670
Tabloids				101	660	233	66630
Web-based Publications				101	698	260	70472
Blogs				101	613	182	61933
The Whole Reference Corpus	**72**	**88**	**244**	**404**	**677**	**241**	**273705**

6.5. Analytical Procedure

The aim of this study is to identify the stylistic patterns that characterise a range of different news types and identify which pattern distinguishes false news from other types of news. To accomplish this, the study utilized Biber's (1988) multidimensional analysis procedure, as outlined in Section 2. This involved identifying dimensions of linguistic variations in both the focus and reference corpora, where each pole in each dimension was interpreted as representing a specific stylistic pattern. The study then examined whether false news articles exhibited distinct usage of these identified stylistic patterns. However, some modifications were made to Biber's procedure to align with the objectives of this study. The modified procedure consisted of four

main phases: (1) selecting features for factor analysis, (2) extracting dimensions and calculating dimension scores, (3) interpreting the dimensions, and (4) comparing the false news corpus to the four reference sub-corpora across each dimension.

6.5.1. Selecting Linguistic Features for the Factor Analysis

The data were tagged using the MAT, which tags texts according to 67 lexico-grammatical features listed in Nini (2015). The tagger produces multiple files that replicate Biber's (1988) analysis. For the purpose of this study, the spreadsheets with the normalized frequencies of each feature in each article were used for the factor analysis, and the grammatically annotated versions of the original texts were used for the qualitative analysis. The features whose normalized frequencies were less than 1 (per 1000 words) were excluded from the analysis as these infrequent occurrences would only add unnecessary noise to the data, making the data meaningless and corrupted. In addition, the feature *type/token ratio* (TTR) was excluded because it depends on the text's length, and it cannot be used to compare texts of different lengths (De Cock et al., 1998). Although the MAT allows users to choose the first X number of words to calculate the TTR, which could be the length of the shortest text in the corpus, some texts might have a higher/lower TTR in their first X words than in the remaining words. This inconsistency indicates that TTR may not accurately represent the texts. Therefore, TTR was excluded, given the problematic issues with this measure. Consequently, 52 features were included in the analysis (see Appendix 6.1).

6.5.2. Extracting the Dimensions and Calculating the Dimension Scores

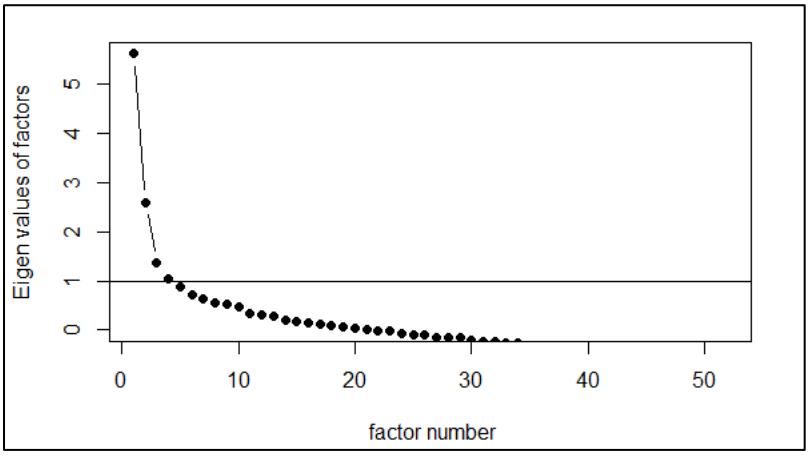
The factor analysis used in this study was conducted using the 'psych' package (Revelle, 2022) in R (R Core Team, 2022), following the guidelines provided by Egbert and Staples (2019). The R code is available in Appendix 6.2. To determine the factorability of the data, the KMO Measure of Sampling Adequacy was used. The KMO statistic was 0.64, which, according to Kaiser (1974), is mediocre. Because MDA is exploratory, an exploratory factor analysis (using the extraction method: principal axis factoring) was conducted. The associated scree plot (see Figure 6-1) shows a clear break between factors 3 and 4; therefore, a three-factor solution was selected. Promax rotation, the most commonly used type of rotation in MDA, was performed. It allows for oblique structure; that is, it allows for correlations among the factors, a desirable

feature given that linguistic variables are usually correlated with each other (Egbert & Staples, 2019).

The cumulative percentage of the shared variance accounted for by the three-factor solution was 19 %, indicating a relatively modest amount of variance accounted for. Along with the mediocre KMO, the cumulative percentage suggests that the dataset does not exhibit clear patterns of stylistic variation. This uncertainty could be attributed in part to the relatively small corpus size, text length, and the homogeneity of the registers analysed, all of which pertain to forms of news writing. Furthermore, these results are not incomparable with previous MDA studies, which often successfully analyse and interpret dimensions that account for a relatively limited amount of variation explained (cf. Egbert & Staples's (2019) review of previous MDA studies). As is the case in these previous studies, the three dimensions identified in this study are interpretable from a stylistic standpoint, displaying familiar patterns of variation. Therefore, these three dimensions have been retained for further analysis.

Only variables with an absolute factor loading above .30 were included in the analysis, and the variables that reached this threshold were included only on the factor where they loaded the most strongly. Then, the features that loaded positively on each factor were separated from those that loaded negatively. The regression method was used to calculate the dimension scores for each news register[8]. The regression scores for each text were computed using R, and the mean score for each register was subsequently calculated (see Appendix 6.3).

Figure 6-1 Scree Plot Representing the Extracted Factors



[8] Details about how regression factor scores are calculated can be found in DiStefano et al. (2009).

6.5.3. Interpreting the Dimensions

To ensure a systematic interpretation of the three dimensions, the following steps were taken:

1) The first step involved comparing the extracted dimensions with Biber's (1988) dimensions to identify any similarities. This comparative analysis was instrumental in gaining a general understanding of these dimensions, considering Biber's dimensions are widely recognized and have been utilized in studying linguistic variation across diverse texts (cf. studies outlined in Conrad & Biber's (2001) book).
2) The second step consisted of reviewing previous studies, some of which were on news writing, to understand the communicative functions shared by the features within each dimension. This also helped identify text types or registers that extensively used these features.
3) Wordlists corresponding to specific features were created using regular expressions in AntConc (Anthony, 2022) and NotePad++. This step aimed to localize the interpretation, i.e., gain an accurate understanding of the communicative functions of these features in news.
4) Concordance analysis was conducted to further understand the communicative functions of certain features in their context. Steps 3 and 4 were necessary in instances where the features deviated from those identified in Biber's (1988) dimensions or were typically uncommon in news articles.
5) The final step encompassed qualitatively analysing texts with the highest and lowest dimension scores. This evaluation helped assess whether the interpretation of each dimension was supported by these texts.

6.5.4. Comparing the False News Corpus to the Reference Sub-Corpora Along the Dimensions

A one-way ANOVA was conducted to determine whether the differences among the five registers were significant. Tukey post-hoc tests were used to make specific follow-up comparisons to identify which reference sub-corpus false news was significantly different from. In order to understand the relationship between the false news corpus and other reference sub-corpora, false news texts with the median dimension scores were qualitatively analysed, as they are the most representative samples.

6.6. Analysis

This section will present the three dimensions separately. First, the communicative functions of the features of each dimension will be discussed. Then, the distribution of the corpora along each dimension will be examined, along with a qualitative analysis of texts with the highest and lowest dimension scores. Finally, the focus will shift towards false news's position on the three dimensions.

6.6.1. Dimension 1: Involved vs. Informational Styles

Table 6.5. illustrates Dimension 1's salient positive and negative features. Dimension 1 in this study is similar to Biber's (1988) Dimension 1: *involved vs. informational.* Most of the positive and negative features overlap with those that are highly loaded on Biber's Dimension 1. The features that overlap with Biber's Dimension 1 are in bold. Biber (1988) states that the opposition between the positive and negative features in Dimension 1 represents the opposition between involved, affective, interactional, and generalised content on the one hand and informationally dense content on the other hand. An informational register, like academic prose, has negative or low scores in this dimension, while positive or high scores indicate a more involved register, like a face-to-face conversation. According to Biber (1988), registers associated with high Dimension 1 scores are typically produced under greater real-time constraints, while those with low Dimension 1 scores are generally those that allow for relatively more careful editing.

Table 6.5 Dimension 1

Features with Positive Loadings	
'Be' as main verb	0.72
Present tense verbs	0.55
Predicative adjectives	0.54
Demonstrative pronouns	0.48
Total adverbs	0.48
Existential 'there'	0.40
Private verbs	0.40
Analytic negation	0.40
Contractions	0.38

Emphatics	0.37
Causative adverbial subordinators	0.33
Second-person pronouns	0.33
Pronoun 'it'	0.33
First-person pronouns	0.33
Features with Negative Loadings	
Average word length	-0.66
Total other nouns	-0.59
Nominalisation	-0.43
Phrasal coordination	-0.41

6.6.1.1. Communicative Functions of Dimension 1 Features

The positive features comprised different types of *verbs* and *pronouns, contractions, adverbs, emphatics, the causative adverbial subordinator, analytical negations*, and *predicative adjectives*. The first verbal feature is ***'be' as main verb***, which functions as a copula that links some attributes, expressed by the subject predicative, to a subject (Biber et al., 2021). Thus, the high loading of 'be' as main verb can be associated with the high loading of *predicative adjectives*, which follow 'be' as main verb. It is also associated with the high loading of *existential there,* which often precedes it. 'Be' as main verb belongs to the category of *current copular verbs*, including 'seem' and 'appears'. Unlike *resulting copular verbs*, such as 'become', 'get', and 'grow', current copular verbs identify attributes that are in a continuing state of existence (Biber et al., 2021; Hinkel, 2002).

A related linguistic feature is ***present tense verbs***, which are used to refer to immediate and present topics and actions or those that are generally true (Biber, 1988; Hinkel, 2002). Unlike past tense, present tense focuses on the information being presented rather than the sequence of actions. Present tense is more common in unplanned speech (Biber, 1988), and it is often associated with generalisations (Hinkel, 2002).

A feature that often appears in the present tense is ***private verbs***, which are used to refer to intellectual states (e.g., 'know', 'think', 'learn') or non-observable intellectual acts (e.g., 'discover', 'realize', 'hope') (Biber, 1988). They are used to report the writer's and others' opinions, wants, and feelings (Biber, 1988; Biber et al., 2021). Overall, *'be' as main verb, present*

tense verbs, and *private verbs* indicate an interactive or involved style (Biber, 1988; Hinkel, 2002), and they also both indicate a verbal, as opposed to nominal style (Biber, 1988).

The positive pole comprises five types of pronouns: *pronoun 'it', demonstrative pronouns, first-person pronouns, second-person pronouns,* and *existential there*. It should be noted that a text with a high frequency of pronouns indicates high internal or external deixis. This characteristic often occurs in registers that are produced in real-time, have little lexical content, involve immediate interaction, require fast production, and lack editing (Biber, 1988; Biber et al., 2021).

Pronoun 'it' (including *it*, *its*, *itself*) is the most generalised form of pronouns as it refers to animate and inanimate objects as well as abstract concepts (Hinkel, 2002). In fact, 'it' can be ambiguous when it refers to whole segments of previous texts (Biber, 1988). 'It' is also used in hedging verb constructions; when it is used with 'seem' or 'appear', it depersonalizes the text by reflecting hedged objectivity or marking less reliable evidentiality (Hinkel, 2002). Moreover, 'it' can mark evaluation when it is followed by a linking verb, an adjective, and a clause (Bednarek, 2007).

Demonstrative pronouns (*this, that, these, those* followed by a verb or punctuation mark) play a role in text cohesion as they have indexical or referential, deictic, and experiential functions (Hinkel, 2002). They can refer to entities outside of the text, known as exophoric referents, or entities previously mentioned in the text, called anaphoric referents (Biber, 1988; Biber et al., 2021). Sometimes, *demonstrative pronouns* can be ambiguous in their referential function when they refer to a preceding clause, sentence, or more vaguely to a preceding text (Biber et al., 2021). Hence, they perform the function of hedges because of their lack of specificity (Hinkel, 2002).

The positive pole includes two personal pronouns that are not typically used in news reports (Biber et al., 2021): *first-person pronouns* (*I, me, us, my, we, our, myself, ourselves*) and *second-person pronouns* (*you, your, yourself, yourselves, thy, thee, thyself, thou*) which refer directly to the addressor and addressee. These two types of pronouns are highly used in interactive discourse with an interpersonal focus (Biber, 1988; Hinkel, 2002). ***First-person pronouns*** indicate ego involvement in a text (Biber, 1988), and they are used to indicate personal testimony. Fahnestock (2011) states that 'I' is used to foreground claims, especially those made by authoritative people. To elucidate the function of *first-person pronouns* in news reports,

random concordances of the pronoun ‘I’ extracted from the whole corpus were examined (see Figure 6-2). The concordances show that many instances of ‘I’ are used in quotations and followed by private verbs like ‘think’, ‘thought’, and ‘know’. Its co-occurrence with private verbs confirms previous research that showed that *first-person pronouns* are often the subjects of private/mental verbs, indicating that mental processes are often correlated with high ego involvement (Biber, 1988; Fahnestock, 2011).

Figure 6-2 Sample Concordances of the First-person Pronoun ‘I’

	File	Left Context	Hit	Right C
1	B Web (27).txt	to make sure that I had bub's health at the top of my list. "	I	know that as a mother, having the vaccinatio
2	B Web (13).txt	of Oregon in January that killed one and sickened six more. "Every story is heartbreaking.	I	know what they are going through," says Na
3	C Web (47).txt	th professionals feel like the pandemic has sent things spinning out of control, Hick said. "	I	think one of the biggest contributors to burn
4	A Blog (19).txt	with Pope Francis.<nl/> 'There's a strong moral dimension to this effort,' said Ban. '	I	think this should be first time for any secreta
5	A Broad (1).txt	find out about astrophysics first. Before he wrote about climate change for If There Is	I	Haven't Found It Yet, his first play, which ran
6	A Blog (17).txt	specially in the case of young researchers without job security. Many of the scientific talks	I	attended addressed the current unprecedent
7	C Blog (43).txt	for many long-haulers coping with a situation that didn't exist a year ago. "	I	still have trouble accepting that things chang
8	C Tab (37).txt	his "terrible buoyancy". He insisted: "It would be wrong to say that at any stage	I	thought, 'Oh my goodness, this is it'. "Some
9	A Broad (4).txt	to keep us awake at night" - speak for yourself, I can lose sleep worrying if	I	turned the sprinkler off - and "there are othe
10	C Tab (37).txt	kept convincing me that everything would almost certainly be all right in the end. But	I	was just frustrated. I remember seeing a lot c

The use of ***second-person pronouns*** indicates that there is a specific addressee and that the text has a high degree of involvement with that addressee (Biber, 1988). In other words, the writer acknowledges the reader’s presence directly and calls on them, which makes the text automatically compelling (Fahnestock, 2011). A random concordance list of the pronoun *‘you’* from the whole corpus was used to examine its function in news reports (see Figure 6-3). As demonstrated by the concordances, many instances of ‘you’ are generic (i.e., not addressing a particular person and meaning anyone) and are found within quotations from public figures. The use of ‘you’ in these quotations allows the readers to participate and engage as members of the news event being reported on, and it elicits an emotional response from them. In other words, including quotations that directly address the audience makes the news report more personal, relatable, and interesting.

Figure 6-3 Sample Concordances of the Second-person Pronoun 'You'

	File	Left Context	Hit	Right Context
1	B Tab (6).txt	›w they're sick! YOU SHOULD AVOID DAIRY WHEN	YOU	HAVE A COLD False While a recent UK study discov
2	B Web (13).txt	›se version and $347 for the three-dose version. "If	you	break down the cost of the vaccine for three to five
3	A Broad (1).txt	›ed If There Is ... so many times, Constellations too.	You	do reach that crushing point where you think, it's ju
4	B Broad (9).txt	›eople don't believe the government's advice, then	you	have created a real problem. Vaccinations are not fa
5	B Broad (15).txt	›n to your body. Though most cases aren't serious,	you	might need to visit a doctor. If you're worried abou
6	C Web (23).txt	st say get as many people vaccinated as quickly as	you	possibly can," Fauci said. "And every day that goes
7	A Web (16).txt	›nent subsidies have been disappearing. "Now what	you'	re finding with rooftop solar is ... it's still being take
8	C Tab (70).txt	one group as a whitewash, another will say, "I told	you	so" and a third will be shouted down for saying the
9	B Tab (15).txt	›LL AND HOW MUCH IS IT? If you get chicken pox,	you	usually get better quickly with no need for medical
10	C Broad (22).txt	a 68-bed field hospital. So this is the kind of thing	you	will see now as this crisis develops and deepens," h

Another type of pronouns that has a high positive loading is ***existential there***. This pronoun has multiple functions: (a) introducing new information (Biber et al., 2021), (b) emphasising new information or the occurrence of something by placing it in the last position of a sentence (the rheme) (Biber et al., 2021; Hinkel, 2002), and (c) expressing a stance or opinion if it is followed by a linking verb, 'something,' 'anything', or 'nothing', and then an adjective (Bednarek, 2007). Biber (1988) states that the frequent occurrence of *existential there* indicates a non-complex structure with a minimum load of information (Biber, 1988) because sentences beginning with *existential there* are relatively simple and do not contain many complex components.

Another feature that is not expected to be found in news reports is ***contractions*** (including all contractions on pronouns and auxiliary verbs). They are markers of informality and shortened forms of expression (Biber, 1988). Previous research shows that *contractions* are used more frequently in conversations and informal writings due to fast and immediate production. Thus, they are less frequently used in formal writing that usually go through the editing process (Biber, 1988). A sample of random concordances of *contractions*, represented by the tag [CONT] is represented in Figure 6-4. The sample shows that four instances of *contractions* follow the first-person pronoun 'we'. This grammatical relation may explain the high frequency of *contractions* on this dimension since *first-person pronouns* also have high loadings on this dimension. In addition, three instances of *contractions* in the sample were found in negation which is also a highly loaded feature on this dimension. More importantly, the sample shows that many instances of the *contractions* are found within quotations. In this way, *contractions* add a conversational tone to the news articles and enhance the flow of the content. Such an informal

and conversational tone makes the article sound more natural and relatable, which in turn makes the article more appealing and engaging.

Figure 6-4 Sample Concordances of the Tag [CONT] in the Corpus

	File	Left Context	Hit	Right Context
1	C Broad (19)_MAT.txt	worse_PRED than_PIN anything_QUPR we_FPP1 've_VPRT [	CONT]	seen_VBN [PRIV] so_RB far_PLACE '_'' ,_ says_VPRT [
2	C Tab (40)_MAT.txt	I_FPP1 feel_VPRT [PRIV] [THATD] we_FPP1 've_VPRT [	CONT]	now_TIME got_VBN arguably_RB the_DT safest_
3	A Blog (10)_MAT.txt	percent_NN growth_NN ._. So_IN we_FPP1 're_VPRT [	CONT] [	BEMA] much_QUAN less_RB willing_JJ to_TO
4	C Web (57)_MAT.txt	VPRT that_DEMO people_NN do_VPRT n't_XX0 [	CONT]	want_VB to_TO be_VB [BEMA] outside_PLACE
5	C Blog (12)_MAT.txt	PUBV] ._. ``_`` We_FPP1 just_EMPH do_VPRT n't_XX0 [	CONT]	have_VB the_DT answer_NN to_PIN that_
6	C Broad (24)_MAT.txt	STPR] ,_ ''_'' she_TPP3 said_VBD [PUBV] ._. ``_`` They_TPP3 're_VPRT [	CONT]	doing_VBG [PROD] this_DEMO without_PIN training_GER ._. ''_''
7	A Broad (1)_MAT.txt	NN and_CC it_PIT was_VBD n't_XX0 [	CONT]	even_RB them_TPP3 driving_VBG ._. ''_'' He_TPP3 had_
8	A Web (4)_MAT.txt	matter_NN of_PIN weeks_NN ._. ``_`` We_FPP1 'd_PRMD [	CONT]	like_VB to_TO see_VB [PRIV] this_DEMO
9	A15_MAT.txt	licensed_JJ psychiatrist_NN ,_ or_CC she_TPP3 'll_PRMD [	CONT]	dose_NN your_SPP2 head_NN full_JJ of_
10	C Tab (19)_MAT.txt	VB [BEMA] infectious_PRED when_RB they_TPP3 've_VPRT [	CONT]	already_RB got_VBN symptoms_NN ,_ but_CC we_

Adverbs, a highly loaded positive feature, are used to elaborate and expand upon the information presented in a text as they can include a time or place reference, specify a manner or degree, express the speaker's or writer's attitude, and link ideas (Biber, 1988; Biber et al., 2021). Hinkel (2002) states that *adverbs* have a variety of roles and functions, including adjuncts, conjuncts, disjuncts, cohesive and referential devices, hedges, evidentials, intensifiers, amplifiers, diminutives, and downtoners. Because of their referential functions, Biber (1988) associates the high frequency of *adverbs* with context-dependent texts in which the addressees can identify the text-internal or text-external referents from the co-text or the physical context of the texts.

Another highly loaded positive feature is ***emphatics***, which include 'for sure', 'a lot', 'such a', 'real + ADJ', 'so + ADJ', 'Do + V', 'just', 'really', 'most', and 'more'. Emphatics are used to show certainty, and they often occur in informal and real-time discourse that has high involvement with the topic (Biber, 1988). Similarly, the ***causative adverbial subordinator*** 'because' is often used in real-time registers in which there is no time to choose exact lexical phrases to express reasoning. Moreover, as Biber (1988) states, 'because' can mark affect or stance when it is used to justify actions or attitudes. Likewise, ***analytical negation*** (not) is generally found in informal registers. Unlike synthetic negation ('no', 'neither', 'nor') which creates integrated structures, *analytical negations* make more fragmented structures. Negation generally occurs more in speech than in writing. This can be attributed to the high frequency of repetitions, denials, rejections, questions, and mental verbs in speech (Biber, 1988).

Dimension 1 includes one distinctive feature that is not found in Biber's (1988) Dimension 1: *involved vs. informational:* ***predicative adjectives****.* They are usually used to elaborate on information presented in the text, lending the text a descriptive nature. Unlike attributive adjectives, which create integrated structures, *predicative adjectives* fragment structures (Biber, 1988). Biber et al. (2021) state that *predicative adjectives* are used in news to both make judgements and supply information, and they mostly occur within quoted speech. To examine the communicative functions of these adjectives in the corpus, a list of the 100 most frequently used *predicative adjectives* was generated. It was found that most of the adjectives in the list were evaluative in that they reflected attitudes and judgments. Table 6.6 categorises these adjectives according to the domains of evaluation. The last row in the table includes adjectives that are descriptive and not evaluative. This emphasises that the positive pole of Dimension 1 in this study, like the positive pole of Dimension 1 in Biber's (1988) analysis, represents an *involved style*.

Table 6.6 Top 100 Predicative Adjectives in the Corpus Categorized based on Their Evaluative Meaning

Ability	able, unable
Comparing or grading	low, first, worse, better, high, higher, less, small, lower, close, quick, slow, greater, huge, minimal, consistent, different, similar
Comprehensibility	clear, unclear, hard, difficult, easy
Emotion	happy, hopeful, worried, cautious, concerned, pleased, proud
Epistemic	likely, unlikely, possible, impossible, sure, confident, certain
Expectedness	common, unknown, rare
Genuineness	true, wrong, correct, real, right, honest, fake
Impact	great, successful, ineffective, effective, interesting, beneficial, good, positive, negative, bad
Importance	necessary, vital, critical, crucial, important, worth, key, essential
Safety	dangerous, fatal, safe, serious, severe, deadly, dire
Topical/others	Sick, infectious, allergic, contagious, ill, infected, immune, unvaccinated
	responsible, aware, ready, vulnerable, willing, subject, open, free, present, active, fine, full, asymptomatic, new, reviewed, available, due

To conclude, the features in the positive pole of Dimension 1 represent two main patterns: a) interpersonal, involved, and affective interaction; and b) fragmented, generalised, and real-time generated content. As the discussion above shows, *present tense verbs, private verbs, adverbs,*

emphatics, predicative adjectives, pronoun it, existential there, first-person pronouns, second-person pronouns, and *the causative adverbial subordinator* can reflect interpersonal and involved interaction. Indeed, some of these features can be used to express attitude like *private verbs, adverbs, emphatics, predicative adjectives, pronoun it, existential there,* and *the causative adverbial subordinator*. Some features indicate a fragmented, generalised, non-complex, real-time constrained, or uncertain presentation of information. These include '*be' as main verb, present tense verbs, demonstrative pronouns, pronoun it, existential there, contractions, analytical negation, causative adverbial subordinates, adverbs,* and *emphatics*. Given that the analysed texts are news articles, which are typically less involved than conversations, the positive pole of this dimension can be described as a *relatively involved style*.

One important consideration regarding the positive pole is that features co-occur not only because of their shared communicative functions but also because of their grammatical relatedness. The analysis shows that *private verbs* are often expressed in the *present tense* and often follow *first-person pronouns*. *Contractions* often follow personal pronouns. In addition, *predicative adjectives* often follow '*be' as main verb*, and *existential there* often precede '*be' as main verb*.

Out of the negative features in Dimension 1, two features are also negatively loaded on Biber's Dimension 1: ***average word length*** (which refers to the existence of long words) and ***total other nouns***. These features are associated with texts that are geared towards providing information. *Nouns* generally make the texts more informative because of their referential specifications, which are used to specify who and what the text is about (Biber et al., 2021). *Long words* convey specific and specialised meanings. Thus, these features mark precise lexical choices (Biber, 1988).

Two other features, ***nominalisation*** (i.e., any noun ending in '-tion', '-ment', '-ness', or '-ity') and ***phrasal coordination*** (i.e., 'and' that combines same grammatical phrases*)*, although not highly loaded on Biber's Dimension 1, are also used in informationally dense texts. *Nominalisations* are used to expand ideas and reduce information to forms that require fewer words (Biber et al., 2021), and they usually convey highly abstract concepts (Biber, 1988). Similarly, *phrasal coordination* is employed in the integration and expansion of information units (Biber, 1988) and in achieving textual cohesion and information flow (Hinkel, 2002).

In conclusion, these negative features are utilized to meticulously integrate and accurately present information. Given that the analysed texts are news articles, which are typically less informationally dense compared to academic prose, the negative pole of this dimension can be interpreted as reflecting a *relatively informational style*.

In Biber's (1988) *involved vs. informational* dimension, conversational texts had high positive scores, while news-based registers, such as editorials, press reviews, and press reportage, had low scores. In other words, news-based registers were more informational than other spoken registers. However, by examining only news articles, the present study found that some news articles were more involved, while others are more informational (see Figure 6-5). One possible explanation for the existence of the involved features in news is the high number of quotations. Previous studies and the examined concordances showed that many of the positive features (e.g., *first-* and *second-person pronouns*, *contractions*, and *predicative adjectives*) often exist in quotations. Thus, it can be assumed that news reports with high numbers of quotations tend to be marked positively on Dimension 1. Another possible reason for the presence of involved features in news texts could be attributed to less stringent editing or less professional writing style.

6.6.1.2. The Distribution of News Types along Dimension 1

Figure 6-5 displays boxplots illustrating the distribution of Dimension 1 scores for the texts from each of the five news types. Table 6.7 provides the mean Dimension 1 scores and standard deviations for each news type. Notably, a clear trend can be recognized, with tabloids showing the highest mean score and being the most involved news type, followed by web-based publications. In contrast, blogs exhibited the lowest negative score, positioning them as the most informational news type. Broadsheets and false news were in the middle, each scoring around zero.

A possible explanation for tabloids being the most involved is the extensive use of quotations (Coltman-Patel, 2023). On the other hand, blogs being the most informational might be due to their reporting of specialised news. This observation supports the findings of previous studies (e.g., Daems et al., 2013; Grieve et al., 2010; Sardinha, 2018), which indicate that blogs that report or comment on news tend to have an informational nature.

Figure 6-5 The Distribution of News Types on Dimension 1

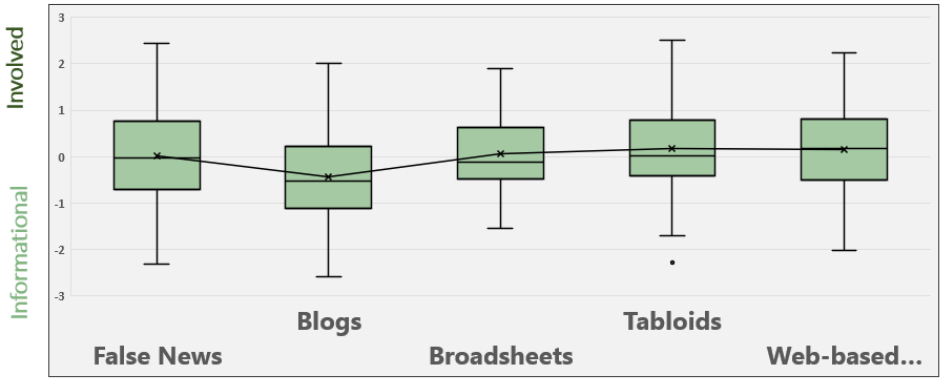


Table 6.7 Dimension 1 Scores for the Five News Types

News Types	Mean	Standard Deviation
Tabloids	.1870	.97942
Web-based Publications	.1623	.93617
Broadsheets	.0638	.79045
False News	.0169	.91682
Blogs	-.4300	.94965

The one-way ANOVA results showed that there was a statistically significant difference between at least two news types with respect to Dimension 1 ($F(4,500) = [7.528]$, $p < .001$). Tukey post-hoc test showed that there was no significant difference between false news and broadsheets ($p = .996$), tabloids ($p = .680$), and web-based publications ($p = .792$). On the other hand, false news was significantly different from blogs ($p = .005$, 95% CI = [.0937, .08002]).

To illustrate the opposition between the *informational* and the *involved styles*, the tabloid (Example 1) and the blog (Example 2) articles with the highest and lowest Dimension 1 scores, respectively, are examined below. The features corresponding to the positive and negative poles of the dimension are indicated by boldface.

Example 1 is an extract from the tabloid article having the highest Dimension 1 score. It contains extensive use of the positive features, which result in the text having an involved, affective, and interactional style. The text quotes Boris Johnson, the previous Prime Minister (PM) of the United Kingdom, who is describing his feelings and thoughts after being infected with COVID-19. The use of *private verbs* like ‘recalled’, ‘feeling’, ‘see’, ‘remember’, and ‘thought’ as well as

predicative adjectives such as ‘honest’, ‘anxious’, ‘adamant’, ‘right’, ‘lucky’, and ‘frustrated’ reflects the involved expression of emotions and concerns surrounding the PM’s infection, making the text affective in style. Moreover, the frequent use of the *pronoun*s ‘I’ and ‘you’ accentuates the interpersonal interaction in the text. In addition, the frequent use of *emphatics* such as ‘really’, ‘just’, ‘did’, and ‘most’ as well as *adverbs* such as ‘pretty’, ‘suddenly’, and ‘incredibly’ makes the text sound both informal and affective. Furthermore, these two features, in addition to the use of *existential ‘there’, pronoun ‘it’, demonstrative pronouns, the causative adverbial subordinator, analytical negation, present tense verbs, ‘be’ as main verb,* and *adverbs* like ‘then’, ‘never’, and ‘back’ lend the text a style similar to that employed in spoken registers. Thus, the text presents information in a style that is timely, generalised, non-complex, or uncertain. This similarity to spoken registers might be attributable to the use of a large amount of quoted material from the PM.

(1) The PM **recalled**: "The thing was, **I was** in denial **because I** was working and **I** kept doing these meetings by video link.

"But **I was really feeling pretty** groggy, to **be** totally **honest** with **you**. **I** was **feeling pretty** wasted - **not** in an intoxicated way, but **just**, **you know**, **pretty** rough."

He **then stops suddenly** and **asks**: "**Have you** had this thing? Well, **don't** get **it**. **You don't** want **it**. **I** was**n't** struggling to breathe but **I just** was**n't** in good shape and **it** was**n't** getting better. **Then** the doctors got **anxious because** they **thought** that **my** readings were **not** where they wanted them to be.

"**Then I** was told **I** had to go into St Thomas'. **I** said **I really** did**n't** want to go into hospital.

"**It** did**n't** seem to me to **be** a good move but they **were pretty adamant**. Looking **back**, they **were right** to force me to go. **I did** have the **most** fantastic care. **It** was **awe-inspiring** to **see how** they look after people and **I was** very **lucky**.”

….

"But **I've never** had anything as serious as **this**.

"Well, no. All **I remember** feeling **was just** frustration. **I** could**n't** see why **I** was**n't** getting better.

> "**I** was **just incredibly frustrated because** the bloody indicators kept going in the wrong direction and **I thought**, '**There's no** medicine for this thing and **there's no** cure'. (*C Tab 37*, Dim 1 score = 2.51)

Example 2 is taken from a blog post having the lowest Dimension 1 score. This blog post provides a detailed overview of the vaccine market, and it makes noticeably extensive use of the negative features, including *long words, nominalisations,* and *nouns* in general, with some noun phrases being combined with 'and'. These features reflect the dense integration of information in the text and suggest that the writer is knowledgeable about the subject matter. In contrast, the positive features are notably absent in this text.

> (2) The **market overview section** comprises **impact factors** such as **drivers, restraints,** and **opportunities** for the global human **vaccines market**. These **factors** would aid the **stakeholders** in establishing a strong **foothold** in the global human **vaccines market**. Furthermore, the **market overview section** comprises key **industry events, product overview, market attractiveness analysis** and emerging **trend** in the human **vaccines**. The **market attractiveness analysis** provides a graphical **view** comparing the **growth and market dynamics** in various segments and country wise to identify the most attractive **market**. **Market share analysis** is also provided in the competitive **landscape section** of the **report** for the **year** 2015 in terms of **value** (%). (*B Blog 14*, Dim 1 score = -2.58)

These examples show that the *involved style* emphasises several news values, as opposed to the *informational style* (cf. Bednarek & Caple, 2017). Firstly, the *present tense* emphasises that the events are happening right now, thereby reflecting immediacy and urgency as well as accentuating the news value of *timeliness*. Secondly, by their descriptive and, most importantly, evaluative functions, the *predicative adjectives, adverbs,* and *emphatics* accentuate the news values of *negativity*, *positivity*, *impact, superlativeness,* and *unexpectedness*. The *private verbs*, reflecting news actors' emotions and experiences, emphasise the news value of *personalisation*, which attracts readers more than providing abstract portrayals of events. In addition, reporting the PM's experience accentuates the news value of *eliteness* as it refers to elite people's experience. Consequently, the communicative functions of the features of the *involved style* can be linked to promoting news values and making news reports related and interesting to the readers.

To summarize, the *involved style* is characterised by: (a) reflecting news actors' experiences, emotions, and judgements; (b) engaging the readers by addressing them directly; and (c) accentuating the news value of *timeliness, eliteness, personalisation, impact, negativity, positivity, superlativeness,* and *unexpectedness*. By contrast, the *informational style* is characterised by providing detailed, specific, and specialised information about a news event. It is important to note that both styles are relative, as the analysed texts belong to written registers reporting news, which typically lean towards being less informational than academic prose and less involved compared to conversations.

6.6.1.3. False News and Dimension 1

As illustrated in the previous section, false news falls in the middle on Dimension 1, showing both involved and informational features, mostly resembling broadsheets. Example 3 is an excerpt from a false news article that has a median Dimension 1 score. The involved features are written in bold green while the informational features are written in bold black.

Example 3 is informationally dense; it describes in detail the scientific evidence that challenges the scientific consensus that CO2 emission is the main reason for climate change. The text uses many scientific nouns (e.g., 'emissions', 'NAO', 'ENSO'), nominalizations (e.g., 'uncertainties', 'oscillations', variability'), phrasal coordination (e.g., 'past and present'), and long words (e.g., 'multi-decadal, 'oceanic-atmospheric'). These features help portray the writer as knowledgeable; he/she presents a large amount of information to support his claim. The text expresses a sense of involvement by using the first-person pronoun 'our' in 'our understanding of climate and climate changes'. The pronoun 'our' is inclusive, suggesting that the reader and the writer share the same uncertainties about climate and climate changes, and exclusive, suggesting that scientists might not provide the full evidence to 'us', i.e., the writer and the reader. The aspect of involvement is also apparent in the texts' emphasis on the *timeliness* of the information reported by using existential 'there' and by using the present and present perfect tenses for reporting information instead of the past tense.

(3) A Growing **Volume** Of **Evidence** **Undercuts** **'Consensus' Science**

During the first 10 **months** of 2017, 400 scientific **papers** **have** been published that **cast doubt** on the **position** that **anthropogenic CO2 emissions** **function** as the climate's

> fundamental **control knob**…or that otherwise **question** the **efficacy** of **climate models** or the related **"consensus" positions commonly** endorsed by **policymakers and mainstream media.**
>
> These 400 new **papers support** the **position** that **there are** significant **limitations and uncertainties** inherent in **our** understanding of **climate and climate changes**. **Climate science is not** settled.
>
> Modern **temperatures, sea levels,** and extreme **weather events are** neither unusual nor unprecedented. Many **regions** of the **Earth are cooler** now than they **have** been for **most** of the last 10,000 **years**.
>
> Natural **factors** such as the **Sun** (106 **papers**), **multi-decadal oceanic-atmospheric oscillations** such as the **NAO, AMO/PDO, ENSO** (37 **papers**), **decadal-scale cloud cover variations**, and internal **variability** in general **have** exerted a significant **influence** on **weather and climate changes** during both the **past and present**. (*A False 6,* Dim 1 score = -0.02)

To summarize, the analysis of the false news sample highlights a strategic combination of informational and involved features. Through the use of specialized vocabulary and dense informational content, the sample challenges mainstream scientific consensus on climate change causation with ample evidence. Despite its information-rich style, the excerpt effectively engages and involves the audience, making them as part of the discussed topic.

6.6.2. Dimension 2: Expository vs. Narrative Styles

Table 6.8 presents the salient positive and negative features of Dimension 2. This dimension is similar to Biber's (1988) Dimension 2: *'narrative vs. non-narrative discourse'* aside from the positive pole of the extracted dimension being similar to the negative pole of Biber's dimension and vice versa. Biber describes the two poles of the dimension as having narrative versus descriptive or expository concerns. In Table 6.8, the features that are highly loaded on Biber's dimension are in boldface. The features written between parentheses have higher loading on Dimension 1; therefore, they are disregarded in the interpretation of this dimension.

Table 6.8 Dimension 2

Features with Positive Loadings	
Attributive adjectives	0.48
Conjuncts	0.33
(**Average word length**	0.41)
(Phrasal coordination	0.31)
Features with Negative Loadings	
Past Tense	-0.83
Third-person pronouns	-0.57
Public verbs	-0.54
Subordinator-that deletion	-0.54

6.6.2.1. Communicative Functions of Dimension 2 Features

The positive pole of Dimension 2 includes *attributive adjectives* and *conjuncts*. ***Attributive adjectives*** (adjectives followed by other adjectives or nouns) are associated with informationally dense content (Biber, 1988) because they are used to provide additional information in noun phrases (Biber et al., 2021). Thus, attributive adjectives are markers of elaborated nominal style (Biber, 1988). Unlike predicative adjectives, attributive adjectives create more integrated structures of nominal elaboration as they elaborate upon the concept more in fewer words (Biber, 1988). Biber et al. (2021) found that attributive adjectives are more frequently used in news stories than predicative adjectives, and they mostly take the form of affiliative classifiers that are used to identify the national or religious group to which referents belong.

Attributive adjectives play attitudinal as well as classificatory functions in texts as they add information to abstract nouns. These functions, along with creating integrative structures, enhance text flow as they supply sufficient or even exclusive descriptions to referents (normally, nouns) and boost their contextual importance (Hinkel, 2002). Table 6.9 presents the top 100 attributive adjectives used in the corpus. The table shows that attributive adjectives serve two functions: description and evaluation (i.e., denoting judgement, affect, or emphasis). In addition, the table shows that the majority of the adjectives was descriptive although considering the context is important for deciding whether the rest of the adjectives are evaluative.

Table 6.9 Top 100 Attributive Adjectives in the Corpus

Descriptive	new, other, last, first, global, such, public, covid, medical, same, least, human, immune, next, recent, high, social, local, big, own, second, due, past, federal, current, clinical, different, old, Chinese, young, national, further, additional, unvaccinated, economic, former, previous, scientific, free, infectious, latest, political, available, financial, small, American, older, pandemic, full, respiratory, sick, total, similar, infected, term, daily, biggest, large, Australian, pregnant, highest, top, viral, natural, little, personal, strong, potential, long, single, general, private, based, related
Evaluative	positive, good, higher, serious, adverse, best, severe, effective, better, safe, significant, common, important, negative, normal, chief, key, bad, healthy, major, likely, possible, real, less, clear, critical

The second salient positive feature is ***conjuncts*** (e.g., 'namely', 'for instance', 'as a result', 'in summary', 'in contrast'), which serve five functions: listing, clarifying, giving results, summarising, and contrasting (Quirk, 1985). These functions indicate the logical and complex relations between clauses in a text, and they are usually associated with texts having highly informational content, such as abstract and technical texts (Biber, 1988).

To conclude, the positive pole of this dimension can be related to informationally dense texts, especially due to the average word length and phrasal coordination having high positive loading on this dimension and having higher loading on the informational pole of Dimension 1. However, the two highly loaded features of this dimension: *attributive adjectives* and *conjuncts* share two communicative functions: (a) They are both used to add precision to a text by providing additional or detailed information, and (b) they play a role in text cohesion by logically connecting information in the text. Together, these two features work to explain, inform, or describe a topic in a clear and organized manner. Therefore, the positive pole of this dimension is not only informational and can be better interpreted as an *expository style*.

Alternatively, the negative features include *past tense, public verbs, third-person pronouns* and *subordinator-that deletion*. Biber (1988) states that *past tense*, *third-person pronouns*, and *public verbs* are associated with the narration of past events. Therefore, they are common in narrative discourse such as fiction (Biber et al., 2021; Hinkel, 2002).

The most distinctive feature of narrative discourse is ***past tense***, which sets up the past timeframe in discourse. It is used to describe sequential past events or to recount past experiences and

activities (Hinkel, 2002). A related feature is that of ***public verbs***, which refer to actions that can be noticed publicly. Public verbs are mainly speech act verbs (e.g., 'admit', 'assert', 'declare'), and they are used to introduce indirect and reported speech (Biber, 1988).

Third-person pronouns are used to refer to animate, usually human, referents who participated in the events and who are not the speaker or the addressees (Biber, 1988). Biber et al. (2021) found that third-person pronouns are most common in news stories when compared to academic writing, conversations, and fiction. Using a third-person perspective serves to present the news report as an unbiased representation of reality, as opposed to using the first-person perspective (Biber et al., 2021). Thus, one function of third-person pronouns is to accentuate formality, detachment, and objectivity (Hinkel, 2002).

Subordinator-that deletion is a form of syntactic reduction and therefore is not preferred in edited writing (Biber, 1988). Although it is not among the narrative features in Biber's (1988) Dimension 2, its high occurrence in this dimension can be attributed to its frequent usage after *public verbs*. For instance, the sentence 'But Cuomo says the state wants anyone who visited between December 18th and 24th to get a COVID-19 test' (*C Blog 39*) demonstrates the deletion of the word 'that' after the verb 'says'. In the corpus, there are 1,424 instances of subordinator-that deletion, with 803 occurrences following public verbs.

To conclude, the negative features identified in Dimension 2 point towards a *narrative style* commonly found in narrative discourse. However, given that the analysed texts are news articles, it is crucial to interpret this dimension as *relatively narrative*, focusing on narrating events, actors, and their statements. This *narrative style* in news articles differs from that found in fictional registers, as found in Biber's (1988) study.

6.6.2.2. The Distribution of News Types along Dimension 2

Figure 6-6 presents boxplots that compare the distribution of Dimension 2 scores for the texts from each of the five news types. Table 6.10 presents the mean dimension scores and the standard deviations for each news type. Overall, we can see a clear trend in the true news, with blogs having the highest positive Dimension 2 score, indicating that they were the most expository. On the other hand, tabloids obtained the lowest negative score, indicating that they were the most narrative. Web-based publications had a negative score, so they tended to be

narrative. Broadsheets had a negative score of approximately zero while false news had a positive score of approximately zero, thereby indicating that both had features of the *narrative* as well as *expository* styles.

The blog's preference for the *expository style* could be attributed to the fact that the examined blogs tend to focus on reporting specialized news or information. By contrast, the tabloid's preference for the *narrative style* can be attributed to the fact that tabloids predominantly target the general public (Breuer & Napthine, 2008) and aim to provide engaging storytelling rather than focusing solely on informative content.

Figure 6-6 The Distribution of News Types on Dimension 2

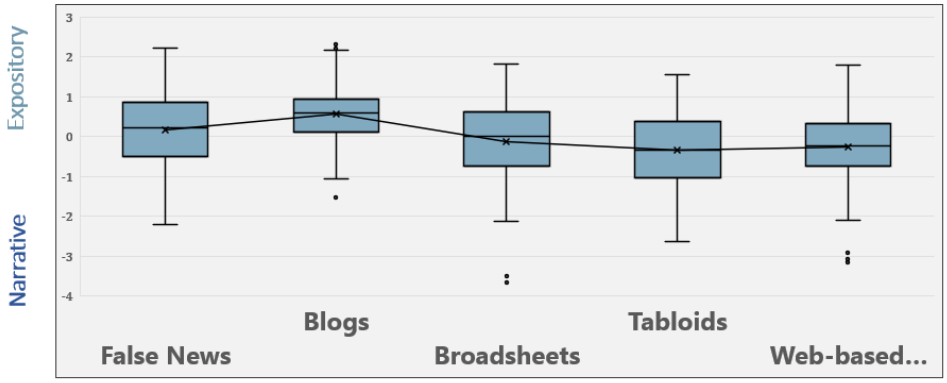


Table 6.10 Dimension 2 Scores for the Five News Types

News Type	**Mean**	**Standard Deviation**
Blogs	.5645	.74789
False News	.1566	.93324
Broadsheets	-.1316	.96985
Web-based publications	-.2511	.87787
Tabloids	-.3383	.93289

The one-way ANOVA results showed that the difference between at least two news types was statistically significant (F [4, 500] = [16.931], p = .000). According to the Tukey post-hoc test, false news differs significantly from blogs (p = .001, 95% CI = [-.7530, -.0628]), tabloids (p = .001, 95% CI = [.1498, .8400]) and web-based publications (p = .011, 95% CI = [.0626, .7528]). However, there was no significant difference between false news broadsheets (*p* = .151).

To illustrate the opposition between the *expository* and the *narrative styles*, the blog (Example 4) and the tabloid (Example 5) articles with the highest and lowest Dimension 2 scores, respectively, are examined below. The positive and negative features of the dimension are indicated by boldface.

Example 4 is an extract from a blog article having the highest Dimension 2 score. It provides a detailed explanation of the authority granted to the government under the DPAS and the Bayh-Dole Act, which are two laws affecting contractors, and what they could have done during the COVID-19 pandemic. The text employs the two features of the *expository style*: *attributive adjectives* such as 'prime', 'extreme', 'critical medical', 'critical pharmaceutical', and 'other essential' and *conjuncts* such as 'therefore' and 'for example'. These features provide additional information and enhance the clarity and cohesion of the text. Based on the complexity of the text and the use of technical terminology, it appears to be aimed at readers with expertise in law. In addition, the author's knowledge and understanding of the subject matter are evident in the text.

> (4) It is **therefore** important for contractors to be aware of which of their **prime** contracts, subcontracts and purchase orders are rated. While it is unlikely that the government would exercise the more **extreme** authority granted under DPAS and effectively commandeer manufacturing and **other** types of facilities in order to produce **critical medical** supplies or to support **other** emergency relief efforts, contractors should further be mindful that the government could also authorize companies to use **certain** goods and technology that is patented by **other** companies without the consent of the patent holder. Specifically, pursuant to the Bayh-Dole Act, 35 U.S.C. ch. 18, or 28U.S.C. 1498, the federal government may authorize a company to produce, **for example**, **critical pharmaceutical** products, testing equipment or **other essential** items that are patented by another company to combat CoVID-19. (*C Blog 34*, Dim 2 score = 2.32)

Example 5, taken from the tabloid article having the lowest Dimension 2 score, employs *past tense verbs* to indicate that the events described have already taken place, establishing a retrospective narrative. This storytelling approach highlights the author's knowledge and assurance of the discussed events. Additionally, the inclusion of public verbs like 'revealed' and 'said,' attributed to 'McPake' and the pronoun 'he,' contributes to the narrative quality of the text, creating the sense that the events are being recounted from the perspective of an observer or a third-person omniscient narrator.

(5) McPake's men **were** unable to train on Friday after Charlie Adam **tested** positive for coronavirus and a member of **his** backroom staff **had** to self isolate as a close contact.

McPake **revealed he** didn't know what his team selection would be as **he waited** to find out who **was** available. **He said**: "We've come away with a point when we **wanted** three but, ultimately, it's been a tough 24 hours.

"We **lost** one of our most influential players through Covid and we've no physio. That alone **was** tough and then we **were** told we couldn't train on Friday, we **turned** up and **were told** to stay in our cars. (*C Tab 5,* Dim 2 score = -2.65)

To summarize, the *expository style* is used in news articles concerning complex topics or argumentation. So, this style provides technical, detailed, or descriptive information. In comparison, the *relatively narrative style*, emphasises the story-like flow of the news text, making it feel more like an interesting story than an informational article. In addition, the *narrative style* reflects the author's high level of familiarity and confidence in conveying the mentioned events.

6.6.2.3. False News and Dimension 2

False news, as mentioned earlier, falls between the positive and negative poles on Dimension 2. Similar to broadsheets, it showcases characteristics of both *expository* and *narrative styles*. To better understand how the style of both false news and broadsheets are so similar with respect to Dimension 2, the false news article that has a median Dimension 2 score (Example 6) and the broadsheet article that has the median Dimension 2 score (Example 7) are examined bellow. The expository features are written in bold blue while the narrative features are written in bold black.

Example 6 is extracted from a false news report entitled ‘Did China Steal Coronavirus from Canada and Weaponize It?’ The text is narrative and expository at the same time because it combines a story-like structure and provides details about the Chinese military’s use of biological technology to gain an advantage in warfare. The *narrative style* is indicated by the chronological overview of the Chinese biotechnical activities, which is in *past tense*, giving the impression that the author is knowledgeable about the military history of China. On the other hand, the *expository style* is indicated by the *attributive adjectives,* such as these in the following phrases: ‘biological and biometric materials’, ‘biomimetic systems’, ‘human performance

enhancement', 'strategic commanding', and 'genetic information', which provide detailed descriptions of China's attempt at 'bloodless victory'.

> (6) Weaponizing Biotech - China's War for **Biological** Dominance …
>
> Since 2016, the **Central** Military Commission has funded projects on **military** brain science, advanced **biomimetic** systems, **biological** and **biomimetic** materials, **human** performance enhancement, and "**new** concept" biotechnology.
>
> In 2016, an AMMS **doctoral** researcher **published** a dissertation, "Research on the Evaluation of **Human** Performance Enhancement Technology," which **characterized** CRISPR-Cas as one of three **primary** technologies that might boost troops' combat effectiveness.
>
> …
>
> In 2016, the **potential strategic** value of **genetic** information **led** the **Chinese** government to launch the **National** Genebank, which intends to become the world's **largest** repository of such data. It aims to "develop and utilize China's **valuable genetic** resources, safeguard **national** security in bioinformatics, and enhance China's capability to seize the **strategic commanding** heights" in the domain of Biotechnology Warfare.
>
> **Chinese** military's interest in biology as an emerging domain of warfare is guided by strategists who talk about **potential** "**genetic weapons**" and the possibility of a "**bloodless** victory." (*C False 34,* Dim 2 score = 0.21*)*

Example 7 is taken from a broadsheet article entitled 'Disasters with Twice the Misery: When Global Warming Collides with a Pandemic'. It uses both *narrative* and *expository styles*. The *narrative style* is demonstrated through the narration of the dangerous situations in India that were caused by extreme weather events, economic instability, and the COVID-19 pandemic. This is achieved through the use of the *past tense* and the *public verb* 'said'. Additionally, the *attributive adjectives* and *conjuncts* lend the text an *expository style*. The author explained the misery of this situation using the adjectives in the following examples: 'bigger threat', 'ordinary people', 'rapid succession', 'extreme weather' 'more frequent' and 'more intense' as well as the conjunct 'for instance'.

(7) **Extreme** weather presents an even **bigger** threat when economies are crashing and **ordinary** people are stretched to their limits.

The hits **came** this week in **rapid** succession: A cyclone **slammed** into the **Indian** megacity of Kolkata, pounding rains **breached** two dams in the **Midwestern** United States, and on Thursday **came** warning that the Atlantic hurricane season could be severe.

It all **served** as a reminder that the coronavirus pandemic, which has killed 325,000 people so far, is colliding with another **global** menace: a fast-heating planet that acutely threatens millions of people, especially the world's **poor**.

Climate change makes **extreme** weather events more **frequent** and more **intense**. Now, because of the pandemic, they come at a time when **national** economies are crashing and **ordinary** people are stretched to their limits.

Relief organizations working in **eastern** India and Bangladesh, **for instance**, **say** the lockdown **had** already forced people to rely on food aid by the time the storm, Cyclone Amphan, **hit**. Then, the high winds and heavy rains **ruined** newly sown crops that **were** meant to feed communities through **next** season. "People have nothing to fall back on," Pankaj Anand, a director at Oxfam India, **said** in a statement Thursday. (*A Broad 8,* Dim 2 score = 0.01)

The analysis of Examples 6 and 7 shows that false news is virtually identical to broadsheets in their relative use of *narrative* as well as *expository styles*. Both examples argue for claims stated in the titles of the articles: China stole the coronavirus from Canada and weaponised it, and extreme weather events can pose a bigger threat when they occur in tandem with pandemics. Through the combination of both styles, the authors could provide a comprehensive and cohesive understanding of the issue, bolster their claims, and add to their credibility.

6.6.3. Dimension 3: Overtly Advocating Style

Table 6.11 shows the highly loaded positive features of Dimension 3. There are no negative features on this dimension, so texts with negative scores on this dimension are characterised simply by their lack of positive features. This dimension is highly similar to Biber's (1988) Dimension 4: *overt expression of persuasion/argumentation*. In Biber's analysis, professional letters were marked as overtly argumentative while broadcasts were marked as non-overtly argumentative. The positive features of the extracted dimension (*infinitives, conditional*

adverbial subordinators, necessity modals, *predictive modals*, and *suasive verbs*) all load highly on Biber's Dimension 4. *Contractions* on the positive pole and *analytical negation*, *nouns*, and *past tense* on the negative pole have higher loadings on other dimensions. Therefore, they were disregarded in the interpretation of this dimension.

Table 6.11 Dimension 3

Features with Positive Loadings	
Infinitives	**0.59**
Conditional adverbial subordinators	**0.38**
Necessity modals	**0.36**
Predictive modals	**0.35**
Suasive verbs	**0.34**
(Contractions	0.32)
Features with Negative Loadings	
(Analytical negation	-0.33)
(Nouns	-0.33)
(Past tense	-0.33)

6.6.3.1. Communicative Functions of Dimension 3 Features

Infinitives include clauses that begin with the infinitive 'to', which is followed by an optional adverb and a verb. These clauses function as complements to nouns, verbs, and adjectives. When preceded by verbs and adjectives, *infinitives* are often used to indicate the speaker's or writer's stance towards the proposition stated in the infinitival clause. Duffley (1992) indicated that *infinitives* are often preceded by verbs that convey meanings of desire, endeavour, command, and requirement (e.g., 'success', 'manage', 'persuade'). *Infinitives* can also function as an adverbial clause that expresses purpose, such as 'in order to' or 'so as to' (Biber, 1988; Biber et al., 2021). *Infinitives* can also be used to achieve integration and expansion of ideas because they allow for expressing ideas concisely (Biber, 1988). These functions result in *infinitives* being highly associated with argumentative discourse that is aimed at persuasion (Biber, 1988).

A related feature is ***suasive verbs*** (e.g., 'agree', 'ask', 'insist', 'recommend', 'beg', 'decide', 'pledge', 'suggest', 'urge'). Suasive verbs imply directives or intentions to bring about some change in the future (Biber, 1988; Hinkel, 2002). Thus, they are associated with persuasion. *Suasive verbs* are often followed by *infinitives*; this indicates that these two features are related

grammatically in addition to their common communicative function: persuasion. However, in the context of news reporting, which prioritizes information dissemination over persuasion, the use of *suasive verbs* is typically uncommon. To further understand the exact communicative function of suasive verbs in news articles, a random sample of their concordances is examined below in Figure 6-7.

Figure 6-7 Sample Concordances of Suasive Verbs in the Corpus

ıe face of international opinion and scientific analysis, I	suggest	that this attitude is born of weakness to withstand the pr
›"Boris Johnson's government has pointedly refused to	ensure	that UK companies are not aiding and abetting the destr
se.</s><s>Snopes is a fact checker – it checks facts to	ensure	they match the party line.</s><s>In such cases, the part
ENDING: British PM Theresa May Other health experts	agree	with Teague statement, including Dr. Mercola, who is on
/ to our attention here at Whiteout Press.</s><s>When	asked	what her connection to the vaccine-autism battle was, th
cided to lend a hand in pushing up vaccination rates by	requiring	certain employees to show proof of vaccination.</s><s>
ake the vaccine.</s><s>In August, Mayor Bill de Blasio	required	the city's 150,000 public-school employees to get the va
ccinate 300 million Americans, given that each vaccine	requires	two doses, by the end of the summer or the beginning o
ied some confusion and he had argued the UK "should	move	at the pace of the slowest country".</s><s>Following the
<s>On July 15, 2021, Pfizer and two of its subsidiaries	agreed	to pay $345 million under a proposed settlement to resol

The concordances of the *suasive verbs* reveal that they were used to report the opinions of certain organizations and public figures (underlined in Figure 6-7). These opinions express what should be done and how it should be achieved. Therefore, it can be inferred that the *suasive verbs* used in the corpus were *attributed*; that is, they were constructed as spoken by someone else rather than averred, i.e. not constructed as spoken by the writers themselves (cf. Hunston, 2011). The purpose of using these *suasive verbs* is to inform and influence public opinion, as organizations and public figures often hold significant influence over the public's views. In addition, references to public figures increase newsworthiness as it emphasises the news value of *eliteness*.

The conditional adverbial subordinators 'if' and 'unless' express direct conditions on which some actions and events are explicitly contingent (Hinkel, 2002). They are used to frame discourse and can refer to real and counterfactual present or past conditions (Biber, 1988). *Conditionals* can be used to hedge propositions or to elicit agreement with propositions the writers or speakers consider risky (Hinkel, 2002). Therefore, they are common in argumentative discourse (Biber, 1988).

Two types of modals are among the highly loaded features: *Predictive modals* ('will, 'would', 'shall') and *necessity modals* ('should', 'must', 'ought'). ***Predictive modals*** can mark future actions or events (volition) or discuss future possibilities and hypothetical or presuppositional situations (prediction) (Biber et al., 2021). The *predictive modal* of 'would' can also function as a hedge by reducing the writer's or speaker's responsibility for the reliability of evidence (Hinkel, 2002). ***Necessity modals*** are used to express personal obligation (Biber et al., 2021). They can also be used to indicate logical reasoning and conclusion-making, and they can indicate inference with a high degree of reliability (Biber et al., 2021; Chafe, 1986).

Taken together, these features are often used when structuring an argument in which purposes are identified, reasons are clarified, and stances towards certain ideas are expressed. These features can be used to assess the likelihood and advisability of an event, predict consequences, anticipate potential outcomes, stress obligation, and emphasise the urgency of a situation.

6.6.3.2. The Distribution of News Types along Dimension 3

Figure 6-8 shows the distribution of Dimension 3 scores for the texts of the five news types, and Table 6.12 provides the mean Dimension 3 scores and standard deviations for each news type. Tabloids exhibited the highest positive score, positioning them as the most overtly argumentative or persuasive news type, followed by web-based publications. Broadsheets and blogs had mean scores close to zero, with broadsheets scoring positively and blogs negatively. Notably, tabloids consistently achieved either the highest or lowest scores across all three dimensions, followed by web-based publications, broadsheets, and finally blogs.

False news on this dimension recorded the lowest mean score, indicating it is the least overtly argumentative or persuasive news type. Figure 6-8 illustrates that the false news corpus had the smallest spread (SD = .64). Thus, the false news articles' Dimension 3 scores were the most consistent.

Figure 6-8 The Distribution of the News Types along Dimension 3

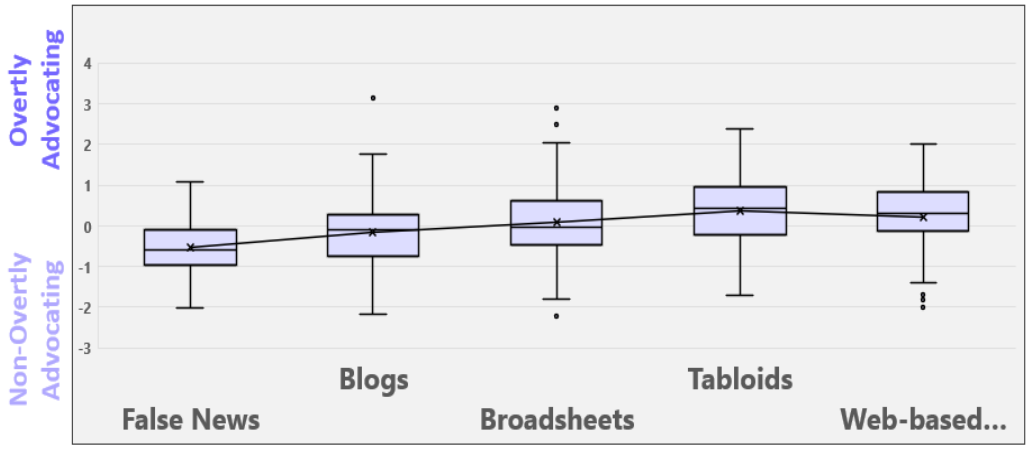


Table 6.12 Dimension 3 Scores for the Five News Types

News Types	Mean	Standard Deviation
Tabloids	.3808	.83405
Web-based Publications	.2248	.83815
Broadsheets	.1042	.93908
Blogs	-.1675	.83815
False News	-.5423	.63697

The one-way ANOVA result shows that the difference between at least two news types was significantly different (F (4,500) = [19.865], p = .000). According to Tukey post-hoc tests, false news was significantly different from blogs (p = .011, 95% C.I. = [-.6906, -.0591]), broadsheets (p = .000, 95% C.I. = [-.9622, -.3307]), tabloids (p = .000, 95% C.I. = [-1.2389, -.6074]), and web-based publications (p = .000, 95% C.I. = [-1.0828, -.4513]).

As stated in Chapter 3, the reference corpus contained both opinion as well as news articles, thereby allowing it to match the focus corpus of false news. Opinion articles, like editorials, are inherently persuasive or argumentative (Biber & Conrad, 2019). The news article (Example 8) and the opinion article (Example 9) that had the highest Dimension 3 scores from the tabloid corpus were examined[9]. This analysis was necessary for two reasons: (a) to better interpret the dimension, and (b) to verify the assumption that opinion articles tend to be persuasive while

[9] The metadata of the tabloid articles provided by Nexis.uk indicate the section (news or opinion) in which a certain article was written. This was used to identify the opinion and news article that have the highest Dimension 3 scores.

news articles do not. The positive features in both examples are indicated by boldface. The false news article that had the lowest Dimension 3 score (Example 10), representing the negative pole of this dimension was also examined.

The news piece presented in Example 8 supports the bill that would bar parents from not vaccinating their children. The *suasive verbs* 'propose' and 'vote' show that important figures such as state lawmakers and the Senate support this bill. The argument here is clearly *attributed* and not *averred*. The *infinitive* used in 'to protect the greater good' explains the purpose of the bill: ensuring public safety. Moreover, the *predictive* 'will' shows that the bill is expected to be approved by the Assembly by obtaining enough votes. Additionally, the *conditional adverbial subordinator* 'if' shows the potential consequences if parents choose to not vaccinate their children. Although *necessity modals* are not used in the text, the verbs 'has to' and 'is needed' express obligation and stress the necessity of mandating vaccines. All in all, these features emphasise the importance of the bill and encourage the readers to accept it.

> (8) New York remains the epicenter of the epidemic, with 66 of the 75 new cases in the city and nearby Rockland County, prompting state lawmakers **to propose** a bill that **would** bar parents from not vaccinating their children.
>
> Sen. Brad Hoylman (D-Manhattan) and Assemblyman Jeff Dinowitz (D-Bronx), the sponsors of the bill, argue that mandating vaccines, other than for medical reasons, is needed **to** protect the greater good.
>
> "The health of people has **to** come first, period," Dinowitz said. "**If** parents believe they have the right **to** decide what's best for their child, maybe they're right up to a point, but they don't have the right **to** make decisions that could adversely impact other children."
>
> The Senate is expected **to vote** on the bill next week. Dinowitz said he believes the Assembly **will** get on board soon, despite some reservations among the members.
>
> "I believe **if** the bill does come to the floor we **will** have enough votes **to** pass it," he added. (*B Tab 20,* Dim 3 score = 1.54)

In the opinion piece presented in Example 9, the author is calling for the blocking of an anti-vaccination film and for disseminating the truth about vaccination. The author effectively employs *suasive verbs* such as 'demand,' 'decide,' and 'agree' to assert their viewpoint on vaccination alongside that of influential figures, such as the mayor and medical experts.

Moreover, the author utilizes the *necessity modal* 'should' to underscore the idea that complete freedom regarding vaccination should not be absolute and that action needs to be taken. Furthermore, the author employs *infinitives* to clarify the film's objective, which is to 'spread unjustified fear about vaccination.' The *infinitives* also were used to express the author's feeling that they were 'proud to reveal that [they] helped block' the film and did not 'hesitate to' do so. The use of these linguistic features conveys the author's stance, which is clearly *averred* in this opinion piece and encourages readers to support vaccination.

> (9) We believe in freedom of speech. But on the issue of vaccination, we believe that freedom **should** not be absolute. So today, we're proud **to** reveal that we helped block the Australian premiere of an unethical and dangerous "documentary" intended **to** spread unjustified fear about vaccination.
>
> When we found out an anti-vaccine group was boasting they had found a venue in Logan, Brisbane, **to** air the film - which has already been dumped by film festivals in the US and Australia - we did not hesitate **to** find out the name of the council-owned venue, ring the federal and state MPs, the Queensland Health Minister and the Australian Medical - Association, and present their unanimously horrified reaction to the mayor of Logan City. We **demanded to** know why the Council was allowing this piece of propaganda **to** be aired in a ratepayer-funded venue.
>
> Sensibly, the mayor **decided to** pull the film.
>
> …
>
> So although the truth about vaccination is that the entire world's medical experts **agree** that it saves lives, there is a persistent and dangerous conspiracy theory that's been running since the 1990s. It's utterly false - a claim, by a disgraced doctor in a long-retracted journal article, that childhood vaccinations somehow caused autism. You**'d** be amazed how many Australians still believe the lie, even though there has never been a shred of real evidence. (*B Tab 22,* Dim 3 score = 1.42)

Both Examples 8 and 9 are calling for some sort of action, although they represent two distinct sub-registers: news stories and opinion. They both promote the importance of vaccination and condemn the anti-vaccination movement. However, the news article is more objective as it reports public figures' stances while the opinion article is more obviously subjective as it

presents the author's own stance. In other words, the degree of the author's involvement and their expression of their feelings and attitudes are more explicit in the opinion article than in the news article. This can be reflected in both articles' Dimension 1 scores: The news article scored 0.17, while the opinion article scored 1.48, indicating a higher orientation towards the *involved style* in the opinion piece.

Example 10 is an extract from a false news report having the lowest Dimension 3 score. The positive features of Dimension 3 are notably absent in the report. It does not explicitly call for action, express obligations or logical reasoning, identify purposes, or predict consequences. The report states facts and figures as reported by the CDC website. Instead of overtly arguing for a point of view, the news report quotes sources and presents facts to substantiate the claim that COVID-19 cases were overcounted in hospitals.

> (10) BOOM! CDC Director Finally Admits that COVID Cases were Over-Counted in Hospitals– Just as Gateway Pundit and Donald Trump Reported in August.
>
> In August 2020 The Gateway Pundit's Joe Hoft posted an article about COVID-19 that sent shockwaves throughout the internet — The CDC website admitted that only around 6% of COVID recorded deaths were due entirely to the Coronavirus-
>
> The gist of the report was that COVID-19 is not nearly as deadly as first projected by the WHO and then by Dr. Tony Fauci and Dr. Debra Birx. Based on CDC numbers in August only 6% of all deaths attributed to COVID-19 were instances where the only factor in the individual's death was due to COVI9-19.
>
> From the CDC website.
>
> For all the other deaths reported by the CDC linked to COVID-19, the individuals who passed away had 2-3 other serious illnesses or co-morbidities. (*C False 59,* Dim 3 score = -2.01)

After analysing the given examples, it is evident that the label of *overtly advocating vs. non-overtly advocating style* is a more appropriate description for this Dimension than the labels of 'persuasive' or 'argumentative.' This conclusion is based on three key reasons. Firstly, the analysis revealed that the positive features of Dimension 3 serve a greater purpose than simply persuading readers to adopt a specific viewpoint. These features are utilized to inform the public about a particular issue, raise awareness, create a sense of urgency, and ultimately influence public

opinion and policy. In essence, they are employed to advocate for a particular viewpoint. Secondly, the authors employ the positive features to make arguments, but they do so within the norms and conventions of journalism. As the analysis showed, the authors extensively use attributed materials to maintain objectivity in their reporting. Thirdly, this label is derived from journalism; *'advocacy journalism'* is a recognized genre within journalism. As Jensen (2008) puts it, "advocacy journalism describes the use of journalism techniques to promote a specific political or social cause. The term is potentially meaningful only in opposition to a category of journalism that does not engage in advocacy, so-called *objective* journalism" (p.1).

6.6.3.3. False News and Dimension 3

While false news appeared to share similarities with other news forms in the first two dimensions, which represent dominant stylistic patterns, a clear contrast emerged upon further examination. False news stood out from all other news types in this third final extracted dimension. Although the false news article that had the lowest Dimension 3 score was examined in the previous section, examining the false news article that had the median score (Example 11) was necessary to understand the position of false news on this dimension.

Example 11 lacks the features of the *overtly advocating style*, except for the words: 'would', 'unless', 'to', and 'beg' found in one segment of the text. Although the text argues that the National Institute of Allergy and Infectious Diseases (NIAID) and Moderna had known that the COVID-19 outbreak was coming and that they might have caused the outbreak, it does not explicitly advocate for this perspective and does not ask readers to accept it throughout the text. Instead, it simply reports events, provides visual evidence, uses emphasis by capitalizing the words 'BEFORE' and 'CAUSE', and uses the rhetorical question: 'Did they CAUSE the outbreak?'. In this way, the text is meant to be suggestive, allowing the reader to draw their own conclusions by themselves and encouraging further investigation.

> (11) Confidential documents have surfaced in a Confidentiality Agreement between the U.S. National Institutes of Allergies and Infectious Diseases (NIAID) and Moderna showing they transferred "potential coronavirus vaccine candidates" to the University of North Carolina, on December 12, 2019
>
> . . . which was nineteen (19) days BEFORE the outbreak of novel coronavirus in China!

> How could they know they **would** need a Coronavirus Vaccine BEFORE the outbreak, **unless** they had knowledge the outbreak was going **to** happen?
>
> This **begs** the question: Did they CAUSE the outbreak?
>
> Above is the header portion of Page 105 of the Confidentiality Agreement. It deals with the actual physical transfer of "mRNA coronavirus vaccine candidates" being sent to the University of North Carolina at Chapel Hill.
>
> It is signed on December 12 by Ralph Baric, PhD, at the University of North Carolina at Chapel Hill, who is listed as the "Recipient" of the mRNA Coronavirus Vaccine candidates, Image below: (*C False 60,* Dim 3 score = 0.60)

The analysis of Example 11 underscores the unique position of false news in Dimension 3. Despite its suggestive nature, the article refrains from *overt advocacy* while hinting at pharmaceutical companies' role in the COVID-19 pandemic. By reporting events and providing visual evidence, the article encourages readers to draw their own conclusions but in a relatively objective manner. Indeed, the avoidance of *overt advocacy* in such a conspiracy theory context highlights the intricacies of false news discourse.

6.7. Discussion

This case study examines whether false news articles circulating online demonstrate distinct stylistic patterns when compared to broadsheets, tabloids, web-based publications, and news blogs. To address this question, multidimensional analysis was employed to initially identify the stylistic patterns in these various types of news and then to analyse the differences and similarities among them based on these stylistic patterns. This section first summarizes these dimensions, discusses the similarities and differences between false news and other news types in terms of these dimensions, and offers interpretations for this comparison. Second, the section discusses the limitations of MDA in answering the research question and suggests improvements.

6.7.1. Stylistic Similarities and Differences Between False News and Other News Types

The analysis identified three dimensions. Dimension 1 differentiates between the *involved style* and the *informationally dense style*. The *involved style* emphasises a range of news values, engages readers, and heavily relies on quotations to convey the experiences, attitudes, judgments,

or emotions of news actors. On the other hand, the *informationally dense style* focuses on providing detailed information on a news event. Dimension 2 distinguishes between the *expository style* and the *narrative style*. The *expository style* aims to provide technical, detailed, or descriptive information, while the *narrative style* narrates a sequence of past events. Dimension 3 represents the contrast between the *overtly advocating style* and the *non-overtly advocating style*. The *overtly advocating style* explicitly advocates for a specific viewpoint and encourages readers to make informed decisions, whereas the *non-overtly advocating style* does not explicitly advocate for any particular viewpoint.

On Dimension 1, false news articles demonstrated similarities with traditional news types, including tabloids and web-based publications, aligning closely with broadsheets by incorporating the *involved* and *informational styles*. Notably, false news articles diverged significantly from news blogs, which were found to lean towards the *informational style*. On Dimension 2, false news articles were found to exhibit resemblances to broadsheets, as both utilized characteristics of the *narrative* and *expository styles*.

Emulating the style of traditional, reliable, and reputable news sources can lead to multiple benefits for false news writers. Firstly, through the *involved style*, false news enhances the newsworthiness of its content and captivates its readers. Secondly, through the *informational*, *narrative*, and *expository*, false news writers establish a sense of expertise and knowledge, bolstering their claims and enhancing their credibility.

The striking resemblance between false news and broadsheets indicates the successful adoption of a professional news writing style by false news writers. This underscores the concerning implications of the widespread dissemination and acceptance of false news, which can be attributed to its skilful emulation of the style commonly found in broadsheets. Moreover, these parallels highlight the challenges of distinguishing false news from true news based solely on style and underscore the limitations of style-based detection algorithms in effectively addressing this issue, as argued by Bondielli and Marcelloni (2019) and Castelo et al. (2019).

Despite these similarities, false news differed significantly from all other news types on Dimension 3. Unlike other news types, false news tended not to overtly advocate for a specific viewpoint. This pattern is supported by a study conducted by Grieve and Woodfield (2023), who examined fake news and real news stories written by Jayson Blair. They found that Blair's real

news articles often included linguistic elements related to stance expression and evidentiality such as *suasive verbs*, *infinitives*, and *predictive modals*. In contrast, Blair's fake news articles lacked these features, further supporting the idea that false news lacks features of overt persuasion, argumentation, or advocacy.

One possible explanation for this is the insincerity in false news. False news writers, lacking reliable information, may refrain from openly advocating for their claims to avoid taking responsibility for potentially misleading information. This notion was confirmed by Grieve and Woodfield (2023), attributing the absence of persuasive features in Blair's fake news to a lack of conviction.

Furthermore, false news may not overtly advocate because it is primarily targeted at an audience that already agrees with the argument being discussed. Research has shown that readers who consume and believe fake news often share the same ideological perspectives as the writers of fake news (Osmundsen et al., 2021b; Pennycook & Rand, 2021; Rini, 2017). This alignment may explain the lack of overt advocacy in false news, as writers may cater to an audience predisposed to agree with the content.

Another possible explanation for the absence of overt advocacy in false news could be viewed as a defensive strategy. Since the analysed false news articles touch on controversial topics and diverge from mainstream narratives, the writers may anticipate backlash against their claims and choose to present them as factual rather than opinionated. This defensive approach may be aimed at building credibility, highlighting impartiality, and maintaining objectivity.

6.7.2. Limitations and Future Research Directions

This study has three notable limitations that raise questions for future research. First, while the study identified various stylistic patterns used in false news, it did not explore when and in what contexts these patterns are employed. Future studies should investigate the presence or absence of these identified patterns in false news across different contexts and topics, such as political campaigns or public health crises. By considering the contextual factors that influence the usage of these patterns, researchers can gain valuable insights into the mechanisms that shape the production and dissemination of false news. For example, exploring how false news employs

specific stylistic approaches in response to different contexts or topics can shed light on the rhetorical strategies employed to make it appealing, convincing, and viral.

Second, while this study revealed the stylistic patterns used in various news types, and whether false news tended to utilize them, it did not provide a conclusion on the overall writing style of false news. To fully understand what sets the style of false news apart from other news types, it is necessary to identify the lexico-grammatical features that are more prevalent in false news. and analyse how these features contribute to the overall style of false news. To achieve this, alternative corpus methods such as key features analysis developed by Biber and Egbert (2018) or the framework proposed by Grieve and Woodfield (2023) can be employed. These approaches might help uncover the unique style of false news.

Another limitation to consider is that the commonly employed list of lexico-grammatical features, proposed by Biber (1988) and available in Nini's (2015) MAT does not directly facilitate the analysis of meanings which are more reflective of communicative functions than lexico-grammatical features. For example, while MDA includes *suasive verbs* as an analytical variable, it does not allow for the identification of the specific meaning conveyed by these verbs. This is because these verbs can express a range of meanings, including requests, decisions, permissions, recommendations, preferences, agreements, and intentions. To overcome this limitation, an alternative approach to MDA involves using semantic taggers instead of POS taggers to find out co-occurrence patterns of meanings. Semantic taggers could allow for the identification of dimensions of variation based on meanings. Accordingly, this approach can provide a more in-depth understanding of the communicative functions within a discourse.

6.8. Conclusion

In conclusion, this study examined the stylistic patterns used in online false news articles in comparison to other news types. The findings showed that false news differed from other news types in the sense that it did not overtly advocate or argue for a specific viewpoint. There are three possible interpretations for this finding. Firstly, false news writers may lack confidence in advocating for their claims due to insincerity or a desire to avoid responsibility for disseminating misleading information. Secondly, false news might primarily be aimed at a receptive or already convinced audience. Thirdly, the avoidance of overt advocacy may be seen as a defensive

strategy; false newswriters may anticipate challenges to their controversial claims and seek to present them as facts rather than opinions. By doing so, they emphasise credibility, neutrality, and objectivity. However, the analysis revealed that false news exhibited several stylistic similarities with traditional reliable news, particularly broadsheets, as false news writers employed *involved, informational, expository*, and *narrative styles* in their articles. These communicative styles serve multiple purposes, such as constructing newsworthiness, establishing the writer's expertise, reinforcing claims, and enhancing credibility. The stylistic similarities between false news and other traditional reliable news sources create a discourse camouflage, potentially enabling false news to be disseminated widely and accepted by readers.

It is important to acknowledge that this study specifically examined the surface presentation of false news, with a focus on the stylistic aspect. However, in order to develop a more comprehensive understanding of false news discourse, it is crucial to delve deeper into the meanings negotiated and rhetorical strategies utilized within false news. Such an analysis can provide insights into how false news distinguishes itself from true news and effectively persuades its readers with a particular viewpoint.

Chapter 7: General Discussion

7.1. Introduction

This study explores the discourse characteristics of online false news. To investigate this question, I utilized a mixed-method approach involving three corpus-assisted analytical methods: move analysis (MA), key semantic domains analysis (KSDA), and multidimensional analysis (MDA). These methods were used to examine the discourse structure, news values, and stylistic choices of false news respectively. The goal of triangulating these different methods was to gain a comprehensive understanding of false news discourse and increase confidence in interpreting the findings.

This chapter is divided into five main sections. The first section provides a summary of the results of the three case studies and their proposed interpretations. The second section discusses the implications of the results of the three case studies in relation to each other, considering how these studies collectively enhance our understanding of the discourse of false news. The third section evaluates the corpus used in this study, focusing on its relevance in analysing the discourse of false news and its broader implications, especially when compared to corpora utilized in previous studies on fake news. The fourth section discusses the benefits and challenges of using the mixed-method approach in this thesis. The final section offers recommendations for future research on false news. It is important to note that the detailed discussions and interpretations of the findings, limitations of the employed methods, and future research directions for each case study are provided in their corresponding chapters. This chapter serves as a synthesis of and a reflection on the overall findings from the thesis.

7.2. Summary of the Case Studies

In this section, a summary of the three case studies is presented, highlighting their research questions, methods, the corpus analysed, and the findings, along with providing interpretations for each.

7.2.1. The First Case Study

The first case study (presented in Chapter 4) aimed to analyse the discourse structure of online false news articles. It utilized genre analysis, specifically the analytical method of MA proposed by Swales (1990), to examine 20 false news articles on climate change. The MA yielded three significant findings.

Firstly, the rhetorical moves which the analysis identified as being distinctive of false news articles suggested that these texts fulfil one primary communicative purpose: challenging mainstream narratives. In addition to this purpose, some articles also aimed to delegitimize political and religious figures. Secondly, the analysis revealed that while the structure of false news articles shared some common elements with traditional hard news reports, it aligned more closely with argumentative essays. Like traditional hard news reports, false news articles demonstrated elements of narrative and expository genres through the inclusion of lead sections, main story events, background information, and descriptions of consequences. However, false news articles also exhibited structural elements typical of an argumentative genre by expressing and restating a central thesis, presenting and refuting opposing views, drawing conclusions, offering recommendations, calling to action, and citing references. The case study suggested that the combination of narrative, expository, and argumentative elements in false news articles seems to enhance their appeal, as they resemble typical hard news reports with an argumentative touch.

Lastly, the analysis uncovered a notable lack of a fixed move structure in false news articles. Moves were observed in various positions across articles and were sometimes repeated within one article, resulting in incoherence. This finding can be explained in several ways. Firstly, it suggests that the structure of false news articles is often customized to serve the writers' specific objectives. Moreover, the lack of coherence can be attributed to the absence of factual accuracy. It reflects the deliberate intention of writers to manipulate readers by employing a range of strategies repeatedly within a single article, such as promoting alternative theories, providing background information, describing consequences, and issuing calls to action, in order to support their main thesis. Additionally, the rapid production of false news articles, driven by the fast-paced and viral nature of online news, may contribute to the incoherence. Lastly, the fact that

false news writers are generally not professional journalists may further contribute to the lack of coherence in their articles.

7.2.2. The Second Case Study

The second case study (presented in Chapter 5) aimed to analyse the news values constructed in online false news articles. Unlike previous studies that used corpus analysis to validate the presence of pre-theorized news values (e.g., Bednarek & Caple, 2017), this study utilized the KSDA, proposed by Rayson (2008), to identify potential news values that might be distinctive of false news discourse. To identify key semantic domains, the study compared the corpus of false news on climate change, vaccinations, and COVID-19 (a total of 137 articles) to the reference corpus of broadsheets, tabloids, web-based publications, and news blogs on the same topics (a total of 548 articles). The goal was to identify prevalent meanings and patterns of usage of these meanings in false news articles and determine whether they constructed news values similar to or different from news values established in the literature.

The analysis revealed seven major semantic categories that are distinctive of the false news corpus: (a) connecting ideas, (b) evidentiality, (c) evaluation, (d) religion and the supernatural, (e) science, (f) time, and (g) news topics. The qualitative analysis of the concordances of the lexis under these semantic categories showed that false news articles constructed news values that have been discussed in previous literature, such as *negativity*, *unexpectedness*, and *consonance*.

The analysis also identified news values that appear to be unique to false news articles. These values included *scepticism*, where credible sources were discredited, as indicated by the frequent use of lexis related to the evaluation of credibility, morality, and rationality. Another value was *corroboration*, where lexis denoting different types of evidence were prevalent, such as sensory, spoken, written, or deductive evidence. *Causality* was another value, where the cause and effect of events were explained, often using lexis involving cause and effect relations and scientific explanations. *Mysticism* was also a value, as indicated by frequent references to religious, spiritual, or mythical elements. Finally, the news value of *historicity* was emphasised, highlighting the historical relevance of current events and claiming historical authenticity, through the frequent use of words related to time.

The prominence of these news values in false news articles can be attributed to their objective of challenging mainstream narratives. They adopt a sceptical stance, requiring various forms of evidence and persuasive tactics that appeal to logic and religious beliefs and demonstrate historical knowledge to refute mainstream narratives. The qualitative analysis showed that these values function as rhetorical devices, enabling false news writers to make their articles more captivating, convincing, and credible.

Moreover, the analysis demonstrated that these news values reflect certain ideologies, such as reinforcing distrust in mainstream knowledge and influencing religious beliefs. Furthermore, these news values allow us to speculate about the types of audiences commonly affected by false news, these include sceptics, religious individuals, and emotionally vulnerable individuals. When considered collectively, these news values have the potential to significantly contribute to the widespread dissemination of false news.

7.2.3. The Third Case Study

In the third case study (presented in Chapter 6), the objective was to analyse whether online false news articles (a total of 101 articles) on climate change, vaccinations, and COVID-19 displayed unique stylistic patterns when compared to articles on the same topics from reliable and reputable sources. These reputable sources included printed broadsheet and tabloid publications, web-based articles, and news blogs, with a total of 504 articles being examined. The study employed MDA, proposed by Biber (1988), to identify dimensions of stylistic variation (i.e., patterns of co-occurring lexico-grammatical features which collectively perform a communicative function) in these types of news and compare them according to these patterns.

The analysis identified three dimensions. Dimension 1 differentiates between the *involved style*, which extensively utilizes quotations to convey the experiences and attitudes of the news actors, engages the readers, and emphasises several news values, and the *informational style*, which focuses on providing detailed information about an event. Dimension 2 distinguishes between the *expository style*, which offers technical and detailed descriptive information about complex topics, and the *narrative style*, which involves narrating a sequence of events. Dimension 3 represents the contrast between the *overtly advocating style* and the *non-overtly advocating style*.

The features associated with the *overtly advocating style* are consistent with the features of the *overt expression of persuasion/overtly argumentative* dimension identified in Biber's (1988) MDA. However, the qualitative analysis revealed that texts with these features did not only aim to persuade readers but also aimed to raise awareness about a particular issue, influence public opinion, and encourage action. Authors employed this style to make arguments while adhering to journalism norms and conventions, often using attributed materials to maintain objectivity in their reporting. Thus, the term 'overt advocacy', derived from journalism, best describes this dimension.

On Dimension 1, false news articles were similar to the traditional and reputable news types: broadsheets, tabloids, and web-based publications, as they incorporated elements of both the *involved* and *informational styles*. On Dimension 2, false news resembled broadsheets as both utilized characteristics of both the *narrative* and *expository styles*. In this way, false news writers effectively mimic the style of reliable and reputable news sources.

Contrary to expectations, false news stood out from all other news types on Dimension 3 in that it does not tend to overtly advocate for a specific viewpoint. The study provided three explanations for this finding. Firstly, false news writers may lack confidence in advocating for their claims due to insincerity or a desire to avoid responsibility for disseminating misleading information. Secondly, false news might primarily be aimed at a receptive audience, eliminating the need for explicit advocacy. A third explanation could be that the avoidance of overt advocacy may be seen as a defensive strategy. While false news is inherently persuasive, writers may anticipate challenges to their controversial or non-mainstream claims and seek to present them as facts rather than opinions. By refraining from overt advocacy, writers may aim to emphasise credibility, impartiality, and objectivity.

7.3. Comparing the Case Studies: Insights and Connections

This section examines the similarities and differences between false and true news, and it discusses the complexity of false news discourse as highlighted by the three case studies. It further investigates the broader implications arising from the collective findings of these case studies.

7.3.1. Discourse Camouflage: The Alarming Resemblance between False and True News

Online false news discourse mirrors some aspects of true news discourse, as indicated by the MA and MDA. Specifically, false news articles incorporate structural elements commonly seen in traditional hard news reports, including the lead, description of main story events and consequences, and the supplement of background information and comments. Given that these structural elements are typically found in reliable news sources, the study proposes that false news articles share some structural similarities with true news articles. Additionally, false news has also been found to share stylistic similarities with reliable news sources, particularly broadsheets, with both false news and broadsheets using *involved, informational, narrative*, and *expository styles*.

The similarities between false and true news discourse, particularly reliable broadsheets, have important implications as these resemblances underscore that false news can adopt a journalistic format that closely resembles credible sources. This similarity is concerning and poses a significant threat due to it making false news appear more credible, thereby increasing the likelihood of its widespread dissemination and acceptance as accurate information. Scholars have emphasised this danger, noting fake news' tendency to adopt a journalistic format similar to that of reliable sources (Bondielli & Marcelloni, 2019; Dentith, 2016; Rini, 2017).

Recognising these parallels is crucial for understanding the challenges faced by current detection methods, guiding future research on false news detection, and improving efforts to combat false news. Indeed, these findings offer an explanation for why the style-based algorithms used to detect fake news have been of limited value up until now (cf. Castelo et al., 2019), false news is increasingly adopting the style of true news, and automated detection algorithms struggle to accurately differentiate between the two. MDA has been used in this study to comprehensively examine a wide range of linguistic features and identify multiple distinct dimensions, with one explainable difference being identified: False news does not tend to overtly advocate for a position. This indicates that the stylistic differences between false and true news are highly complex and cannot be determined by investigating a simple set of grammatical features. Therefore, studies focusing on fake news detection should explore deeper discourse levels in the development of more advanced automatic detection algorithms.

Moreover, this finding highlights the limitation inherent in relying *solely* on superficial stylistic features to differentiate false news from true news. As false news frequently camouflages itself as professional journalism, it is essential to conduct thorough analyses of news articles to identify falsehoods, including meticulous examinations of their communicative functions, meanings, values, ideologies, and biases. This perspective aligns with multiple researchers' conclusions (e.g., Grieve & Woodfield, 2023; Sabbah, 2023; Vamanu, 2019), all of whom argue that carefully reading news items and applying critical discourse analysis methods can be valuable for detecting fake news.

7.3.2. The Unique Characteristics of False News Discourse

While the study has identified similarities between false news and true news discourse, distinct characteristics of false news have also been revealed by the three case studies. First, false news leans to adopt an argumentative discourse structure while also incorporating the narrative and expository elements that are typical of traditional news genres. Second, false news constructs distinct news values, such as *scepticism, causality, corroboration, mysticism*, and *historicity*. Lastly, false news does not explicitly advocate for a position, unlike trustworthy news sources. The combination of these findings allows for a more comprehensive understanding of false news discourse and its impact on its audience. Below, a possible rationale for why these traits are characteristic of false news is offered, and the way in which these traits contribute to false news's influence on its audience is elucidated.

7.3.2.1. Challenging Mainstream Narratives and the Post-Truth Era

A possible explanation for the three characteristics discussed above is false news articles aiming to challenge established beliefs, potentially in an effort to sow doubt among readers. For instance, articles which assert that the flu shot spreads cancer, climate change is a hoax, and COVID-19 is a Chinese bioweapon[10] directly contradict mainstream narratives. This rationale is supported by the findings of the three case studies. The MA has illustrated that online false news articles employ an argumentative discourse structure to challenge predominant narratives while endorsing alternative viewpoints. The KSDA has indicated that false news incorporates the news value of *scepticism*, aligning with this objective, and that false news leverages various news

[10] These claims are derived from the corpus (Articles' ID: *B False 11, A False 6, C False 1*).

values, including *consonance, unexpectedness, negativity, causality, corroboration*, and *historicity,* to bolster the persuasiveness of such scepticism. The MDA has underscored that the absence of overt advocacy may stem from authors knowingly challenging mainstream media perspectives by presenting their claims as factual without explicit promotion, foreseeing potential pushback and controversy.

These findings suggest that false news writers are actively pursuing the shaping of public opinion and influencing media discourse. This underscores a power struggle within the news sphere, in which the production and dissemination of information are fiercely contested. Moreover, these findings suggest that we are currently living in a post-truth era, which is characterised by a growing scepticism towards established knowledge and decreased trust in reliable media and governmental organisations. This ideological climate of distrust is concerning as it challenges the foundations of truth and knowledge, making it increasingly difficult for individuals to make well-informed decisions. As Brennen et al. (2020) point out:

> Governments have not always succeeded in providing clear, useful, and trusted information to address pressing public questions. In the absence of sufficient information, misinformation … may fill in gaps in public understanding, and those distrustful of their government or political elites may be disinclined to trust official communications on these matters. (p. 6)

7.3.2.2. Audience-Centric Crafting of False News Discourse

The three case studies indicate that false news's primary goal of challenging mainstream theories results in it being inherently argumentative or persuasive. In this context, false news writers must have meticulously crafted their articles to engage audiences who may disagree with their content or who may be open to considering it but are seeking additional supportive evidence. Although the MDA has shown that false news discourse does not use overt advocacy, indicated by the lack of lexico-grammatical features associated with overt argumentation or persuasion, the findings indicate that false news writers use other persuasive tactics, as explained below.

The findings from the KSDA suggest that false news strategically emphasises distinct news values to fulfil persuasive intentions effectively. False news writers provide written, spoken, deductive, and historical evidence through *corroboration* and *historicity*. In addition, false news

writers appeal to individuals' logic by explaining the reasoning through *causality* and by providing analogical arguments through *unexpectedness* and *consonance*. Furthermore, false news tactics extend to exploiting emotions or religious beliefs through *negativity* and *mysticism* in an effort to maximise false news articles' persuasive impact on readers. By emphasising these news values, false news writers can craft narratives that not only challenge prevailing beliefs but also resonate with readers on a logical and emotional level with the ultimate goal of changing readers' perspectives.

Moreover, the narrative and expository elements identified in the case studies may also serve persuasive intentions. For example, the findings from the MDA have shown that false news, similar to broadsheets, employs both *expository* and *narrative styles*. Additionally, the KSDA has highlighted the coexistence of expository elements, characterized by scientific and technological vocabulary emphasising cause-and-effect relationships, and narrative elements, exemplified by the inclusion of time-related words in these stories. These collective observations suggest that false news often uses storytelling techniques and scientific-sounding language to present an illusion of credibility. By narrating both recent and historical events, false news writers attempt to claim historical authenticity and establish a connection between current and past events, thereby providing apparent historical evidence. Similarly, by leveraging scientific and technical terminology and concepts, false news writers aim to convince readers that their claims are grounded in empirical evidence or expert knowledge. These findings suggest that false news writers purposely present themselves as knowledgeable and well-informed by projecting a sense of expertise that is beyond their readers' reach.

These findings challenge studies which have suggested that fake news writers predominantly employ heuristic persuasion techniques while avoiding logical reasoning and technical language. Existing research has also implied that individuals who lack critical thinking skills and those who readily trust news sources are more susceptible to believing false news (Bronstein et al., 2019; Horne & Adalı, 2017; Pennycook & Rand, 2021; Rini, 2017). However, technical and scientific language being used in false news articles along with their emphasis on various types of evidence, including historical evidence, suggest that even individuals with critical thinking skills can still come to believe false information. These findings challenge the notion that only certain

groups are vulnerable to false information and highlight the complexity of the factors that influence individuals' susceptibility to false news.

The avoidance of overt advocacy can also be viewed as a strategic choice: By refraining from explicit lexico-grammatical features that indicate a clear stance, false news authors adopt a journalistic style that emphasises objectivity and neutrality. This approach presents their claims as factual rather than subjective opinions, bolstering the perceived credibility of false news for readers and potentially increasing its influence.

Bell's (1991, 2002) framework of *audience design* can elucidate why false news exhibits these unique characteristics. Bell contends that speakers or writers design their style primarily for and in response to their audience. According to Bell, communicators can be both *responsive*, adjusting their style based on their changing audience, and *initiative*, using various language resources creatively to effectively engage the audience and establish a clear identity. Thus, *audience design* "informs all levels of a speaker's linguistic choices" (Bell, 1991, p. 105). In the context of false news discourse, writers aim to challenge mainstream narratives and engage audiences with differing viewpoints. Anticipating the audience's reactions, false news writers craft their content in a way that enhances persuasiveness and credibility. They manipulate news genre conventions in a way that diverges from true news reports, ultimately aiming to influence their audience's perspectives. Indeed, false news discourse's use of an argumentative structure, adoption of news values that cater to scepticism and appeal to emotions or religious convictions, utilisation of technological and scientific explanations, and deliberate avoidance of overt advocacy are all strategies intended for audiences with a low level of credulity. Therefore, the distinctiveness of false news compared to true news lies in its communicative goal of questioning mainstream theories as well as targeting the audience's credulity.

7.3.3. The Complexity of False News Discourse

In addition to revealing the general patterns found in false news discourse, including those that are similar or different to those of true news discourse, this study shows that false news discourse is a complex and multifaceted phenomenon. For example, the MA has demonstrated that false news articles do not follow a consistent move structure. Instead, false news writers seem to tailor their articles' discourse organisation to fit their content and motivations.

Another complexity is that while the MA has indicated that false news articles follow an argumentative structure, the MDA has found that false news does not explicitly advocate or argue for a position. These apparently differing results can be attributed to each analysis being focused on different aspects of false news discourse. The MDA has focused on lexico-grammatical features as units of analysis, thus capturing more surface-level presentational or stylistic aspects, while the MA has entailed an investigation of the moves, which are larger discourse segments that perform rhetorical or communicative functions. However, the combination of the results indicates that false news uses an argumentative discourse structure to challenge mainstream narratives without explicitly adopting an argumentative or advocating style. False news articles do state a main thesis, refute mainstream theories, and promote alternative ones, albeit with fewer *suasive verbs*, *infinitives*, *conditional adverbials*, and *predictive* and *necessity modals* (i.e., features associated with overt argumentation). In other words, false news aims to persuade readers to believe in theories that contradict conventional wisdom by adopting a journalistic style that emphasises objectivity and neutrality, more often than true news sources, and by using various persuasive tactics, outlined in Section 7.3.2.2.

Nevertheless, it should be noted that the dataset used in the MA was small as it only focused on one topic, while the MDA dataset comprised a larger corpus that encompassed three different topics and was comparative in nature. This difference in the corpora's sizes could explain these complex findings. Therefore, future research should analyse a larger false news corpus from the genre perspective.

In conclusion, the three case studies' findings complement one another and provide a comprehensive understanding of false news discourse. The analyses demonstrate that false news, while appearing similar to reliable news sources, employs unique discourse structures, news values, and styles to achieve its persuasive and viral goals. However, false news discourse is a complex and multifaceted phenomenon that needs further investigation.

7.4. Evaluating the Corpus: Comparative Analysis with Grieve and Woodfield's Corpus

The creation of a corpus of false news has sparked considerable debate (Asr & Taboada, 2019; Grieve & Woodfield, 2023). This section reflects on the corpus used in this study, its relevance to the analysis of false news, and the implications that can be drawn from analysing it, particularly in comparison to the corpus used by Grieve and Woodfield (2023).

In this thesis, I employed a common research practice for studying false news, which included collecting verifiably false news articles from fact-checking sites and comparing them to news articles obtained from reputable and reliable sources. While this method allows for the collection of a large dataset, a major limitation of this method is the inability to ascertain the authors' intentions behind their creation of false news. However, determining intention can be challenging, and there are a few cases of proven deceptive intent. By focusing only on these cases, the prevalence of partisan and harmful false news spread online may be overlooked.

Grieve and Woodfield (2023) argue that linguistic differences between real and fake news cannot simply be caused by the falsehood of the information, instead proposing that different communicative intents behind writing fake or real news (i.e. to deceive versus to inform) lead to these differences in style. In some instances, fake news authors may genuinely believe in the content they are producing, which challenges the assumption that their language would naturally differ from that of real news. Grieve and Woodfield also emphasise the importance of examining the most pernicious type of fake news: authors knowingly writing articles to deceive their readers. This was exemplified in their comparison of the language used in real and fake news written by Jayson Blair, who knowingly fabricated and plagiarised content in some of his articles for personal reasons.

As stated earlier, the false news articles analysed in this study were selected based on fact-checking. While it is evident that the information presented in these articles is untrue, it is unclear whether the authors genuinely believe the falsehoods they are disseminating. For example, the individuals who deny climate change in their articles may sincerely hold the belief that it is not occurring or may be feigning denial for various reasons. Moreover, the truth can change over time, in that some information initially verified as false may eventually be proven to

be true, or vice versa. Consequently, it cannot be assumed that the primary communicative intent of the analysed false news articles is to deceive. Despite a possible absence of deceptive intents, several differences between false and true news were revealed in the analysis. This implies, in accordance with Grieve and Woodfield's (2023) claim, that there must be some other differences in the communicative intent of the two groups of authors that result in these linguistic variations. As discussed in Section 7.3.2, the intention to challenge mainstream narratives is the distinguishing factor between false and true news. In essence, the differences between false and true news arise not from variations in the underlying truth of these two types of news but from variance in how authors perceive and present the information they share in relation to mainstream narratives and the audience's credulity.

The communicative purpose of challenging mainstream narratives raises two crucial considerations. First, this purpose shapes the genre of false news, in that refuting mainstream views and promoting alternative ones necessitates a discourse structure that combines expository, narrative, and argumentative elements; a focus on meanings that convey *scepticism, corroboration, causality*, and *historicity*; and a tendency to avoid explicit advocacy. Consequently, training automatic fact-checking algorithms using articles that aim to challenge certain views and promote alternative ones may result in these algorithms being biased towards detecting this specific genre. Thus, the effectiveness of these algorithms can be impacted. For example, the fact-checking articles produced by reputable sources such as Snopes or PolitiFact, which explain why certain information is false, may belong to the same genre as false news, as these articles narrate events, expose facts and causes, and provide evidence to refute false claims while offering alternative information, just like false news. Thus, fake news detection algorithms might rate these fact-checking articles as false.

Second, the findings of this study may be influenced by the specific communicative intent of challenging mainstream narratives. As the analysed articles mostly focus on controversial and polarising topics, further investigations need to be conducted to determine whether false news articles that are driven by different intents and on different topics exhibit similar discourse structures, news values, and stylistic patterns. It would also be valuable to explore whether false news articles authored by professional journalists for non-ideological purposes, such as meeting

tight deadlines, share the same discourse characteristics as false news produced with the intent of refuting conventional wisdom.

For example, a comparison of the results from the MDA herein and Grieve and Woodfield's (2023) study on Blair's fake and real news articles provides interesting insights. The MDA has found that some false news articles tended to be informationally dense, whereas Blair's fake news was found to be less informationally dense compared to his real news. Therefore, additional research is needed to explore the variations and complexities of false news discourse across various contexts, purposes, and a wider spectrum of false news topics.

Nevertheless, there are notable similarities between the distinct patterns identified in this study's false news corpus and Blair's fake news: The decreased use of *suasive verbs*, *necessity* and *predictive modals*, *conditional adverbials*, and *infinitives* reflect the absence of explicit advocacy in the false news corpus, while the lack of conviction in Blair's deceptive news is demonstrated by the reduced usage of *suasive verbs*, *predictive modals*, and *infinitives*. Both of these patterns indicate the writers being relatively restrained in terms of their stance, persuasion, or argumentation.

Both Blair and the authors of the analysed false news articles demonstrate a shared concern about their content being perceived as fake or untrue by their audience, which is manifested in their similar language usage patterns. Blair's apprehension stems from his concern that his articles might be identified as fake, revealing his lack of conviction, while the apprehension of the false news authors arises from their act of challenging mainstream narratives. This overlap suggests that this shared stylistic pattern is motivated by these authors' desire to prevent their content from appearing deceptive or false. Consequently, it can be inferred that avoiding explicitly persuasive or opinionated language may serve as a general marker of fake or false news, regardless of its intended purpose. However, to generalise this finding, a comprehensive analysis of stance and persuasion across various types of false news is needed.

7.5. Exploring the Mixed-Method Approach

This section examines the effectiveness of the mixed-method approach implemented in this study, with a focus on the benefits and challenges associated with this methodology as well as the implications that arise from its use.

7.5.1. The Benefits of This Study's Mixed-Method Approach

The mixed-method approach has demonstrated four main benefits in this thesis: (a) gaining a comprehensive understanding of false news discourse, (b) validating the results, (c) evaluating the corpus, and (d) reflecting on the three corpus-assisted analytical methods. These benefits are discussed in detail below.

7.5.1.1. Gaining a Comprehensive Understanding of False News Discourse

The three case studies shed light on various aspects of false news discourse, including its discourse structure, communicative purposes, news values, and stylistic choices. Additionally, this approach has provided insights into the appeal and contagion of false news, thereby offering tentative hypotheses on the characteristics of those who engage with it. The studies have also highlighted the complexity of false news discourse and indicated that further investigation is imperative since the multifaceted nature of false news makes it challenging to draw generalised conclusions from a single method of analysis. As Baker and Egbert (2016) argue, human behaviour, particularly language, is complex, and no single methodological approach can offer a complete and comprehensive understanding. Relying on a single method risks overlooking or misunderstanding crucial information, especially when dealing with a complex and multifaceted phenomenon, such as false news.

7.5.1.2. Validating the Results

The MA has indicated that an argumentative element is included in false news discourse that is geared towards challenging mainstream theories. However, the subsequent MDA, for which a larger dataset was used, has revealed that false news tends not to overtly argue or advocate for a position. As discussed in Section 7.3.3, this apparent discrepancy provides a broader understanding of false news. By examining different units of analysis through these two methods, it has been be suggested that while false news articles have an argumentative structure, they do not use explicitly argumentative language. Therefore, generalisations should not be made from a single method as the chosen approach can significantly influence the results obtained. Indeed, Baker and Egbert (2016) contend that researchers should be cautious of potentially achieving findings that are an "artifact" of the employed method, even when analysing a corpus that is claimed to be representative of the target domain of interest (p. 201).

7.5.1.3. Evaluating the Corpus

To lessen the potential influence of register variation on the comparison between online false news and true news, this study's reference corpus was carefully selected. It comprises four sub-corpora representing various news sources. This diverse reference corpus has proved advantageous in two ways. First, in the MDA, a separate comparison was conducted between the false news corpus and each of the four reference sub-corpora, which has yielded valuable insights into the similarities and differences between false news and these different types of news sources. Second, the KSDA entailed comparing the entire false news corpus to the entirety of the reference corpus, thus enabling the identification of distinct meanings and, consequently, the identification of the news values constructed in false news discourse.

7.5.1.4. Reflecting on the Employed Methods

The mixed-method approach has also allowed for the identification of both the strengths and limitations of the employed methods. While each corpus-assisted analytical method's limitations in terms of answering the research questions are discussed in each case study chapter, this section provides a general reflection on the usefulness of these methods in terms of analysing false news discourse and the effectiveness of applying them in a combined framework to analyse the corpus.

The MA approach has proved to be a valuable tool in understanding the structure and objectives of online false news articles. It has also provided important insights into the context in which false news is generated and the strategies employed by false news writers to achieve their objectives. In addition to it previously being used in English as a second language research, this study indicates that the MA approach can offer insights into emerging and problematic genres, such as false news.

The KSDA is a highly corpus-driven approach that can yield unexpected findings. By analysing keywords or key semantic domains, researchers can develop theories about the discourse under investigation, even when the research is being conducted without a specific hypothesis guiding it. In this study, the method of KSDA was initially used to identify the rhetorical strategies employed in false news articles. However, the qualitative analysis of the concordances of the key semantic domains has indicated that these domains reflect news values. As suggested by

McEnery (2016), the effectiveness of the KDSA method relies on the researcher's contextualisation and interpretation. For instance, Dance (2019) used KSDA to explore the motivations behind the sharing of disinformation online, while KSDA has proved valuable for identifying the news values associated with false news in this study.

Similar to KSDA, MDA is a corpus-driven approach, in that it offers a comprehensive linguistic overview of corpora by considering a wide range of lexico-grammatical features and analysing their communicative functions (Baker & Egbert, 2016). Originally developed by Biber (1988), MDA has been primarily used to examine linguistic variation across different spoken and written registers in English. However, as shown in the third case study, MDA can also be applied to identify distinct communicative styles within a single register, such as news articles. Indeed, current research practices often entail MDA being employed to analyse stylistic variation within a specific language variety, such as investigations of variation across different time periods or different topics (e.g., Clarke & Grieve, 2017, 2019; Conrad & Biber, 2001; Gray, 2013; Grieve et al., 2010).

A comparison of the three methods used herein yields interesting conclusions. The MDA has been effective in identifying the discourse style or outer packaging of false news while the MA and KSDA have provided insights into the discourse functionality and rhetoric. The collective findings from the three case studies highlight the complementary nature of these methods in comprehensively deciphering the intricate layers of false news discourse, even though each method has its limitations and may not capture the complexity of the data individually.

7.5.2. The Challenges of This Study's Mixed-Method Approach

While employing multiple methods in this study provided numerous benefits, it also presented certain challenges. First, implementing multiple methods requires a significant investment of time due to the additional analyses needed to ensure rigorous and reliable findings. Due to the time constraints of this PhD project, additional analyses to further enhance the findings could not be conducted. These analyses include investigating the move structure of false news articles on vaccinations and COVID-19, comparing the move structures between the false news corpus and the reference corpus, analysing the parts of speech and lexis of each move, examining whether false news differs from other news types in terms of specific word choice (e.g. 'think', 'know',

‘expect’, or ‘believe’) despite their similarities in the use of lexico-grammatical features (e.g. ‘private verbs’ in general), and exploring collocations of key semantic domains to deeply understand the news values they construct. Second, applying multiple methods requires the researcher, especially one working solo, to invest considerable effort in developing the necessary expertise in the chosen methods. This process can be particularly demanding if the chosen methods are unfamiliar to the researcher. Third, different approaches have various data requirements. For instance, a qualitative analysis such as MA is best suited for smaller datasets, while MDA necessitates a larger corpus of sufficiently lengthy texts. As a result, it may not always be feasible to triangulate different corpus methods using the same dataset. Overall, these challenges highlight the complexity involved in implementing a mixed-method approach and emphasise that researchers need to carefully consider their approaches and plan accordingly when employing a mixed-method approach.

7.6. Recommendations for Future Research

This study offers several future research avenues. Firstly, the results can contribute to the advancement of false news detection efforts. By delving into the nuanced aspects of false news discourse, this study offers insights that can complement the surface-level linguistic features typically relied upon in previous research on detecting fake news. It is worth noting that while the findings may not be universally applicable to all instances of fake news, they do provide valuable insights for identifying fake news. The diverse corpus used in this study, encompassing various topics, authors, news outlets, and publication dates, enhances the applicability and relevance of the results for both manual and automated fact-checking initiatives.

One approach to improve the detection of fake news is the development of algorithms that can identify the lexico-grammatical features typically associated with overt advocacy or the meanings associated with the news values specific to false news in unverified articles. These algorithms can then compare the detected features against their prevalence in reliable sources, with high or low prevalence of these features in these articles serving as an indicator of their falsity. Asr et al.'s (2023) study on automatic fake news detection found that semantic features outperformed other traits such as readability or surface textual characteristics in terms of distinguishing fake news from real news. Therefore, the semantic categories identified in the second case study herein could be used to automate the detection of fake news. A similar

approach was used in Miani and Lewandowsky's (2023) study in which a lexicon of conspiratorial language was created to automatically identify texts containing conspiracy theories. This lexicon included words belonging to various semantic domains, such as evaluation of veracity (e.g. ‘truth’) and hiding/hidden (e.g. ‘secret’). These domains are similar to the prevalent meanings found in the false news corpus examined in the second case study.

In addition to improving detection techniques, future research should also identify the motives that lead to individuals engaging with false news. While the analysis of false news discourse in this study helps generate hypotheses about engagement with false news, experimental discourse analysis, as proposed by Fuoli and Hart (2018), can be used to test and validate these hypotheses. Therefore, to gain insights into why individuals read, believe, and share fake news, future research could adopt experimental discourse analysis.

By applying the mixed-method approach in this study, several recommendations emerge that can inform future research. Firstly, it is important to continue examining fake news discourse from different perspectives and using diverse methodologies. This study emphasises that employing various methods on the same corpus can yield complementary results, providing a comprehensive understanding of the complexity of false news discourse. Secondly, collaborative discourse analysis is highly valuable, particularly for the investigation of problematic texts or instances of harmful language, such as fake news. Involving multiple researchers who have diverse perspectives and methodological expertise can lead to the discovery of interesting findings and robust implications. This collaborative approach can also address challenges related to time and expertise due to the involvement of multiple researchers potentially accelerating the analysis process (Baker & Egbert, 2016). Lastly, collaboration between discourse analysts and computer scientists is crucial for addressing the issue of fake news. Discourse analysts may not have the computational skills necessary for developing algorithms that detect fake news, but their knowledge of linguistic theories and analytical tools is essential in investigating false news. By collaborating, discourse analysts can contribute theories derived from their analyses, while computer scientists can create fake news detection algorithms based on these theories. This interdisciplinary collaboration will enhance efforts to combat fake news.

7.7. Conclusion

This study's integration of three corpus-assisted methods has provided valuable insights into false news discourse. The study has identified structural and stylistic similarities between false and true news and highlighted the serious threat posed by false information, as false news can easily gain traction and credibility when resembling credible journalism. These parallels also underscore the difficulties faced by automated fake news detection systems that focus on surface-level features. Nonetheless, future research should focus on investigating fake news discourse in more depth to strengthen detection methods.

The study has also identified differences between false and true news and emphasised that the underlying factor for these distinctions is false news's primary objective of questioning mainstream narratives. Aware of the contradictory nature of this content compared to mainstream narratives and its potential lack of audience credulity, false news creators strategically tailor their discourse to appeal to their audience in a way that is distinct from true news discourse. This intentional challenging of mainstream narratives highlights the prevailing issues of scepticism and distrust in the post-truth era in which we currently reside.

The mixed-method approach employed in this study has proved valuable in terms of providing a thorough understanding of false news discourse, validating the results, evaluating the corpus, and reflecting on the employed methods. However, the multifaceted nature of fake news suggests that interdisciplinary research is needed to effectively address this complex phenomenon, improve fake news detection methods, and devise strategies to combat its proliferation.

Chapter 8: Conclusion

8.1. Summary of the Thesis

This thesis investigates the discourse characteristics of online false news. By applying a mixed-method corpus-assisted discourse approach, the study delves into three aspects of false news discourse: discourse structure, discursive news values, and stylistic patterns to explore their impact on the audience. The research has utilized three corpus-assisted analytical methods: move analysis (MA), key semantic domains analysis (KSDA), and multidimensional analysis (MDA) to examine these aspects, respectively. The corpus consists of a focus corpus of 137 online verifiably false news articles, covering three topics: climate change, vaccination, and COVID-19, and a comparative corpus of 548 news articles from reputable sources discussing the same subjects.

The MA findings have revealed that false news articles aim to challenge mainstream narratives. This objective is evident in their discourse structure, which often mirrors that of argumentative essays while incorporating structural elements akin to traditional hard news reports. However, the analysis has indicated that there is no consistent discourse structure that can be deemed typical of online false news.

Through the KSDA, it was observed that false news discourse accentuates specific meanings that could construct news values. The analysis found that false news emphasises news values like *consonance*, *unexpectedness*, and *negativity* more than true news does. In addition, unique news values were also identified that distinguish false news from true news. These include discrediting credible sources (*scepticism*), providing various forms of evidence (*corroboration*), explaining cause and effect (*causality*), referencing religious, spiritual, or mythical elements (*mysticism*), and emphasising the relevance between past and relevant events while claiming historical authenticity (*historicity*).

The MDA has shown that false news shares stylistic similarities with traditional, credible news sources, particularly broadsheets. Both false news and broadsheets employ *informational,*

involved, narrative, and *expository styles*. However, false news tends to avoid using an *explicitly advocating style*, setting it apart from other news sources.

The identified structural and stylistic similarities between false and true news hold a number of significant implications. First, they indicate that false news is becoming increasingly threatening. As it closely resembles reliable and reputable journalism, false news has the potential to gain credibility and momentum easily. These similarities also underscore the challenges faced by research dedicated to designing automated systems for false news detection since these systems rely on superficial linguistic features. Thus, the study proposes that future research should focus on a more in-depth analysis of false news discourse, such as meanings and stylistic patterns to enhance detection methods.

Moreover, the identified differences between false news and true news, I argue, stem from false news's primary aim to challenge mainstream narratives. Recognizing the conflicting nature of their content when compared to mainstream narratives and the potentially low credulity of their audience, false news writers strategically craft their discourse to resonate with their target audience, in a style distinct from that of true news discourse. This deliberate challenge to mainstream narratives also highlights the prevalent issues of doubt and distrust in the current post-truth era we live in.

The utilization of a mixed-method approach in this study has proved beneficial in gaining a comprehensive understanding of false news discourse, validating findings, assessing the corpus, and reflecting on the utilized corpus-assisted analytical methods. However, given the multifaceted nature of false news, there is a clear necessity for ongoing interdisciplinary research to effectively tackle this intricate phenomenon and enhance methods for detecting and combatting false news.

8.2. Combatting False News: Practical Applications and Policy Pathways

The study demonstrates that false news, particularly in health and environmental domains, both emerges from and contributes to a lack of trust in governments and media organizations. This, in turn, impacts individual decision-making and societal well-being. Consequently, these findings underscore the urgent need for governments and mainstream media to address these challenges.

Rebuilding trust and credibility is paramount in order to maintain a well-informed and trusting society. Therefore, it is crucial for governments and mainstream media to recognize the significance of their discourse and take proactive measures to enhance its quality, transparency, and reliability. This includes working towards improving communication practices and ensuring that accurate and trustworthy information is readily accessible to the public.

Moreover, the study highlights the value of discourse analysis in uncovering subtle discrepancies between false news and true news. Grieve and Woodfield (2023) advocate for the use of critical discourse analysis as a valuable tool to navigate the complexities of fake news discourse. They argue that effectively addressing the problem of fake news requires more than large-scale detection methods. Instead, they stress the need for careful and sceptical engagement with news content. Therefore, style-based automatic detection algorithms may not be the most effective strategy for combating the spread of false news. Instead, education can play a vital role in equipping individuals to identify false news and avoid exploitation and deception by understanding the persuasive tactics employed by false news, such as news values, discursively constructed objectivity, and seemingly logical arguments.

Not only do individuals need to cultivate critical thinking skills but also media literacy skills. The study has shown that false news provides various types of evidence and logical reasoning as well as in-depth scientific and technical explanations. To safeguard themselves from false information, individuals need to learn how to evaluate the credibility and reliability of news sources, recognize bias and propaganda in news articles, interpret different forms of media messages, distinguish between fact and opinion in news reporting, and compare and contrast information from multiple sources.

8.3. Contributions of the Thesis

This thesis makes several important contributions. Firstly, it enhances our understanding of false news discourse and adds to the growing body of linguistic research on false news (discussed in Chapter 2) by conducting three case studies analysing its discourse structure, news values, and stylistic choices. These findings have the potential to improve current manual and automatic false news detection methods.

Furthermore, the thesis suggests some insights into the situational characteristics of false news. By examining different aspects of false news discourse, conclusions about the appeal of false news, its purposes, audiences, and ideological underpinnings could be drawn. Understanding these situational characteristics is crucial in tackling the spread and acceptance of false news.

Moreover, the thesis makes several valuable methodological contributions. Firstly, it addresses the challenges and opportunities in constructing representative false news corpora. It also provides recommendations for building a focus corpus that encompasses verifiably false news articles and for considering register variation while building the reference corpus (refer to Chapter 3). This is particularly important for researchers developing false news detection algorithms, as limitations can arise from using unrepresentative data. A second methodological contribution is that the thesis explores the effectiveness of employing a mixed-method approach in discourse analysis, especially when studying problematic discourses like false news. It discusses the benefits and challenges of using multiple methods and highlights the value of collaborative work in discourse analysis. Another methodological contribution is that the thesis introduces a novel approach to analysing news values inductively (outlined in Chapter 5). This approach showcases the potential effectiveness of using KSDA to analyse news values and discover new ones. By employing this approach, the study underscores the dynamic nature of news values, which can evolve over time and vary based on the target audience. The results of this method also emphasise the significance of analysing meanings in false news discourse, demonstrating how meanings can reflect news values, construct persuasive arguments, underscore ideological beliefs, and offer insights into the intended audience of false news.

In terms of practical implications, the thesis raises awareness of a significant societal issue: the widespread distrust towards the discourse of governments and mainstream media. It emphasises the importance of addressing the public's epistemological needs and working towards building trust in these discourses. Furthermore, it underscores the importance of educating individuals on how to scrutinize news and avoid falling prey to the persuasive tactics employed in false news discourse.

8.4. Concluding Remarks

Given that false news is communicated through language, studying its discourse is essential in addressing this problem. Thus, the main objective of this thesis was to gain a better understanding of online false news discourse by examining its structure, news values, and stylistic choices. Through this investigation, it has become evident that these aspects contribute to the persuasive and viral power of false news, making it linguistically captivating to consumers. This thesis not only enhances our understanding of false news discourse but also provides valuable insights into its situational characteristics, offers methodological recommendations for its analysis, and suggests practical strategies to combat its dissemination. However, it is essential to acknowledge the complexity of false news discourse, which involves multiple authors, purposes, topics, and contexts. Therefore, further research is necessary to fully grasp the multifaceted nature of this phenomenon.

References

About Newstex | We distribute authoritative content. (2022). Newstex. https://www.newstex.com/about-us

Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. *Complexity*, *2020*, 1–11. https://doi.org/10.1155/2020/8885861

Aikhenvald, A. Y. (2012). The essence of mirativity. *Linguistic Typology*, *16*(3), 435–485. https://doi.org/10.1515/lity-2012-0017

Alba-Juez, L., & Mackenzie, J. L. (2019). Emotion, lies, and "bullshit" in journalistic discourse: The case of fake news. *Iberica*, *2019*, 17–50.

Allan, S. (2011). Reweaving the Internet: Online News of September 11. In B. Zelizer & S. Allan (Eds.), *Journalism after September 11* (2nd ed., pp. 119–140). Routledge.

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Anthony, L. (2016). *SarAnt (Version 1.1.0)* [Computer software]. Waseda University. https://www.laurenceanthony.net/software

Anthony, L. (2022). *AntConc (Version 4.2.0)* [Computer software]. Waseda University. https://www.laurenceanthony.net/software/antconc/

Archer, D., McEnery, T., Rayson, P., & Hardie, A. (2003). Developing an automated semantic analysis system for Early Modern English. *Proceedings of the Corpus Linguistics 2003 Conference.*, 22–31. https://www.semanticscholar.org/paper/Developing-an-automated-semantic-analysis-system-Archer-McEnery/969e93bfe764c97a6a1030381bb375ff18ec9d1f

Archer, D., Wilson, A., & Rayson, P. (2002). *Introduction to the USAS category system*. http://ucrel.lancs.ac.uk/usas/

Askehave, I., & Swales, J. (2001). Genre Identification and Communicative Purpose: A Problem and a Possible Solution. *Applied Linguistics*, *22*(2), 195–212.

Asr, F., Mokhtari, M., & Taboada, M. (2023). Misinformation Detection in News Text. In S. M. Maci, M. Demata, M. McGlashan, & P. Seargeant, *The Routledge Handbook of Discourse and Disinformation* (1st ed., pp. 79–102). Routledge. https://doi.org/10.4324/9781003224495-7

Asr, F., & Taboada, M. (2019). Big Data and quality data for fake news and misinformation detection. *Big Data & Society*, *6*, 1–14. https://doi.org/10.1177/2053951719843310

Baker, P. (2004). Querying Keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis. *Journal of English Linguistics*, *32*(4), 346–359. https://doi.org/10.1177/0075424204269894

Baker, P. (2006). *Using Corpora in Discourse Analysis*. A&C Black.

Baker, P. (2020). Corpus-assisted discourse analysis. In C. Hart (Ed.), *Researching Discourse* (1st ed., pp. 124–142). Routledge. https://doi.org/10.4324/9780367815042-7

Baker, P., & Egbert, J. (2016). Research Synthesis. In P. Baker & J. Egbert (Eds.), *Triangulating Methodological Approaches in Corpus Linguistic Research* (1st ed., pp. 1–19). Routledge. https://doi.org/10.4324/9781315724812

Baker, P., & Ellece, S. (2011). *Key Terms in Discourse Analysis*. Continuum International Publishing Group.

Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge University Press.

Bakir, V., & McStay, A. (2018). Fake News and The Economy of Emotions: Problems, causes, solutions. *Digital Journalism*, *6*(2), 154–175. https://doi.org/10.1080/21670811.2017.1345645

Baptista, J. P., & Gradim, A. (2020). Understanding Fake News Consumption: A Review. *Social Sciences*, *9*(185), 1–22. https://doi.org/10.3390/socsci9100185

Barclay, D. A. (2018). *Fake news, propaganda, and plain old lies: How to find trustworthy information in the digital age*. Rowman & Littlefield.

Bawarshi, A. S., & Reiff, M. J. (2010). *Genre: An Introduction to History, Theory, Research, and Pedagogy*. Parlor Press.

Bayram, A. B., & Shields, T. (2021). Who Trusts the WHO? Heuristics and Americans' Trust in the World Health Organization During the COVID-19 Pandemic. *Social Science Quarterly*, *102*(5), 2312–2330. https://doi.org/10.1111/ssqu.12977

Bednarek, M. (2006). Epistemological positioning and evidentiality in English news discourse: A text-driven approach. *Text & Talk - An Interdisciplinary Journal of Language, Discourse Communication Studies*, *26*(6), 635–660. https://doi.org/10.1515/TEXT.2006.027

Bednarek, M. (2007). Local grammar and register variation: Explorations in broadsheet and tabloid newspaper discourse. *English Language Research*, *1*(1), 1–19.

Bednarek, M. (2016). Voices and values in the news: News media talk, news values and attribution. *Discourse, Context & Media*, *11*, 27–37. https://doi.org/10.1016/j.dcm.2015.11.004

Bednarek, M., & Caple, H. (2012). *News Discourse*. Bloomsbury Publishing Plc. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=967751

Bednarek, M., & Caple, H. (2014). Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society*, *25*(2), 135–158. https://doi.org/10.1177/0957926513516041

Bednarek, M., & Caple, H. (2017). *The Discourse of News Values: How News Organizations Create Newsworthiness*. Oxford University Press. https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780190653934.001.0001/acprof-9780190653934

Bell, A. (1991). *The language of news media*. Basil Blackwell.

Bell, A. (2002). Back in style: Reworking audience design. In J. R. Rickford & P. Eckert (Eds.), *Style and Sociolinguistic Variation* (pp. 139–169). Cambridge University Press. https://doi.org/10.1017/CBO9780511613258.010

Bell, A. (2005). News Stories as Narratives. In M. Inderjeet, P. James, & G. Robert (Eds.), *The Language of Time: A Reader* (pp. 397–409). Oxford University Press.

Bhatia, V. K. (2014). *Analysing Genre: Language Use in Professional Settings*. Routledge.

Bhatia, V. K. (2017). Methodological Issues in Genre Analysis. *Journal of Language and Communication in Business*, *9*(16), 39–59. https://doi.org/10.7146/hjlcb.v9i16.25383

Biber, D. (1988). *Variation across Speech and Writing* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Biber, D. (1989). A typology of English texts. *Linguistics*, *27*, 3–43.

Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, *8*(4), 243–257. https://doi.org/10.1093/llc/8.4.243

Biber, D. (2013). Introductory comments by the author. *Linguistics*, *51*, 15–17. https://doi.org/DOI 10.1515/ling-2013-004

Biber, D., Connor, U., & Upton, T. A. (2007). *Discourse on the Move: Using corpus analysis to describe discourse structure*. John Benjamins Publishing Company. https://doi.org/10.1075/scl.28

Biber, D., & Conrad, S. (2019). *Register, Genre, and Style* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781108686136

Biber, D., & Egbert, J. (2018). *Register Variation Online*. Cambridge University Press. https://doi.org/10.1017/9781316388228

Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of Spoken and Written English*. John Benjamins Publishing Company. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=6747888

Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit* (1st edition). O'Reilly Media. https://api.

Bondielli, A., & Marcelloni, F. (2019). A survey on fake news and rumour detection techniques. *Information Sciences*, *497*, 38–55. https://doi.org/10.1016/j.ins.2019.05.035

Brennen, J. S., Simon, F. M., Howard, P. N., & Nielsen, R. K. (2020). *Types, Sources, and Claims of COVID-19 Misinformation* (FACTSHEET, pp. 1–13). University of Oxford. https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation

Breuer, I., & Napthine, M. (2008). *Persuasive Language in Media Texts*. Insight Publications.

Brezina, V. (2018). *Statistics in Corpus Linguistics: A Practical Guide* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781316410899

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in Fake News is Associated with Delusionality, Dogmatism, Religious Fundamentalism, and

Reduced Analytic Thinking. *Journal of Applied Research in Memory and Cognition*, *8*(1), 108–117. https://doi.org/10.1016/j.jarmac.2018.09.005

Brown, G., & Yule, G. (1983). *Discourse analysis*. University Press. https://doi.org/10.1017/CBO9780511805226

Burki, T. (2019). Vaccine misinformation and social media. *The Lancet Digital Health*, *1*(6), 258–259. https://doi.org/10.1016/S2589-7500(19)30136-0

BuzzFeed News: Top 50 Fake News for 2016. (2016). [dataset]. BuzzFeed News. https://docs.google.com/spreadsheets/d/1sTkRkHLvZp9XlJOynYMXGslKY9fuB_e-2mrxqgLwvZY/edit?usp=sharing&usp=embed_facebook

BuzzFeedNews/2017-12-fake-news-top-50. (2017). [dataset]. BuzzFeed News. https://github.com/BuzzFeedNews/2017-12-fake-news-top-50 (Original work published 2017)

BuzzFeedNews/2018-12-fake-news-top-50. (2018). [dataset]. BuzzFeed News. https://github.com/BuzzFeedNews/2018-12-fake-news-top-50 (Original work published 2018)

Caple, H., & Bednarek, M. (2013). *Delving into the discourse: Approaches to news values in journalism studies and beyond*. https://www.semanticscholar.org/paper/Delving-into-the-discourse%3A-approaches-to-news-in-Caple-Bednarek/4700e1b8e8525d6c9013599c06bb32df52732853

Castelo, S., Almeida, T., Elghafari, A., Santos, A., Pham, K., Nakamura, E., & Freire, J. (2019). A Topic-Agnostic Approach for Identifying Fake News Pages. *Companion Proceedings of The 2019 World Wide Web Conference*, 975–980. https://doi.org/10.1145/3308560.3316739

Catenaccio, P. (2008). Press releases as a hybrid genre: Addressing the informative/promotional conundrum. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *18*(1), 9–31. https://doi.org/10.1075/prag.18.1.02cat

Chafe, W. (1986). Evidentiality in English Conversation and Academic Writing. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The linguistic coding of epistemology* (pp. 261–272). Ablex Publishing Corporation.

Chavanachid, P., & Sripicharn, P. (2019). A Corpus-informed Study of Move Structures and Linguistic Features of Press Release and Corresponding News Reports of Products and Services in the Thai Business Context. *Journal of Studies in the English Language*, *14*(1), 25–92.

Chen, C., & Liu, R. (2023). Discrepancies in the portrayal of the COVID-19 vaccine in Chinese and US international media outlets: A corpus-based discursive news values analysis. *Global Public Health*, *18*(1), 1–15. https://doi.org/10.1080/17441692.2023.2201315

Clarke, I., & Grieve, J. (2017). Dimensions of Abusive Language on Twitter. *Proceedings of the First Workshop on Abusive Language Online*, 1–10. https://doi.org/10.18653/v1/W17-3001

Clarke, I., & Grieve, J. (2019). Stylistic variation on the Donald Trump Twitter account: A linguistic analysis of tweets posted between 2009 and 2018. *PLOS ONE*, *14*(9), 1–27. https://doi.org/10.1371/journal.pone.0222062

Coan, T. G., Boussalis, C., Cook, J., & Nanko, M. O. (2021). Computer-assisted classification of contrarian claims about climate change. *Scientific Reports*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-01714-4

Coltman-Patel, T. (2023). *(Mis)Representing Weight and Obesity in the British Press: Fear, Divisiveness, Shame and Stigma*. Springer Nature.

Conley, D., & Lamble, S. (2006). *The Daily Miracle: An Introduction to Journalism*. Oxford University Press.

Conrad, S., & Biber, D. (2001). Multi-dimensional methodology and the dimensions of register variation in English. In S. Conrad & D. Biber (Eds.), *Variation in English: Multi-dimensional studies* (pp. 13–42). Longman.

Conroy, N. K., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1–4. https://doi.org/10.1002/pra2.2015.145052010082

Cook, J. (2017). Understanding and countering climate science denial. *Journal & Proceedings of the Royal Society of New South Wales*, *105*(2), 207–219.

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE*, *12*(5), 1–21. https://doi.org/10.1371/journal.pone.0175799

Cotter, C. (2010). *News Talk: Investigating the Language of Journalism* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811975

Culpeper, J., & Demmen, J. (2015). Keywords. In D. Biber & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 90–105). Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.006

Daems, J., Speelman, D., & Ruette, T. (2013). Register analysis in blogs: Correlation between professional sector and functional dimensions. *Leuven Working Papers in Linguistics*, 1–27.

Dance, W. (2019). Disinformation online: Social media user's motivations for sharing "fake news." *Science in Parliamen*, *75*(3), 20–22.

Davies, M. (2014). *Oppositions and Ideology in News Discourse*. Bloomsbury Publishing Plc. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=1748587

De Cock, S., Granger, S., Leech, G., & McEnery, T. (1998). An automated approach to the phrasicon of EFL learners. In S. Granger (Ed.), *Learner English on computer* (pp. 67–79). Longman.

Demestichas, K., Remoundou, K., & Adamopoulou, E. (2020). Food for Thought: Fighting Fake News and Online Disinformation. *IT Professional*, *22*(2), 28–34. https://doi.org/10.1109/MITP.2020.2978043

Dentith, M. R. X. (2016). The Problem of Fake News. *Public Reason*, *8*(1–2), 65–79.

Dentith, M. R. X. (2018). What is fake news? *University of Bucharest Review*, *8*(1), 24–34.

Dey, A., Rafi, R. Z., Hasan Parash, S., Arko, S. K., & Chakrabarty, A. (2018). Fake News Pattern Recognition using Linguistic Analysis. *2018 Joint 7th International Conference on Informatics, Electronics Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision Pattern Recognition (icIVPR)*, 305–309. https://doi.org/10.1109/ICIEV.2018.8641018

DiStefano, C., Zhu, M., & Mîndrilã, D. (2009). Understanding and Using Factor Scores: Considerations for the Applied Researcher. *Practical Assessment, Research, and Evaluation*, *14*, 1–11. https://doi.org/10.7275/DA8T-4G52

Douglas, C. (2018). Religion and Fake News: Faith-Based Alternative Information Ecosystems in the US and Europe. *The Review of Faith & International Affairs*, *16*(1), 61–73. https://doi.org/10.1080/15570274.2018.1433522

Duffley, P. J. (1992). *The English infinitive*. Longman.

Dunlap, R., & McCright, A. (2015). Challenging Climate Change: The Denial Countermovement. In R. E. Dunlap & R. J. Brulle (Eds.), *Climate Change and Society: Sociological Perspectives* (pp. 301–332). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199356102.003.0010

Dunlap, R., & McCright, A. M. (2011). Organized Climate Change Denial. In J. Dryzek, R. Norgaard, & D. Schlosberg (Eds.), *The Oxford Handbook of Climate Change and Society* (pp. 144–161). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199566600.003.0010

Egbert, J. (2019). Corpus Design and Representativeness. In T. B. Sardinha & M. V. Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 27–42). Bloomsbury Academic.

Egbert, J., & Baker, P. (2019). Introduction. In J. Egbert & P. Baker (Eds.), *Using Corpus Methods to Triangulate Linguistic Analysis* (pp. 1–21). Routledge. https://doi.org/10.4324/9781315112466

Egbert, J., & Biber, D. (2019). Incorporating text dispersion into keyword analyses. *Corpora*, *14*(1), 77–104. https://doi.org/10.3366/cor.2019.0162

Egbert, J., & Staples, S. (2019). Doing Multidimensional Analysis in SPSS, SAS, and R. In T. B. Sardinha & M. V. Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 125–144). Bloomsbury Academic.

Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, *43*(2), 97–116. https://doi.org/10.1080/23808985.2019.1602782

Fahnestock, J. (2011). *Rhetorical style: The uses of language in persuasion*. Oxford University Press.

Fakhruddin, W. F. W. W., & Hassan, H. (2015). A Review of Genre Approaches within Linguistic Traditions. *LSP International Journal*, *2*(2), 53–68. https://doi.org/10.11113/lspi.v2n2.28

Fernández-López, M., & Perea, M. (2020). Language does not modulate fake news credibility, but emotion does. *Psicológica Journal*, *41*(2), 84–102. https://doi.org/10.2478/psicolj-2020-0005

Ferreira, W., & Vlachos, A. (2016). Emergent: A novel data-set for stance classification. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1163–1168. https://doi.org/10.18653/v1/N16-1138

Fowler, R. (1991). *Language in the News: Discourse and Ideology in the Press*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=1460820

Fruttaldo, A., & Venuti, M. (2017). A cross-cultural discursive approach to news values in the press in the US, the UK and Italy: The Case of the Supreme Court ruling on same-sex marriage. *ESP Across Cultures*, *14*, 81–97.

Fuoli, M., & Hart, C. (2018). Trust-building strategies in corporate discourse: An experimental study. *Discourse & Society*, *29*(5), 514–552. https://doi.org/10.1177/0957926518770264

Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse analysis: A critical review* (pp. 225–258). Routledge.

Galasiński, D. (2000). *The language of deception: A discourse analytical study*. Sage Publications. https://ebookcentral.proquest.com/lib/bham/detail.action?docID=996475

Galtung, J., & Ruge, M. H. (1965). The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research*, *2*(1), 64–90. https://doi.org/10.1177/002234336500200104

Gardner, L., Dong, E., Khan, K., & Sarkar, S. (2020). Persistence of US measles risk due to vaccine hesitancy and outbreaks abroad. *The Lancet. Infectious Diseases*, *20*(10), 1114–1115. https://doi.org/10.1016/S1473-3099(20)30522-3

Garside, R. (1987). The CLAWS word-tagging system. In R. Garside, N. L. Geoffrey, & S. Geoffrey (Eds.), *The Computational Analysis of English: A Corpus-based Approach.* (pp. 30–41). Longman.

Gelfert, A. (2018). Fake News: A Definition. *Informal Logic*, *38*(1), 84–117.

Gillings, M. (2023). *Corpus-assisted discourse studies*. University Press. https://doi.org/10.1017/9781009168144

Goffman, E. (1975). *Frame analysis: An essay on the organization of experience*. Penguin.

Goffman, E. (1981). *Forms of talk*. Blackwell.

Golbeck, J., Mauriello, M., Auxier, B., Bhanushali, K. H., Bonk, C., Bouzaghrane, M. A., Buntain, C., Chanduka, R., Cheakalos, P., Everett, J. B., Falak, W., Gieringer, C., Graney, J., Hoffman, K. M., Huth, L., Ma, Z., Jha, M., Khan, M., Kori, V., … Visnansky, G. (2018). Fake News vs Satire: A Dataset and Analysis. *Proceedings of the 10th ACM Conference on Web Science*, 17–21. https://doi.org/10.1145/3201064.3201100

Gravanis, G., Vakali, A., Diamantaras, K., & Karadais, P. (2019a). Behind the cues: A benchmarking study for fake news detection. *Expert Systems with Applications*, *128*, 201–213. https://doi.org/10.1016/j.eswa.2019.03.036

Gravanis, G., Vakali, A., Diamantaras, K., & Karadais, P. (2019b). Behind the cues: A benchmarking study for fake news detection. *Expert Systems with Applications*, *128*, 201–213. https://doi.org/10.1016/j.eswa.2019.03.036

Graves, L. (2017). Anatomy of a Fact Check: Objective Practice and the Contested Epistemology of Fact Checking: Anatomy of a Fact Check. *Communication, Culture & Critique*, *10*(3), 518–537. https://doi.org/10.1111/cccr.12163

Gray, B. (2013). More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles. *Corpora*, *8*(2), 153–181. https://doi.org/10.3366/cor.2013.0039

Grieve, J., Biber, D., Friginal, E., & Nekrasova, T. (2010). Variation Among Blogs: A Multi-dimensional Analysis. In A. Mehler, S. Sharoff, & M. Santini (Eds.), *Genres on the Web* (Vol. 42, pp. 303–322). Springer Netherlands. https://doi.org/10.1007/978-90-481-9178-9_14

Grieve, J., & Woodfield, H. (2023). *The Language of Fake News*. Cambridge University Press. https://doi.org/10.1017/9781009349161

Groom, N., & Grieve, J. (2019). Chapter 9. The evolution of a legal genre: Rhetorical moves in British patent specifications, 1711 to 1860. In T. Fanego & P. Rodríguez-Puente (Eds.), *Studies in Corpus Linguistics* (Vol. 91, pp. 201–234). John Benjamins Publishing Company. https://doi.org/10.1075/scl.91.09gro

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.

Hancock, J. T., & Bailenson, J. N. (2021). The Social Impact of Deepfakes. *Cyberpsychology, Behavior, and Social Networking*, *24*(3), 149–152. https://doi.org/10.1089/cyber.2021.29208.jth

Hansson, S., Orru, K., Torpan, S., Bäck, A., Kazemekaityte, A., Meyer, S. F., Ludvigsen, J., Savadori, L., Galvagni, A., & Pigrée, A. (2021). COVID-19 information disorder: Six types of harmful information during the pandemic in Europe. *Journal of Risk Research*, *24*(3–4), 1–14. https://doi.org/10.1080/13669877.2020.1871058

Harcup, T., & O'Neill, D. (2001). What Is News? Galtung and Ruge revisited. *Journalism Studies*, *2*(2), 261–280. https://doi.org/10.1080/14616700118449

Hinkel, E. (2002). *Second Language Writers' Text: Linguistic and Rhetorical Features*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=234143

Horne, B., & Adalı, S. (2017). This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News. *Proceedings of the International AAAI Conference on Web and Social Media*, *11*(1). https://doi.org/10.1609/icwsm.v11i1.14976

Hunston, S. (2011). *Corpus approaches to evaluation: Phraseology and evaluative language*. Routledge.

Hussain, A., Ali, S., Ahmed, M., & Hussain, S. (2018). The Anti-vaccination Movement: A Regression in Modern Medicine. *Cureus*, *10*(7), 1–8. https://doi.org/10.7759/cureus.2919

Hyland, K. (1990). A Genre Description of the Argumentative Essay. *RELC Journal*, *21*(1), 66–78. https://doi.org/10.1177/003368829002100105

Hyon, S. (1996). Genre in Three Traditions: Implications for ESL. *TESOL Quarterly*, *30*(4), 693–722. https://doi.org/10.2307/3587930

Ibrahim, N. (2020, May 8). *Is It Dangerous to Wear a COVID-19 Protective Mask for Too Long?* Snopes. https://www.snopes.com//fact-check/masks-dangerous-health/

Igwebuike, E. E., & Chimuanya, L. (2021). Legitimating falsehood in social media: A discourse analysis of political fake news. *Discourse & Communication*, *15*(1), 42–58. https://doi.org/10.1177/1750481320961659

Jabbari, M. J., & Farokhipour, S. (2014). A Contrastive Genre Analysis of Iranian and American English News Reports. *Global Journal of Interdisciplinary Social Sciences*, *3*(1), 23–30.

Jankowski, N. W. (2018). Researching Fake News: A Selective Examination of Empirical Studies. *Javnost - The Public*, *25*(1–2), 248–255. https://doi.org/10.1080/13183222.2018.1418964

Jensen, R. (2008). Advocacy Journalism. In *The International Encyclopedia of Communication*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405186407.wbieca029

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kanestion, A., Sarjit Singh, M. K., & Shamsudin, S. (2017). Developing A Framework for Writing Skill: A Corpus-Based Analysis Of The Written Argumentative Essays. *Sains Humanika*, *9*(4–2), 39–47. https://doi.org/10.11113/sh.v9n4-2.1358

Karimi, H., & Tang, J. (2019). Learning Hierarchical Discourse-level Structure for Fake News Detection. *arXiv:1903.07389*, 1–11.

Kisielinski, K., Wagner, S., Hirsch, O., Klosterhalfen, B., & Prescher, A. (2023). Possible toxicity of chronic carbon dioxide exposure associated with face mask use, particularly in pregnant women, children and adolescents – A scoping review. *Heliyon*, *9*(4), 1–18. https://doi.org/10.1016/j.heliyon.2023.e14117

Labov, W. (1999). The transformation of experience in narrative. In A. Jaworsky & N. Coupland (Eds.), *The discourse reader* (pp. 221–235). Routledge.

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Levy, N. (2017). The Bad News About Fake News. *Social Epistemology Review and Reply Collective*, *6*(8), 20–36.

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

Lynas, M., Adams, S., & Stockert, K. (2023). *Climate Misinformation in the Media*. Alliance for Science. https://allianceforscience.org/wp-content/uploads/2023/09/Climate-Misinformation-Study-FINAL.pdf

Maci, S. M. (2019). Discourse Strategies of Fake News in the Anti-vax Campaign. *Lingue Culture Mediazioni*, *6*(1), 15–43. https://doi.org/10.7358/lcm-2019-001-maci

Mann, W., & Thompson, S. (1988). Rhetorical Structure Theory: Towards a Functional Theory of Text Organization. *Text*, *8*(3), 243–281.

Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, *5*(1), 1–47. https://doi.org/10.1186/s41235-020-00252-3

Martin, J. R. (1984). Language, register and genre. *Children Writing: Reader*, *1*, 21–30.

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.

Masterton, M. (2005). Asian journalists seek values worth preserving. *Asia Pacific Media Educator*, *16*, 41–48.

McEnery, T. (2005). Corpus Linguistics. In R. Mitkov (Ed.), *The Oxford Handbook of Computational Linguistics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199276349.013.0024

McEnery, T. (2016). Keywords. In P. Baker & J. Egbert (Eds.), *Triangulating Methodological Approaches in Corpus Research* (pp. 20–32). Routledge.

McIntire, G. (2017, March 22). How to Build a "Fake News" Classification Model. *Open Data Science*. https://opendatascience.com/how-to-build-a-fake-news-classification-model/, https://opendatascience.com/how-to-build-a-fake-news-classification-model/

Mian, A., & Khan, S. (2020). Coronavirus: The spread of misinformation. *BMC Medicine*, *18*(1), 89. https://doi.org/10.1186/s12916-020-01556-3

Miani, A., & Lewandowsky, S. (2023). *Towards automatic detection of conspiratorial language in texts*. British Association of Applied Linguistics, University of York, york, UK. https://www.baalconference2023.com/programme/

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013, January 16). *Efficient Estimation of Word Representations in Vector Space*. International Conference on Learning Representations. https://www.semanticscholar.org/paper/Efficient-Estimation-of-Word-Representations-in-Mikolov-Chen/330da625c15427c6e42ccfa3b747fb29e5835bf0

Miller, C. (1984). Genre as Social Action. *Quarterly Journal of Speech*, *70*, 151–167. https://doi.org/10.1080/00335638409383686

Mitra, T., & Gilbert, E. (2015). CREDBANK: A Large-scale Social Media Corpus With Associated Credibility Annotations. *Proceedings of the Ninth International AAAI Conference on Web and Social Media*, 258–267.

Montgomery, M. (2007). *The Discourse of Broadcast News: A Linguistic Approach*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=325107

Motta, M. (2018). The enduring effect of scientific interest on trust in climate scientists in the United States. *Nature Climate Change*, *8*(6), 485–488. https://doi.org/10.1038/s41558-018-0126-9

Mourão, R. R., & Robertson, C. T. (2019). Fake News as Discursive Integration: An Analysis of Sites That Publish False, Misleading, Hyperpartisan and Sensational Information. *Journalism Studies*, *20*(14), 2077–2095. https://doi.org/10.1080/1461670X.2019.1566871

Myers, G. (2010). *Discourse of blogs and wikis*. Continuum.

Newman, O. (2008). *Online Business Sourcebook*. Walter de Gruyter.

Nini, A. (2015). *Multidimensional Analysis Tagger (v. 1.3) – Manual*. https://drive.google.com/file/d/11BAw-Dl5FDLTJFXqixc2uIgQQtU0g0no/view?usp=sharing

Nini, A. (2017). Register variation in malicious forensic texts. *International Journal of Speech Language and the Law*, *24*(1), 99–126. https://doi.org/10.1558/ijsll.30173

Nini, A. (2019). The Multi-dimensional Analysis Tagger. In T. B. Sardinha & M. V. Pinto (Eds.), *Multi-dimensional analysis: Research methods and current issues* (pp. 67–94). Bloomsbury Academic.

NVivo 11 for Windows Help—How are cluster analysis diagrams generated? (n.d.). Retrieved September 1, 2021, from https://help-

nv11.qsrinternational.com/desktop/deep_concepts/how_are_cluster_analysis_diagrams_generated_.htm

Oshikawa, R., Qian, J., & Wang, W. Y. (2020). A Survey on Natural Language Processing for Fake News Detection. *arXiv:1811.00770 [Cs]*. http://arxiv.org/abs/1811.00770

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021a). Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter. *American Political Science Review*, *115*(3), 999–1015. https://doi.org/10.1017/S0003055421000290

Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021b). Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter. *American Political Science Review*, *115*(3), 999–1015. https://doi.org/10.1017/S0003055421000290

Palma, B. (2020, January 31). *Did Health Experts "Predict" New Coronavirus Could Kill 65 Million People?* Snopes. https://www.snopes.com/fact-check/coronavirus-65-million-people/

Parikh, S. B., & Atrey, P. K. (2018). Media-Rich Fake News Detection: A Survey. *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 436–441. https://doi.org/10.1109/MIPR.2018.00093

Patel, M. (2019). National Update on Measles Cases and Outbreaks—United States, January 1–October 1, 2019. *MMWR. Morbidity and Mortality Weekly Report*, *68*. https://doi.org/10.15585/mmwr.mm6840e2

Pavlik, J. (2000). The Impact of Technology on Journalism. *Journalism Studies*, *1*(2), 229–237. https://doi.org/10.1080/14616700050028226

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology. General*, *147*(12), 1865–1880. https://doi.org/10.1037/xge0000465

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, *88*(2), 185–200. https://doi.org/10.1111/jopy.12476

Pennycook, G., & Rand, D. G. (2021). The Psychology of Fake News. *Trends in Cognitive Sciences*, *25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017, August). *Automatic Detection of Fake News*. International Conference on Computational Linguistics. https://www.semanticscholar.org/paper/Automatic-Detection-of-Fake-News-P%C3%A9rez-Rosas-Kleinberg/81581d81b508ee2ae15a6b835468283c8278c058

Pian, W., Chi, J., & Ma, F. (2021). The causes, impacts and countermeasures of COVID-19 "Infodemic": A systematic review using narrative synthesis. *Information Processing & Management*, *58*(6), 102713. https://doi.org/10.1016/j.ipm.2021.102713

Põldvere, N., Uddin, Z., & Thomas, A. (2023). The PolitiFact-Oslo Corpus: A New Dataset for Fake News Analysis and Detection. *Information*, *14*(12), Article 12. https://doi.org/10.3390/info14120627

Potts, A., Bednarek, M., & Caple, H. (2015). How can computer-based methods help researchers to investigate news values in large datasets? A corpus linguistic study of the construction of newsworthiness in the reporting on Hurricane Katrina. *Discourse & Communication*, *9*(2), 149–172. https://doi.org/10.1177/1750481314568548

QSR International Pty Ltd. (2020). *Nvivo* (released in March 2020) [Computer software]. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Quirk, R., Grenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman.

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2931–2937. https://doi.org/10.18653/v1/D17-1317

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, *13*(4), 519–549. https://doi.org/10.1075/ijcl.13.4.06ray

Rayson, P. (2009). *Wmatrix: A web-based corpus processing environment.* [Computer software]. Computing Department, Lancaster University. http://ucrel.lancs.ac.uk/wmatrix/

Recasens, M., Danescu-Niculescu-Mizil, C., & Jurafsky, D. (2013). Linguistic Models for Analyzing and Detecting Biased Language. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 1650–1659. https://aclanthology.org/P13-1162

Revelle, W. (2022). *Procedures for Personality and Psychological Research* (2.2.9) [Computer software]. https://CRAN.R-project.org/package=psych

Richardson, J. E. (2006). *Analysing newspapers: An approach from critical discourse analysis*. Palgrave Macmillan.

Ricketson, M. (2014). *Telling True Stories: Navigating the challenges of writing narrative non-fiction*. Allen & Unwin.

Rini, R. (2017). Fake News and Partisan Epistemology. *Kennedy Institute of Ethics Journal*, *27*(2), 43–64. https://doi.org/10.1353/ken.2017.0025

Royal Society for Public Health. (2019). *Moving the needle: Promoting vaccination uptake across the life course.* https://www.rsph.org.uk/uploads/assets/uploaded/3b82db00-a7ef-494c-85451e78ce18a779.pdf

Rubin, V., Conroy, N., Chen, Y., & Cornwell, S. (2016). Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News. *Proceedings of the Second Workshop on Computational Approaches to Deception Detection*, 7–17. https://doi.org/10.18653/v1/W16-0802

Rubin, V., & Lukoianova, T. (2015). Truth and deception at the rhetorical structure level. *Journal of the Association for Information Science and Technology*, *66*(5), 905–917. https://doi.org/10.1002/asi.23216

Sabbah, F. (2023). Critical Discourse Analysis Approaches to Investigating Fake News and Disinformation. In S. M. Maci, M. Demata, M. McGlashan, & P. Seargeant, *The Routledge Handbook of Discourse and Disinformation* (1st ed., pp. 33–51). Routledge. https://doi.org/10.4324/9781003224495-4

Samraj, B. (2014). Move Structure. In K. P. Schneider & A. Barron (Eds.), *Pragmatics of Discourse*. De Gruyter, Inc. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=1209369

Sardinha, T. B. (2018). Dimensions of variation across Internet registers. *International Journal of Corpus Linguistics*, *23*(2), 125–157. https://doi.org/10.1075/ijcl.15026.ber

Scott, M. (1997). PC analysis of key words—And key key words. *System*, *25*(2), 233–245. https://doi.org/10.1016/S0346-251X(97)00011-0

Scott, M. (1998). *WordSmith Tools Help Manual. Version 3.0.* https://lexically.net/wordsmith/version3/manual.pdf

Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2019). FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. *arXiv:1809.01286 [Cs]*.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36. https://doi.org/10.1145/3137597.3137600

Shu, K., Wang, S., & Liu, H. (2018). Understanding User Profiles on Social Media for Fake News Detection. *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 430–435. https://doi.org/10.1109/MIPR.2018.00092

Shu, K., Wang, S., & Liu, H. (2019). Beyond News Contents: The Role of Social Context for Fake News Detection. *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, 312–320. https://doi.org/10.1145/3289600.3290994

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. University Press.

Sousa-Silva, R. (2022). Fighting the Fake: A Forensic Linguistic Analysis to Fake News Detection. *International Journal for the Semiotics of Law - Revue Internationale de Sémiotique Juridique*, *35*(6), 2409–2433. https://doi.org/10.1007/s11196-022-09901-w

Swales, J. (1990). *Genre Analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. (2004). *Research Genres: Explorations and Applications*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524827

Tacchini, E., Ballarin, G., Della Vedova, M., Moret, S., & Alfaro, L. (2017, September 18). *Some Like it Hoax: Automated Fake News Detection in Social Networks*.

Tandoc, E., Lim, Z. W., Ling, R., & Jr. (2018). Defining "Fake News." *Digital Journalism*, *6*(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tannen, D., Hamilton, H. E., & Schiffrin, D. (2015). *The Handbook of Discourse Analysis*. John Wiley & Sons, Incorporated. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=1895501

Tardy, C. M., & Swales, J. (2014). Genre Analysis. In K. P. Schneider & A. Barron (Eds.), *Pragmatics of Discourse* (pp. 165–188). De Gruyter Mouton.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, *61*(12), 2544–2558. https://doi.org/10.1002/asi.21416

Thorne, J., Vlachos, A., Christodoulopoulos, C., & Mittal, A. (2018). FEVER: A Large-scale Dataset for Fact Extraction and VERification. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, 809–819. https://doi.org/10.18653/v1/N18-1074

Titak, A., & Roberson, A. (2013). Dimensions of web registers: An exploratory multi-dimensional comparison. *Corpora*, *8*(2), 235–260. https://doi.org/10.3366/cor.2013.0042

Toolan, M., Well, A. D., & Myers, J. L. (2001). *Narrative: A Critical Linguistic Introduction*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=1099361

Toutanova, K., Klein, D., Manning, C. D., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - NAACL '03*, *1*, 173–180. https://doi.org/10.3115/1073445.1073478

Treen, K. M. d'I., Williams, H. T. P., & O'Neill, S. J. (2020). Online misinformation about climate change. *WIREs Climate Change*, *11*(5), e665. https://doi.org/10.1002/wcc.665

Trimarco, P. (2014). *Digital Textuality*. Macmillan International Higher Education.

Upton, T. A., & Cohen, M. A. (2009). An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies*, *11*(5), 585–605. https://doi.org/10.1177/1461445609341006

Vamanu, I. (2019). Fake News and Propaganda: A Critical Discourse Research Perspective. *Open Information Science*, *3*(1), 197–208. https://doi.org/10.1515/opis-2019-0014

Van Dijk, T. (1983). Discourse Analysis: Its Development and Application to the Structure of News. *Journal of Communication*, *33*(2), 20–43. https://doi.org/10.1111/j.1460-2466.1983.tb02386.x

Van Dijk, T. (1985). Structures of News in the Press. *Discourse and Communication: New Approaches to the Analysis of Mass Media Discourse and Communication*, *10*, 69–93.

Van Dijk, T. (1988). *News as Discourse*. Hillsdale, NJ ; Hove : L. Erlbaum Associates.

Van Dijk, T. (1998). *Ideology: A Multidisciplinary Approach*. SAGE Publications. http://ebookcentral.proquest.com/lib/bham/detail.action?docID=1024022

Van Dijk, T. (2004). Knowledge and News. *Revista Canaria De Estudios Inglesis*, *49*, 71–86.

Van Dijk, T. (2017). Power in the News Media. In D. L. Helfert (Ed.), *Political Communication in Action* (1st ed.). Lynne Rienner Publishers.

Van Leeuwen, T. (2007). Legitimation in discourse and communication. *Discourse & Communication*, *1*(1), 91–112. https://doi.org/10.1177/1750481307071986

Volkova, S., Shaffer, K., Jang, J. Y., & Hodas, N. (2017). Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 647–653. https://doi.org/10.18653/v1/P17-2102

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wall, M. (2005). 'Blogs of war': Weblogs as news. *Journalism*, *6*(2), 153–172. https://doi.org/10.1177/1464884905051006

Walton, D. (2007). *Media Argumentation: Dialectic, Persuasion and Rhetoric*. Cambridge University Press. https://doi.org/10.1017/CBO9780511619311

Wang, W. Y. (2017). "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection. *arXiv:1705.00648 [Cs]*. http://arxiv.org/abs/1705.00648

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe Report*, *9*.

https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html

White, P. R. R. (1997). Death, Disruption and the Moral Order: The narrative impulse in mass-media 'hard news' reporting. In F. Christie & J. R. Martin (Eds.), *Genres and Institutions: Social Processes in the Workplace ad School* (pp. 101–133). Cassell.

World Health Organization. (2017). *Vaccination and trust: How concerns arise and the role of communication in mitigating crises*. http://www.euro.who.int/__data/assets/pdf_file/0004/329647/Vaccines-and-trust.PDF

World Health Organization. (2023). *Coronavirus disease (COVID-19) pandemic*. https://www.who.int/europe/emergencies/situations/covid-19

Zafarani, R., Zhou, X., Shu, K., & Liu, H. (2019). Fake News Research: Theories, Detection Strategies, and Open Problems. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3207–3208. https://doi.org/10.1145/3292500.3332287

Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *Journal of Data and Information Quality*, *11*(3), 1–37. https://doi.org/10.1145/3309699

Zhang, A. X., Robbins, M., Bice, E., Hawke, S., Karger, D., Mina, A. X., Ranganathan, A., Metz, S. E., Appling, S., Sehat, C. M., Gilmore, N., Adams, N. B., Vincent, E., & Lee, J. (2018). A Structured Response to Misinformation: Defining and Annotating Credibility Indicators in News Articles. *Proceedings of the The Web Conference 2018. International World Wide Web Conferences Steering Committee*, 603–612. https://doi.org/10.1145/3184558.3188731

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, *57*(2), 102025. https://doi.org/10.1016/j.ipm.2019.03.004

Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2020). Fake News Early Detection: A Theory-driven Model. *Digital Threats: Research and Practice*, *1*(2), 1–25. https://doi.org/10.1145/3377478

Zhou, X., & Zafarani, R. (2018). Fake News: A Survey of Research, Detection Methods, and Opportunities. *arXiv:1812.00315 [Cs]*. http://arxiv.org/abs/1812.00315

Zubiaga, A., Liakata, M., Procter, R., Wong Sak Hoi, G., & Tolmie, P. (2016). Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. *PLoS ONE*, *11*(3), 1–29. https://doi.org/10.1371/journal.pone.0150989

Appendices

Appendix 4.1. Coding Manual for Analysing the Move Structure of False News Articles on Climate Change

Code	*Moves*	**Definitions**	*Strategies/Steps*	**Definitions**
H	***1. Headline***	This move contains the main story event and sometimes the main thesis adopted by the news writer/outlet.		
G	***2. Grab Attention/ Summarize***	This move functions as an explanation of the main title or as a summary of the article and captures readers' attention. It can appear in various forms as subtitles or tags after the headline; bullet summaries, subheadings, and pull quotes within the text; or legends after images.		
W	***3. Byline***	This move indicates the author's name, sometimes with a short description. It usually appears after the headline.		
D	***4. Dateline***	This move indicates the publication date. It usually appears after the headline along with the byline.		
I	***5. Image***	This move consists of an image usually located before the lead of the news article. The image is either a depiction of real-world events/persons related to the main story (See figure 1) or symbolic (See figure 2), explaining the main topic of the news story, sometimes in a sarcastic way. It is sometimes followed by captions.		
L	***6. Lead***	This move constitutes the beginning of the news article itself; it usually follows the image. It can range from one sentence to two paragraphs. The main function of this move is summarizing the main story event and construing the newsworthiness or establishing the significance of the story. This move contains one or more of the following strategies →	***a. Report the main story event***	This strategy describes an event (e.g., Australian bushfires, 2016 USA presidential election, Paris Agreement on climate change) which is used as a starting point for delegitimizing mainstream theories.
			b. Report mainstream theories	This strategy reports what has already been established by mainstream media or governmental organizations about the topic of the news article.
			c. State a main thesis	This strategy expresses the main thesis adopted by the news writer/outlet towards the topic of the news article by stating that mainstream theories are incorrect and/or proposing alternative theories.

			d. Summarize the Refutation	This strategy refutes mainstream theories in a concise manner.
M	***7. Detail the main story event***	This move gives more details about the main story event that has been reported in the lead. It may contain information about the actions, place, time, and people involved. This move is either located at the beginning of the body text after the lead or occurs multiple times in the middle of the body text.		
E	***8. Explain mainstream theories***	This move elaborates on what has already been established by the mainstream media or governmental organization about the topic of the news article. It can be a textual or visual explanation; sometimes the visual explanation contains caricature. It tends to be located at the beginning of the body text, after the lead. However, in many articles, it occurs in multiple places in the middle of the body text.		
R	***9. Refute mainstream theories***	This move refutes mainstream theories. This move is located either at the beginning of the body text after the lead or occurs multiple times throughout the body text. It has one or more of the following strategies: →	***a. Challenge mainstream theories***	This strategy disputes mainstream theories by citing authoritative sources (textual or audio-visual), using logical reasoning, reporting unattributed facts, or using evaluation.
			b. Promote alternative theories	This move supports alternative theories by citing authoritative sources that positively evaluate them or describe positive reactions to them.
			c. Present opposing views	This move presents views that do not support the theories adopted by the news writer/outlet, and at the same time, they do not accord with mainstream theories. This strategy can enhance objectivity and provide support to refute mainstream theories.
B	***10. Give Background***	This move provides background information or tells previously occurred stories that are necessary to understand the current news story or to support the main thesis. This move is located in the middle of the body text, and it can occur once or multiple times.		
K	***11. Describe Consequences***	This move describes the implications or consequences of the main story event. It further involves the reporting of insights from authoritative sources that underscore		

		the implications or consequences of such events or theories. It can be found at various points within the news article and may appear multiple times.		
C	***12. Comment***	It is usually located at the end of the news article; however, there are many instances of comments spread throughout the body text. It establishes the newsworthiness or the significance of the news story and the writer's position through the following strategies →	***a. Restate the main thesis***	This strategy re-expresses the main thesis adopted by the news writer/outlet towards the topic of the news article by stating that the mainstream theories are wrong and/or stating that alternative theories are true. It sometimes includes a summary of the refutation.
			b. Make a conclusion	This strategy interprets the analysis and makes a judgement about mainstream theories (e.g., inferring the motivation behind the governmental behaviors or declarations).
			c. Predict the future	This strategy makes predictions about what is likely to happen or states what is going to happen in the future as a result of mainstream theories.
			d. Recommend	This strategy gives advice on what the reader should do/believe or recommends reading other related or non-related news reports.
			e. Call to action	This strategy invites people in power to take some desired action.
S	***13. Cite sources***	This move cites the sources of the information in the news article, usually links to other websites. It usually occurs at the end of the news article.		
A	***14. Advertise***	This move involves advertisements. They can be textual or audio-visual, and they can appear in different locations throughout the news article.		
N	***15. Note***	This move contains notes that comment on something outside the news text (for example, notes about the author, website/blog, or the comments on the news report itself). Notes usually appear at the beginning or end of the news article.		
P	***16. Promote engagement***	This move promotes readers' engagement with the news article's content by presenting links that enable readers to post comments or share the article on social media		

		platforms. This move is usually located at the beginning or end of the news article and sometimes in the middle.		

Examples:

1. ***Headline:***

DELINGPOLE: 'Global Warming' Is a Myth, Say 58 Scientific Papers in 2017

2. ***Grab attention/summarize:***

- *The Mail on Sunday can reveal a landmark paper exaggerated global warming*
- *It was rushed through and timed to influence the Paris agreement on climate change*
- *America's National Oceanic and Atmospheric Administration broke its own rules*
- *The report claimed the pause in global warming never existed, but it was based on misleading, 'unverified' data*

3. ***Byline:***

By DAVID ROSE FOR THE MAIL ON SUNDAY

4. ***Dateline:***

6 Jun 2017

5. Image:

Figure 1. a real image

Figure 2. a symbolic image

6. Lead:

a. Report the main story event:

The raging inferno that has engulfed a large swath of Australia is horrifying. More than 1,500 homes have been destroyed, 25 people have died and millions of kangaroos and koala bears and other animals may have been lost.

b. Report mainstream theories:

Radical environmental extremists and Hollywood celebrities have blamed the fires on climate change.

c. State a main thesis:

But it turns out — climate change has nothing to do with the humanitarian crisis unfolding Down Under.

It turns out — many of the fires were set by arsonists.

d. Summarize the Refutation:

"Global warming" is a myth — so say 80 graphs from 58 peer-reviewed scientific papers published in 2017.

7. Explain mainstream theories:

By "global warming" these papers don't, of course, mean the mild warming of around 0.8 degrees Celsius that the planet has experienced since the middle of the 19th century as the world crawled out of the Little Ice Age. Pretty much everyone, alarmists and skeptics alike, is agreed on that.

Rather, they mean "global warming" in the sense that is most commonly used today by grant-troughing scientists, and huxter politicians, and scaremongering green activists, and brainwashed mainstream media (MSM) environmental correspondents. "Global warming" as in the scary, historically unprecedented, primarily man-made phenomenon which we must address urgently before the icecaps melt and the Pacific islands disappear beneath the waves and all the baby polar bears drown.

8. Explain the main story event:

The Northwest Passage from the Atlantic to the Pacific has remained blocked by pack-ice all year. More than 20 yachts that had planned to sail it have been left ice-bound and a cruise ship attempting the route was forced to turn back.

8. Refute mainstream theories:

a. Challenge mainstream theories:

- **Citing authoritative sources:** involves showing that information, opinions, images, or figures are taking from authoritative sources such as governmental organizations, scientists, newspapers, or journal articles …etc.

The Australian reports more than 180 people have been arrested since the start of the bushfire season.

It was the year 1958, to be precise, when NASA first observed that changes in the solar orbit of the earth, along with alterations to the earth's axial tilt, are both responsible for what climate scientists today have dubbed as "warming" (or "cooling," depending on their agenda).

- **Logical reasoning**: involves reasoning that are based on deduction, using a given set of facts to deduce other facts from by logical reasoning, or induction, looking for a trend or pattern and then generalizing it.

If you're still inclined to believe what Al Gore has to say about global warming, please consider the fact that since he embarked on his crusade, his wealth has grown from $2 million in 2001 to $100 million in 2016 – largely due to investments in fake "green tech" companies and the effective embezzlement of numerous grants and loans.

- **Unattributed facts:** involves reporting facts that are not attributed to a certain source.

NOAA not only failed, but it effectively mounted a cover-up when challenged over its data. After the paper was published, the US House of Representatives Science Committee launched an inquiry into its Pausebuster claims. NOAA refused to comply with subpoenas demanding internal emails from the committee chairman, the Texas Republican Lamar Smith, and falsely claimed that no one had raised concerns about the paper internally.

- **Evaluation:** involves evaluating the information given or its source in a way that delegitimizes the mainstream theories.

But the whistleblower, Dr John Bates, a top NOAA scientist with an impeccable reputation, has shown The Mail on Sunday irrefutable evidence that the paper was based on misleading, 'unverified' data.

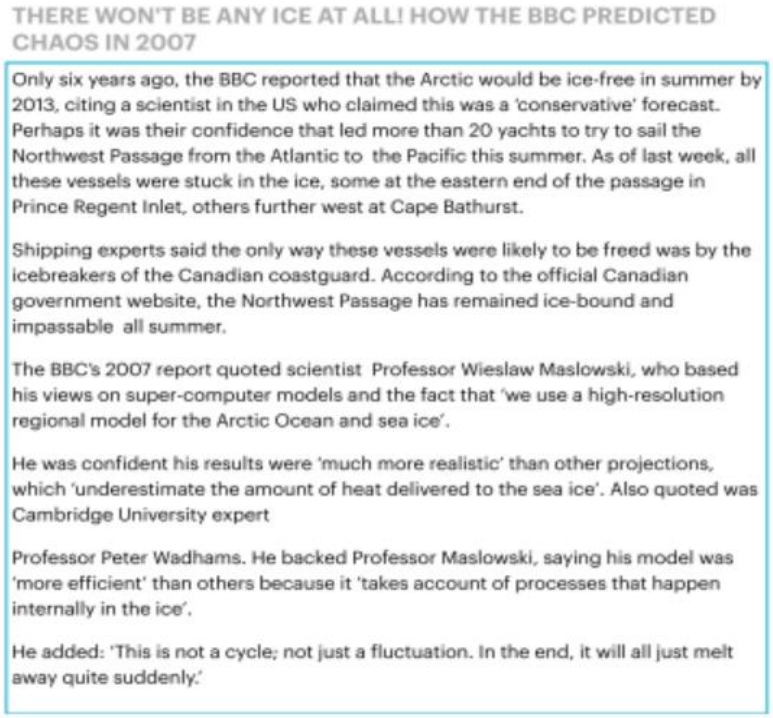
b. Promote alternative theories:

Milankovitch first put forward his model, it went ignored for nearly half a century. Then, in 1976, a study published in the journal Science confirmed that Milankovitch's theory is, in fact, accurate, and that it does correspond to various periods of climate change that have occurred throughout history.

In 1982, six years after this study was published, the National Research Council of the U.S. National Academy of Sciences adopted Milankovitch's theory as truth, declaring that:

"... orbital variations remain the most thoroughly examined mechanism of climatic change on time scales of tens of thousands of years and are by far the clearest case of a direct effect of changing insolation on the lower atmosphere of Earth."

9. Give background:



THERE WON'T BE ANY ICE AT ALL! HOW THE BBC PREDICTED CHAOS IN 2007

Only six years ago, the BBC reported that the Arctic would be ice-free in summer by 2013, citing a scientist in the US who claimed this was a 'conservative' forecast. Perhaps it was their confidence that led more than 20 yachts to try to sail the Northwest Passage from the Atlantic to the Pacific this summer. As of last week, all these vessels were stuck in the ice, some at the eastern end of the passage in Prince Regent Inlet, others further west at Cape Bathurst.

Shipping experts said the only way these vessels were likely to be freed was by the icebreakers of the Canadian coastguard. According to the official Canadian government website, the Northwest Passage has remained ice-bound and impassable all summer.

The BBC's 2007 report quoted scientist Professor Wieslaw Maslowski, who based his views on super-computer models and the fact that 'we use a high-resolution regional model for the Arctic Ocean and sea ice'.

He was confident his results were 'much more realistic' than other projections, which 'underestimate the amount of heat delivered to the sea ice'. Also quoted was Cambridge University expert

Professor Peter Wadhams. He backed Professor Maslowski, saying his model was 'more efficient' than others because it 'takes account of processes that happen internally in the ice'.

He added: 'This is not a cycle; not just a fluctuation. In the end, it will all just melt away quite suddenly.'

10. Present opposing views:

Others are more cautious. Dr Ed Hawkins, of Reading University, drew the graph published by The Mail on Sunday in March showing how far world temperatures have diverged from computer predictions. He admitted the cycles may have caused some of the recorded warming, but insisted that natural variability alone could not explain all of the temperature rise over the past 150 years.

11. Comment:

a. Restate the main thesis

The truth, however, is much more along the lines of what Serbian astrophysicist Milutin Milankovitch, after whom the Milankovitch Climate Theory is named, proposed about how the seasonal and latitudinal variations of solar radiation that hit the earth in different ways, and at different times, have the greatest impact on earth's changing climate patterns.

b. Make a conclusion

But the global warming crowd continues to push their agenda on the public while lining their pockets in the process.

c. Predict the future

Like Climategate, this scandal is likely to reverberate around the world, and reignite some of science's most hotly contested debates.

d. Describe consequences

The impact was huge and lasting. On publication day, the BBC said the pause in global warming was 'an illusion caused by inaccurate data'.

One American magazine described the paper as a 'science bomb' dropped on sceptics.

Its impact could be seen in this newspaper last month when, writing to launch his Ladybird book about climate change, Prince Charles stated baldly: 'There isn't a pause… it is hard to reject the facts on the basis of the evidence.'

e. Recommend:

For more related news about climate change and global warming from an independent, non-establishment perspective, be sure to check out ClimateScienceNews.com.

e. Call to action

Whatever takes its place, said Dr Bates, 'there needs to be a fundamental change to the way NOAA deals with data so that people can check and validate scientific results. I'm hoping that this will be a wake-up call to the climate science community – a signal that we have to put in place processes to make sure this kind of crap doesn't happen again.
'I want to address the systemic problems. I don't care whether modifications to the datasets make temperatures go up or down. But I want the observations to speak for themselves, and for that, there needs to be a new emphasis that ethical standards must be

maintained.'

12. Cite sources:

Sources for this article include:

HalTurnerRadioShow.com

NaturalNews.com

BBC.com

13. Advertise:

Support our mission and protect your health: Organic Seeds of Life combines Red Raspberry Seed Power, Black Cumin Seed Power and Red Grape Seed Powder into the most potent nutrient-rich supplemental superfood powder you've ever experienced. Loaded with flavonoids, antioxidants, anthocyanins, OPCs, ALA and a vast array of vital nutrients. Learn more here.

14. Note:

EDITOR'S NOTE: Facebook is cracking down on Conservative content. Many of you have complained that you never see our content in your news feeds. There's only one way to fight back — and that's by subscribing to my FREE weekly newsletter. Click here.

15. Promote engagement:

Appendix 4.2. Semantic Analysis of the Moves

This appendix provides an overview of the 20 most frequently semantic domains found in the moves of false news articles on climate change. It is important to mention that this analysis focused solely on moves that are typically longer in terms of word count and do not involve images or links. Consequently, certain moves such as *byline*, *dateline, image, cite sources, advertise, note*, and *promote engagement* were excluded from this analysis.

Top Frequent Semantic Tags in *Headline*	Grammatical bin (20.68%), Unmatched (5.06%), Other proper names (4.64%), Weather (3.38%), Science and technology in general (2.95%), Temperature: Hot/fire (2.95%), Pronouns (2.53%), Geographical names (2.53%), Numbers (2.53%), Personal names (2.11%), Change (2.11%), The universe (2.11%), Geographical terms (2.11%), Existing (1.69%), Cause & effect/Connection (1.69%), Helping (1.27%), Speech acts (1.27%), Location and direction (1.27%), Time: Present; simultaneous (1.27%), Moving, coming and going (1.27%), Evaluation: False (1.27%), Degree: Approximators (1.27%)	**Top Frequent Semantic Tags in *Grab attention/Summarize***	Grammatical bin (24.71%), Unmatched (4.1%), Geographical terms (3.46%), Pronouns (3.2%), Change (2.69%), Personal names (2.43%), Weather (2.43%), Knowledge (2.3%), Existing (2.3%), Numbers (2.05%), Temperature: Hot /on fire (2.05%), Geographical names (2.05%), Location and directions (1.92%), Other proper names (1.79%), Speech acts (1.66%), Cause & effect/Connection (1.66%), Open; Finding; Showing (1.41%), Paper documents and writing (1.41%), General actions / making (1.28%), The universe (1.28%)

Top Frequent Semantic Tags in *Lead*	Grammatical bin (29.75%), Pronouns (4.45%), Unmatched(2.44%), Existing (2.44%), Numbers (2.1%), Weather (2.02%), Speech acts (1.85%), Location and direction (1.6%), Evaluation: True (1.51%), Change (1.43%), Geographical names (1.34%), Personal names (1.34%), Geographical terms (1.34%), Temperature: Hot / on fire (1.26%), General actions /making (1.18%), Time: Period (1.18%), Science and technology in general (1.09%), Quantities: many/much (1.01%), Speech: Communicative (1.01%), Other proper names (1.01%)	**Top Frequent Semantic Tags in *Detail the Main Story Event***	Grammatical bin (31.85%), Pronouns (4.92%), Unmatched (3.86%), Geographical names (2.58%), Personal names (2.46%), Numbers (1.99%), Speech acts (1.87%), Location and direction (1.87%), Existing (1.76%), Speech communicative (1.64%), Geographical terms (1.41%), Weather (1.29%), Religion and supernatural (1.17%), Quantities: many/much (1.05%), In power (1.05%), Sailing, swimming, etc. (1.05%), Politics (1.05%), Government (0.94%), Time: Present; simultaneous (0.82%), Other proper names (0.82%), The universe (0.82%), Law and order (0.82%), Wanted (0.82%), Likely (0.82%), Cause & effect/Connection (0.82%), Violent/Angry (0.82%), Moving, coming and going (0.82%)

Top Frequent Semantic Tags in *Explain Mainstream Theories*	Grammatical bin (31.44%), Pronouns (4.23%), Unmatched (3.17%), Geographical terms (3.04%), Temperature: Hot/fire (2.64%), Existing (2.51%), Numbers (1.98%), Time: Period (1.72%), Location and direction (1.59%), Speech acts (1.59%) Cause & effect/Connection (1.45%), The Media: Newspapers etc. (1.32%), Other proper names (1.19%), The Media (1.19%), Weather (1.19%), Speech: Communicative (1.19%), Fear/shock (1.06%), Geographical names (1.06%), Objects generally (0.92%), Comparing: Usual (0.92%), Science and technology in general (0.92%), Moving, coming and going (0.92%), The universe (0.92%).	**Top Frequent Semantic Tags in *Refute Mainstream Theories***	Grammatical bin (31.03%), Pronouns (5.04%), Unmatched (3.87%), Existing (2.84%), Geographical terms (1.79%), Numbers (1.69%), Time: period (1.48%), Personal names (1.41%), Location and directions (1.33%), Speech acts (1.27%), Change (1.26%), Paper documents and writing (1.14%), Geographical names (1.1%), Temperature: Hot /on fire (1.06%), Speech: Communicative (1.05%), Weather (1.01%), Cause & Effect/Connection (1%), Other proper names (0.97%), knowledge (0.96%), Open; Finding; Showing (0.87%)

Top Frequent Semantic Tags in *Give Background*	Grammatical bin (32.88%), Unmatched (4.54%), Pronouns (3.93%), Speech acts (3.19%), Existing (2.33%), Numbers (2.09%), Geographical terms (1.84%), Other proper names (1.84%), Substances and materials: Solid (1.72%), Location and direction (1.47%), Weather (1.47%), Time: Period (1.23%), Objects generally (1.1%), Science and technology in general (1.1%), Change (1.1%), Degree: Boosters (1.1%), Cause & Effect/Connection (0.86%), Quantities: many/much (0.86%), General actions / making (0.86%), Investigate, examine, test, search (0.86%).	**Top Frequent Semantic Tags in *Describe Consequences***	Grammatical bin (32.03%), Pronouns (3.91%), Unmatched (3.42%), Geographical names (2.69%), Speech acts (2.2%), Likely (1.71%), Personal names (1.71%), The universe (1.71%), General actions / making (1.71%), Government (1.71%), Existing (1.71%), Location and direction (1.47%), Weather (1.47%), Geographical terms (1.22%), Giving (1.22%), Quantities: many/much (1.22%), Speech: Communicative (1.22%), Cause & Effect/Connection (1.22%), Change (1.22%), The Media: Newspapers etc. (1.22%)

Top Frequent Semantic Tags in *Comment*	Grammatical bin (28.79%), Pronouns (7.26%), Unmatched (3.15%), Existing (2.91%), Speech acts (1.56%), Personal names (1.44%), geographical terms (1.41%), General actions /making (1.32%), Negative (1.32%), Location and direction (1.29%), Weather (1.14%), Other proper names (1.11%), Likely (1.11%), Science and technology in general (1.11%), Getting and possession (1.08%), In power, (1.02%), Speech: Communicative (0.99%), Cause & Effect/Connection (0.93%), Change (0.9%), Temperature: Hot / on fire (0.84%)

Appendix 5.1. Keywords List

	Type	Freq Target	Freq Reference	Range Target	Range Reference	Norm Freq Target	Norm Freq Reference	Norm Range Target	Norm Range Reference	Keyness (Likelihood)	Keyness (Effect)
1	the	6646	19891	137	536	65646.64	58058.11	1	1	78.419	0.104
2	that	1537	4049	131	519	15181.9	11818.28	0.956	0.968	68.097	0.029
3	is	1333	3548	132	521	13166.86	10355.95	0.964	0.972	54.43	0.025
4	this	640	1530	114	440	6321.674	4465.784	0.832	0.821	52.071	0.012
5	was	620	1678	106	412	6124.122	4897.769	0.774	0.769	21.934	0.012
6	vaccine	522	1067	67	196	5156.116	3114.374	0.489	0.366	83.783	0.01
7	vaccines	338	390	57	129	3338.634	1138.337	0.416	0.241	195.933	0.007
8	data	230	279	40	126	2271.852	814.349	0.292	0.235	123.605	0.005
9	deaths	230	338	37	123	2271.852	986.559	0.27	0.229	88.263	0.005
10	dr	193	309	47	122	1906.38	901.913	0.343	0.228	61.709	0.004
11	these	202	336	71	196	1995.278	980.721	0.518	0.366	59.069	0.004
12	pfizer	136	66	18	26	1343.356	192.642	0.131	0.049	181.027	0.003
13	warming	127	136	16	47	1254.457	396.959	0.117	0.088	81.599	0.003
14	study	169	241	38	107	1669.317	703.434	0.277	0.2	68.75	0.003
15	vaccinated	164	275	34	113	1619.929	802.674	0.248	0.211	47.015	0.003
16	years	142	276	56	173	1402.622	805.592	0.409	0.323	26.94	0.003
17	global	142	283	29	112	1402.622	826.024	0.212	0.209	24.811	0.003
18	peanut	94	1	3	1	928.496	2.919	0.022	0.002	267.35	0.002
19	gates	87	15	13	8	859.353	43.782	0.095	0.015	179.804	0.002
20	adverse	83	21	25	14	819.842	61.295	0.182	0.026	151.629	0.002
21	wuhan	79	48	17	16	780.332	140.103	0.124	0.03	89.993	0.002
22	fact	99	85	48	70	977.884	248.099	0.35	0.131	82.684	0.002
23	published	98	99	42	74	968.006	288.963	0.307	0.138	67.888	0.002
24	cdc	123	153	33	48	1214.947	446.578	0.241	0.09	63.503	0.002
25	women	86	93	14	47	849.475	271.45	0.102	0.088	54.522	0.002
26	used	107	149	45	105	1056.905	434.903	0.328	0.196	45.501	0.002

27	why	100	136	43	94	987.762	396.959	0.314	0.175	44.392	0.002
28	china	99	146	25	48	977.884	426.147	0.182	0.09	37.693	0.002
29	scientists	87	121	28	75	859.353	353.176	0.204	0.14	37.076	0.002
30	reported	123	209	43	124	1214.947	610.032	0.314	0.231	34.115	0.002
31	science	81	114	30	57	800.087	332.745	0.219	0.106	33.764	0.002
32	following	82	121	37	89	809.965	353.176	0.27	0.166	31.181	0.002
33	same	95	160	53	116	938.374	467.01	0.387	0.216	26.93	0.002
34	death	107	206	40	108	1056.905	601.276	0.292	0.201	20.918	0.002
35	incl	67	0	1	0	661.8	0	0.007	0	198.085	0.001
36	disorders	68	10	7	3	671.678	29.188	0.051	0.006	146.475	0.001
37	noaa	47	0	5	0	464.248	0	0.036	0	138.948	0.001
38	corona	43	3	9	2	424.738	8.756	0.066	0.004	106.494	0.001
39	aluminum	36	0	5	0	355.594	0	0.036	0	106.425	0.001
40	vaers	38	1	11	1	375.349	2.919	0.08	0.002	103.555	0.001
41	ice	59	22	9	12	582.779	64.214	0.066	0.022	91.07	0.001
42	bates	33	1	2	1	325.961	2.919	0.015	0.002	89.05	0.001
43	th	62	27	25	18	612.412	78.808	0.182	0.034	88.036	0.001
44	paper	65	33	11	23	642.045	96.321	0.08	0.043	84.035	0.001
45	article	58	27	25	17	572.902	78.808	0.182	0.032	79.182	0.001
46	aluminium	26	0	1	0	256.818	0	0.007	0	76.861	0.001
47	pregnancy	40	9	6	7	395.105	26.269	0.044	0.013	76.172	0.001
48	pcr	54	26	7	11	533.391	75.889	0.051	0.021	72.209	0.001
49	pharma	36	7	13	4	355.594	20.432	0.095	0.007	71.841	0.001
50	mainstream	33	5	21	5	325.961	14.594	0.153	0.009	70.551	0.001
51	unvaccinated	72	53	11	30	711.188	154.697	0.08	0.056	69.91	0.001
52	fda	63	40	16	20	622.29	116.753	0.117	0.037	69.349	0.001
53	shots	65	43	13	25	642.045	125.509	0.095	0.047	69.218	0.001
54	dhs	28	2	1	1	276.573	5.838	0.007	0.002	69.113	0.001
55	video	55	32	32	24	543.269	93.402	0.234	0.045	64.71	0.001
56	nasa	33	7	8	5	325.961	20.432	0.058	0.009	64.081	0.001
57	chart	29	4	8	4	286.451	11.675	0.058	0.007	63.424	0.001

58	chinese	68	53	13	26	671.678	154.697	0.095	0.049	62.593	0.001
59	mrna	43	18	11	6	424.738	52.539	0.08	0.011	62.428	0.001
60	adjuvants	27	3	6	3	266.696	8.756	0.044	0.006	61.865	0.001
61	sars	63	46	15	24	622.29	134.265	0.109	0.045	61.62	0.001
62	biological	28	4	12	4	276.573	11.675	0.088	0.007	60.731	0.001
63	allergy	46	23	6	17	454.37	67.133	0.044	0.032	60.057	0.001
64	below	49	27	36	23	484.003	78.808	0.263	0.043	59.938	0.001
65	reactions	36	12	13	7	355.594	35.026	0.095	0.013	58.653	0.001
66	al	33	9	8	8	325.961	26.269	0.058	0.015	58.57	0.001
67	influenza	60	45	9	25	592.657	131.347	0.066	0.047	57.265	0.001
68	studies	69	62	27	38	681.556	180.966	0.197	0.071	54.855	0.001
69	allergies	27	5	5	4	266.696	14.594	0.036	0.007	54.668	0.001
70	peanuts	27	5	3	1	266.696	14.594	0.022	0.002	54.668	0.001
71	fake	36	14	12	10	355.594	40.863	0.088	0.019	54.377	0.001
72	japan	44	24	7	18	434.615	70.052	0.051	0.034	54.203	0.001
73	brain	40	19	11	15	395.105	55.457	0.08	0.028	53.936	0.001
74	bill	60	48	18	33	592.657	140.103	0.131	0.062	53.845	0.001
75	sea	46	28	9	18	454.37	81.727	0.066	0.034	52.321	0.001
76	hospitalisations	30	9	2	6	296.328	26.269	0.015	0.011	51.209	0.001
77	queen	27	8	5	4	266.696	23.351	0.036	0.007	46.33	0.001
78	cov	48	36	12	18	474.126	105.077	0.088	0.034	45.809	0.001
79	proteins	34	17	8	9	335.839	49.62	0.058	0.017	44.388	0.001
80	reports	67	73	37	58	661.8	213.073	0.27	0.108	42.041	0.001
81	oil	41	29	5	18	404.982	84.646	0.036	0.034	41.246	0.001
82	laboratory	38	25	10	20	375.349	72.97	0.073	0.037	40.644	0.001
83	effects	66	74	34	58	651.923	215.992	0.248	0.108	39.802	0.001
84	viral	43	34	15	25	424.738	99.24	0.109	0.047	39.03	0.001
85	document	28	13	15	11	276.573	37.945	0.109	0.021	38.282	0.001
86	military	30	18	9	14	296.328	52.539	0.066	0.026	34.494	0.001
87	century	27	14	15	12	266.696	40.863	0.109	0.022	34.42	0.001
88	trend	36	27	13	20	355.594	78.808	0.095	0.037	34.355	0.001

89	lab	45	45	9	17	444.493	131.347	0.066	0.032	31.561	0.001
90	evidence	62	78	34	66	612.412	227.667	0.248	0.123	31.421	0.001
91	experimental	33	25	14	12	325.961	72.97	0.102	0.022	31.199	0.001
92	st	39	35	11	28	385.227	102.158	0.08	0.052	31.043	0.001
93	big	75	108	28	74	740.821	315.232	0.204	0.138	29.925	0.001
94	truth	31	23	20	16	306.206	67.133	0.146	0.03	29.879	0.001
95	media	74	106	42	74	730.944	309.394	0.307	0.138	29.828	0.001
96	scientist	26	16	16	12	256.818	46.701	0.117	0.022	29.323	0.001
97	earth	39	37	13	23	385.227	107.996	0.095	0.043	29.142	0.001
98	actually	51	60	32	45	503.758	175.129	0.234	0.084	28.683	0.001
99	page	36	33	23	25	355.594	96.321	0.168	0.047	27.984	0.001
100	pharmaceutical	26	17	15	14	256.818	49.62	0.109	0.026	27.948	0.001
101	shot	56	72	21	40	553.147	210.155	0.153	0.075	27.384	0.001
102	post	44	49	27	39	434.615	143.022	0.197	0.073	26.785	0.001
103	respiratory	45	51	18	35	444.493	148.859	0.131	0.065	26.724	0.001
104	injection	28	21	11	17	276.573	61.295	0.08	0.032	26.72	0.001
105	answer	29	23	16	19	286.451	67.133	0.117	0.035	26.243	0.001
106	cause	69	104	35	78	681.556	303.557	0.255	0.146	25.125	0.001
107	effectiveness	34	33	9	22	335.839	96.321	0.066	0.041	24.728	0.001
108	authors	28	23	16	19	276.573	67.133	0.117	0.035	24.47	0.001
109	pregnant	38	41	8	22	375.349	119.671	0.058	0.041	24.158	0.001
110	body	59	84	26	65	582.779	245.18	0.19	0.121	24.059	0.001
111	scientific	51	67	20	49	503.758	195.56	0.146	0.091	24.048	0.001
112	events	52	71	16	46	513.636	207.236	0.117	0.086	22.914	0.001
113	known	57	82	38	65	563.024	239.343	0.277	0.121	22.784	0.001
114	showing	36	39	28	37	355.594	113.834	0.204	0.069	22.762	0.001
115	protein	31	30	11	13	306.206	87.564	0.08	0.024	22.625	0.001
116	viruses	39	46	15	26	385.227	134.265	0.109	0.049	21.849	0.001
117	humans	32	34	21	23	316.084	99.24	0.153	0.043	20.765	0.001
118	hcq	25	0	1	0	246.94	0	0.007	0	73.904	0
119	bartlett	21	0	1	0	207.43	0	0.007	0	62.079	0

120	raggedy	21	0	1	0	207.43	0	0.007	0	62.079	0
121	ukhsa	21	0	2	0	207.43	0	0.015	0	62.079	0
122	bats	24	2	7	2	237.063	5.838	0.051	0.004	57.881	0
123	canadian	21	1	8	1	207.43	2.919	0.058	0.002	54.461	0
124	doll	21	1	2	1	207.43	2.919	0.015	0.002	54.461	0
125	menstrual	18	0	6	0	177.797	0	0.044	0	53.21	0
126	snopes	17	0	1	0	167.919	0	0.007	0	50.254	0
127	adjuvant	22	3	5	3	217.308	8.756	0.036	0.006	48.242	0
128	fetal	16	0	2	0	158.042	0	0.015	0	47.298	0
129	marcella	16	0	1	0	158.042	0	0.007	0	47.298	0
130	schumer	16	0	2	0	158.042	0	0.015	0	47.298	0
131	newsom	25	6	2	2	246.94	17.513	0.015	0.004	46.548	0
132	cooling	22	4	7	2	217.308	11.675	0.051	0.004	44.781	0
133	asd	15	0	3	0	148.164	0	0.022	0	44.341	0
134	bryson	15	0	1	0	148.164	0	0.007	0	44.341	0
135	gruelle	15	0	1	0	148.164	0	0.007	0	44.341	0
136	nml	14	0	1	0	138.287	0	0.007	0	41.385	0
137	narrative	19	3	10	3	187.675	8.756	0.073	0.006	40.194	0
138	papers	21	5	7	5	207.43	14.594	0.051	0.009	39.211	0
139	adjustments	13	0	3	0	128.409	0	0.022	0	38.429	0
140	bell	13	0	2	0	128.409	0	0.015	0	38.429	0
141	luciferase	13	0	1	0	128.409	0	0.007	0	38.429	0
142	sterilization	13	0	5	0	128.409	0	0.036	0	38.429	0
143	allergic	21	6	7	4	207.43	17.513	0.051	0.007	36.581	0
144	bat	21	6	6	4	207.43	17.513	0.044	0.007	36.581	0
145	saint	16	2	3	2	158.042	5.838	0.022	0.004	35.775	0
146	warfare	16	2	2	2	158.042	5.838	0.015	0.004	35.775	0
147	doshi	12	0	1	0	118.531	0	0.007	0	35.473	0
148	pausebuster	12	0	1	0	118.531	0	0.007	0	35.473	0
149	qiu	12	0	1	0	118.531	0	0.007	0	35.473	0
150	whistle	12	0	6	0	118.531	0	0.044	0	35.473	0

151	injected	24	10	13	8	237.063	29.188	0.095	0.015	34.931	0
152	karl	18	4	2	3	177.797	11.675	0.015	0.006	34.419	0
153	spontaneous	18	4	3	1	177.797	11.675	0.022	0.002	34.419	0
154	arctic	21	7	3	6	207.43	20.432	0.022	0.011	34.212	0
155	ncov	15	2	5	2	148.164	5.838	0.036	0.004	33.062	0
156	berenson	11	0	2	0	108.654	0	0.015	0	32.517	0
157	carole	11	0	1	0	108.654	0	0.007	0	32.517	0
158	excipients	11	0	1	0	108.654	0	0.007	0	32.517	0
159	myocarditis	11	0	4	0	108.654	0	0.029	0	32.517	0
160	psmsl	11	0	1	0	108.654	0	0.007	0	32.517	0
161	toner	11	0	1	0	108.654	0	0.007	0	32.517	0
162	flawed	16	3	10	3	158.042	8.756	0.073	0.006	32.277	0
163	bourla	13	1	4	1	128.409	2.919	0.029	0.002	31.742	0
164	abortion	19	6	5	4	187.675	17.513	0.036	0.007	31.719	0
165	magazine	19	6	7	6	187.675	17.513	0.051	0.011	31.719	0
166	tweet	19	6	7	6	187.675	17.513	0.051	0.011	31.719	0
167	pause	18	5	5	5	177.797	14.594	0.036	0.009	31.714	0
168	surveillance	18	5	7	5	177.797	14.594	0.051	0.009	31.714	0
169	hospitalisation	22	10	2	9	217.308	29.188	0.015	0.017	30.463	0
170	nobody	21	9	15	7	207.43	26.269	0.109	0.013	30.086	0
171	gbs	10	0	2	0	98.776	0	0.015	0	29.561	0
172	peralta	10	0	1	0	98.776	0	0.007	0	29.561	0
173	purification	10	0	1	0	98.776	0	0.007	0	29.561	0
174	sequences	10	0	2	0	98.776	0	0.015	0	29.561	0
175	xiangguo	10	0	1	0	98.776	0	0.007	0	29.561	0
176	documents	19	7	10	6	187.675	20.432	0.073	0.011	29.501	0
177	aerial	12	1	2	1	118.531	2.919	0.015	0.002	28.94	0
178	archive	12	1	3	1	118.531	2.919	0.022	0.002	28.94	0
179	miscarriage	12	1	3	1	118.531	2.919	0.022	0.002	28.94	0
180	palsy	12	1	2	1	118.531	2.919	0.015	0.002	28.94	0
181	et	14	3	4	3	138.287	8.756	0.029	0.006	27.094	0

182	exact	14	3	11	3	138.287	8.756	0.08	0.006	27.094	0
183	mouse	14	3	1	2	138.287	8.756	0.007	0.004	27.094	0
184	reproductive	14	3	8	2	138.287	8.756	0.058	0.004	27.094	0
185	victor	14	3	2	3	138.287	8.756	0.015	0.006	27.094	0
186	ann	17	6	1	5	167.919	17.513	0.007	0.009	26.958	0
187	eating	16	5	7	4	158.042	14.594	0.051	0.007	26.833	0
188	peer	24	15	15	13	237.063	43.782	0.109	0.024	26.744	0
189	anaphylactic	9	0	4	0	88.899	0	0.029	0	26.604	0
190	blaylock	9	0	1	0	88.899	0	0.007	0	26.604	0
191	bnt	9	0	3	0	88.899	0	0.022	0	26.604	0
192	charité	9	0	1	0	88.899	0	0.007	0	26.604	0
193	cq	9	0	1	0	88.899	0	0.007	0	26.604	0
194	inserts	9	0	2	0	88.899	0	0.015	0	26.604	0
195	marines	9	0	1	0	88.899	0	0.007	0	26.604	0
196	penicillin	9	0	2	0	88.899	0	0.015	0	26.604	0
197	rag	9	0	1	0	88.899	0	0.007	0	26.604	0
198	simulation	9	0	3	0	88.899	0	0.022	0	26.604	0
199	thermo	9	0	1	0	88.899	0	0.007	0	26.604	0
200	aden	11	1	1	1	108.654	2.919	0.007	0.002	26.15	0
201	cbs	11	1	4	1	108.654	2.919	0.029	0.002	26.15	0
202	intact	11	1	4	1	108.654	2.919	0.029	0.002	26.15	0
203	readings	11	1	3	1	108.654	2.919	0.022	0.002	26.15	0
204	seafood	11	1	7	1	108.654	2.919	0.051	0.002	26.15	0
205	spraying	11	1	2	1	108.654	2.919	0.015	0.002	26.15	0
206	image	21	12	10	8	207.43	35.026	0.073	0.015	25.03	0
207	journals	12	2	5	2	118.531	5.838	0.036	0.004	25.025	0
208	ingredients	17	7	6	4	167.919	20.432	0.044	0.007	24.903	0
209	premature	16	6	10	6	158.042	17.513	0.073	0.011	24.622	0
210	x	13	3	4	3	128.409	8.756	0.029	0.006	24.54	0
211	alarmists	14	4	6	3	138.287	11.675	0.044	0.006	24.387	0
212	melinda	14	4	6	3	138.287	11.675	0.044	0.006	24.387	0

213	aflds	8	0	1	0	79.021	0	0.007	0	23.648	0
214	depopulation	8	0	4	0	79.021	0	0.029	0	23.648	0
215	experiment	8	0	3	0	79.021	0	0.022	0	23.648	0
216	interaction	8	0	2	0	79.021	0	0.015	0	23.648	0
217	pirbright	8	0	1	0	79.021	0	0.007	0	23.648	0
218	rost	8	0	2	0	79.021	0	0.015	0	23.648	0
219	wireless	8	0	1	0	79.021	0	0.007	0	23.648	0
220	wnv	8	0	1	0	79.021	0	0.007	0	23.648	0
221	yunnan	8	0	2	0	79.021	0	0.015	0	23.648	0
222	reviewed	25	19	16	18	246.94	55.457	0.117	0.034	23.564	0
223	listed	23	16	15	14	227.185	46.701	0.109	0.026	23.472	0
224	fraser	10	1	2	1	98.776	2.919	0.015	0.002	23.376	0
225	mercury	10	1	6	1	98.776	2.919	0.044	0.002	23.376	0
226	conclusion	16	7	15	7	158.042	20.432	0.109	0.013	22.655	0
227	agents	11	2	3	2	108.654	5.838	0.022	0.004	22.39	0
228	autistic	11	2	3	1	108.654	5.838	0.022	0.002	22.39	0
229	eua	11	2	3	1	108.654	5.838	0.022	0.002	22.39	0
230	oils	11	2	4	2	108.654	5.838	0.029	0.004	22.39	0
231	purified	11	2	2	2	108.654	5.838	0.015	0.004	22.39	0
232	pelosi	15	6	3	3	148.164	17.513	0.022	0.006	22.32	0
233	dangers	14	5	11	4	138.287	14.594	0.08	0.007	22.073	0
234	gore	14	5	3	4	138.287	14.594	0.022	0.007	22.073	0
235	injections	14	5	6	5	138.287	14.594	0.044	0.009	22.073	0
236	hilton	12	3	1	3	118.531	8.756	0.007	0.006	22.014	0
237	deliberately	13	4	5	4	128.409	11.675	0.036	0.007	21.95	0
238	doubled	13	4	8	4	128.409	11.675	0.058	0.007	21.95	0
239	tissue	13	4	5	4	128.409	11.675	0.036	0.007	21.95	0
240	words	23	18	16	15	227.185	52.539	0.117	0.028	21.083	0
241	mice	21	15	3	7	207.43	43.782	0.022	0.013	20.943	0
242	injury	16	8	8	6	158.042	23.351	0.058	0.011	20.887	0
243	elliott	7	0	1	0	69.143	0	0.007	0	20.692	0

244	fabricated	7	0	3	0	69.143	0	0.022	0	20.692	0
245	milankovitch	7	0	1	0	69.143	0	0.007	0	20.692	0
246	odinga	7	0	2	0	69.143	0	0.015	0	20.692	0
247	patron	7	0	3	0	69.143	0	0.022	0	20.692	0
248	ppv	7	0	1	0	69.143	0	0.007	0	20.692	0
249	prophet	7	0	2	0	69.143	0	0.015	0	20.692	0
250	specificity	7	0	1	0	69.143	0	0.007	0	20.692	0
251	yeadon	7	0	3	0	69.143	0	0.022	0	20.692	0
252	johnny	9	1	1	1	88.899	2.919	0.007	0.002	20.621	0
253	kansas	9	1	2	1	88.899	2.919	0.015	0.002	20.621	0
254	pregnancies	9	1	3	1	88.899	2.919	0.022	0.002	20.621	0
255	roman	9	1	4	1	88.899	2.919	0.029	0.002	20.621	0
256	stranger	9	1	2	1	88.899	2.919	0.015	0.002	20.621	0
257	unborn	9	1	2	1	88.899	2.919	0.015	0.002	20.621	0
258	consent	18	11	9	7	177.797	32.107	0.066	0.013	20.409	0
259	rna	20	14	1	5	197.552	40.863	0.007	0.009	20.301	0

Appendix 5.2. Keywords Using Text-dispersion Keyness

	Type	Freq Target	Freq Reference	Range Target	Range Reference	Norm Freq Target	Norm Freq Reference	Norm Range Target	Norm Range Reference	Keyness (Likelihood)	Keyness (Effect)
1	fact	99	85	48	70	977.884	248.099	0.35	0.131	31.989	0.002
2	reports	67	73	37	58	661.8	213.073	0.27	0.108	20.64	0.001
3	below	49	27	36	23	484.003	78.808	0.263	0.043	52.263	0.001
4	cdc	123	153	33	48	1214.947	446.578	0.241	0.09	20.356	0.002
5	video	55	32	32	24	543.269	93.402	0.234	0.045	40.745	0.001
6	actually	51	60	32	45	503.758	175.129	0.234	0.084	20.676	0.001
7	adverse	83	21	25	14	819.842	61.295	0.182	0.026	37.973	0.002
8	article	58	27	25	17	572.902	78.808	0.182	0.032	33.371	0.001
9	th	62	27	25	18	612.412	78.808	0.182	0.034	31.988	0.001
10	mainstream	33	5	21	5	325.961	14.594	0.153	0.009	46.1	0.001
11	truth	31	23	20	16	306.206	67.133	0.146	0.03	23.088	0.001
12	conclusion	16	7	15	7	158.042	20.432	0.109	0.013	24.488	0
13	nobody	21	9	15	7	207.43	26.269	0.109	0.013	24.488	0
14	pharma	36	7	13	4	355.594	20.432	0.095	0.007	25.528	0.001
15	biological	28	4	12	4	276.573	11.675	0.088	0.007	22.758	0.001
16	vaers	38	1	11	1	375.349	2.919	0.08	0.002	29.284	0.001
17	exact	14	3	11	3	138.287	8.756	0.08	0.006	22.468	0

Appendix 5.3. Key Semantic Domains

	Semantic Tag	Target Observed Freq	Target Norma Freq	Reference Observed Freq	Reference Norm Freq	overuse/ underuse	LL	Log Ratio	Semantic Domain
1	A6.2--	6	0.01	0	0	+	24.06	6.27	Comparing: unusual
2	S2-	4	0.01	0	0	+	16.04	5.68	No people
3	Q2.1-	2	0	0	0	+	8.02	4.68	Speech: Not communicating
4	O4.6++	15	0.03	15	0	+	22.91	2.68	Temperature: Hot/on Fire
5	O4.6--	7	0.01	7	0	+	10.69	2.68	Temperature: Cold
6	N3.5+	6	0.01	6	0	+	9.16	2.68	Weight: Heavy
7	S1.1.3+++	8	0.01	9	0	+	11.18	2.51	Participating
8	A5.4-	59	0.1	75	0.02	+	74.47	2.34	Evaluation: Unauthentic
9	X3.4+	6	0.01	8	0	+	7.26	2.27	Seen
10	X3.2-	14	0.02	20	0.01	+	15.86	2.17	Sound: Quiet
11	A3-	7	0.01	10	0	+	7.93	2.17	Non-existing
12	E5+	7	0.01	11	0	+	7.2	2.03	Bravery
13	X2.6-	27	0.05	47	0.01	+	24.77	1.88	Unexpected
14	O4.6	58	0.1	102	0.03	+	52.56	1.87	Temperature
15	A5.2+	205	0.36	367	0.1	+	181.86	1.84	Evaluation: TRUE
16	O1.1	150	0.27	271	0.07	+	131.6	1.83	Substances and materials: Solid
17	K5.2	18	0.03	33	0.01	+	15.51	1.81	Games
18	A3	10	0.02	19	0.01	+	8.24	1.76	Being
19	O4.6+	191	0.34	369	0.1	+	153.98	1.73	Temperature: Hot/on Fire
20	A6.2-	70	0.12	138	0.04	+	54.94	1.71	Comparing: Unusual
21	O1	134	0.24	268	0.07	+	103.16	1.68	Substances and materials generally
22	S1.2.6-	17	0.03	34	0.01	+	13.09	1.68	Foolish
23	L1	19	0.03	39	0.01	+	14.11	1.65	Life and living things
24	Q4.2	181	0.32	396	0.11	+	122.62	1.55	The Media: Newspapers etc.
25	X5.1-	19	0.03	43	0.01	+	12.22	1.51	Inattentive
26	S9	146	0.26	335	0.09	+	91.92	1.49	Religion and the supernatural

27	Y1	356	0.63	827	0.23	+	219.76	1.47	Science and technology in general
28	O1.2	91	0.16	215	0.06	+	54.65	1.44	Substances and materials: liquid
29	Q4.1	74	0.13	178	0.05	+	43.14	1.42	The Media: Books
30	X2.2	122	0.22	299	0.08	+	68.91	1.39	Knowledge
31	O4.6-	36	0.06	89	0.02	+	20.03	1.38	Temperature: Cold
32	F1	255	0.45	635	0.18	+	140.1	1.37	Food
33	A5.2-	75	0.13	196	0.05	+	37.77	1.3	Evaluation: FALSE
34	A6.1+++	59	0.1	170	0.05	+	24.46	1.16	Comparing: Similar
35	L3	38	0.07	109	0.03	+	15.91	1.16	Plants
36	W3	322	0.57	949	0.26	+	127.12	1.13	Geographical terms
37	A7	23	0.04	70	0.02	+	8.44	1.08	Probability
38	N6+++	20	0.04	62	0.02	+	7.03	1.05	Frequent
39	E5-	84	0.15	270	0.07	+	27.03	1	Fear/shock
40	A5.4+	53	0.09	171	0.05	+	16.89	0.99	Evaluation: Authentic
41	G2.2-	41	0.07	136	0.04	+	12.16	0.95	Unethical
42	W1	146	0.26	515	0.14	+	36.44	0.87	The universe
43	A1.5.1	142	0.25	511	0.14	+	33.36	0.84	Using
44	Z4	168	0.3	613	0.17	+	37.79	0.82	Discourse Bin
45	G2.1-	50	0.09	187	0.05	+	10.38	0.78	Crime
46	W4	199	0.35	758	0.21	+	38.84	0.76	Weather
47	Q4	159	0.28	635	0.18	+	26.12	0.69	The Media
48	G3	131	0.23	530	0.15	+	20.46	0.67	Warfare, defence and the army; weapons
49	A6.2+	165	0.29	677	0.19	+	24.36	0.65	Comparing: Usual
50	M4	47	0.08	192	0.05	+	7.06	0.65	Sailing, swimming, etc.
51	A2.2	583	1.03	2399	0.66	+	85.04	0.64	Cause&Effect/Connection
52	Y2	111	0.2	458	0.13	+	16.01	0.64	Information technology and computing
53	A10+	358	0.63	1487	0.41	+	50.22	0.63	Open; Finding; Showing
54	T1	91	0.16	379	0.1	+	12.62	0.63	Time
55	A6.1+	129	0.23	542	0.15	+	17.23	0.61	Comparing: Similar
56	A1.1.2	102	0.18	430	0.12	+	13.43	0.61	Damaging and destroying
57	A1.7-	69	0.12	294	0.08	+	8.66	0.59	No constraint

58	A13.1	61	0.11	262	0.07	+	7.39	0.58	Degree: Non-specific
59	A8	74	0.13	333	0.09	+	7.13	0.51	Seem

Appendix 6.1. Lexico-grammatical Features Used in the Factor Analysis

	Tags	**Features**
1	AWL	average word length
2	AMP	amplifiers
3	ANDC	independent clause coordination
4	BEMA	be as main verb
5	BYPA	by-passives
6	CAUS	causative adverbial subordinators
7	COND	conditional adverbial subordinators
8	CONJ	conjuncts
9	CONT	contractions
10	DEMO	demonstratives
11	DEMP	demonstrative pronouns
12	DWNT	downtoners
13	EMPH	emphatics
14	EX	existential there
15	FPP1	first person pronouns
16	GER	gerunds
17	JJ	attributive adjectives
18	NEMD	necessity modals
19	NN	total other nouns
20	NOMZ	nominalizations
21	OSUB	other adverbial subordinators
22	PASS	agentless passives
23	PEAS	perfect aspect
24	PHC	phrasal coordination
25	PIN	total prepositional phrases

26	PIT	pronoun it
27	PLACE	place adverbials
28	POMD	possibility modals
29	PRED	predicative adjectives
30	PRESP	present participial clauses
31	PRIV	private verbs
32	PRMD	predictive modals
33	PUBV	public verbs
34	RB	total adverbs
35	SERE	sentence relatives
36	SPAU	split auxiliaries
37	SPP2	second person pronouns
38	SUAV	suasive verbs
39	SYNE	synthetic negation
40	THATD	subordinator that deletion
41	THVC	that verb compliments
42	TIME	time adverbials
43	TO	infinitives
44	TOBJ	that relative clauses on object position
45	TPP3	third person pronouns
46	TSUB	that relative clauses on subject position
47	VBD	past tense
48	VPRT	present tense
49	WHSUB	WH relative clauses on subject position
50	WZPAST	past participial WHIZ deletion relatives
51	WZPRES	present participial WHIZ deletion relatives
52	XX0	analytical negation

Appendix 6.2. R Code for the Factor Analysis

```
#Installing the psych package
Install.packages (‘psych’)
library ('psych')
#Data
CN <-  read.csv (file = "CNDataset.csv", header = T)
CNnumeric <- (CN[,3:54])
#KMO Results
KMO (CNnumeric)
#Scree Plot
scree (CNnumeric)
#Factor Analysis
f_loading_results <- fa (CNnumeric, fm = "pa", nfactors = 3, rotate ="promax")
#Factors and Their Loading
print (f_loading_results, cut = .30)
#Calculating Factor Scores Using Regression
f_score_results <- factor.scores (CNnumeric, f_loading_results, method = 'regression')
#Presenting Factor Scores
DF <- data.frame (f_score_results$scores)
```

Appendix 6.3. Dimension Scores

Filename	News Type	Dimension 1 Scores	Dimension 2 Scores	Dimension 3 Scores
A10	false	-1.159164	-0.3398849	-0.2759701
A11	false	0.77681269	0.73544478	-0.2589026
A12	false	0.0068874	-0.5696045	-0.9107714
A13	false	-0.1853011	-0.9744352	-0.9277468
A14	false	0.63878105	-0.2845275	-0.7518954
A15	false	1.09635008	0.98914981	0.35646142
A16	false	0.15642065	-1.8799482	-1.070308
A17	false	0.43231842	1.26171167	-0.8259285
A18	false	-1.6852786	0.68748484	-1.230955
A19	false	1.1949525	-0.5069034	-0.2958731
A20	false	-0.7350231	-0.0647229	-1.1966299
A2	false	0.69055297	0.98742819	-0.5874407
A3	false	0.14196349	0.77348673	-0.3695467
A4	false	1.48853563	1.46384811	-0.8974834
A5	false	1.35569341	1.50216747	-0.9660204
A6	false	-0.0205284	2.22652329	-1.7887931
A7	false	0.7405821	-0.1259124	-0.6582592
A8	false	0.97034471	0.52037024	-0.6771655
B10	false	-0.2622134	-0.841057	-0.3926899
B11	false	0.55902801	-0.1123844	-0.7301558
B12	false	-0.9530882	1.49003	-0.4988912
B13	false	0.5403675	-0.9738573	-1.3322026
B14	false	-0.9808693	-0.0781573	-0.8767065
B16	false	0.80179851	1.07889458	0.89909035
B18	false	0.82578652	-0.9472058	-1.0054241

B19	false	-2.114142	0.40658098	-1.5113245
B1	false	1.62493994	0.5099692	0.245678
B20	false	1.15860483	0.30492767	-0.4185045
B21	false	0.14089728	-0.0355546	-0.9936658
B22	false	0.40971309	0.19196071	-0.8729759
B23	false	0.97979604	0.33011435	-0.9513967
B24	false	-0.5672118	-0.9419951	-1.2308528
B25	false	0.06581919	1.62584972	-0.1308857
B27	false	0.8594008	1.00498426	0.02714909
B2	false	-1.0746699	1.06312145	-1.2941623
B4	false	-0.2449802	0.32602317	-0.2645356
B6	false	-0.7525101	-0.9137168	-1.2959315
B7	false	0.64325132	0.87941125	-1.0110601
B8	false	-0.0480599	-0.1166013	-0.3509284
B9	false	-0.018922	0.34334989	-0.2023653
C12	false	-0.0641668	0.28543767	-0.3808292
C14	false	-0.3492284	-1.8539367	-0.3747601
C15	false	-0.7238113	0.25962408	0.81511022
C16	false	-0.5975646	0.80523022	1.01388909
C18	false	-0.826845	-0.205469	0.11007967
C19	false	1.11555699	0.03180447	0.3384071
C1	false	-1.0093504	-0.9744026	-1.5113754
C20	false	-0.756914	-0.6731646	-1.4283225
C25	false	1.10486266	-1.1736058	-1.0407363
C28	false	0.39198929	-0.3186599	0.40951232
C2	false	0.49649944	-0.5146938	0.315828
C31	false	0.91232568	0.4508898	-0.7094484
C33	false	-0.7798464	-0.0511543	-0.4748848
C34	false	-1.8294647	0.21310033	-1.7683242
C35	false	0.99546436	-0.7978474	-1.495367
C37	false	-1.9602777	0.3968536	-0.1673459

C38	false	0.76316364	-0.3299531	1.08494156
C39	false	-0.0482544	-1.829359	-1.3695633
C42	false	-1.0743243	-0.982028	-0.0854684
C44	false	-0.4035428	0.92799024	-0.498853
C46	false	0.68406369	1.42059205	0.13695999
C47	false	-1.0097731	0.86551317	-1.3833211
C48	false	0.99951684	1.37651216	0.02308979
C49	false	0.77491564	-0.0414665	0.31787937
C50	false	1.17871333	0.14035088	-0.644193
C51	false	0.18346535	0.26799912	0.9357646
C52	false	-0.547101	1.63865344	-0.0344027
C53	false	-0.9872087	-0.0523225	-0.4577453
C54	false	-0.0411814	-0.6904084	-0.8095972
C55	false	-1.4436425	1.22590224	-0.8185701
C57	false	0.30742032	0.7796531	-0.3286583
C58	false	-2.3080814	1.837159	-0.2721653
C59	false	0.33593862	-1.6559531	-2.0138647
C5	false	-0.7538587	-1.1650298	0.15074555
C60	false	-0.5151209	-0.1226883	-0.6049654
C63	false	-0.0257159	0.05852733	-0.0310359
C65	false	1.61792235	-0.7614962	-0.6833249
C66	false	-0.1019035	1.37595375	-0.5197273
C67	false	0.34492295	0.46955326	-0.7943757
C68	false	-0.551563	0.57211956	-0.8510634
C69	false	1.44756952	0.35548962	-1.162105
C6	false	-0.1009407	0.07738197	-1.3569019
C70	false	0.20742729	1.37750311	-0.0199667
C71	false	0.03244498	0.8763066	0.08981925
C72	false	1.13338224	0.0815878	-0.0596475
C73	false	-1.0300635	1.38546937	-0.8979046
C74	false	-1.4667495	0.96947784	-1.1451762

C75	false	1.13744468	-0.4882544	-0.829775
C76	false	-0.1314389	0.32636154	-0.7007511
C78	false	0.09842156	-1.3067128	0.12532119
C79	false	-0.0531776	-0.4609066	-0.2867659
C80	false	-0.6824452	-0.4745098	-0.9131529
C81	false	2.45148762	-0.0547579	-0.010487
C82	false	-0.4334158	-1.1784295	-1.4388577
C83	false	-0.9259979	-2.2170095	-0.7637265
C84	false	-0.0382887	-0.6089999	-0.1363971
C85	false	-0.4984143	0.59885481	-0.6522621
C86	false	-0.5000384	1.9356525	-0.1251188
C87	false	-0.8184891	1.13303914	-0.6307499
C89	false	-0.1453708	0.75306049	0.0419706
C90	false	1.02916693	0.53257561	-0.4742384
ABlog10	blog	0.18939969	0.38991806	0.17997679
ABlog11	blog	1.09773022	0.70121481	-0.0392411
ABlog12	blog	0.65055386	0.94756756	0.02688988
ABlog13	blog	0.06278011	2.20919231	0.29573667
ABlog14	blog	1.0722053	1.27284703	0.03295364
ABlog15	blog	0.85855361	1.27811801	-0.8623263
ABlog16	blog	-0.1459469	0.5900772	-0.6788723
ABlog17	blog	-0.4760407	-0.259002	-0.8762179
ABlog18	blog	-0.4173049	0.70984146	0.41016209
ABlog19	blog	-1.2078286	0.06601967	0.23072919
ABlog2	blog	-0.4852731	0.76726523	0.28145528
ABlog3	blog	1.53085188	-0.6143541	0.40487472
ABlog4	blog	0.18309455	0.71200318	-0.1516739
ABlog5	blog	-0.8492446	-0.2908956	0.53481524
ABlog6	blog	-2.1888932	1.29796144	-1.0352669
ABlog7	blog	0.06043382	1.23721807	0.40155476
ABlog8	blog	-0.1103386	0.30450045	0.36664504

ABlog9	blog	0.71313977	0.93275579	0.34885942
BBlog1	blog	-0.9608996	-1.0534887	-0.5356204
BBlog10	blog	0.40358164	0.34975344	-1.0052863
BBlog11	blog	-2.0308119	1.78663774	-1.0572188
BBlog12	blog	-1.8572756	2.17672084	-0.9896091
BBlog13	blog	0.01860354	1.52336631	-1.0673785
BBlog14	blog	-2.5804062	1.718434	-1.429333
BBlog17	blog	-0.0389532	1.14675589	-0.3717542
BBlog19	blog	-2.1515697	1.23836363	-0.1116398
BBlog2	blog	0.34136863	1.38977847	-0.6337519
BBlog20	blog	-0.0450417	-0.3697467	-1.7879451
BBlog21	blog	1.31033606	0.53765266	-0.6089211
BBlog22	blog	1.24820331	0.14335922	-0.0372435
BBlog23	blog	-0.360248	0.47811087	-0.6588604
BBlog24	blog	-0.7047151	-0.0265489	-2.1607418
BBlog25	blog	-0.9317403	1.0522656	-1.0914348
BBlog26	blog	-1.3443183	1.75804817	-1.4877219
BBlog27	blog	-1.7193899	0.16799493	-1.981592
BBlog3	blog	-2.0402029	1.48380669	-0.3266018
BBlog5	blog	1.63527342	1.73900901	0.48802692
BBlog7	blog	0.04548983	0.8288912	1.16772442
BBlog8	blog	-0.638566	0.86589257	-0.0285228
BBlog9	blog	-1.809226	1.51093868	-0.4732233
CBlog1	blog	-1.1726477	0.26100559	-0.061909
CBlog10	blog	-1.0296532	1.77711499	1.35179441
CBlog11	blog	-0.5157994	0.72123455	-0.1060231
CBlog12	blog	0.89577593	0.23974407	-0.218132
CBlog13	blog	-0.3586553	0.59765838	-0.097941
CBlog14	blog	-0.2984248	-0.8540199	-0.2800383
CBlog15	blog	-1.9313212	0.66782516	-0.525354
CBlog16	blog	-0.5446004	0.75053535	1.762233

CBlog17	blog	0.44910401	0.95583807	-1.1119999
CBlog18	blog	-0.6574692	0.61504734	0.9254267
CBlog19	blog	0.23998007	0.35560529	-1.1254824
CBlog2	blog	-0.535268	0.77978257	-0.3928584
CBlog21	blog	0.27573959	0.68720916	0.29419239
CBlog22	blog	-1.1446749	0.6326099	0.16040749
CBlog23	blog	-1.0906282	0.99871912	0.17973504
CBlog24	blog	-0.2295132	0.76227271	-1.5832291
CBlog25	blog	-1.1751601	0.7154119	-0.2959168
CBlog26	blog	-1.1315325	0.5000082	-0.1563691
CBlog27	blog	-0.7625298	-0.5438446	-0.3049177
CBlog28	blog	-1.1836318	-0.256846	-1.2025969
CBlog29	blog	-1.1630757	0.0929999	0.21393265
CBlog3	blog	0.87424562	-0.5302221	-0.8484025
CBlog30	blog	-1.3204468	0.81140799	-0.742894
CBlog31	blog	-1.743978	0.3343144	0.27958317
CBlog32	blog	0.72418392	0.36939034	-0.3409922
CBlog33	blog	-0.2811966	1.25555692	0.50594713
CBlog34	blog	-0.95038	2.32106306	1.35784625
CBlog35	blog	-0.7142057	0.35606813	-0.0294193
CBlog36	blog	-1.3065135	-1.0492664	-0.0223037
CBlog37	blog	-0.5303544	2.13258472	-0.0990963
CBlog38	blog	1.09485787	0.74214338	3.12384453
CBlog39	blog	-0.6492914	-1.5237584	-0.4041922
CBlog4	blog	-0.7860656	0.88915115	-0.8404712
CBlog40	blog	-0.8029727	-0.2172272	-0.0634126
CBlog41	blog	-1.1983364	0.55649161	0.07911015
CBlog42	blog	-1.1345813	-0.4302408	-0.0435672
CBlog43	blog	2.01905474	-0.8515097	0.65588208
CBlog44	blog	0.57583786	1.03627489	0.61122712
CBlog45	blog	-0.602996	0.25331712	0.7313975

CBlog46	blog	-0.4864134	0.10965631	-0.1881081
CBlog47	blog	-0.8091971	0.58708802	0.43081487
CBlog48	blog	-0.5185268	0.89256498	0.08150298
CBlog49	blog	-0.5365983	0.45168774	-0.7743171
CBlog5	blog	0.94052403	0.32882983	0.03412278
CBlog50	blog	0.73533291	0.00817231	-0.8815464
CBlog51	blog	-0.0476547	0.38940028	0.65082993
CBlog52	blog	-0.8976691	-0.0339737	-0.3186344
CBlog53	blog	-0.4267076	-0.3546229	0.19072295
CBlog54	blog	-2.5076268	0.70695184	0.53858009
CBlog55	blog	0.35137851	0.878328	0.21873406
CBlog56	blog	-1.0197513	0.87843645	1.6070645
CBlog57	blog	-0.5868186	-0.5635677	-0.7548738
CBlog59	blog	0.38691108	0.57332555	0.81097848
CBlog60	blog	0.26560078	-0.1401962	1.08820982
CBlog61	blog	-1.3037507	0.60532681	-0.5276173
CBlog62	blog	-1.4483167	0.41645186	-0.5364735
CBlog63	blog	-0.8095923	0.33746006	-1.2712069
CBlog64	blog	-0.5656822	-0.2675529	-0.9838755
CBlog7	blog	-0.837341	-0.3414418	-1.1401849
CBlog8	blog	0.14122469	1.10122398	-0.120583
CBlog9	blog	0.01565563	0.83959449	-0.0848179
ABroad1	broad	1.72490763	-1.3925405	0.64618952
ABroad10	broad	0.71693713	0.65457276	1.30022152
ABroad13	broad	0.28662743	-0.3259973	-0.3109215
ABroad14	broad	-0.1526012	-0.2383869	0.76625344
ABroad15	broad	0.26940135	0.06726666	-0.3831267
ABroad16	broad	-0.0770896	-0.3405326	-0.9100918
ABroad17	broad	1.19517368	1.30520929	-0.083075
ABroad18	broad	0.29770241	0.68077952	-0.1403048
ABroad19	broad	0.40517112	0.69463453	0.89210231

ABroad2	broad	0.50968772	-0.1781838	-0.0275137
ABroad20	broad	0.93307801	1.05543309	0.79975224
ABroad3	broad	1.43281274	0.04200392	-0.215944
ABroad4	broad	1.48545433	0.48093958	0.94672326
ABroad5	broad	-0.9934052	0.01596122	0.93507857
ABroad6	broad	-0.1777823	-0.1110138	-0.0303759
ABroad8	broad	-0.3657023	0.01290785	-0.6678711
ABroad7	broad	1.4800227	0.0963671	-0.2481427
BBroad10	broad	1.70217147	1.22945107	0.9326416
BBroad11	broad	-0.3308024	-0.5983899	-0.5011814
BBroad12	broad	-0.98768	-1.0275369	-0.7443786
BBroad13	broad	-0.3947089	-0.3998655	-0.6673384
BBroad14	broad	0.33796505	0.99061518	-0.0425218
BBroad15	broad	1.89546146	0.39477461	2.4743091
BBroad16	broad	-0.525106	-2.093241	-0.7259116
BBroad17	broad	-0.1380555	0.61997912	-0.92874
BBroad18	broad	-1.1966038	0.87895505	0.25176909
BBroad19	broad	-0.2248946	-1.2123658	-0.8902416
BBroad20	broad	-1.059052	0.03311958	-0.7077149
BBroad21	broad	-0.2211946	-0.364801	-1.6849801
BBroad22	broad	-0.9286752	0.21085503	-0.4433167
BBroad23	broad	-0.0005369	0.46080301	0.54535535
BBroad24	broad	0.00247653	0.46882	-0.0136863
BBroad25	broad	0.06152592	0.07488972	-0.2546852
BBroad26	broad	0.24501171	0.1661088	-0.8184654
BBroad27	broad	-0.1613821	0.71448674	0.59186943
BBroad6	broad	-0.0696451	0.30430327	-0.2897512
BBroad7	broad	-0.6106051	-0.1417718	-1.8027209
BBroad8	broad	-0.6375182	0.46675774	-0.3235858
BBroad9	broad	1.33118975	0.70832814	0.1585613
CBroad1	broad	-0.9004579	-1.0097705	-1.0150574

CBroad10	broad	-0.4214192	-0.0483552	0.31362949
CBroad11	broad	0.0425846	0.18715736	1.75307177
CBroad12	broad	-0.3068404	-3.6626756	-2.209428
CBroad13	broad	0.43244098	0.81255877	-0.1376761
CBroad14	broad	-0.5131715	-0.4424194	-0.5145269
CBroad15	broad	-0.0884133	-0.0414752	-0.4081191
CBroad16	broad	-0.7356365	-0.1965936	-1.126907
CBroad17	broad	1.04590154	0.82685624	0.52102432
CBroad19	broad	0.26988203	0.13076015	1.91294925
CBroad2	broad	-0.0395962	0.18522365	1.33317937
CBroad20	broad	-1.1664935	-0.8207448	0.34664379
CBroad21	broad	0.68133256	-2.1349965	0.01223156
CBroad22	broad	-0.7578318	-0.6978675	-0.3900018
CBroad23	broad	-0.1630006	-0.9923667	-0.1123908
CBroad24	broad	-0.1625397	-1.2923328	0.40555517
CBroad25	broad	0.89064144	0.60951939	1.94980216
CBroad26	broad	-0.8066753	-0.7605519	-0.2569134
CBroad28	broad	0.68409771	-0.2497179	1.21291375
CBroad29	broad	-0.3506679	-0.6343873	1.24009362
CBroad3	broad	0.40296721	-0.8487307	0.7083055
CBroad30	broad	1.43715016	1.03151374	2.90289103
CBroad31	broad	-0.1945217	0.5475599	-0.3442101
CBroad32	broad	0.77294232	1.81821992	0.21908384
CBroad34	broad	-0.4880069	-0.3262561	-0.8200491
CBroad35	broad	-0.0659773	0.31670233	0.12896366
CBroad36	broad	1.37122248	0.92520052	1.39094977
CBroad37	broad	-0.134663	0.67437208	0.40044282
CBroad38	broad	-0.5051153	0.59659771	0.20154583
CBroad39	broad	-1.0513728	-1.1286766	-0.3603866
CBroad4	broad	-0.1446917	-1.0179903	1.27195705
CBroad40	broad	1.38614431	0.54800811	0.3168038

CBroad41	broad	0.91008904	-0.1082476	2.02357705
CBroad42	broad	0.57337892	-1.0378462	-0.499029
CBroad43	broad	0.58537048	-0.7178976	0.26967836
CBroad44	broad	1.16171259	-3.5066076	-0.0122549
CBroad45	broad	-1.2138176	0.63438114	-0.3965539
CBroad46	broad	1.27458888	-1.1628702	0.45225324
CBroad47	broad	-0.1156811	0.04054911	0.31293819
CBroad48	broad	-0.4597376	0.54173211	-0.5373494
CBroad49	broad	-0.5762611	-1.660031	-0.1157934
CBroad5	broad	-0.6025428	-1.205318	-1.019127
CBroad50	broad	0.65512953	0.71995099	0.35566463
CBroad51	broad	-0.0564804	-1.1455552	-0.2809992
CBroad52	broad	1.69679554	-0.4403803	1.84458988
CBroad53	broad	-0.4936085	-1.6943801	-1.4832317
CBroad54	broad	1.08477239	1.12702065	1.13049918
CBroad55	broad	-0.1451459	1.0221245	0.33331258
CBroad56	broad	-0.5371233	-0.4676307	-1.2140835
CBroad57	broad	0.32866401	-2.1047201	0.94116354
CBroad58	broad	-0.2151282	0.60707914	0.58777666
CBroad59	broad	0.06017181	-0.4936053	-0.210311
CBroad6	broad	0.06635651	0.88920537	2.03795601
CBroad60	broad	-1.1247941	-1.0903703	0.29307086
CBroad61	broad	-0.2888489	-0.7782009	-0.3595116
CBroad63	broad	0.68402324	0.67728456	-0.876286
CBroad64	broad	-0.1702753	-1.0507252	-0.2085614
CBroad65	broad	-1.0867	1.3231713	-0.0282313
CBroad66	broad	-0.8628836	-0.2476791	0.89046832
CBroad69	broad	-1.5436753	1.18077503	-0.8404809
CBroad7	broad	-0.1752728	-0.3654914	-1.4991017
CBroad8	broad	-0.4491295	-0.0889481	0.39738074
ATab10	tab	-2.2742345	0.53540142	0.22300964

ATab11	tab	-1.0118242	0.62611609	0.69485833
ATab12	tab	0.01463921	-0.2034892	0.74614513
ATab13	tab	-0.0501382	-0.2521346	0.49255792
ATab14	tab	0.71761608	-0.3266647	0.46702197
ATab15	tab	0.27104401	-0.0745157	0.02484526
ATab16	tab	0.43256883	0.37765706	1.27962046
ATab17	tab	-0.9198146	-0.891556	0.28273526
ATab18	tab	-0.1583066	0.62518429	-0.4000345
ATab19	tab	-0.1578751	-1.618701	0.67454775
ATab20	tab	1.28318516	1.41935153	0.42831524
ATab3	tab	-0.0632871	0.22463197	0.18416131
ATab4	tab	-1.6106237	-0.4152141	-1.182953
ATab5	tab	0.00850379	-1.1002832	1.35562998
ATab6	tab	-1.6987265	-0.1289412	-0.8039625
ATab7	tab	-0.4140952	-1.4399842	-0.2395803
ATab8	tab	-0.1020778	0.24114934	1.02850781
ATab9	tab	0.36699283	0.38124263	1.08662445
BTab10	tab	0.18726444	1.13212852	1.98364736
BTab11	tab	1.27980739	-0.8246785	0.07891821
BTab12	tab	-1.183778	0.39414948	0.9921392
BTab13	tab	1.90915195	0.74810693	0.86322844
BTab14	tab	0.10462442	-0.1851154	-0.0656984
BTab15	tab	1.953396	1.55333939	0.69833187
BTab16	tab	0.97987853	-0.4976306	1.29397501
BTab17	tab	0.99110269	1.0535135	1.05083119
BTab18	tab	1.34115451	0.68467209	-0.7148694
BTab19	tab	-0.0634043	0.43766326	0.16765801
BTab20	tab	0.16661258	-0.5270853	1.53657665
BTab21	tab	0.47377284	0.63670173	0.23522692
BTab22	tab	1.47862422	-0.0455915	1.42286249
BTab23	tab	1.50225259	0.83956161	0.97310365

BTab24	tab	1.62206864	-1.897616	2.07507544
BTab25	tab	0.0189568	-0.0117266	0.56651119
BTab26	tab	0.05037216	0.2286345	0.52966627
BTab4	tab	0.18179392	0.46375105	-0.6311224
BTab5	tab	-0.041438	-2.3346004	-0.2278583
BTab6	tab	2.18125092	1.27025018	1.30458538
BTab8	tab	0.79509185	1.00715701	0.97534728
CTab10	tab	-0.1048976	-1.033652	-0.5300764
CTab11	tab	-0.668052	-0.3669717	0.57923031
CTab12	tab	2.25748876	-0.3367736	2.38263633
CTab13	tab	-0.3479053	-1.3039098	2.10152349
CTab14	tab	0.59552767	-1.5075891	1.6102039
CTab15	tab	-1.1744233	-0.5784394	0.11557333
CTab16	tab	0.63799737	-0.5235889	1.79270012
CTab17	tab	0.7577109	0.39879749	0.56924144
CTab18	tab	-0.0837803	1.13457057	0.3773811
CTab19	tab	1.35286799	0.31439526	-0.8360229
CTab2	tab	-0.2634228	-1.0037896	-0.2573027
CTab20	tab	-0.2791704	-0.5854242	-0.7345855
CTab22	tab	-0.1200146	-2.1747713	-0.5278088
CTab23	tab	0.12876393	-1.9410761	-0.6396293
CTab24	tab	-1.0533799	-1.0320397	-1.6945032
CTab26	tab	-0.0573804	-1.2117122	0.94290767
CTab27	tab	-0.7517791	-0.4778963	0.41152968
CTab28	tab	-0.4768601	0.52450304	-0.1364068
CTab29	tab	-0.2431622	0.32005355	-1.2129929
CTab3	tab	0.09872386	-1.2174906	-0.2210123
CTab31	tab	-0.4182412	-0.0484254	0.63248971
CTab32	tab	1.05613047	0.43449399	0.9491597
CTab33	tab	0.8050784	-0.1871373	1.15528127
CTab35	tab	2.00505833	-0.7160094	0.64412986

CTab36	tab	-0.3132171	-1.1694432	0.57756207
CTab37	tab	2.50706506	-2.3906618	0.44478572
CTab38	tab	-1.1342777	-0.5323649	-0.4574209
CTab39	tab	-0.4071247	-1.7101353	-0.4316307
CTab40	tab	2.49528028	-1.3187109	0.49024978
CTab41	tab	0.59093822	-1.5476692	0.67342846
CTab42	tab	-0.0876902	-0.3738773	-0.202811
CTab43	tab	0.80555465	-1.1190461	0.75455643
CTab44	tab	0.11155324	-1.1119608	0.57002119
CTab45	tab	-0.2087922	-0.6320579	0.12447092
CTab46	tab	0.62913114	1.32930551	0.78875797
CTab47	tab	1.54594085	0.09308131	0.60128562
CTab48	tab	-1.0502413	1.0303754	0.04564887
CTab49	tab	0.21733287	0.57226477	0.37837698
CTab5	tab	1.7329181	-2.6463952	0.20530139
CTab50	tab	1.77520047	0.21766899	0.29344445
CTab51	tab	0.53615243	-0.0227513	0.10926958
CTab53	tab	-0.2239772	-0.8795552	-0.6506159
CTab54	tab	-1.2472414	-0.3465801	-0.3282348
CTab55	tab	-1.3412833	-0.7080243	0.30454984
CTab56	tab	-0.4031752	-2.1479636	0.24907512
CTab57	tab	-0.7649181	-1.5431133	-1.3074182
CTab58	tab	-0.3523362	-0.8500275	-1.5300627
CTab59	tab	-1.4239815	-0.7500993	-0.0425227
CTab6	tab	-0.0500834	0.21885056	0.92610693
CTab60	tab	1.19013828	0.12102184	1.55181538
CTab61	tab	0.10008626	-0.6019163	-0.3114841
CTab62	tab	-0.7266137	-0.2292152	-0.0882101
CTab63	tab	-1.02878	-0.5496437	-0.3416007
CTab65	tab	0.19117374	0.10124689	0.87865823
CTab66	tab	-0.636577	-0.1845719	1.35447025

CTab67	tab	-0.874058	-1.5550365	-1.215085
CTab68	tab	-0.0372857	-0.610008	1.45642126
CTab70	tab	0.65778752	0.22576977	1.18998794
CTab71	tab	0.53882923	-1.1539062	0.65180336
CTab72	tab	-0.4937332	-1.3762596	-0.6236319
CTab75	tab	1.07076487	0.76329866	1.08654447
CTab8	tab	0.78597862	0.23311406	1.3661524
AWeb10	web	-0.1053874	0.76694718	1.3222205
AWeb11	web	-0.7570925	-0.2637788	0.91783915
AWeb12	web	-1.1713244	-0.2335568	1.65654857
AWeb13	web	1.35150523	1.81013655	-0.6159352
AWeb14	web	0.55134587	-0.0744296	-0.6307834
AWeb15	web	0.05631644	-0.5816651	-0.1067669
AWeb16	web	1.37906314	0.35402307	1.08567028
AWeb17	web	1.24235762	0.50498057	1.35997045
AWeb18	web	-0.7091054	-1.4567324	-0.1195865
AWeb19	web	-0.7307765	-1.0472879	0.73681669
AWeb2	web	0.38858405	0.25442656	1.41747985
AWeb20	web	0.65492636	1.07782141	0.87616404
AWeb3	web	-0.7264925	-3.0821553	-0.0193846
AWeb4	web	1.02973495	-0.6569792	0.57351452
AWeb6	web	-0.738633	0.79642885	-0.4679143
AWeb7	web	0.707224	-0.8365896	0.10929366
AWeb8	web	-1.1723855	0.45750808	-0.6442012
AWeb9	web	-0.1109659	-0.0451242	0.89358708
BWeb10	web	-1.9408788	0.56691232	-1.8079468
BWeb11	web	-1.6934964	1.39427871	-2.0065726
BWeb12	web	-1.3436691	-0.5629809	0.63729353
BWeb13	web	1.32649485	-0.1696189	0.51919856
BWeb14	web	0.10196338	1.22331881	-1.6921044
BWeb15	web	1.51353105	0.06463879	1.10654259

BWeb16	web	1.0565625	-2.0954705	0.69746244
BWeb17	web	0.33588291	0.32647292	-0.0537863
BWeb18	web	1.34335423	0.05770456	1.09443243
BWeb19	web	-2.0264044	0.71752038	-1.0421879
BWeb20	web	-0.5291012	0.9389407	0.09756467
BWeb21	web	-0.4320684	-0.1416541	0.26861222
BWeb22	web	-1.8353763	1.57537545	-0.0804527
BWeb23	web	0.40077306	-0.7278418	-0.1036423
BWeb24	web	0.79567463	0.25176128	1.29791653
BWeb25	web	1.7535941	-0.7429733	-0.1246058
BWeb26	web	-0.6444545	-0.1328346	0.32902964
BWeb27	web	1.12452497	-0.7341344	0.05633034
BWeb6	web	-0.1027941	-1.2056452	1.19932131
BWeb7	web	-0.8467293	-0.1964615	-0.4299595
BWeb8	web	0.35990747	-0.3355722	-0.1339466
BWeb9	web	-1.2773852	1.28782908	-0.8243943
CWeb1	web	0.13987231	-0.095614	0.55444946
CWeb10	web	0.56058446	-0.0826652	-0.9251378
CWeb11	web	-0.4933123	-0.3521267	0.31221595
CWeb12	web	-0.8783427	-2.914666	-1.3934981
CWeb13	web	0.94321551	0.63296507	-0.9337127
CWeb14	web	-0.6438992	-0.9245921	0.57293578
CWeb15	web	-0.215267	0.43024096	0.92579064
CWeb16	web	0.51694005	-0.2887378	0.42013519
CWeb17	web	0.22773176	0.55957662	-0.7337377
CWeb18	web	-0.8533971	0.69561813	-0.0667383
CWeb19	web	-0.3580507	-0.2886934	0.79188505
CWeb2	web	-0.0233994	-0.2629234	-1.1063386
CWeb20	web	-0.4410371	-0.8731969	-0.8511331
CWeb21	web	0.41742031	-0.4446133	-0.5280906
CWeb22	web	0.59086298	0.59467761	-0.5520247

CWeb23	web	0.52160062	-0.2162654	-0.0994049
CWeb24	web	1.6324651	-0.4924903	0.79304469
CWeb25	web	2.13477246	1.20674857	0.30181182
CWeb26	web	0.54401944	-0.9884065	0.9755027
CWeb27	web	-0.1816735	0.35304948	0.24421658
CWeb28	web	0.84030407	-0.7482243	1.02096369
CWeb29	web	1.28700985	-3.16617	0.25676334
CWeb3	web	-0.2931489	-0.0392324	0.22251184
CWeb30	web	-0.5157926	-0.9101019	-0.9436429
CWeb31	web	0.1659983	-0.9594501	-0.1150891
CWeb32	web	0.25188591	-0.5301654	0.80028389
CWeb33	web	-0.4171118	-0.4611251	0.64064432
CWeb34	web	2.2324714	-1.7215605	0.70511912
CWeb35	web	-0.8447137	-0.6409474	-0.0415088
CWeb36	web	0.15035395	0.34505398	0.49242079
CWeb37	web	-0.5257363	-0.9523223	1.08313435
CWeb38	web	0.3405873	-1.2201388	1.56203738
CWeb39	web	1.12442262	-0.1348022	0.45095365
CWeb4	web	-0.504035	-0.1819094	-1.3812157
CWeb40	web	-1.1220253	0.44033403	0.2259923
CWeb41	web	1.31048644	0.07389923	0.84920565
CWeb42	web	0.76955421	0.36459757	0.4574294
CWeb43	web	-0.9705456	-0.6134064	-0.5580313
CWeb44	web	-0.3112827	-0.4715998	0.48640167
CWeb45	web	-0.0193825	-0.3236688	0.98917125
CWeb46	web	-0.4851959	-1.535549	0.81441101
CWeb47	web	1.29354281	-0.4597685	0.2955831
CWeb48	web	1.49562454	-0.8167967	1.91966579
CWeb49	web	-0.1128948	-0.1848137	1.39429877
CWeb5	web	1.24828234	0.108736	0.56480269
CWeb50	web	0.25976781	-0.2183476	1.28315125

CWeb51	web	-0.2889544	-0.2780171	0.66461661
CWeb52	web	0.47045572	-0.31222	-0.1897233
CWeb54	web	0.22456665	-0.2503171	0.2994116
CWeb55	web	0.51801988	-0.8795186	1.14402478
CWeb57	web	0.24222152	-1.1403129	0.14692356
CWeb58	web	-0.6580712	-0.7040234	-0.4664145
CWeb59	web	1.70353332	-0.9494408	2.01996158
CWeb6	web	2.13703602	0.08002702	0.02732427
CWeb60	web	-0.1161741	0.31394738	-1.82974
CWeb61	web	0.61092743	-1.1067439	-0.0516687
CWeb62	web	0.93515429	-0.6312449	0.11976665
CWeb63	web	0.35957509	-0.2077218	0.55555259
CWeb7	web	1.41489294	1.02917418	-0.1339384
CWeb8	web	-0.4650579	-0.6049597	-0.042836
CWeb9	web	-0.0651475	-0.1154983	0.91074618