Series Preface

The *Elements in Forensic Linguistics* series from Cambridge University Press publishes across four main topic areas (1) investigative and forensic text analysis; (2) the study of spoken linguistic practices in legal contexts; (3) the linguistic analysis of written legal texts; (4) explorations of the origins, development and scope of the field in various countries and regions. *The Language of Fake News* by Jack Grieve and Helena Woodfield is situated in the first of these and examines whether there are observable linguistic differences between fake news and genuine news articles.

Jack Grieve is best known for bringing quantitative and corpus methods to a variety of linguistic questions such as dialectology, language change, and authorship analysis methods. His quantitative work always brings linguistic insights and understanding to the fore and here with Helena Woodfield, whose principal area of research is fake news, they together bring this approach to the natural experiment provided by the Jayson Blair episode at *The New York Times.*

Jayson Blair was accused of and admitted falsifying a large number of news stories at *The New York Times,* and a subsequent inquiry by the paper identified the bad, and by implication, the good stories for the relevant period of his employment. For Grieve and Woodfield this creates parallel corpora ripe for exploration. Their principal insight is that as fake news and real news have distinctive communicative functions, respectively to deceive and to inform, the language used to carry these functions will also differ. In this Element they set out to identify and describe those differences. The implication of this approach is that linguistic analysis, independently from fact-checking approaches, can make an important contribution to fake news detection.

This sets up a new research agenda for linguistic fake news detection, which can be further explored, perhaps in future Cambridge Elements.

Tim Grant
Series Editor

1 Introduction

1.1 The Problem of Fake News

There is no simple definition of fake news. The term can be used to refer to any news that is suspected to be inaccurate, biassed, misleading, or fabricated. This includes news originating from across the news media landscape, from anonymous blogs to mainstream newspapers. The term is often used by the public, politicians, and the news media to attack news, journalists, and news outlets deemed to be

problematic. It is even common for allegations of fake news made by one outlet to be labelled as fake news by another. During the 2016 American presidential election, Hillary Clinton, the Democrats, and the mainstream press claimed that fake news from social media accounts, right-wing news outlets, and foreign governments was propelling Donald Trump to victory, while Trump, the Republicans, and the right-wing press claimed that Clinton, the Democrats, and the mainstream press were spreading fake news about these and other scandals to undermine Trump's campaign (Allcott & Gentzkow 2017). Fake news became the focus of the news, with news organisations arguing over whose news was faker.

Given this situation, how can the public judge what news is real and what news is fake? We cannot trust the news media to lead public inquiry into its own practices, nor can we trust the government or industry to monitor the news media, as they are most often the subject of the news whose validity is being debated. Academic research on fake news is therefore especially important, but it is also difficult to conduct (Lazer et al. 2018). Researchers must define fake news in a specific and meaningful way and then apply this definition to identify instances of real and fake news for analysis. This is a challenging task. Any piece of news communicates a wide range of information, some of which can be true, some of which can be false, and all of which can be an opinion. Often the only way to verify if news is real or fake is to conduct additional independent investigation into the events being covered. Crucially, even if fake news is defined precisely and in a way that is acceptable to most people, researchers must still label individual pieces of news as fake that a substantial proportion of the public believe are real. The study of fake news therefore quickly becomes politicised, further eroding public confidence, and encouraging researchers to define fake news in such a way that data can be collected easily and uncontroversially, often moving research further away from the central problem of fake news.

To understand the central problem of fake news, it is important to consider the history of fake news. Although most current research focuses on the very recent phenomenon of online fake news, reviewing the history of deception in the news media can help researchers understand what communicative events are considered fake news, how these different forms of fake news are related, and which types of fake news should be of greatest scholarly concern. The history of fake news can also point us to specific cases for further analysis, depoliticising the study of fake news by allowing researchers to focus on news coverage of events that are of less immediate consequence.

The history of fake news is almost as old as the history of news itself. In Europe, the precursor to the modern newspaper were the *avvisi*, handwritten political newsletters from Italy that circulated across the continent during the sixteenth and seventeenth centuries (Infelise 2002). Unlike personal letters, the

avvisi were intended to report general information and to be widely read. Unsurprisingly, we can find reports almost immediately of authorities questioning the veracity of the information being presented and the motives of their authors, who were generally anonymous. For example, in 1570, Pope Pius V executed one suspected author, Niccolo Franco, for defaming the church. Alternatively, the Italian scholar Girolamo Frachetta considered whether the *avvisi* could be used in wartime to spread false information to the enemy in his 1624 political treatise *A Seminar on the Governance of the State and of War* (Infelise 2002).

Indeed, there are many cases of fake news being used to mislead foreign populations and governments, exactly as Frachetta suggested. Many historical examples come from the Cold War, especially the Soviet use of *dezinformatsiya*, the purposeful spread of false information, which was often spread via the foreign press (Cull et al. 2017). The word *disinformation* only entered the English language in the 1980s due to increased awareness of so-called *active measures*, a wide range of strategies used by the USSR for undermining foreign countries, including fake news. One of the most famous of these initiatives was 'Project Infektion', which involved the Soviets spreading rumours that the United States had engineered AIDS, initiated by a letter published in an obscure Indian newspaper in 1983, titled 'AIDS may invade India: Mystery disease caused by US experiments' (Boghardt 2009).

It is perhaps more common, however, for fake news during wartime to be directed at one's own citizens – to encourage support for war and to manage expectations. For example, during World War I, the Committee on Public Information was established in the United States to influence the media and shape popular opinion, especially as President Woodrow Wilson had campaigned on staying out of the war (Hollihan 1984). Similar strategies were used to promote the Vietnam and Iraq wars. Most notably, we now know that reporting on the presence of weapons of mass destruction in Iraq after 9/11 was fabricated to build support for the war, especially to help Tony Blair justify the United Kingdom joining the coalition (Robinson 2017).

Fake news is not new, but the nature of fake news has shifted in recent years due to the growth of digital communication and social media (Lazer et al. 2018). The Internet has changed the medium over which news is published and accessed by the public. Consequently, people now have access to a much wider range of news sources, which disseminate information continuously throughout the day, often from very specific perspectives, while social media provides a platform for people worldwide to share and discuss the news. One important effect of this new approach to the production and consumption of

news is that people can focus exclusively on the information they want to hear, leading to what has become known as the *echo chamber* (Del Vicario et al. 2016). The rise of blogging and social media has also given the opportunity to people from outside the mainstream news media to spread their own message, including potentially fake news.

This type of online fake news has been the focus of much concern in recent years, including in the lead up to the 2016 US Election. Perhaps the most notorious example was the 'Pizzagate' conspiracy theory, which went viral in 2016, after Wikileaks published the personal emails of John Podesta, Clinton's campaign manager (Kang 2016). Extremist websites and social media accounts reported that the emails contained coded messages related to a satanic paedophile ring involving high-ranking officials, which allegedly met at various locations, including the Comet Ping Pong pizzeria in Washington. Provoked by these reports, a man travelled to the nation's capital from North Carolina, shooting at the pizzeria with a semi-automatic rifle. The Covid pandemic also offers numerous examples of this type of online fake news (see van Der Linden et al. 2020). For example, social media has been used to spread fake news about alternative treatments for Covid that are potentially deadly, including ingesting bleach (World Health Organization 2020).

Although it seems reasonable to assume that the amount of fake news has increased in recent years, we should not assume that the effect of fake news has become worse. Most notably, reporting from the mainstream news media leading up to the second Iraq War, which predates the rise of social media, was arguably far more damaging than anything that has happened since. In some ways, social media has even made it more difficult for certain types of fake news to spread by increasing public scrutiny of the news media and by amplifying alternative perspectives. An important example is coverage of the murder of George Floyd in May 2020. This event was filmed by a teenager named Darnella Frazier, who was walking to the store for groceries. She posted the video on social media, giving rise to widespread public protest to police brutality towards African Americans – a topic often overlooked by mainstream news media, which can be considered an example of fake news by omission (Wenzel 2019). In recognition of the importance of this act, which was only made possible by the existence of these non-traditional platforms for sharing information, Frazier received a Pulitzer Prize in 2021.

Overall, the problem of fake news is long-standing, pervasive, and potentially of great consequence, even leading to war. The study of fake news is therefore of true societal importance. Fake news is also very diverse, driven by a wide range of specific political, social, economic, and individual factors. In addition, it is clear that fake news, at least in its most troubling instantiations, is not simply

characterised by inaccurate reporting: it is intentionally dishonest, designed to deceive as opposed to inform the public.

In this Element, we therefore adopt the view that fake news is most productively analysed as *deceptive news*, in contrast with most academic research on fake news, which focuses on *false news*. In other words, we define fake news based on the intent of the author: as opposed to real news, whose primary goal is to inform readers about new and important information that the journalist believes to be true, the goal of fake news is to deceive the public, to make them believe information that the journalist believes to be false. This approach not only forces us to concentrate on the most problematic forms of fake news, but, as we argue, it provides a more meaningful basis for the analysis of the *language* of fake news, which is the subject of this Element.

1.2 Fake News and Linguistics

Understanding the language of fake news is key to understanding the problem of fake news because most cases of fake news *are* language. Fake news can involve pictures and other media, but usually an instance of fake news consists primarily of a *news text* – an article in a newspaper, a report on the radio, a post on social media, an interview on television. The news text is the basic communicative unit of journalism and consequently the basic unit of analysis in most research on fake news. The main questions we pursue in this Element are therefore how can the language of fake news be analysed in a meaningful way? How can we describe linguistic differences between news texts that are real and news texts that are fake? And how can we understand why this variation exists?

Crucially, however, we should not assume that the language of real and fake news differs systematically. There has been considerable research in natural language processing (Oshikawa et al. 2020) where machine learning models are trained to automatically distinguish between real and fake news based on patterns of language use, often achieving relatively high levels of accuracy. This may seem like evidence of variation between the language of real and fake news, but it is important to consider these results with care, especially as this research prioritises the maximisation of classification accuracy over the explanation of patterns of language use. There are two basic reasons for caution. First, the data upon which these systems are trained and evaluated may not isolate variation between real and fake news, especially given the inherent challenges associated with defining these terms and identifying cases of each. Second, these systems often focus more on variation in language *content* than variation in language *structure*: the identification of *topical* trends that tend to distinguish

between real and fake news is different from the identification of *stylistic* variation in the language of real and fake news regardless of topic. In other words, research in natural language processing focuses on the *language* of fake news, but it does not necessarily focus on the *linguistics* of fake news.

Although linguistic perspectives on fake news are limited, fake news is fundamentally a linguistic phenomenon, and its analysis should therefore be grounded in linguistic theory. To address this basic limitation with fake news research, we propose a framework for the linguistic analysis of fake news in this Element. Our framework is based on functional theories of language use, drawing especially on research on register variation (Biber & Conrad 2019), which has repeatedly demonstrated that differences in communicative purpose and context are reflected in linguistic structure. In addition, our framework is based on the distinction between misinformation and disinformation (Rubin 2019), which we believe is crucial for understanding what fake news is and why the language of real and fake news should differ. By bringing together these two perspectives for the first time, we provide a basis for the linguistic analysis of fake news – for collecting real and fake news texts, for comparing their grammatical structure, and for understanding why this structure varies depending on whether their author intends to inform or deceive.

To demonstrate how our framework can be used to better understand the language of fake news, in this Element, we focus on one especially famous episode drawn from the history of the news media. This case involves Jayson Blair, a young reporter at *The New York Times*, who published a series of fabricated news articles in the early 2000s (Hindman 2005). In addition to its notoriety, this case is especially well suited for the linguistic analysis of fake news for three reasons. First, there is a relatively large amount of real and fake news available from one author and from the same time period, which has been validated through an extensive investigation by *The New York Times* (Barry et al. 2003) and acknowledged by Blair himself (Blair 2004). This gives us a controlled context for the study of fake news, where we have substantial amounts of comparable and valid real and fake news data, allowing us to effectively isolate the effects of deception on the language of one journalist. Second, we know much about why Blair wrote fake news, including from his own account and the account of *The New York Times,* giving us a basis for explaining differences in language use that we observe. Third, this case is relatively uncontroversial, as it is old enough that everyone can agree that the articles in question were faked, regardless of their political outlook – an important factor that has often limited the societal impact of fake news research. In addition, the case reminds us that fake news can be found across the news media, including in one of the most respected newspapers in the world.

The analysis of the language of fake news grounded in linguistic theories and methods also opens up the possibility for a wide range of applications. Although our goal is not to develop systems for fake news detection, the most obvious application of our research is to support the language-based identification of fake news. Most notably, this includes considerable current research in natural language processing concerned with developing systems for automatically classifying real and fake news at scale via supervised machine learning (Oshikawa et al. 2020). As noted above, these systems can achieve good results, but they are not designed to explain why the language of real and fake news differs, and they appear to focus more on the content of fake news than its linguistic structure (Castelo et al. 2019). Our framework is not intended to supplant these types of systems, but it can offer an explanation for why they work, or why they might appear to work, which is necessary to justify the real-world application of such tools. Furthermore, the identification of a principled set of linguistic features for the analysis of real and fake news can be used to enhance existing machine learning systems, which tend to be based on relatively superficial feature sets like the use of individual words and word sequences. These types of insights can be especially useful to improve performance on more challenging cases, which also seem likely to be the most important cases of fake news. In addition, our framework can directly inform how fake news corpora should be compiled in a principled manner for the robust training and evaluation of fake news detection systems, which is a major limitation in much current research on fake news (Asr & Taboada 2018).

The framework we propose is also of direct value to the detailed discursive analysis of individual cases of fake news of sufficient importance to warrant close attention. For example, in a legal context, empirical analysis presented as evidence in court is often required to be based on accepted scientific theory (Allen 1993). Until now, however, there has been no clear explanation for why the linguistic structure of real and fake news should be expected to differ systematically. Our framework also potentially provides a basis for extending discourse analytic methods for deception detection more generally in forensic linguistics, which is relevant across a wide range of areas, including for police interviews (Picornell 2013). In addition, our framework can be of value for supporting work in *investigative linguistics*, which is an emerging field of applied linguistics that focuses on the application of methods for the study of language use to make sense of real-world issues currently in the news (Grieve & Woodfield 2020).

More generally, understanding the language of fake news, and how it differs from the language of real news, is important for understanding the language of the news media, and, through this language, the biases and ideologies that

underlie any act of journalism. The current fake news crisis reflects a growing and general distrust of the news media that cannot be rectified simply by developing systems for automatically detecting real and fake news with a reasonable degree of accuracy. An article that obliquely expresses the editorial view of a newspaper in a context that appears to be purely informational is not necessarily fake, but it has real societal consequences. Being able to recognise the motivations of journalists and news outlets through the analysis of their language is an important part of reading the news intelligently and holding the news media accountable. Studying the discourse of fake news is therefore part of the greater enterprise of understanding the expression of information, opinion, and prejudice in the news media – understanding how the language of the news media shapes the world around us and our perceptions of it. We therefore hope that our framework will also be valuable for the critical analysis of the news media (van Dijk 1983).

Finally, our framework and its application can also help us better understand the psychology of fake news (Pennycook & Rand 2019, 2021). Why do people create, share, and believe fake news? These are basic questions whose answers are central to understanding the phenomenon of fake news in the modern world. Most notably, as we demonstrate through our analysis of the case of Jayson Blair and *The New York Times*, variation in the linguistic structure of fake news reflects the specific communicative goals of authors who consciously write fake news and the production circumstances in which fake news is produced. Appreciating the linguistic structure of fake news can also potentially help us understand why some fake news is more likely to be believed and to be shared, which may be especially important for combating the spread of fake news online.

1.3 Overview

Fake news is a long-standing problem, but it is receiving unprecedented attention today due to the rise of online news and social media, as well as growing distrust of the mainstream news media. Although fake news most commonly involves news texts, the study of the language of fake news has been limited, with researchers focusing more on the automated classification of true and false news than on explaining why the structure of real and fake news differs. It is therefore crucial to extend our understanding of the language of fake news, especially through the detailed analysis of real and fake news texts collected in a principled manner and grounded in linguistic theory.

Given this background, the goal of this Element is threefold. First, we introduce a new linguistic framework for the analysis of the language of fake

news, focusing on understanding how the linguistic structure of fake news differs from real news, drawing especially on the distinction between misinformation and disinformation and the concept of register variation. Second, based on this framework, we conduct a detailed analysis of the language of fake news in the famous case of Jayson Blair and *The New York Times* to identify and explain systematic differences in the grammatical structure of his real and fake news.[1] Third, we consider how our results can help address the problem of fake news, including by informing research in natural language processing and psychology.

The remainder of this Element is organised as follows. In Section 2, we present a critical review of research on the language of fake news, before presenting our theoretical framework for the linguistic analysis of fake news, which directly addresses limitations with previous research. In Section 3, we review the case of Jayson Blair, including the background, the scandal, the investigation, and the aftermath. In Section 4, we describe the corpus of Jayson Blair's writings that we collected, which is the basis of this study. In Section 5, we present our main linguistic analysis, discussing a range of grammatical features that vary across Blair's real and fake news. We find that Blair's fake news is written in a less dense style than his real news and with less conviction. We then offer explanations for these findings based on specific factors that led Blair to write fake news. Finally, in Section 6, we consider the implications of our research for our understanding of fake news more generally.

2 Analysing the Language of Fake News

The language of fake news has received considerable attention in recent years, especially in natural language processing, where the focus has been on the development of machine learning systems for the automatic classification of real and fake news based on language content. In this section, we critically review recent research on the language of fake news, arguing that it has been limited by the definition of fake news as false news and the lack of control for other sources of linguistic variation. To address these issues, we propose a framework for the linguistic analysis of fake news that is grounded in theories of disinformation and register variation. This framework provides a basis for describing the linguistic differences between real and fake news and explaining why these differences exist.

[1] This study was approved by the University of Birmingham's ethics review panel. All data analysed are published and publicly available, including via in the online archives of *The New York Times*.

2.1 Defining Fake News

The first major challenge in the study of the language of fake news is to define fake news in such a way that instances of fake news texts can be identified and collected (Tandoc et al. 2018; Asr & Taboada 2019). There is, however, no simple or standard definition of fake news, which is better understood as the product of a range of practices that are related to the validity of information being shared by the news media. Researchers must therefore define the specific form of fake news they are interested in studying. Any coherent definition of fake news can be the starting point for meaningful empirical research, but researchers naturally tend to focus on certain types of fake news, depending both on the perceived societal importance of that type of fake news and the feasibility of collecting news texts of that type in a reliable and efficient manner – considerations that are often at odds with each other.

The vast majority of research on the language of fake news has been conducted in natural language processing and has focused on the development of tools for automatically distinguishing real and fake news (e.g. Conroy et al. 2015; Rubin et al. 2015; Shu et al. 2017; Asr & Taboada 2018; Bondielli & Marcelloni 2019; Oshikawa et al. 2020; Zhou & Zafarani 2020). In general, this research defines fake news as *false news* – untrue information disseminated by the news media. This has also been the definition that has been adopted in the very limited amount of linguistic research on this topic in discourse analysis (e.g. Igwebuike & Chimuanya 2021). Crucially, this definition of fake news is based on the underlying truth of the information being conveyed: to study fake news from this perspective, comparable corpora of *true* and *false* news must be compiled. For example, to develop a machine learning system capable of distinguishing between true and false news requires that many true and false news texts be collected so that the system can be trained and tested on this dataset.

A major advantage of this *veracity-based approach* to fake news research is that it allows fake news to be collected with relative ease. Most commonly this involves drawing on the work of fact-checking organisations and mainstream news media organisations that identify fake news, including both instances of fake news and sources of fake news (Asr & Taboada 2018). This information is then used as a basis for compiling a corpus of fake news texts. These texts most commonly include passages from news articles (e.g. Vlachos & Riedel 2014; Wang 2017), social media posts (e.g. Shu et al. 2017, 2020; Wang 2017; Santia & Williams 2018), and complete news articles (e.g. Rashkin et al. 2017; Horne & Adali 2017; Santia & Williams 2018; Castelo et al. 2019; Lin et al. 2019; Bonet-Jover et al. 2021). Alternatively, some studies have used crowdsourcing