

**Borys Grinchenko Kyiv Metropolitan University
Faculty of Romance and Germanic Philology
English Language and Communication Department**

**KATE MIDDLETON'S AND MEGHAN MARKLE'S DISCURSIVE
PORTRAITS IN THE MEDIA**

MA Paper
Oleksandra Katasonova
Group MLAm-1-23-2.0d
Research supervisor
Professor G.Yu. Tsapro

Approved for defense
at the Linguistics and
Translation Department
25, May, 2023, Minutes 11

_____ Andrii KOZACHUK

*Цим підписом засвідчує, що
копія на факс, рукопис та електронний
документ є ідентичні!
30.05.2025*

Kyiv, 2025

Contents	
ABSTRACT	3
Анотація	5
INTRODUCTION	7
1.1. Media Discourse as a Tool for Shaping Public Opinion.	9
1.2. The Role of Multimodality in Shaping Media Discourse.....	11
1.3. Theoretical Foundations of Critical Discourse Analysis in the Context of Media Bias.	13
1.4. Interpreting Communicative Intent: Pragmatic Tools for Identifying Media Bias. 17	
RESEARCH METHODS	19
2.1.Theoretical concepts of corpus linguistics.....	19
2.2.Corpus analysis tool applied in the research.....	21
2.3.Utilization of Sketch Engine and Stages of Corpus-Based Research.....	24
2.4.Multimodal Analysis Procedure.	28
RESULTS	32
3.1.Overview of the Analytical Framework.	32
3.2.Critical Discourse Analysis (CDA) Results, integrating the tools of corpus analysis	33
3.3.Corpus and pragmatic analysis results.....	42
1.5. Multimodal discourse analysis.....	53
CONCLUSIONS.....	62
REFERENCES.....	65

ABSTRACT

The thesis explores the linguistic construction of media bias in the representation of Catherine, Princess of Wales, and Meghan, Duchess of Sussex, using a multi-method approach grounded in Critical Discourse Analysis (CDA), Speech Act Theory, corpus linguistics, and Multimodal Discourse Analysis (MDA). Media bias is treated as a discursive phenomenon encoded in lexical selection, pragmatic strategies, and multimodal resources (mostly images), all of which contribute to shaping ideological narratives and influencing public perception.

The empirical material consists of 170 articles and 200 visual elements (images, headlines, and layouts) drawn from major British and American media outlets, including the *BBC*, *The Guardian*, *The New York Times*, *USA Today*, *People*, *Vogue*, and *The Daily Mail*. The texts include the articles covering major events of the life of each of the personalities within the royal family: engagement, wedding, and public appearances. The articles were divided evenly between the two personalities and equal corpora were constructed for each figure to ensure methodological reliability in the subsequent analysis.

CDA is applied via Fairclough's three-dimensional model, mostly focusing on textual features and lexical units employed to construct the biased portraits of the personalities. Pragmatic analysis follows Searle's classification of speech acts, identifying communicative functions (assertives, directives, expressives, etc.) to reveal bias at the level of linguistic intention. CDA was performed on 70 articles covering the royal weddings. Articles about Meghan Markle frequently used emotionally charged vocabulary and metaphors, including *fairy-tale wedding*, *loverbirds*, and *melt our hearts*. The word "Markle" appeared in reference to the Duchess of Sussex, unlike articles about Catherine, which more often used formal titles. Royal protocol elements (e.g., *titles*, *carriages*, *tiaras*) were emphasized in both corpora. Meghan's background as a *divorcee* and former "*Suits*" actress was recurrently mentioned. In contrast, articles about Catherine featured more formal vocabulary and highlighted institutional protocol. Significant attention was paid to her dress, accessories, and makeup, with detailed references to designer names and historical tiaras. Audience size was frequently quantified (e.g., "*watched by an estimated one billion people*").

Corpus analysis was conducted using Sketch Engine. Tools such as wordlists, keyword analysis, and thesauri were employed to identify frequency patterns, semantic salience, and collocational behavior of nouns, adjectives, and verbs. These findings back up the pragmatic and discursive analysis and used individually as another method for identifying media bias. Corpus analysis showed that in articles about Meghan, her name and surname ranked higher in frequency than Prince Harry's. Wedding-related words like *newlyweds*, *nuptials*, and *bridesmaid* appeared frequently, as did celebrity guest names (e.g., *Winfrey*, *Clooney*). The wordlist revealed frequent mentions of royal terms (*duke*, *duchess*, *queen*) and references to Diana and mother. In Catherine's

corpus, keywords included *tiara*, *palace*, and *throne are present*. There is a higher frequency of regal lexical items. Articles were centered around the bride's attire, including specific materials, designers, and accessories.

Multimodal Discourse Analysis, based on Kress and van Leeuwen's framework, examines visual modes, such as images and layout to reveal how multimodal strategies support or reinforce bias in media representation. Articles about Meghan included more visual images, especially informal ones, showcasing her past and some photos made by paparazzi, at the same time, those images for Catherine were strictly formal.

Keywords: *media discourse, media bias, Critical Discourse Analysis, Speech Act Theory, corpus analysis, Sketch Engine, multimodal discourse analysis, verbal portrait, discourse strategies, the British royal family, Catherine Middleton, Meghan Markle.*

Анотація

У роботі досліджується лінгвістичне конструювання упередженості ЗМІ у висвітленні образів Кетрін, принцеси Уельської, та Меган, герцогині Сассекської в медіа. Для цього використано багатометодичний підхід, що поєднує критичний дискурсивний аналіз (КДА), теорію мовленнєвих актів, корпусну лінгвістику та мультимодальний дискурсивний аналіз (МДА). Упередженість ЗМІ розглядається як дискурсивне явище, реалізоване через добір лексики, прагматичні стратегії та мультимодальні ресурси (переважно зображення), які формують ідеологічні наративи та впливають на громадське сприйняття.

Емпіричний матеріал складається зі 170 статей і 200 візуальних елементів (зображення, заголовки), відібраних із провідних британських та американських медіа, таких як *BBC*, *The Guardian*, *The New York Times*, *USA Today*, *People*, *Vogue* та *The Daily Mail*. Тексти охоплюють головні події з життя кожної з персон у королівській родині: заручини, весілля та публічні виступи. Статті були рівномірно розподілені між двома фігурами, а корпуси сформовано у відповідності до принципу рівності для забезпечення методологічної надійності аналізу.

КДА було застосовано за тривимірною моделлю Нормана Ферклау, зосередившись переважно на текстуальних особливостях та лексичних одиницях, що використовуються для побудови упередженого образу особистостей. Прагматичний аналіз ґрунтувався на класифікації мовленнєвих актів Сьорля (асертиви, директиви, експресиви тощо) з метою виявлення упередженості на рівні мовленнєвого наміру. Аналіз охопив 70 статей, присвячених королівським весіллям. У текстах про Меган Маркл часто використовувалася емоційно забарвлена лексика та метафори, зокрема “казкове весілля” (*fairy-tale wedding*), “закохані пташки” (*loverbirds*), “розтопили наші серця” (“*melt our hearts*”). До герцогині Сассекської нерідко зверталися як до “Маркл”, на відміну від Кетрін, яка частіше згадувалася з використанням офіційного титулу. В обох корпусах підкреслювалися елементи королівського протоколу (титули, карети, тіари). Біографія Меган як “розлученої” та “колишньої акторки серіалу *Suits*” згадувалась неодноразово. Водночас статті про Кетрін відзначалися більш формальною лексикою й наголосом на протокол. Значну увагу приділено опису сукні, аксесуарів і макіяжу, часто з посиланнями на дизайнерів та історичні тіари. Розмір аудиторії зазвичай кількісно зазначався (наприклад: “*переглядали близько одного мільярда людей*”).

Корпусний аналіз здійснювався за допомогою *Sketch Engine*. Використовувалися інструменти створення списків слів, аналізу ключових слів і тезаурусів для виявлення частотності, семантичної значущості та колокаційної поведінки іменників, прикметників та дієслів. Ці результати підтверджують прагматичний та дискурсивний аналіз і виступають окремим методом виявлення

медіабайасу. У статтях про Меган її ім'я та прізвище зустрічаються частіше, ніж ім'я принца Гаррі. Часто вживалися весільні слова (*“молодята”, “шлюбна церемонія”, “подружка нареченої”*), а також імена відомих гостей (*“Вінфрі”, “Клуні”*). У списках слів були поширеними слова *“герцог”, “герцогиня”, “королева”, “Діана”, “мати”*. У корпусі Кетрін ключовими були слова *“тіара”, “палац”, “трон”*. Частота лексики, пов'язаної з королівським статусом і інституційною роллю, була вищою. Більше уваги приділялося опису вбрання, матеріалів та дизайнерських деталей.

Мультимодальний дискурс-аналіз, заснований на моделі Кресса та ван Левена, розглядає візуальні засоби (зображення, макети) для виявлення стратегій, що підсилюють медіаупередження. У статтях про Меган було більше візуального контенту, зокрема неформальних фото з її минулого та знімків папараці, тоді як візуальний супровід Кетрін мав суто офіційний характер.

Ключові слова: *медіадискурс, медіабайас, критичний дискурс-аналіз, теорія мовленнєвих актів, корпусний аналіз, Sketch Engine, мультимодальний дискурс-аналіз, вербальний портрет, дискурсивні стратегії, британська королівська родина, Кетрін Міддлтон, Меган Маркл.*

INTRODUCTION

The main **objectives** of this paper are to outline theoretical studies of media discourse, corpus linguistics and speech theory acts. To develop a complex methodology for the study of Catherine, the Princess of Wales and Meghan, the Duchess of Sussex discursive portraits in the media and study the types of linguistic analysis which can be employed to identify media bias. To single out passages, containing Catherine`s and Meghan`s descriptions in the media. To characterize discursive portraits understudy in the media. To create 2 corpora and process them with Sketch Engine to identify key words in both corpora as well as verbs for further pragmatic analysis and thesaurus to back up the results obtained with the help of CDA analysis. To outline specific features of the Princess of Wales and the Duchess of Sussex discursive portraits in the media.

The object of the research is media discourse presenting Kate Middleton's and Meghan Markle's engagements, weddings, first royal engagements, and relationship announcements.

The subject of the research is the linguistic means used to construct verbal portrayals of Kate Middleton's and Meghan Markle's engagements, weddings, first royal engagements, and relationship announcements.

The **aim** of the research is to analyze the linguistic means used in media discourse to construct verbal portrayals of Kate Middleton`s and Meghan Markle`s engagements, weddings, first royal engagements, and relationship announcements.

Research method. Achieving the aim and objectives of the research requires a complex methodology, which includes: **Critical Discourse Analysis (CDA)** – to uncover the discursive portraits of the two personalities and reveal media bias; **corpus analysis** – to identify keywords, wordlists for specific parts of speech, and thesaurus entries in the two corpora; **pragmatic analysis** – to classify the verbs used in constructing the two discursive portraits; and **multimodal discourse analysis** – to examine how these portraits are supported by visual elements.

The **research material** includes the articles published by the most influential media sources in the UK and the USA: *BBC, the New York Times, ABS News, Vogue, Cosmopolitan, Birdes, Daily Mail, the Guardian, etc.* 170 articles covering major events in the lives of Catherine and Meghan, including 200 visual images.

The **practical significance** of the obtained results can be used in the course of corpus linguistics, in the course of English language discourse, in the section "pragmatic types of statements" of the theoretical grammar of English, and practical classes in English.

The findings of the research were presented at the following academic conferences:

1. “*Philological Gems*”, Borys Grinchenko Kyiv Metropolitan University (Katasonova O., May 15, 2024. *Deconstructing Articles on Meghan Markle’s Engagement: A Structural Analysis*);
2. “*New Trends in Translation Studies, Philology, and Language Didactics in the Context of Globalization*” (Katasonova O., November 7, 2024. *Meghan Markle’s Wedding in Media Discourse*).

Selected results were also published in the scholarly journal *Mundus Philologiae*: Katasonova O. (2025). *Kate Middleton’s and Meghan Markle’s Image in the Media: Discursive and Corpus Analysis*. *Mundus Philologiae*, No. 4.

THEORETICAL BACKGROUND

1.1. Media Discourse as a Tool for Shaping Public Opinion.

The scholar Anne O’Keeffe defines discourse as the study of how the language is used or the study of language beyond the sentence (implying that only by combining the sentences, meaning and coherence may occur) (O’Keeffe, 2006). As this notion is rather broad, scholars have managed to point to certain types of discourse based on the spheres of communication, modality and function (Carvalho, 2008). Teun A. van Dijk and Ruth Wodak tend to link discourse to specific social context. They identified personal discourse, public discourse and institutional discourse. (van Dijk, 1983) According to this classification, media discourse may be referred to both institutional and public discourse. As institutional discourse involves the verbal exchange between the representatives of work-related institutions, having an endeavor to achieve a particular professional goal (Freed, 2015). Whereas public discourse refers to the communication addressed to the general public (Speeches, media articles, advertisements, social media posts) (Schollon, 2007).

Media discourse can combine various types of discourse, depending on the aim of the author and the style of the article. The definition of this notion is broad. Media discourse encompasses interactions conveyed in written or spoken form and disseminated to reach audiences who are not physically present (Janks, 1997). Moreover, it constructs representations of reality through both broadcast and print media (O’Keeffe, 2006). Media discourse may also be observed through the lens of beliefs underlying the media content. In this case one may characterize it by the terms of objectivity along with ideological bias (Bonnot, 2024).

Different scholars identify various functions of media discourse. According to Wodak and Busch (2004), the informative function of language is fundamental to media texts, as they reflect topics and aspects of extra-linguistic reality, such as real-life events and narratives. Primarily, media texts and discourses are created to inform the public by conveying different kinds of messages. Van Dijk (1998) highlights mainly the persuasive function of such a discourse, mentioning that it shapes ideologies and influences what the public sees as normal (Van Dijk, 1998). Norman Fairclough (1995) emphasizes its ideological function, claiming that media discourse is capable of constructing versions of reality (Fairclough, 1995). Apart from ideological and persuasive functions, media also performs entertaining, interpretive, socializing ones (Tavadze, Diasamidze, Katamadze, Davitadze, 2024). Yet the focus of prominent scholars of the first two functions, may indicate the extent of influence media has on public minds.

Dividing media discourse into types can be a challenge, as every scholar has a different approach to this process. One way involves the division based on the type of publication (Tavadze, Diasamidze, Katamadze, Davitadze, 2024). As content of

different media sources may be distributed in multiple ways, including printed editions, online blogs, radio, and television broadcasts, social media content (Tavadze, Diasamidze, Katamadze, Davitadze, 2024). Other researchers prefer to classify media discourse by focusing on its impact on society or by analyzing the underlying principles and values held by the authors. Much of this analysis is region-specific and heavily influenced by the customs and traditions of both media producers and their audiences (Bonnot. 2024).

Another classification for media discourse is its division by modality into oral (interviews, podcasts, TV talk shows) and written one (articles, news reports, blogs) (O`Keeffe, 2006).

Although written media is often perceived as neutral, it frequently uses emotional language capable of shaping public opinion (O`Keeffe, 2006), particularly during significant events. Nevertheless, it is important to note that the relationship between the media and the public is far from linear. It can be better described as "mutually constitutive," which implies that the media tends to reinforce views that are already prevalent within society (O`Keeffe, 2006).

Scholars Biber and Finegan (1994) studied the notion of language used in the newspapers (written media source) in depth and identified it as one of the main registers present in the English language. Swales, 1990, defines the language used in newspapers as a separate genre. Oral media is mostly studied by means of Conversational analysis. CA has been the dominant approach to studying spoken media discourse, particularly in radio and television contexts. Originating from a branch of sociology rather than linguistics, CA has shaped the way researchers examine spoken interactions. The seminal works of Sacks, Schegloff, and Jefferson have significantly contributed to and influenced the development of research on spoken media discourse (Schegloff, 1968; Sacks, 1974; Schegloff, 1977; Sacks, 1992).

The main peculiarity of this type of media discourse is that it is researched with the emphasis on social interaction rather than language. Scholars pay attention to how the interlocutors move to the mutual goal, shape their conversation and negotiate (Carpentier, De Cleen, Van Brussel, 2019).

The most recent addition to media discourse is the appearance of digital or social media discourse. With the rise of social networks, such as Instagram, Snapchat, Facebook, long with various messengers, media discourse has expanded beyond traditional television and print media (Mammadov, Tomaszczyk, 2022). This development has allowed individuals to actively comment on and engage with online content, thereby becoming an integral part of media discourse themselves and either reinforcing or challenging the images constructed by magazines, television channels, or social media platforms (Wodak, 2019).

Even before the mass appearance and popularity of social networks, Kevin Pearce would define mass media as the process by which a person, group of people, or large organization creates a message and transmits it through some type of anonymous

audience of various sizes and types (Pearce, 2009). And with the growing popularity of social media, the process of conveying and receiving such messages has become available to the vast majority of people, contributing to spreading biased information in public.

Susan Herring had largely contributed to the study of social media discourse, as she developed the approach known as Computer-Mediated Discourse Analysis (CMDA). Herring emphasizes that online communication, including that found on social media platforms blurs the line between oral and written discourse. She presents it as a hybrid form that combines features of both modes (Herring, 2013). Apart from that, she defines the identity construction and power dynamics as two pivotal elements of social media interactions. According to Herring, the public image is shaped through language and multimodal resources such as emojis, memes, and images (Herring, 2013). This new development has contributed to making the media even more biased than ever before.

Moreover, language is not the only means the media applies to disguise the prejudice for objectivity. One of the key tendencies of media discourse is the presence of the notion of framing in most of the articles. Framing refers to the tendency of media resources to select and highlight certain aspects of the event, character trait of the personality, quote from the speech, while omitting others. This contributes to shaping the way the audience interprets various events and perceives public personalities (Entman, 1993). Through framing, the media can have a subtle influence on the perception of the situation and conceal the prejudice by doing so.

That leads to an integral notion of media discourse. Bias in media discourse is reflected through the preference for particular perspectives, individuals, or ideologies. All of the mentioned factors often lead to unequal or unbalanced representations (Fairclough, 1995; van Dijk, 1998). Even though bias is not always an intentional step, it undoubtedly influences the public perception of certain events.

It is its subjective, emotional nature that has made media discourse so attractive for the scholars throughout the years. By studying the issue under scrutiny, scholars have eventually concluded that the bias may be skillfully concealed by means of using the “objective language” that yet just purports to be neutral (White, 1997). That leads to the point that with time and the development of journalism, identifying and uncovering bias in the media discourse has become more difficult.

1.2. The Role of Multimodality in Shaping Media Discourse.

The notion of multimodal discourse already was mentioned in this article. As multimodality is a part of social media discourse and media discourse in general (Herring, 2013).

Multimodal discourse analysis examines how meaning is constructed through the combination of various modes of communication, rather than relying solely on language (Jones, 2012). Multimodality refers to the analytical strategies used to examine different semiotic modes, including visual imagery, sound, and various forms

of digital technology utilized in media (Luca, 2020). In conducting multimodal analysis of media resources, it is essential to consider multiple elements of an article, such as "layout, color, writing, imagery, and typography."

Certain strategies are required to be applied when interpreting how meaning is constructed across different communicative modes. According to Kress and van Leeuwen (1996), the first step is to identify the semiotic modes involved. These may include text, image, layout, color, and sound. Each mode should then be analyzed separately to understand its individual contribution to meaning-making. Apart from this, it is pivotal to examine the interaction between modes to see the connection between them and the role of each separately (Kress & van Leeuwen, 2001). Jewitt (2009) emphasizes the importance of viewing modes as semiotic resources, where choices of visual imagery, font style, and color are socially and culturally shaped. Furthermore, multimodal analysis requires viewing the discourse in the framework of its social context, as all the modes are to some extent influenced by the propaganda, public opinion, and the expectations of the audience (Chouliaraki & Fairclough, 1999). Thus, by performing multimodal analysis one may detect the concealed prejudice and the implication behind each mode.

It's worth mentioning that multimodal discourse may as well be observed in oral media discourse. As multimodality stands for all sorts of discourse apart from the spoken one, it may include intonation, voice, gestures, facial expressions and even choice of clothing (van Leeuwen, 2015). Therefore, the multimodal analysis involves strategies from various research fields, and makes multimodality an interdisciplinary notion (Kress's, 2012).

Multimodal discourse demands an interdisciplinary approach, as it encompasses several fields of study. One of these fields is social semiotics. Social semiotics is a theoretical approach that studies how humans use various semiotic resources (that include language, images, gestures, sounds, and spatial arrangements). Social semiotics is meant to create and interpret meaning within social contexts. The framework was developed and popularized by Gunther Kress and Theo van Leeuwen. These scholars extend the traditional meaning of semiotics, highlighting the social and cultural aspects of communication (Tan, O'Halloran, Wignell, 2020).

One of the core principles of semiotics is the belief that meaning is always realized through multiple modes that interact with each other. According to Gunther Kress and Theo van Leeuwen, the prominent scholars in this realm, meaning is not derived from a single mode (e.g., text) but rather from the combination of multiple modes. For instance, a magazine advertisement tends to combine visuals, color schemes, text, and layout to convey one single message (Kress, van Leeuwen, 2001). Therefore, when conducting the media discourse analysis along with multimodal analysis, one shall be vigilant and pay attention to both the internal features of each mode and the broader ones. Thus, effective multimodal analysis is not limited to a single theoretical framework; it requires a successful combination of perspectives to

fully grasp how complex meanings are achieved across different semiotic resources (Kress, van Leeuwen, 2001).

Multimodality plays a significant role in shaping and reinforcing bias within media discourse. Kress and van Leeuwen (1996) argue that "no mode is innocent". By that, they imply that every choice of image, color, font, and layout contributes to the creation of a certain bias. By thoroughly selecting and combining various modes, media can subtly alter audience perceptions without even tapping into the means of verbal language. As Machin and Mayr (2012) explain, "visual communication is a powerful tool for ideological purposes, precisely because its meanings often appear natural and transparent." It means that prejudice is even harder to uncover when it's concealed with the help of multimodal means, the public seems not to realize the influence that every minute detail the article's design has on it, as even typographical emphasis can create implicit evaluations that words alone might not express.

Thus, multimodal analysis is essential for uncovering concealed messages and exposing the role of media discourse in contributing to the creation of public prejudice and the shift of the opinions.

1.3. Theoretical Foundations of Critical Discourse Analysis in the Context of Media Bias.

Media discourse analysis alongside multimodal discourse are closely linked to the concept of Critical Discourse Analysis (CDA). For many years, CDA has been used to examine a wide range of social issues, ideologies, and, importantly, power dynamics and dominance within specific contexts (Janks, 1997). As noted by M. Bitl, CDA has proven to be an effective tool for uncovering the underlying intentions of authors who convey biased viewpoints through their articles (O'Keeffe, 2006).

CDA is a type of discourse analysis that focuses on social problems, such as abuse and domination. The scholar Ruth Wodak describes it as an analysis "with an attitude" (Wodak, Meyer, 2015). It was dubbed so, as it addresses significant social issues, and helps to make the analytical approach more profound. CDA claims that language, power and ideology are interconnected, and that the way the world is portrayed in texts influences our perception and view of reality. (Stubbs, 1997). Scholars who work in the realm of CDA are sometimes referred to as dissidents, considering that one of the primary tasks of discourse is to expose and defy inequality, which makes CDA some kind of a social movement (van Dijk, 2015).

Between the mid-1960s and early 1970s, there was notable development in areas shared by emerging disciplines within the humanities and social sciences. In spite of the discrepancies in their academic roots, the diversity of methods and research focuses applied by the scholars, along with several newly established fields (including semiotics, pragmatics, psycholinguistics, sociolinguistics, and discourse studies) contributed to the emersion of a common interest in discourse: a field of linguistics that shared at least seven core dimensions (Wodak, Meyer, 2009).

CDA was pioneered by a group of scholars who aspired to bridge the gulf between linguistic analysis and critical social theory (Janks, 1997). Norman Fairclough (1995) is often considered one of its founders and most prominent figures. He dedicates his words to emphasizing the dynamic relationship between discourse, power, and ideology. His research established a foundation for exploring the ways language may contribute to the construction of social hierarchies and the expansion of inequality. Fairclough studies CDA primarily through the lens of sociolinguistics (Rafiah, 2018). His approach to CDA is entitled textually-oriented discourse analysis (TODA). This method was designed to mark the combination of philosophical with linguistic methodology (Rafiah, 2018). Some of its key directions involve the research into the connection between various linguistic features of the text, sociolinguistic speech genres, and broader sociological practices. It's Fairclough whose theory on CDA has remained prominent for the last decades (Janks, 1997).

Another outstanding personality in the realm of CDA is Teun A. van Dijk (1993). This scholar managed to expand the scope of CDA by incorporating cognitive and social dimensions into discourse studies. He is particularly attentive to the issues of social inequality and media discourse (Rafiah, 2018). The main peculiarity of this scholar's contribution is the ability to formulate the discourse elements in a way that makes them applicable (Eriyanto, 2001). This method was dubbed "social cognition". Van Dijk views discourse as three interconnected dimensions: text, social cognition, and social context. It is the interaction between these components that reflects the underlying direction or the hidden agenda (such as bias) of a given discourse (Bhatia, 2012). One of the key strengths of such an approach to discourse analysis lies in the ability to connect textual content with contextual factors by means of the social cognition of various discourse producers (van Dijk, 1988).

Ruth Wodak (2001) was a scholar who managed to enrich the realm of CDA by creating and applying Discourse-Historical Approach, which observes discourse within its broader sociopolitical and historical contexts. Such a method enables one to receive a deeper understanding of meaning-making over time. An essential principle of Critical Discourse Analysis (CDA) is the idea that "all discourses are historical and can therefore only be understood with reference to their context" (Wodak, 2009; Meyer, 2001). This suggests that discourse is intertwined both synchronically, with other ongoing communicative events, and diachronically, with events from the past. For Wodak, the concept of context is central to CDA, as it incorporates socio-psychological, political, and ideological elements, thus promoting an interdisciplinary approach to analysis (Amoussou, Allagbe, 2022).

Ultimately, Gunther Kress and Theo van Leeuwen (2001) introduced the concept of multimodality to CDA. This advancement highlighted the role of visual and other non-linguistic elements in communication (Wodak, Meyer, 2015). Together, these scholars have broadened CDA's theoretical and methodological horizons and manage

to turn it into a powerful tool for analyzing both textual and multimodal forms of discourse.

There are several key principles in CDA that primarily focus on the relationship between language, power, and society. Fairclough came up with a fundamental assumption of CDA, claiming that language is not neutral (Fairclough, 1995). Conversely, scholars see it as a powerful tool able to influence and shape perceptions, ideas, and common opinions within the society. It contributes to the construction and maintenance of social power relations, which can either reinforce or challenge existing social hierarchies (van Dijk, 2015).

Another core principle of CDA is that discourse reflects and reinforces ideologies (van Dijk, 1993). Language in this case is presented in a form of a vehicle, which enables transmission of dominant ideologies. It contributes to shaping people's beliefs, behaviors, and attitudes (Wodak, Meyer, 2015). CDA aims to uncover the hidden forms in which the ideologies are represented within the discourse (van Dijk, 1993). The role of ideology in discourse can't be underestimated, as it is a significant part of shaping social consciousness and establishing power dynamics.

Scholars also consider discourse and social structures being interconnected (Wodak, 2001). Discourse does not merely mirror society; it plays an active role in creating and perpetuating social realities. By analyzing discourse, CDA reveals how social structures such as race, class, and gender are reproduced through language. This principle highlights the interconnection between language and society. It is noteworthy that both of these aspects have an influence on one another (van Dijk, 2015).

Another cornerstone of CDA claims that context is essential for understanding discourse (Wodak, 2001). Every discourse exists in a certain context. CDA emphasizes the need to consider these contexts in order to fully understand the meaning and implications of discourse. Texts are analyzed not just for their linguistic content but also for the broader contexts in which they are produced and consumed (Wodak, Meyer, 2015).

Moreover, CDA advocates for a critical approach to texts, which involves analyzing not only the obvious linguistic features but also the underlying assumptions and power relations concealed within the speech (van Dijk, 1993). This critical analysis uncovers biases, ideologies, and hidden agendas within texts, which might otherwise go unnoticed.

Another important aspect of CDA is its interdisciplinary nature, drawing from fields such as linguistics, sociology, philosophy, and political science (Fairclough, 1995). Norman Fairclough divided the framework of Critical Discourse Analysis to three-part of frameworks, known as "three dimensional frameworks" which sees the discourse into three levels of analysis: text, discourse practice, socio-cultural practice (Rafiah, 2018). By means of an interdisciplinary approach, the analysis enables one to see a more comprehensive picture and detect bias more effectively.

All the three mentioned scholars agree on the primary principle of CDA which claims that language is a social practice, through which the world is represented (Fairclough, 1995).

The process of performing discourse analysis was presented by Norman Fairclough Fairclough's (1989, 1995) and consisted of s three inter-related processes of analysis tied to three inter-related dimensions of discourse. These processes include: text analysis (description), processing analysis (interpretation) and social analysis (explanation) (Janks, 1997)

Van Dijk (1995) suggests that the analyst must first explore which structures and strategies in texts. One more step to take involves the examination and identification of patterns of bias (dominance, manipulation). Essentially, the discourse analyst must focus on how inequalities are expressed, enacted, legitimated, and perpetuated through language (van Dijk, 1995). It implies that linguistic theories must be exploited in the process of performing the analysis. Van Dijk (2001) argues that CDA requires a strong linguistic foundation, where "linguistic" is understood in a broad, structural-functional sense.

The scholars Kazemian and Hashemi (2014) express their assumption on the interconnection between discourse and Systemic Functional Linguistics (SFL) (Kazemian, Hashemi, 2014). CDA is known to borrow from Systemic Functional Grammar. Furthermore, Fairclough (2003) asserts that SFL is a valuable resource for CDA, with many major contributions to the field having emerged from SFL (Amoussou, Allagbe, 2022).

However, as previously mentioned, CDA requires a combination of methods, meaning that a linguistic approach should be complemented by historical, socio-political, sociological, anthropological, and sociolinguistic perspectives (Amoussou, Allagbe, 2022). To guide text analysis, one may discover the answers for the raise several key questions on the following topics: transitivity (what transitivity patterns appear?) mood and modality (how is mood expressed?), vocabulary (how do words reveal ideological stances?), interactional control features (what interactional controls are present in the text?), topicality (which topics are highlighted in the theme position of the sentence, or which are foregrounded?), presuppositions (are there assumptions in the discourse that are not explicitly stated but taken for granted by the author or speaker?), vagueness (are there vague expressions that lack clarity or do not convey precise meaning?), implication (what implicit information can be inferred from the discourse, based on the context in which it is used?) (Fairclough, 1989, 1992, 1995; Huckin,1997; van Dijk, 1997).

CDA requires complex approach and attention to both linguistic and non-linguistic features of text, yet, when applied properly, it is able to provide the most effective way of unraveling bias or manipulation in the media and is widely used as an integral part of media discourse analysis (Sriwimon, Zilli, 2017).

1.4. Interpreting Communicative Intent: Pragmatic Tools for Identifying Media Bias.

Pragmatics is the branch of linguistics that discovers how meaning is conveyed and interpreted in both verbal and nonverbal communication. It plays an indispensable role in language study, due to its emphasis on how speakers convey their communicative intention and how listeners interpret its meaning based on context and social environment (Green, Bach, Harnish, 1983). Pragmatics involves the analysis of communicative meaning by examining the actual utterances. In this sense, it investigates the speaker's intended meaning in relation to situational and sociocultural contexts (Green, Bach, Harnish, 1983). Unlike semantics, which examines the literal meaning of words and sentences, pragmatics investigates how meaning is constructed through interaction, intention, and situational factors (Huang, 2016). Consequently, the field of pragmatics explores how language functions effectively to convey appropriate meaning but also how it can be used respectfully to prevent miscommunication between interlocutors.

The concept of "pragmatics" first emerged in the United States during the late 1930s. In 1938, Charles W. Morris, a philosopher at the University of Chicago, introduced pragmatics as a component of semiotics (the study of signs) dividing it into three areas: semantics (the relationship between signs and objects), syntactics (the relationships among signs), and pragmatics (the relationship between signs and their users) (González-Lloret, 2012). This contributed to the shift from communication being viewed as an abstract system to the system that considered the participants engaged in the conversation (González-Lloret, 2012).

By the late 1960s, generative semantics has appeared in response to Chomsky's transformational grammar. This new approach aimed to incorporate pragmatic elements within formal grammatical theory. It was influenced by scholars J.L. Austin (1962), John Searle (1969), and H.P. Grice (1975). Later it contributed to the advancements of some of the foundational concepts of modern pragmatics. The core ideas of which were the following: language is no longer perceived as an abstract system governed solely by truth conditions; meaning is established by the context and shaped by the roles of speaker and listener (Austin, 1962); and the functions of language functions involve performing action along with conveying information (Austin, 1962).

One of the most prominent achievements of Austin and Searle was the development of Speech Act theory. The core component of this theory is speech act—a spoken utterance endowed with a communicative purpose (Nordquist, 2024). According to Austin, every utterance consists of three components: the locutionary act (the literal meaning of the sentence), the illocutionary act (the speaker's intended function, such as requesting, promising, or apologizing), and the perlocutionary act (the effect on the listener, such as persuading or frightening) (Harris, Fogal, Moss, (2018).

From the previous notion stems another pillar of the theory. Searle categorized illocutionary acts into five main types: assertives (statements of fact), directives

(requests or commands), commissives (promises or offers), expressives (expressions of emotions or attitudes), and declaratives (utterances that bring about a change in reality, such as pronouncing someone married) (Zaib, Mahmood, 2023).

This theory facilitates the analysis of each word beyond its literal meaning and helps to understand how every word functions during communication (Nordquist, 2024).

Pragmatic analysis involves a collection of linguistic and logical methods, applied to create structured interpretations of various discourse interactions (Harris, Fogal, Moss, (2018). The theory of Speech Acts may contribute to the successful performance of this kind of analysis. In order to conduct such analysis, one shall examine how utterances function within specific social and communicative contexts) (Harris, Fogal, Moss, (2018). To conduct a pragmatic analysis grounded in Speech Act Theory, the first step is to identify the type of speech act being performed. According to Searle's (1969) classification, it is possible to categorize the utterance as one of five main types (assertives, commissives, expressives, declaratives, descriptives). The next step the researchers shall take involves the analysis of illocutionary force. Illocutionary force stands for the intention the speaker had at the moment of utterance (Searle, Vanderveken, 1985). This includes understanding the speaker's goal, the social norms governing the interaction, and the power dynamics between participants (Nordquist, 2024). The final step involves examination of the perlocutionary effect. This term stands for the listener's listener response or is expected response and assesses the efficiency of speech act in achieving its intended function (Piskorska, 2016). Speech Act theory facilitates the analysis of each word beyond its literal meaning and helps to understand how every word functions during communication (Nordquist, 2024). Pragmatic analysis based on speech act theory may prove to be particularly valuable in uncovering implicit intentions of the speakers, their social positioning, bias and prejudice in the media, and power relations in both spoken and written communication.

RESEARCH METHODS

2.1. Theoretical concepts of corpus linguistics.

Corpus linguistics plays a pivotal role in identifying media bias and is an essential step in conducting the analysis. Corpus linguistics is grounded in empirical, inductive analysis. It is based on authentic language use and enables scholars to uncover patterns and formulate rules based on how people actually communicate (Litosseliti, 2024). The rationale for this method is well-supported: individuals often struggle to make accurate introspective judgments about language, as their perceptions are influenced by cognitive and social biases (Litosseliti, 2010). Corpus linguistics can be viewed as both as a separate discipline and solely a tool for the further analysis, it only depends on the sphere of its application (Kuebler, Zinsmeister, 2014). *Corpus Linguistics and Linguistically Annotated Corpora*. Bloomsbury Publishing). It is widely acknowledged that corpus linguistics is often linked to a perspective on language which implies that linguistic rules emerge from their application. The idea that language develops via the interaction between speakers is the foundation for this perspective. Therefore, if one aims to understand how a language like English functions, it is both logical and effective to examine it in real-life contexts; corpus methodology provides all the essential means for conducting this examination (Lindquist, 2014). When simplified to its core, corpus linguistics may seem to be a ‘theory-neutral’ study; yet, in the majority of cases, it is shaped by the researcher’s theoretical orientation (Chapelle, 2020). Each scholar, working in the realm of corpus analysis, aims to determine what constitutes a ‘feature’ and which frequencies are relevant, based on their underlying assumptions about what is significant in language. Approaches that apply pre-existing linguistic categories—developed independently of corpus data—are often described as ‘corpus-based’ (Tognini-Bonelli, 2001). While computational and statistical tools now play a vital role in corpus linguistics, it is noteworthy that the field fundamentally revolves around the investigation of linguistic questions and the use of corpora (Wright, 2015). In this regard, Charles Fillmore’s well-known phrase “computer-aided armchair linguistics” (1992) remains relevant. All the corpus based analyses are guided by the linguistic theory and reasoned argumentation (Fillmore, 2011).

In the late 1950s a punched-card technology was first employed. Back in the day, it took more than 24 hours to process a corpus of 60,000 words. This is when the concept of applying computer-based storage in order to perform the analysis of authentic language data first emerged (Vaughan, O’Keeffe, 2015). The increasing prominence of corpora in empirical linguistic research has closely mirrored advancements in computer storage and processing capabilities. Today, corpora containing over a billion words can be searched almost instantaneously, and with the advent of cloud computing, storage limitations have largely been eliminated (Vaughan, O’Keeffe, 2015). Corpus studies started its rapid development in the 1980s, with the new arguments in favor of using corpora tools in linguistics sparking a frenzy. This

research method is still relevant, as of the 'schools' of corpus linguistics keep developing. Corpus linguistics is undergoing methodological advancement, and the number of languages explored by corpus researchers continues to expand each year (McEnery, 2019). The term corpus linguistics was introduced by Aarts and Meijs in 1984 and became common since then (McEnery, Hardie, 2012). It is during that period that scholars began to explore the boundaries of the emerging field, debating what qualified as a corpus and whether corpus linguistics should be observed as a distinct discipline of linguistics or solely as a methodological approach to linguistic analysis (Vaughan, O'Keeffe, 2015). Today, a wide range of linguistic disciplines increasingly incorporate corpus analysis tools, either partially or fully, and corpus linguistics itself, centered on viewing language as data, has become a well-established field within academia (Vaughan, O'Keeffe, 2015).

One of the key strengths of the corpus linguistics approach is its ability to quantify linguistic patterns. It allows researchers to draw more robust and data-driven conclusions. For instance, rather than making a general claim such as "women are more prone to using emotional language than men," one may perform corpus analysis and be able to test this hypothesis empirically (Litosseliti, 2010). It can reveal not only whether the claim holds true, but also the proportional frequency of using emotional language by gender, the variety of words used, their relative frequencies, and contextual or functional differences in usage (Litosseliti, 2010).

A noteworthy insight regarding corpus-based research is that linguistic patterns are rarely absolute; they tend to occur along a continuum. Furthermore, large-scale corpora tools are able to make it possible to examine both unique linguistic phenomena and frequently repeated patterns, which offers a more profound understanding of language in use (Litosseliti, 2024).

Corpus can be approached through two main methodologies. It can be categorised as consisting of qualitative and quantitative analysis. It is common practice to integrate both quantitative and qualitative methods when performing corpus analysis. Typically, research begins with quantitative observations and gradually progresses toward qualitative interpretation (Nordquist, 2019). However, this process is often cyclical in nature. Quantitative findings are ideally examined through qualitative analysis — for instance, to uncover the reasons behind a specific frequency pattern. Conversely, qualitative insights, drawn from the researcher's contextual understanding of language samples, may serve as the basis for categorizing data within a corpus. (Leech, 2012). Qualitative analysis aims to provide a rich and detailed description of linguistic data without focusing on the frequency of features. In this approach, scholars pay the same amount of attention to the rare phenomena as well as some more common ones. This type of analysis allows to establish nuanced distinctions that might be overlooked in more rigid classifications. Peculiarities of various languages are taken into consideration rather than forced into predefined categories (Rheindorf, Springerlink, 2019). However, a key limitation of qualitative approaches is that their

findings cannot be generalized to larger populations with the same statistical reliability as quantitative methods. Because the results are not subjected to tests of significance, it remains unclear whether observed patterns are meaningful or simply due to chance (Olejniczak, 2018).

Quantitative research involves categorizing linguistic features, counting them, and often constructing statistical models to explain observed patterns. This approach enables generalization of findings to broader populations, provided that appropriate sampling methods and statistical significance testing are employed (Chapelle, 2020). It also facilitates direct comparisons between corpora, helping to determine whether certain linguistic phenomena are representative of a language or variety, or merely random occurrences (Chapelle, 2020). Even when focused on a single language variety, quantitative analysis provides precise insights into the frequency and rarity of features, indicating what is typical or atypical for the researched language (Gries, 2014). However, this method offers a less nuanced view of language than qualitative analysis. Additionally, quantitative methods often treat rare phenomena as insignificant. This can lead to the collapsing of categories and, consequently, a loss of descriptive richness. Thus, while powerful, quantitative analysis represents an idealized view of linguistic data (Gries, 2014).

To summarize, despite being relatively young, corpus analysis has evolved into a powerful and methodologically mature approach actively applied within linguistic research.

2.2. Corpus analysis tool applied in the research.

There is an ongoing debate within corpus linguistics, which concerns the appropriate methodological starting point for corpus analysis. One perspective advocates for a 'corpus-driven' approach (Tognini-Bonelli, 2002). According to this perspective, analysis begins with direct observation of corpus data without relying on pre-established linguistic theories (Tognini-Bonelli, 2002). This method is mostly associated with the examination of concordance lines in texts using tools such as Key Word in Context (KWIC) (Anthony, 2013). It stemmed from the empirical investigation and was boosted by technological advancements. now it allows scholars to explore the authentic language use. Apart from quantitative insights, it is able to challenge existing theories of language. This feature makes it both a methodological tool and a theoretical lens that bridges the gulf between data-driven analysis and the interpretive linguistic inquiry.

Conversely, there is an opposing view. According to which it is impossible to completely reject prior linguistic knowledge when analyzing corpora. Based on this notion, all the scholars, applying corpus analysis, test existing theoretical models based real language data. Subsequently, such type of analysis may lead to the refinement of those models or the development of new ones (Anthony, 2013). Such a point of view on corpus analysis is known as the 'corpus-based' approach (McEnery, Hardie, 2011).

In order to perform the high-quality corpus analysis, specialized software—commonly referred to as a 'corpus tool' is required. It enables scholars to search, operate, and store linguistic data. Such tools are essential for the successful performance of lexicographic analysis, as the diverse resources make corpus extremely valuable for scholars. (Maachi, Khamar, 2025).

Over the past decades, both corpus compilation and the development of corpus tools have advanced significantly. Undoubtedly, such advancement was enabled with the help of technological progress. The process of compiling and storing corpora has become much faster and more efficient, which resulted in the possibility of creation of much larger corpora (Maachi, Khamar, 2025). The primary impact of corpus software tools lies in the way we perceive language (Vaughan, O'Keeffe, 2015). These tools allow for rapid display of language data and can compute and present information in mere milliseconds, changing the way linguistic patterns are understood (Hunston, 2002).

In order to perform a high-quality corpus analysis, the researcher is supposed to select and collect linguistic data that successfully represents the language variety or phenomenon under investigation. Sources can include newspapers, literature, academic writing, social media posts, spoken transcripts, legal documents, or learner language (McEnery, Hardie, 2012).

A pivotal step in conducting corpus analysis involves thorough preparation for this process. The functionality of many corpus tools mainly depends on the proper preparation of corpus data. (Kilgarriff, Kosem, 2012). The collected texts are supposed to be cleaned and meticulously formatted. This includes standardizing file formats, omitting additional content, and ensuring consistency in punctuation and encoding. Simultaneously, metadata (e.g., date, source, author, genre, region) is attached to each file to enable more refined searches and filtering during analysis (Kilgarriff, Iztok Kosem, 2012).

If the researcher is willing to facilitate deeper linguistic analysis, corpora may sometimes enriched with linguistic annotations. This includes part-of-speech tagging, lemmatization, syntactic parsing, or semantic annotation (Brezina, 2020). These tasks are rarely performed manually, typically the researchers apply natural language processing (NLP) tools and corpus software such as Sketch Engine, AntConc, or LancsBox (Anthony, 2023).

With the prepared corpus loaded into a corpus tool, researchers can conduct various types of searches and analyses. This may include the creation of frequency lists, concordance lines, n-grams, keyword analysis, distributional thesaurus, etc. (Fendel, 2024).

One of the most common types of analysis involves the creation of word list. Creating a word list enables users to load a corpus and explore basic frequency trends (Fendel, 2024). This frequency analysis reveals which words appear most frequently within a text or across a collection of texts. Although different concordance tools may

vary in how they define a "word," the key distinction for word lists lies in the difference between tokens (individual instances of words) and types (the number of unique word forms present in the corpus) (Vaughan, O’Keeffe, 2015).

Another widespread tool applied when performing the corpus analysis involves the concordance. The use of computational methods has drastically improved linguists’ ability to determine which words frequently occur together and how these patterns of co-occurrence influence meaning—an advancement that has notably shaped lexical studies (Vaughan, O’Keeffe, 2015). The concordance view is the software feature that has made this approach to corpus data possible (Brezina, 2020). Among the various tools used in corpus analysis, the concordance view perhaps requires the most active human interpretation of automatically generated data. It displays a selected search term within its surrounding linguistic context, creating a collocation, though the context can be adjusted for various perspectives (Vaughan, O’Keeffe, 2015). The work of scholar John Sinclair work has become especially influential in this area, as it contributed to forming an explanation for certain word combinations and what makes them sound more or less natural (Vaughan, O’Keeffe, 2015).

While the concept of collocation was initially introduced by Firth (1935), it was the development of corpus methodologies that enabled researchers to explore it in depth.

One more beneficial tool is the creation of keyword list. The researchers tend to define the target word, when exploiting concordances. Those words may be selected frequency lists, which commonly serve as the initial point of entry into corpus data (Kim, 2007). Frequency lists are capable of offering a broad picture of how often words appear within a corpus, however, occasionally researchers may also feel the need to compare wordlists from different corpora to produce a set of keywords. These are not merely the most frequent words, but rather those that appear with unusual frequency, compared to a reference corpus. In this way, wordlists reflect raw frequency, whereas keyword analysis highlights the relative prominence of certain words in a specific text or dataset and how actively these words are used (Kim, 2007).

Finally, a pivotal tool for performing the profound corpus analysis is thesaurus. This tool is provided by software such as Sketch Engine, which generates automatic thesauri by analyzing collocational and grammatical behavior across a corpus (Kilgarriff, 2014). A thesaurus in corpus analysis is not merely a list of synonyms. It is an elaborate and efficient tool that identifies semantically related words based on distributional similarity (words that tend to occur in similar grammatical and lexical contexts) (Kilgarriff, 2014). This technique is based on the distributional hypothesis, which claims that words with similar meanings tend to appear in similar contexts.

Modern technologies offer a wide range of software for performing corpus analysis. This involves Sketch engine (developed in 2003, a program that analyzes a text displays the frequency and context of word usage) (Kilgarriff, 2014), AntConc (a free, multiplatform corpus analysis toolkit, developed in 2002) (Anthony, 2023),

LancsBox (introduced in 2015 as a new-generation corpus analysis tool, meant to facilitate the study of language data across various disciplines).

All of these tools are bound to contribute to the performance of high-quality analysis and efficiently show the results of it almost instantaneously.

2.3.Utilization of Sketch Engine and Stages of Corpus-Based Research.

Sketch Engine is an intuitive, web-based corpus analysis platform that supports more than 90 languages. It offers users access to both contemporary and historical corpora, enabling a wide range of applications—from text proofreading to in-depth diachronic linguistic research (Kilgarriff, 2014). Sketch Engine enables users to generate and examine word sketches, identify thesaurally related terms, and analyze 'sketch differences.' These word sketches are fully integrated with concordance functions. Users are able to select a specific collocation within a word sketch, after which they are directed to a concordance view that displays the corpus evidence underlying the collocates grammatical relation (Kilgarriff, 2014).

Sketch Engine was developed by Adam Kilgarriff and his colleagues at Lexical Computing Ltd. The technology was released in 2003 (Wikipedia Contributors, 2025). The tool's primary goal was to be able to perform the automated creation of "word sketches" (one-page summaries of a word's grammatical and collocational behavior derived from corpus data) (Kilgarriff, 2015).

Over the years, Sketch Engine has begun to support over 90 languages and has become an essential tool for scholars, translators, and teachers (Wikipedia Contributors, 2025). Its features include concordance searches, frequency lists, and term extraction, among others, making it a comprehensive resource for corpus-based language analysis (Wikipedia Contributors, 2025).

In this research, Sketch Engine was applied as the primary tool for conducting a corpus-based analysis. The process was carried out in several stages, each aimed at systematically extracting and interpreting lexical data that could contribute to identifying media bias. Alongside with that, Sketch engine tools were also employed with an aim to create the foundation for the further pragmatic analysis.

The first step involved the creation of corpora itself. In order to perform it, the articles on a certain topic (e.g. engagement, first public appearance, wedding, etc.) were collected and organized in two separate folders. The articles were taken from the most well-known and credible media sources, including BBC, Vogue, The Guardian, Brides, The Daily Mail, USA today, etc. After that, the URLs of the selected articles were uploaded on the Sketch Engine platform and two corpuses (for articles about Catherine and about Meghan) were compiled with the help of "create corpus tool" available homepage of the website (Figure 1). Both corpora had approximately the same number of words to ensure comparability and maintain the objectivity of the future analysis. Once the corpora were successfully compiled, various tools within Sketch Engine were employed to conduct individual corpus analyses.

RECENTLY USED CORPORA			NEW CORPUS
Meghan`s first public engagement	English	13,545	
Catherine`s first public engagement	English	13,606	
Meghan`s relationship timeline	English	71,932	
Catherine relationship timeline	English	61,910	

Figure 1. Corpora compilation on Sketch Engine website

The first tool utilized when performing the corpus analysis was the wordlist generator. This tool in Sketch Engine is capable of generating various types of frequency lists. Those may be organized by parts of speech (nouns, verbs, and adjectives) by specific character patterns (e.g., words that begin with, end with, or contain certain sequences), and by linguistic attributes (word forms, tags, and lemmas). In order to employ this tool, one has to select the part of speech which is required for the research and press “go” button (Figure 2). These criteria can be applied individually or in combination. The tool also provides three distinct frequency metrics: raw frequency, frequency per million words, and Average Reduced Frequency (ARF) (Sketch Engine, 2019). Wordlist operates at the token level, meaning it analyzes each word occurrence separately.

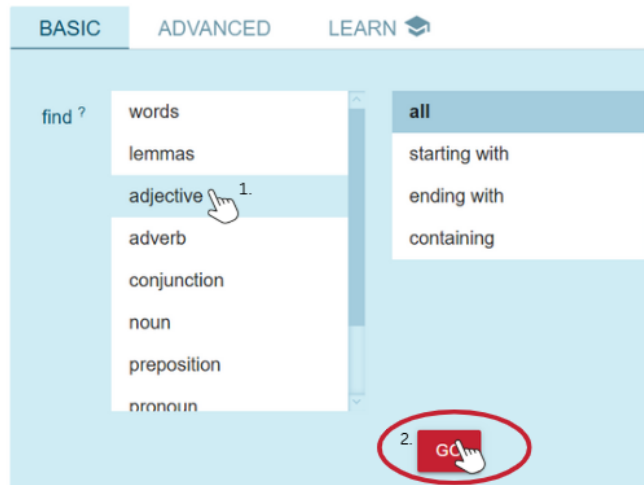


Figure 2. Utilizing the Wordlist tool on Sketch Engine website

Under the default settings, the tool automatically excludes non-lexical items, thereby producing a clean list composed exclusively of words (Sketch Engine, 2019). In order to generate a wordlist, one should select a corpora and choose the one of the criteria for the words to be processed (depending on parts of speech, character patterns,

linguistic attributes). To conduct the research, we selected verbs, adjectives and nouns individually, utilized the tool and analyzed the results by examining the most frequent words (up to 80) of each of mentioned parts of speech (Figure 3). The wordlists (of verbs) then played a pivotal role in performing the pragmatic analysis, based on the theory of speech acts.

WORDLIST Catherine relationship timeline (1,088 items | 4,190 total frequency)

adjective

Adjective	Frequency ? ↓	Adjective	Frequency ? ↓	Adjective	Frequency ? ↓	Adjective	Frequency ? ↓	Adjective	Frequency ? ↓
1 royal	198 ***	11 second	40 ***	21 big	32 ***	31 least	25 ***	41 third	17 ***
2 first	132 ***	12 last	39 ***	22 former	32 ***	32 old	24 ***	42 wrong	16 ***
3 other	95 ***	13 future	39 ***	23 close	31 ***	33 different	23 ***	43 sweet	16 ***
4 young	78 ***	14 same	38 ***	24 happy	30 ***	34 little	22 ***	44 married	16 ***
5 good	74 ***	15 private	38 ***	25 great	29 ***	35 romantic	21 ***	45 red	15 ***
6 more	66 ***	16 many	37 ***	26 official	29 ***	36 right	20 ***	46 front	15 ***
7 new	59 ***	17 [number]	36 ***	27 late	28 ***	37 much	19 ***	47 rare	15 ***
8 public	49 ***	18 long	35 ***	28 early	25 ***	38 clear	18 ***	48 blue	15 ***
9 british	49 ***	19 few	34 ***	29 such	25 ***	39 special	18 ***	49 strong	15 ***
10 own	41 ***	20 next	33 ***	30 black	25 ***	40 personal	17 ***	50 joint	14 ***

Figure 3. The Wordlist exposing the adjectives and their frequency in the selected corpora

Apart from wordlist, the research also demanded the employment of keyword selection. Keywords are single-token items that occur more frequently in the focus corpus. They are used to highlight distinctive features of the focus corpus. Keywords can be extracted using the Keywords & Terms tool in Sketch Engine (Cukr, 2024) (Figure 4). Following the creation of the wordlist, key content words were defined based on their relative frequency and relevance to the research questions and provided an overview for each of the compiled corpora (Figure 5).

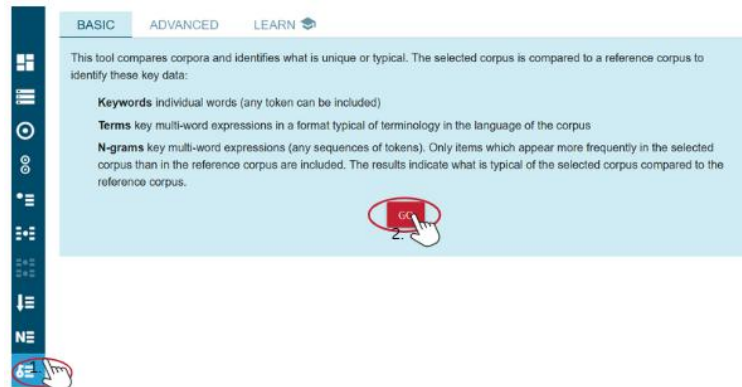


Figure 4. Utilizing the Keywords tool on Sketch engine website

KEYWORDS

SINGLE-WORDS ☒ MULTI-WORD TERMS ☒

reference corpus: English Web 2021 (enTenTen21)

Term	Term	Term	Term
1 prince harry	14 business insider	27 adrian dennis	40 private secretary
2 meghan markle	15 charity fair	28 kate middleton	41 mr mohammad
3 royal engagement	16 queen elizabeth ii	29 personalised advertising	42 queen laugh
4 kensington palace	17 business insider subscriber	30 use profile	43 broke royal protocol
5 innovative story	18 nottingham cottage	31 queen elizabeth	44 first joint engagement
6 nottingham academy	19 joint engagement	32 terrence higgins trust	45 first official royal engagement
7 royal protocol	20 start reading	33 windsor castle	46 community recording studio
8 first royal engagement	21 first official	34 princess diana	47 official royal engagement
9 official engagement	22 first engagement	35 full effect	48 christopher furlong
10 first official engagement	23 world aids day	36 prince william	49 today's engagement
11 nottingham contemporary	24 getty image	37 buckingham palace	50 surprising reaction
12 royal foundation	25 information on a device	38 limited datum	
13 royal family	26 national justice museum	39 access information	

Figure 5. The Keywords (Multi-word terms) from the selected corpora

Lastly, the research demanded the construction of thesaurus. Sketch Engine is capable of generating a distributional thesaurus for a corpus, depending in the common collocations used in the corpora. If two words share numerous collocates, they will appear in each other's thesaurus entry (Kilgarriff, 2014). The Sketch Engine compiles a thesaurus using the strategy of identifying the 'nearest neighbours' for each word. It employs mathematics for discovering this similarity (Kilgarriff, 2003). The thesaurus function in Sketch Engine was employed to explore semantic relations and lexical variation. This feature allowed the identification of words that are related to the selected keywords. This feature allowed the identification of words that are related to the selected keywords, offering insights into lexical clustering within the corpus and contributing to the revealing the media bias. We searched collocations with the words "Kate" and "Meghan" for the sake of creating a comprehensive picture of how the personalities are presented in the media. In order to perform it, we typed in the names when analyzing each corpora and pressed the "go" button (Figure 6). After that, we presented the thesaurus in a form of a bubble chart, which is another tool available on Sketch Engine website (Figure 7). The next step was to create the thesaurus for the articles about the other personality to enable the juxtaposition (Figure 8).

THESAURUS

BASIC ADVANCED LEARN

Search ? 1.
Kate

GO

Figure 6. Utilizing the Thesaurus tool on Sketch Engine website

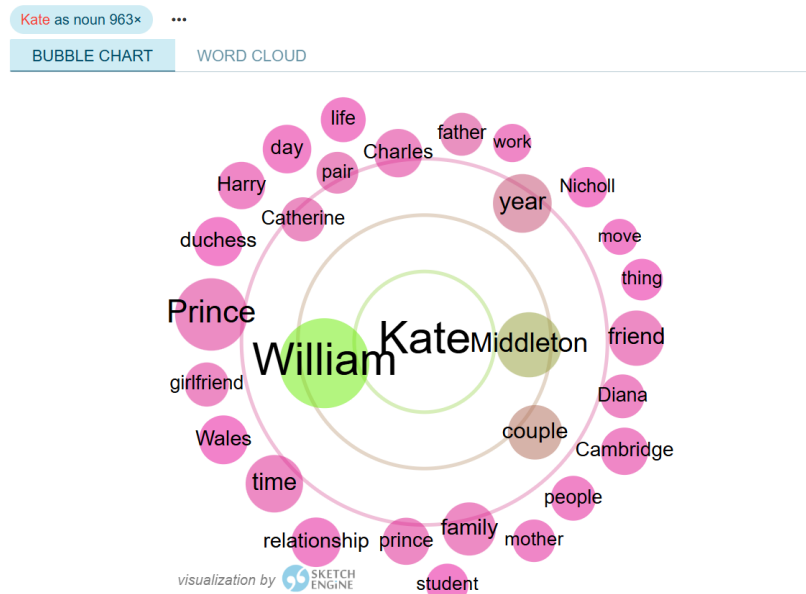


Figure 7. Thesaurus based on *Catherine's relationship timeline* corpora

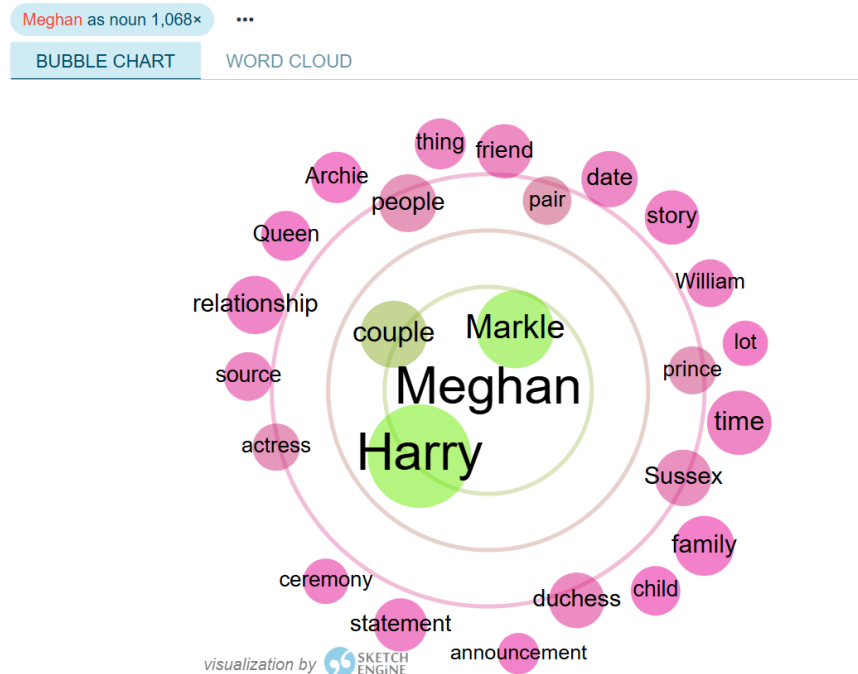


Figure 8. Thesaurus based on *Meghan's relationship timeline* corpora

2.4. Multimodal Analysis Procedure.

Apart from corpus analysis, the study required the performance of a multimodal analysis to examine how meaning is constructed not solely by means of the language but also via visual and contextual elements. In order to perform this kind of analysis, a collection of articles about both personalities on identical topics was compiled and the way each article was presented on the website page was analyzed (Figure 9, Figure 10).

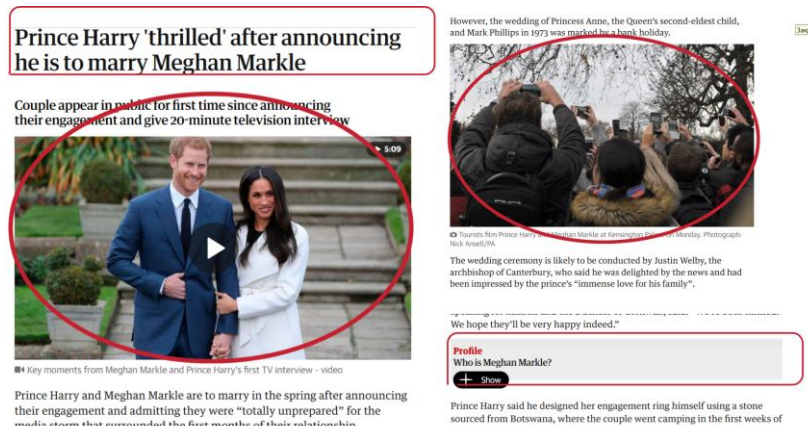


Figure 9. Multimodality in the article covering the engagement of Meghan, The Duchess of Sussex

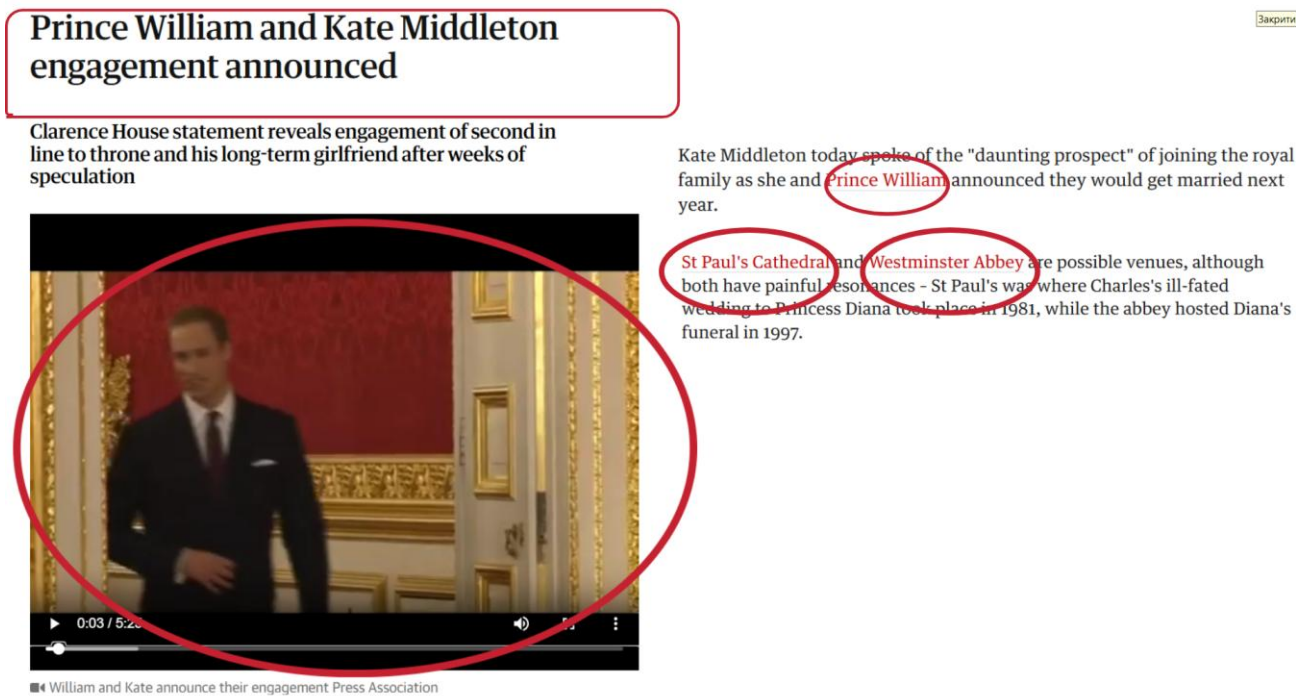


Figure 10. Figure 9. Multimodality in the article covering the engagement of Catherine, The Princess of Wales

This phase of the research focused on analyzing all the aspects of selected media materials, including photographs, headlines, image captions, and layout elements that accompanied the texts in the original articles. The analysis was informed by the principles of multimodal discourse analysis (MDA), which consider the interaction between various semiotic modes (e.g., linguistic, visual, spatial) (Herring, 2013).

After selecting each article, its layout, images, videos and headlines were juxtaposed in order to enable the further visual analysis. More than 30 articles covering

one event regarding each personality ensured a rich dataset for analyzing the multimodal construction.

At first, each image that accompanied the articles was examined for compositional features. Those included framing, gaze, and color usage. This analysis aimed to uncover how visual elements contribute to the portrayal of the subjects (Ghooshchi, 2021) (Figure 11, Figure 12).



Figure 11. Visual elements from the article covering the engagement of Meghan, The Duchess of Sussex



Figure 12. Visual elements from the article covering the engagement of Catherine, The Princess of Wales

Proceeding the first stage, Headlines and captions were analyzed for lexical choices, syntactic structures, and rhetorical devices. This step assessed how language reinforces or contrasts with visual representations.

The headlines of the articles about Meghan's engagement included: Prince Harry 'thrilled' after announcing he is to marry Meghan Markle (The Guardian), Prince Harry and Meghan Markle Are Engaged. Get ready for another royal wedding (Harper's Bazaar), Meghan Markle 'started unpacking her bags in Harry's house' before engagement was announced (Daily Record), Prince Harry says 'stars were aligned' in engagement to Meghan Markle: 'Everything was just perfect' (ABC News), All the details you need to know about Meghan Markle and Prince Harry's fairy-tale wedding (Business Insider), The Look of Love! See Every Photo of Newly-Engaged Prince Harry and Meghan Markle (People), Why Meghan Markle's engagement to Prince Harry is controversial? Hint: She's a divorced biracial American who stars in a TV show (Vox), When Meghan weds Harry, Britain's relationship with race will change for ever (The Guardian), Prince William says he hopes Prince Harry's engagement means he 'will stop scrounging my food' (ABC News).

The headlines of the articles about Catherine's engagement included: *Mr. and Mrs. Wales (Vanity Fair)*, *Prince William and Kate Middleton announce their engagement (The Guardian)*, *Diana's Ring Seals Prince William's Marriage Plans (The New York Times)*, *Royal Wedding Date: Prince William and Kate Middleton to marry on April 29 (The Telegraph)*, *Royal wedding: Prince William to marry Kate Middleton (BBC)*, *Royal wedding next year: Katie's eight-year wait is over as she and Prince William finally announce they ARE engaged and will marry in 2011 (The Daily Mail)*, *Major change Kate made to her looks before engagement to William to be 'more appropriate' (Mirror)*, *Twice The Fun: Kate Middleton May Wear Two Wedding Dresses (Time)*, *Royal engagement (DW)*, *Prince William and Kate Middleton Are Engaged! (People)*, *Prince William Proposes to Kate Middleton With Princess Diana's Engagement Ring (ABC News)*, *William and Kate's 'crazy' engagement announcement (Hello!)*

Finally, findings from the analyses of Catherine and Meghan's representations were compared to identify patterns, similarities, and differences in their multimodal portrayals, aiming to define the discrepancies in the ways the public image of both personalities was created.

RESULTS

3.1. Overview of the Analytical Framework.

In order to recognize media bias concerning the public images of Catherine, the Princess of Wales and Meghan, the Duchess of Sussex, various research tools and methods were employed. Those included four main analytical approaches: Critical Discourse Analysis (CDA), pragmatic analysis through Speech Act Theory, Corpus analysis and Multimodal Discourse Analysis (MDA). A total of 170 articles were analyzed, along with 200 accompanying videos and visual elements where available. Each method was targeted at different dimensions of meaning and aimed to provide a coherent and comprehensive picture. The employed methods proved to be highly effective in revealing media bias.

For the CDA, Fairclough's three-dimensional model was applied, focusing on textual features (e.g., lexical choices, modality, and transitivity), discursive strategies (e.g., nomination, predication), and broader ideological representations within social contexts (Fairclough, 1995). Fairclough's Critical Discourse Analysis framework was employed to uncover how media discourse benefits the construction of public images of Kate Middleton and Meghan Markle. This approach provided a multi-level analysis of language use. It helped to reveal the interaction between lexical choices, discursive structures, and broader social ideologies in order to reflect the biased perspectives of the media.

Another analysis method applied was the pragmatic one. In order to perform it, Searle's classification of speech acts was followed, categorizing utterances into assertives, directives, commissives, expressives, and declaratives (Searle, 1969). Each speech act was coded according to its function and contextual usage across the texts and videos. Searle's classification of speech acts was enabled examination of the pragmatic functions of verbs in the articles about Kate Middleton and Meghan Markle. This pragmatic layer added depth to the broader discourse analysis, illustrating how bias is not only lexical but also functional in nature.

Multimodal Discourse Analysis was conducted using the framework of Kress and van Leeuwen. Visual data were coded for salience, gaze, framing, vectors, and compositional layout to examine how visual and textual elements interacted to produce meaning (Kress, van Leeuwen, 2001). Multimodal Discourse Analysis (MDA) provided insight into the visual strategies used in the media portrayal of Kate Middleton and Meghan Markle. By analyzing image selection, it was possible to show how visual representations complement and reinforce the linguistic bias identified through CDA.

Finally, corpus analysis was employed to receive data-driven information on the lexical units used by the articles' authors. Corpus analysis provides quantifiable evidence by examining patterns in language that would be difficult to identify manually. This objectivity strengthens the validity of the findings (Fillmore, 2011). Corpus analysis helps identify recurring linguistic patterns that are crucial for revealing

bias. It enables researcher to identify themes, keywords, and collocations. This analysis method improves the understanding of how language choices contribute to the construction of public images.

The combination of all the mentioned analytic methods have highly contributed to acquiring the comprehensive results of the research.

3.2.Critical Discourse Analysis (CDA) Results, integrating the tools of corpus analysis

For the sake of performing critical discourse analysis, 70 articles covering the wedding days of Meghan, the Duchess of Sussex and Catherine, the Princess of Wales were selected, analyzed and compared. The articles were published by the most prominent British and American media sources, including *BBC*, *The New York Times*, *USA Today*, *The Guardian*, *The Daily Mail*, *People*, *Hello*, *Vogue*, *Vanity Fair*, *Brides*, *ABC News*, etc. After compiling the selection of articles about each personality, we performed the CDA, focusing on the following notions: common discursive themes, linguistic choices, the reflection of social norm and ideological bias, along with intertextuality.

The analysis of the articles covering Meghan Markle`s wedding has shown the excessive use of elevated language in order to create romantic idealization: *“Harry mouths ‘You look amazing’ and melts our hearts”* (*USA Today*, 2018). The quote contains the emotionally colored adjective *amazing* and the idiom to *melt one`s heart* to reflect the moving effect of Prince`s remark.

“Prince Harry captured hearts when he became visibly emotional during the ceremony, wiping tears from his eyes on two separate occasions” (*ABC News*, 2018). The word *heart* was used again to highlight the deep feelings of the groom.

“Prince Harry and Meghan Markle have released three official wedding portraits, two days after their fairy-tale wedding in Windsor” (*Vogue*, 2018). The idealized metaphor *fairy-tale wedding* highlights with significance and magnificence of the event.

“The lovebirds broke from palace tradition when they locked lips on the steps of St. George's Chapel, becoming the first royal couple to forego a first smooch on the balcony of Buckingham Palace in nearly 30 years” (*USA Today*, 2018). The emotionally colored metaphor *lovebirds*, accompanied with the romanticized expression *locked lips* and the informal term “smooch” contributed to the heartwarming portrayal of the couple`s first public kiss.

Another prominent feature present in the articles covering the Duke and Duchess`s wedding is the focus on historical and institutional significance often juxtaposed with the bride`s background as a commoner:

“Markle has been welcomed into royal circles, accompanying Harry to family functions including Christmas with the queen at her country retreat” (*NBC News*, 2018). The clear emphasis on the significance of spending Christmas in royal circles

highlights the importance of sticking to this tradition and the honor the bride-to-be was presented with.

“Queen Elizabeth II honored her red-headed, 33-year-old grandson with a new title: the Duke of Sussex, making Markle the Duchess of Sussex (First Post, 2018).”

“I don’t imagine Meghan Markle – the new Duchess of Sussex – will ever forget the sight of her father-in-law lurking by a “cascading hedgerow” of spring flowers halfway along the aisle of St George’s Chapel in Windsor, ready to take her arm (The Guardian, 2018).”

“It’s no secret that the newly named Duke and Duchess of Sussex would like to have children” (Town and Country Magazine, 2018).

“Prince Charles reportedly admired the now Duchess of Sussex all the more for her handling of the difficult situation and stepped up to walk his soon-to-be daughter-in-law partially down the aisle” (Editorial, 2018). Almost each analyzed articles focuses on the titles of the newlyweds, highlighting the eminence of being referred to as duke and duchess.

“The horse-drawn carriage containing the newlyweds will be accompanied by a military escort” (The Guardian, 2018).

“He was given special permission from the Queen to keep his short beard as it is customary to be clean-shaven when dressed in Army uniform” (BBC, 2018).

“She teamed it with the Queen Mary Diamond Bandeau tiara, loaned by the late Queen Elizabeth II (Hello, 2018). The following quotes showcase the prominence of royal symbols (carriage, Army uniform, tiara), elevating the royalty and regal items.

Conversely to the gracious and polished image of the royal family, presented in the media, the image of Meghan Markle is rather down-to-earth, with the immense focus on her background, including the former career and relationship status:

“The new princess is not only a divorcee but, in her own words, a ‘strong, confident mixed-race woman’ (NBC News, 2018).

“Queen is said to have found the gown “flamboyant” and thought it wasn’t fit for a divorcee” (Mirror, 2022). Both quotes focus on Meghan’s previous marriage, using the word “divorcee”, which possesses a social stigma of judgment.

“Wearing a tiara and a flowing white dress, the former ‘Suits’ star began her latest role at a ceremony that brought modern soul to the traditional pomp and pageantry of a royal wedding” (NBC News, 2018). The quote highlights Meghan’s former career as an actress.

“She often hugs people she meets, a break with tradition” (NBC News, 2018).

“Queen Elizabeth II was believed to have felt ‘uncomfortable’ and took issue with a number of things about the big day” (Mirror, 2022). This focuses on differences of Meghan’s habits and personality from the well-mannered royalty. It expresses the juxtaposition between the restrained personalities of the royal family members and the informality of the Americans.

Another noteworthy feature, it how often The Duchess of Sussex was referred to as *Markle*. She was the only personality in all the analyzed articles who was dubbed by her last name:

"Markle has been welcomed into royal circles, accompanying Harry to family functions including Christmas with the queen at her country retreat in Sandringham, Norfolk, and the Commonwealth Day ceremony at Westminster Abbey where she sat in the row behind the queen" (NBC News, 2018).

"The order of service, published on the morning of the wedding, revealed Markle would not promise to obey, and Harry would wear a ring" (The Guardian, 2018).

"The world got its first glimpse as Markle arrived at Windsor Castle by car, accompanied by her mother, Doria Ragland, 61" (The Guardian, 2018). Such reference downplays the status of the Duchess of Sussex, constructing Meghan as different from the members of the royal family. It also may be interpreted as a way of maintaining power hierarchies.

A prominent feature of the articles included multiple references to the couple's past, especially to Princess Diana. The media seemed to intend to juxtapose Prince Harry's wife with his late mother:

"Prince Harry perches atop an emerald green sofa—his mother sat on the exact same seat with baby Harry on her lap" (Cosmopolitan, 2020).

"The couple specifically chose them to be included in Ms. Markle's bouquet to honor the memory of the late Princess on this special day" (Cosmopolitan, 2020).

"The couple honored Harry's late mother, Princess Diana, throughout the ceremony" (ABC News, 2018).

"The Duchess of Sussex chose to remember Harry's mum by picking some of Diana's favourite flowers to use in her wedding bouquet" (Cosmopolitan, 2020).

"A Kensington Palace spokesperson later confirmed that the aquamarine ring was Princess Diana's, and when the Internet unanimously sobbed at the thought of Harry gifting it to his new wife" (Cosmopolitan, 2020). These mentioning are meant to provoke emotional response caused by share associations and memories. Along with that, such references may either legitimize Meghan, or, vice versa, show her failure to keep up with the affection the public has for Princess Diana.

The presence of typical wedding terminology is also worth mentioning. Words like newlyweds, bride and groom were scattered throughout multiple articles.

"The newlyweds followed the kiss with a 25-minute carriage procession through Windsor to greet adoring onlookers, many of whom had camped out for days to catch a glimpse of them" (USA Today, 2018).

"This Little Light of Mine" by Etta James as the newlyweds departed, pausing on the steps for a kiss" (NBC News, 2018).

"An emotional-looking prince and his smiling bride exchanged vows and rings before the Queen and 600 guests at St George's Chapel" (BBC, 2018).

"The royal wedding was full of knowing glances and sweet moments of PDA between the bride and groom" (Town and Country Magazine, 2018).

"Before a word was uttered, Markle's beauty look alone could have predicted the tone of the interview—the bride was effortlessly elegant in her Givenchy gown" (Glamour Magazine, 2018).

"Prince Harry captured hearts when he became visibly emotional during the ceremony, wiping tears from his eyes on two separate occasions as he stood next to his bride" (ABC News, 2018).

"Prince William will act as best man, a role Harry fulfilled for his brother at his wedding in 2011" (Edition, 2018).

"Prince William, the best man, stood beside his brother as Meghan walked down the aisle, escorted part of the way by Prince Charles" (The Guardian, 2018). The use of such language makes the royalty seem closer to the commoners. It emphasizes the traditions and normalizes the image of the royal family in this regard.

Lastly, the mentioning of numbers was a distinguished feature of the articles:

"The historic moment was witnessed by 600 guests—including Oprah Winfrey, George and Amal Clooney, and Elton John—a flag-waving crowd of around 100,000 outside, and millions of TV viewers around the world" (NBC News, 2018).

"The couple also invited 2,640 members of the public to the castle's grounds, including local children and community leaders" (NBC News, 2018).

"The wedding spectacle was witnessed by as many as 100,000 Britishers and foreign visitors, apart from those watching around the globe" (NDTV, 2018).

"She teamed it with the Queen Mary Diamond Bandeau tiara, loaned by the late Queen Elizabeth II, and an intricate five-metre-long silk tulle veil with hand-embroidered flowers which took 3,900 hours to create" (Hello, 2018). The emphasis on the numbers highlights how historically significant the event was and elevates the royal wedding in the eyes of commoners.

To summarize, the prominent association with Cinderella was created by multiple references to the wedding as a fairy-tale event and constant stress on the origin of the bride. Which makes the image of Meghan down-to-earth, nowhere near as sophisticated as the one of the members of the royal family. At the same time, the description of the event itself is meant to highlight the magnificence of the event. In order to support the assumptions made while conducting CDA, we have employed corpus analysis, and the Keywords, Thesaurus and Wordlist tools of Sketch Engine platform in particular.

The results of utilizing the Keywords tool has highlighted the significance of the bride: words *Markle* and *Meghan* occupy the first places on the list, whereas *Prince* and *Harry* are placed on the 13th and 23rd places. Meghan's background is also prominent among the keywords (*Doria* (7th place) and *Suits* (21st place)). Wedding-related vocabulary also played a pivotal role throughout the articles (*newlywed* (9th place), *nuptials* (20th), *bridesmaid* (24th)). Another noteworthy notion is the emphasis on the

star-studded guest list of the royal wedding (*Winfrey, Amal, Elton, Clooney, Beckham*), all of the names belong to A-list celebrities and create the elevated image of the wedding (Figure 13).

reference corpus: English Web 2021 (enTenTen21)

Lemma	Lemma	Lemma	Lemma	Lemma
1 Markle ...	11 Getty ...	21 Suits ...	31 tiara ...	41 Frogmore ...
2 Meghan ...	12 Sussex ...	22 Fellowes ...	32 Chapel ...	42 wedding ...
3 Windsor ...	13 Prince ...	23 Harry ...	33 elderflower ...	43 Ascot ...
4 Waight ...	14 Kensington ...	24 bridesmaid ...	34 Eugenie ...	44 Beckham ...
5 duchess ...	15 Clooney ...	25 WPA ...	35 Middleton ...	45 curtsy ...
6 Ragland ...	16 Welby ...	26 Sentebale ...	36 Kilkeel ...	46 Chelsy ...
7 Doria ...	17 royal ...	27 Kanneh-Mason ...	37 bandeau ...	47 Sheku ...
8 Givenchy ...	18 Amal ...	28 pageboy ...	38 Elton ...	48 Princess ...
9 newlywed ...	19 Elba ...	29 Ptak ...	39 Curry ...	49 Landau ...
10 Winfrey ...	20 nuptials ...	30 Oprah ...	40 horse-drawn ...	50 Mulroney ...

Figure 13. The keywords generated from the articles covering the wedding of Meghan Markle

The generated wordlist of nouns has also managed to support some of the notions assumed earlier (Figure 14). The difference in the usage of names Harry and Meghan is rather slight. A significant quantity of nouns reflects the royal background of the groom (*royal, queen, princess, duchess and duke, carriage*). The mentionings of the words mother and Diana highlight the assumption about Princess Diana's role in the creation of Meghan's public portrait. An unexpected discovery turned out to be the lack of media coverage of the bride's wedding outfit, which is rather unusual for such kind of articles (*dress* was only mentioned 39 times).

noun

(2,156 items | 9,815 total frequency)

🔍

📄

👁

☰

🌿

🕒

🌟

Noun	Frequency ? ↓	Noun	Frequency ? ↓	Noun	Frequency ? ↓	Noun	Frequency ? ↓
1 harry	303 ...	14 castle	71 ...	27 palace	44 ...	40 carriage	36 ...
2 meghan	297 ...	15 duchess	66 ...	28 st	43 ...	41 way	31 ...
3 prince	268 ...	16 princess	66 ...	29 content	42 ...	42 [number]	30 ...
4 wedding	243 ...	17 sussex	62 ...	30 may	41 ...	43 flower	30 ...
5 markle	189 ...	18 family	57 ...	31 mother	41 ...	44 kensington	29 ...
6 windsor	119 ...	19 guest	55 ...	32 diana	40 ...	45 moment	28 ...
7 couple	105 ...	20 duke	54 ...	33 william	40 ...	46 service	27 ...
8 george	98 ...	21 people	49 ...	34 world	39 ...	47 curry	26 ...
9 day	84 ...	22 charles	49 ...	35 dress	39 ...	48 king	26 ...
10 chapel	79 ...	23 image	47 ...	36 year	38 ...	49 father	26 ...
11 royal	76 ...	24 getty	47 ...	37 love	38 ...	50 wife	25 ...
12 ceremony	75 ...	25 time	46 ...	38 reception	37 ...		
13 queen	71 ...	26 bride	45 ...	39 friend	36 ...		

Figure 14. The wordlist of nouns generated from the articles covering the wedding of Meghan Markle

Finally, the thesaurus formed on the basis of the selected corpora for the name *Meghan* has provided a backup for the conclusions drawn from the CDA of the articles (Figure 16). The bubble chart included some regal vocabulary (*duke, carriage, queen*),

vocabulary reflecting the bride's past (*actress*), as well as the vocabulary concerning the guest-list of the ceremony (*Chopra, guest*).

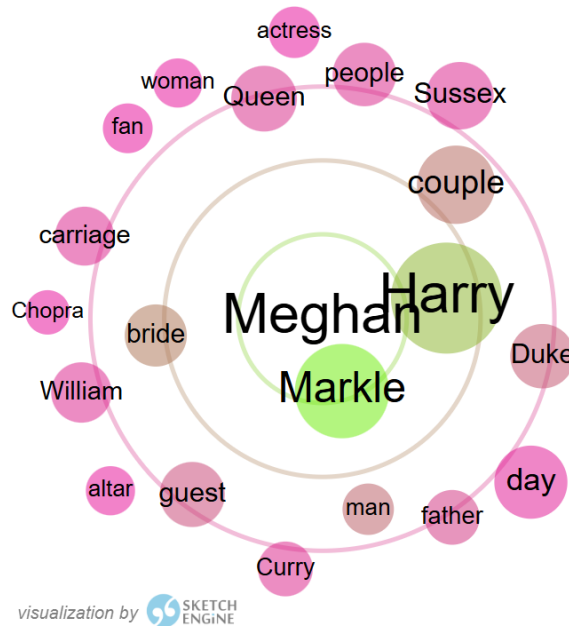


Figure 15. The thesaurus for the word “Meghan” generated from the articles covering the wedding of Meghan Markle

In contrast to the highly emotional language in almost every article covering the wedding of Meghan Markle, those about Catherine's wedding have shown a prominent difference regarding this issue. Instead of the excessive use of poetic expressions and metaphors, in the articles about Princess of Wales the authors mostly opted for formal vocabulary, highlighting the noble status of the couple and following the protocol rules:

“Britain's second-in-line to the throne was married Friday in a ceremony that lived up to all pomp and pageantry that the world has come to expect of Britain's royal family” (Voa News, 2011).

“History was made when Prince William married Kate Middleton in the biggest royal wedding of the century on April 29, 2011” (Brides, 2020).

“Prince William and Catherine Middleton became husband and wife Friday morning—at London's Westminster Abbey” (People, 2011).

“Miss Catherine Middleton of Bucklebury, Berkshire went into Westminster Abbey at 11am and came out an hour and a quarter later, holding the hand of the second in line to the throne as Her Royal Highness the Duchess of Cambridge” (The Guardian, 2011).

“After the church ceremony, the newly married couple—now the Duke and Duchess of Cambridge—traveled in an open-top carriage to Buckingham Palace, from where they waved” (VOA News, 2011). The institutional and ceremonial language in this case was preferred to highlight the royal hierarchy. Conversely to the emotionally

saturated image of Meghan, Catherine is presented as a restrained figure, appreciating tradition and duty.

Another prominent peculiarity of the articles covering Catherine's wedding is the highlight of attention the couple had received from the public on the day of the ceremony:

"The world seemingly ground to a halt to witness the Prince and Princess of Wales' historic wedding on 29 April 2011" (Hello, 2025).

"They were cheered by 500,000 well-wishers who gathered outside the palace, as RAF planes flew past in honour of the new royal couple" (BBC, 2011).

"It was an historic royal wedding, watched by an estimated one billion people around the world" (Hello, 2025).

"The ceremony was viewed live by tens of millions more around the world, including 72 million live streams on YouTube" (BBC, 2011).

"Thousands of people in the streets of London cheered Friday afternoon as the royal newlyweds, Prince William and Kate Middleton, made their way back from a brief rest to Buckingham Palace for a buffet dinner and all-night party with about 300 guests" (ABC News, 2011)

"Fans of the couple have posted edited clips from the day ahead of their 12th wedding anniversary next month, one of which shows a brief encounter between the two as they prepared to leave Westminster Abbey for Buckingham Palace in an open-top horse-drawn carriage" (Newsweek, 2023).

"British flags waved in the air and painted faces, royal hats, frocks, dress jackets swarmed the streets" (Voanews, 2011).

"We're all part of history. It's just a wonderful experience" (Voa news, 2011). Immense numbers and highlight on public excitement creates the impression of mass approval of the event and the unity of the couple with the public, whereas the coverage of Meghan's wedding mostly described the long list of high-profile guests.

Catherine's dress and make up received a lot more attention from the media than the one of Meghan. Each of the articles described the Princess's outfit in detail, focusing even on minute nuances:

"Middleton paired the Alexander McQueen dress with Queen Elizabeth II's Cartier Halo Tiara, which the queen's father, King George VI, commissioned in 1936 for his wife shortly before he assumed the throne" (Women's Wear Daily, 2025).

"Thought to have cost Carole and Michael Middleton £250,000, the bride's gown featured handmade lace created by the Royal School of Needlework and an ivory satin Victorian-inspired corset bodice" (Hello, 2025).

"The bride opted for a simple creation by Sarah Burton at Alexander McQueen with lace sleeves, a swagged train, and a fine veil of Chantilly lace" (The Guardian, 2011).

"Kate wore a now-iconic long-sleeve lace gown by Sarah Burton for Alexander McQueen and had her hair half-up, half-down, secured with a tiara" (Brides, 2023).

“For the momentous event, Kate commissioned Sarah Burton of the London fashion house of Alexander McQueen to create her highly anticipated wedding dress” (Newsweek, 2023).

“Middleton arrived at Buckingham Palace wearing a strapless white satin Sarah Burton evening gown with a circle skirt and diamante embroidered detail round the waist, according to The Associated Press” (ABC News, 2011). The excessive emphasis on bride’s look benefits the creation of Catherine’s aesthetic idealization, contributing to the image of a fairy-tail princess beyond the common person’s reach. Alongside with highlighted historical origin of some items, such features of the articles shape the mold of traditional institutionalized femininity, which Catherine undoubtedly fits.

The multiple descriptions of the Princess’s makeup also bear a hidden agenda:

“For one of the most anticipated events in modern history, it's surprising to discover that the Princess of Wales decided to do her own bridal makeup” (Woman and Home, 2024).

“Kate did her own makeup for her wedding, and she went with dark liner and plenty of blush on the cheeks” (Cosmopolitan, 2018).

“Middleton’s makeup featured soft, natural tones with emphasis on her eyes, complementing her sophisticated bridal look” (Vogue, 2011). The highlighted statement *“the Princess of Wales decided to do her own bridal makeup”* creates the humane and authentic image, exposing Catherine’s propriety and aspiration for natural beauty.

Unlike Meghan Markle, the Princess of Wales didn’t receive much attention towards her background, it was only mentioned in a couple of articles with an aim to contribute to Catherine’s down-to-earth personality: *“For country girl Kate, the floral budget was huge” (Woman and Home, 2023).* However, Catherine was also referred to as “Middleton” in a few articles, which downplayed the personality. Nevertheless, such reference was not significantly common:

“Middleton told William in the carriage, “I’m so happy,” according to a lip reader” (ABC News, 2011).

“Leading up to the event, much speculation surrounded which designer Middleton would tap for her gown” (Women’s Wear Daily, 2025).

The assumptions made while performing the CDA were also backed up with corpus analysis tools. The list contains a significant number of words regal words (*tiara* (4th place), *royal* (10th) *palace* (21st)), proving the point about using the ceremonial language to highlight the couple’s hierarchy in the royal family. A number of words, describing the Princess’s outfit back up the point about its significance for creating her public image (*McQueen* (6th), *lace* (29th), *gown* (36th), *satin* (41st)). The other noteworthy mentionings concern the wedding-related vocabulary and the mentioning of Princess Diana (less common in the articles covering Kate’s wedding), the feature the articles about Meghan and Catherine share.

SINGLE-WORDS ✓ MULTI-WORD TERMS ✓

reference corpus: English Web 2021 (enTenTen21)

Lemma	Lemma	Lemma	Lemma	Lemma
1 Middleton ...	11 Westminster ...	21 Palace ...	31 Catherine ...	41 satin ...
2 Buckingham ...	12 Prince ...	22 Armstrong-Jones ...	32 Pettifer ...	42 Lowther-Pinkerton ...
3 Getty ...	13 Cutsem ...	23 Markle ...	33 diamante ...	43 Sommerlath ...
4 tiara ...	14 abbey ...	24 wedding ...	34 Camilla ...	44 Straubenzee ...
5 duchess ...	15 bride ...	25 Eugenie ...	35 Burton ...	45 Earthshot ...
6 McQueen ...	16 nuptials ...	26 bridesmaid ...	36 gown ...	46 Carrickmacross ...
7 Kate ...	17 Princess ...	27 Cartier ...	37 Carole ...	47 strapless ...
8 Abbey ...	18 Meghan ...	28 Diana ...	38 Tiara ...	48 Lopes ...
9 Pippa ...	19 newlywed ...	29 lace ...	39 pageantry ...	49 restyled ...
10 royal ...	20 well-wisher ...	30 Crawford-Smith ...	40 Newsweek ...	50 rule-breaking ...

Figure 16. The keywords generated from the articles covering the wedding of Kate Middleton

The wordlist of nouns also managed to prove a few presuppositions made earlier (Figure 17). In this case, the words *prince* and *William* turned out to be used significantly more often than the words *Kate* and *Middleton*, making the role of the bride less prominent than the one of Meghan in the articles covering her wedding. The use of regal vocabulary along with the lexical units, referring to the bride's dress still stay the distinct features of the majority of the articles.

noun (1,845 items | 7,632 total frequency)

Noun	Frequency ? ↓	Noun	Frequency ? ↓	Noun	Frequency ? ↓	Noun	Frequency ? ↓
1 prince 213 ...		14 bride 71 ...		27 year 36 ...		40 father 24 ...	
2 wedding 208 ...		15 family 63 ...		28 charles 36 ...		41 world 24 ...	
3 william 196 ...		16 getty 56 ...		29 diana 36 ...		42 mcqueen 23 ...	
4 kate 158 ...		17 westminster 55 ...		30 wales 33 ...		43 balcony 23 ...	
5 middleton 121 ...		18 day 54 ...		31 lace 33 ...		44 burton 23 ...	
6 image 94 ...		19 catherine 51 ...		32 april 33 ...		45 friend 23 ...	
7 couple 89 ...		20 buckingham 46 ...		33 elizabeth 32 ...		46 gown 22 ...	
8 princess 86 ...		21 guest 45 ...		34 credit 32 ...		47 alexander 22 ...	
9 queen 85 ...		22 people 43 ...		35 duke 30 ...		48 wife 22 ...	
10 abbey 80 ...		23 london 43 ...		36 cambridge 28 ...		49 uk 22 ...	
11 royal 78 ...		24 duchess 40 ...		37 crowd 28 ...		50 reception 22 ...	
12 palace 76 ...		25 king 39 ...		38 ceremony 27 ...			
13 dress 72 ...		26 harry 39 ...		39 time 27 ...			

Figure 17. The wordlist of nouns generated from the articles covering the wedding of Catherine Middleton

Finally, the corresponding thesaurus was generated for the name *Kate* (Figure 18). The thesaurus again contained the words referring to the wedding outfit. The peculiar findings were the words related to family (*mother*, *family*, *father*), their number was significantly bigger than in the one of Meghan. Also, the words *Britain* and *world* are noteworthy, as they expose the contrast between the images of Catherine and Meghan in the media.

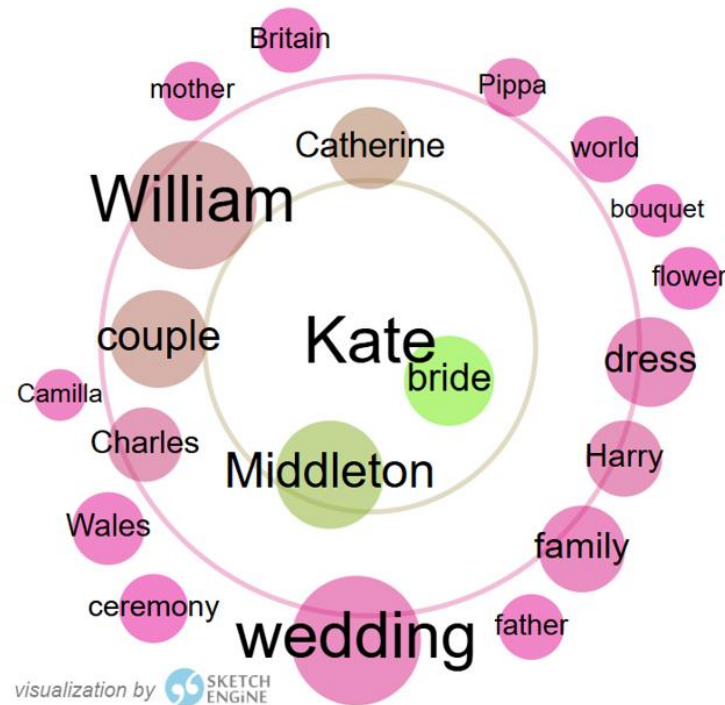


Figure 18. The thesaurus for the word “Kate” generated from the articles covering the wedding of Catherine Middleton.

The findings, extracted from the corpora analysis, significantly helped to support the assumptions made with the help of CDA. It showcased the successful cooperation of both methods for the sake of unveiling the media bias.

All in all, the most prominent discrepancies between the coverage of Meghan’s and Catherine’s wedding included the choice of language: more poetic for Meghan and more formal for Catherine. Meghan’s wedding coverage emphasized modernity, diversity and focused on the couple’s sincere feelings, whereas Kate’s wedding coverage reinforced royal continuity, tradition, and elegance. Meghan’s identity as a mixed-race American actress is frequently highlighted, whereas Kate is framed as a British aristocratic figure, perfect princess and the embodiment of elegance. Regarding the present similarities in the portrayal of weddings, both have an emphasis on the bride and the prominent accent on wedding-related vocabulary (bride, groom, best man, newlyweds). Moreover, both events are described as global spectacles, however, Catherine’s wedding is framed as more universally traditional, with an emphasis on public integration to the event, while Meghan’s is more celebrity-driven, focusing on the guest list and definite personalities from it.

3.3. Corpus and pragmatic analysis results.

Corpus analysis has proved to be an indispensable method for conducting this research. Its remarkable ability to back up any assumptions with numbers was of use

for a multiple number of times. The most widespread tools that showed their highest significance for various aspects of the research included the generation of keywords list, thesaurus and wordlists for the particular parts of speech. Sketch Engine platform has proved to be able to allow the successful employment of all the mentioned tool and provide the reliable data beneficial for the uncovering of the media bias.

Throughout the research, 6 corpora including the articles on 6 different topics were compiled. One of this topic included the articles covering the first royal engagement of both Catherine and Meghan as princes' wives. A corpus analysis was performed on two separate corpora (articles about Catherine's royal engagement and articles about Meghan's first royal engagement), exposing the following results.

The wordlist of adjectives generated from the articles about Meghan's engagement, has showcased the Duchess of Sussex as a modern, innovative and progressive member of the royal family, stressing her ability to bring something fresh to a rather reserved establishment the royal family was considered to be (Figure 19). This interference may be supported by the presence of such adjectives: *first, new, innovative, future, young, etc.*

"Megan and Prince Harry holding hands at a royal engagement is a refreshingly modern approach to their new role both as a couple and as representatives of the royal family" (People, 2017).

"Both are bonding with the newest addition to the family" (Business Insider, 2019).

"It's now home to another young royal family" (Vogue, 2022).

"Harry was figuring out what his future role in the royal family might look like" (Vogue, 2022). All of the quotes underscore the novelty and historical significance of Meghan's appearance, framing it as a milestone in royal history.

Another noteworthy feature of the wordlist concerns the mixture of the adjectives *beautiful, perfect, and natural* typically used to create the standardized, yet superficial female portrait with the adjectives like *excited, keen, and genuine* help to build an image of Meghan as a warm, relatable, and emotionally authentic person.

"She is just as beautiful as she is on TV. She is absolutely perfect, most wonderful and perfect for him" (People, 2017).

"She's beautiful and a modern woman" (The Guardian, 2017).

"Lessons have been learned from the past and everyone is very keen that Meghan should feel included from the beginning" (9 News, 2017)

"In the clip, Markle, excited, violated a major protocol" (The Royal Observer, 2025). Such a choice of adjectives, benefits the creation of the image of an energetic and emotional woman, not ready to walk behind, but eager to take action.

Another pivotal notion, concerning the choice of adjectives in the articles is the presence of words referring to privacy: *private, personal, public*.

"...features personal video clips and photographs that capture their life together in their bolthole" (Vogue, 2022).

visualization, highlights several central themes: institutional framing, relational identity, along with public visibility.

Firstly, Meghan is consistently framed within the royal institution. Adjectives like *royal*, *duchess*, *monarch* appeared most frequently in the corpus, positioning her within a formal, hierarchical structure.

“Queen Elizabeth II and Meghan Markle are the ultimate royal power couple” (9 News, 2017).

“Meghan 'broke royal protocol' on first engagement with the Queen” (Express, 2025).

“...there was a row after a video showed the newly created Duchess entering a car before the monarch” (Daily Mail, 2018). These terms anchor her identity within traditional royal narratives, suggesting that her early public appearances were seen less as individual milestones and more as symbolic steps into the monarchy.

Secondly, Meghan’s identity is strongly relational. The most prominent co-occurrences are with *Harry*, *couple*, *pair* and *home*, which indicates that media coverage tended to emphasize her role as a partner and family member over that of an autonomous figure.

“Chanting 'Harry, Harry!' and 'Meghan, Meghan!' hundreds of people lined the streets of a central English city Friday to welcome Britain's Prince Harry and his American fiancée, actress Meghan Markle.” (China Daily, 2017).

“...the couple were nervous about how the press would cover it” (Daily Mail, 2025)

The pair meet in London through friends and begin a relationship (The Guardian, 2017). Meghan’s portrayal is shaped through her romantic connections and relationship status rather than her personal voice or achievements.

A third noticeable trend is the emphasis on public appearance and engagement. Words like *event*, *visit*, *people*, *fan*, *coat*, and *selfie* suggest that the media spotlight was directed at how she presented herself and interacted with the public.

“Meghan had to follow the plethora of long-established rules and habits that were in place for such events” (Daily Mail, 2025).

“The royal Twitter has thoroughly documented their visit, from what a natural Markle is in front of the crowd to how chic she dressed inside” (Elle, 2017).

“There are fan videos of Markle outside as well, charming everyone” (Elle, 2017).

“Turning your back to the Duke and Duchess and taking a selfie is discouraged if at all possible” (Business Insider, 2017).

These terms, often tied to visual and performative aspects of her role, underscore how Meghan was depicted as part of a carefully managed spectacle. All aspects of her behavior (fashion choices, crowd interactions, and symbolic destinations) all subject to scrutiny.

The identical analysis was performed on the basis of the articles about Catherine's first engagement, showcasing the differences and similarities between the portrayal of both personalities. A wordlist of nouns was generated (Figure 21). The analysis of the most frequent nouns in articles covering Catherine Middleton's first royal engagement reveals several dominant themes that shaped the media narrative. These themes reflect how the public and press constructed her identity at the outset of her royal life.

noun (1,345 items | 4,755 total frequency)

Noun	Frequency ? ↓	Noun	Frequency ? ↓	Noun	Frequency ? ↓	Noun	Frequency ? ↓
1 kate	197 ***	14 crowd	40 ***	27 april	22 ***	40 story	17 ***
2 william	146 ***	15 day	39 ***	28 event	22 ***	41 life	17 ***
3 prince	119 ***	16 diana	38 ***	29 trearddur	22 ***	42 cambridge	17 ***
4 middleton	85 ***	17 anglesey	34 ***	30 visit	21 ***	43 world	16 ***
5 engagement	66 ***	18 image	34 ***	31 duchess	20 ***	44 relationship	16 ***
6 couple	65 ***	19 camilla	33 ***	32 university	20 ***	45 ceremony	15 ***
7 lifeboat	59 ***	20 princess	27 ***	33 insider	19 ***	46 palace	15 ***
8 charles	55 ***	21 appearance	26 ***	34 andrew	18 ***	47 official	15 ***
9 year	52 ***	22 february	26 ***	35 coat	18 ***	48 business	15 ***
10 getty	45 ***	23 bay	25 ***	36 home	18 ***	49 champagne	15 ***
11 royal	43 ***	24 wedding	25 ***	37 endeavour	18 ***	50 rescue	15 ***
12 wales	43 ***	25 people	24 ***	38 hereford	18 ***		
13 time	41 ***	26 family	24 ***	39 mili	18 ***		

Figure 21. The wordlist of nouns generated from the articles covering the first royal engagement attended by Catherine Middleton

The most frequent nouns are *Kate* (196), *William* (145), *prince* (122), and *Middleton* (85). It indicates a strong focus on her relationship with Prince William, yet the spotlight is still on the Princess, even though the difference is insignificant. The prominence of the noun *couple* (65) reinforces this idea, suggesting that their partnership was central to the story, rather than Catherine as an individual.

“There were people in the crowd who had traveled from further afield, but most were from the nearby area that had embraced the royal couple as one of their own” (Town and Country, 2021).

“But within seconds, she appeared to have charmed a crowd of more than 1,000 people, who had been waiting hours in a cold and blustery wind for a glimpse of the couple – and, in particular, for a peek at the bride-to-be” (Daily Mail, 2011).

“And as the couple told The Telegraph in 2010, the breakup ended up being good for their relationship” (Business Insider, 2022).

The frequent mentions of *Charles*, *Diana*, and *Camilla* further situate Kate within the broader royal family narrative.

“.... his relationship with Kate Middleton was seemingly influenced by King Charles' relationships” (Business Insider, 2022).

“Princess Diana was more popular than husband Prince Charles, which reportedly upset him, during their visit to Wales after their honeymoon” (Nine News, 2021).

“...Christmas Day service with Prince Harry, Prince Charles and Camilla , Duchess of Cornwall (Business Insider, 2022). Such mentionings brought Catherine`s closer to the other senior members of the family, benefiting her integration into the royal family.

It`s hard to underestimate the prominence of ceremonial vocabulary in the articles. Words like *engagement* (66), *royal* (43), *duchess* (20), and *princess* (31), *engagement* (66), *royal* (43), *duchess* (20), and *princess* (31) highlight the traditional and ceremonial aspects of Catherine`s role.

“Remember Kate Middleton's first ever official royal engagement? How the Duchess of Cambridge entered royal family life alongside husband Prince William” (OK, 2018).

“...due to William's future role as Prince of Wales, and Kate to likely become Princess of Wales” (Nine News, 2021).

“And the new Duke and Duchess of Cambridge went on to triumph on tour in Canada and California” (Mirror, 2011). This suggests the media focused heavily on her new status and the symbolic importance of her royal duties.

Lastly, nouns such as *family* (23), *home* (18), *life* (17), and *relationship* (16) hint at an underlying narrative about Kate's personal life and her integration into royal domesticity.

“Almost seven years ago, it was Kate Middleton who made her official debut alongside then-fiancé Prince William when she joined him for their formal debut in Anglesey, where they made their first home” (People, 2017)

“Eager to cement the relationship, Middleton, who may on Charles's accession become Princess of Wales, wore a Welch Fusiliers badge in the designer beret she chose for the occasion” (The Guardian, 2011).

“William introduced his fiancé as "someone who is not only about to join the family, but is also about to become an 'Anglesonian'" (Nine News, 2021). Such representation of Catherine hints her subordinate role in the family, yet, the full integration and acceptance in the family. These terms humanize her and foster public relatability.

The thesaurus based on the articles covering Catherine`s first public engagement, has proved the notions, assumed during the analysis of the wordlist (Figure 22). The thesaurus reflects the semantic neighborhood of the word *Kate*, meaning the words that co-occur with it most frequently in context. This provides insight into how she is portrayed and positioned within media discourse.

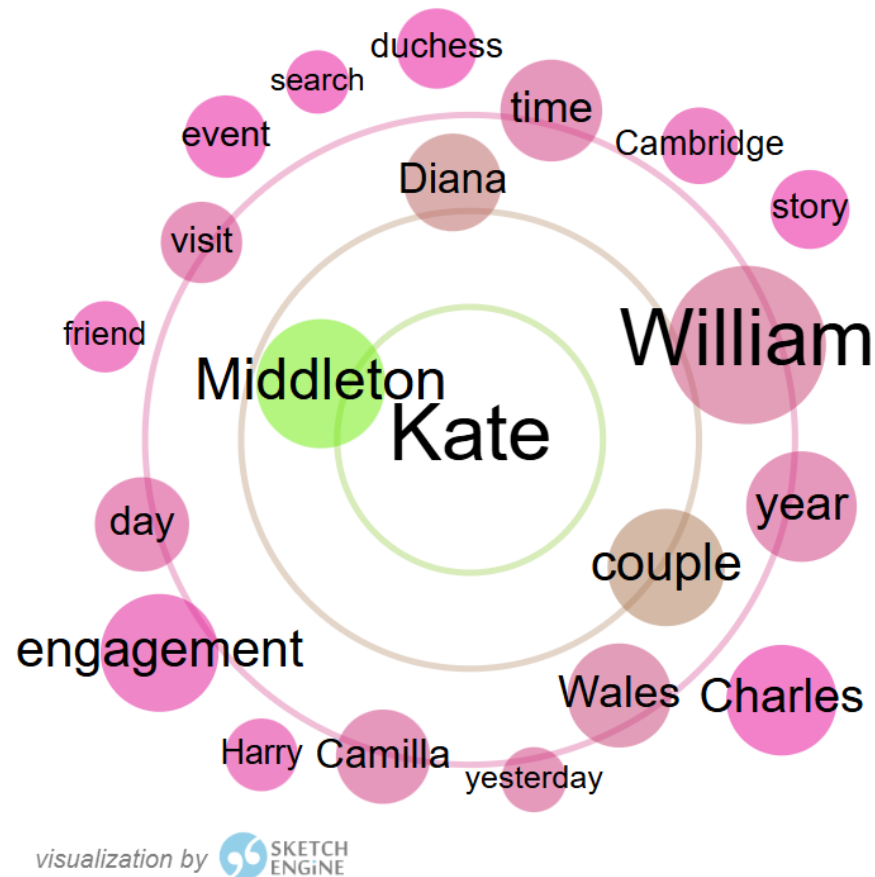


Figure 22. The thesaurus for the word “Kate” generated from the articles covering the first royal engagement of Catherine Middleton

These names clearly embed Kate within the royal lineage, tying her identity closely to the institution. William’s very high frequency especially emphasizes the media’s framing of Kate in relational terms—as a partner, rather than an autonomous figure.

“Kate Middleton, and 29 years old, the soon-to-be royal joined Prince William for a day out in Wales” (Nine News, 2021).

“The Duchess of Cambridge entered royal family life alongside husband Prince William” (OK, 2017)

“After bidding the crowd “prynhawn da” – good afternoon – William introduced his fiancée in a brief speech” (The Guardian, 2011).

A similar role is performed by the heavy use of *couple* and *engagement*, which underlines the media’s emphasis on romantic milestones as defining moments in Kate’s story.

“The bride and groom-to-be formally opened Darwen Aldridge Community Academy together, where they watched a dance performance, and then paid a visit to Witton Country Park, where the sporty couple cheered on sprinters” (Hello, 2021).

“It was their first joint royal event since their engagement was announced months earlier” (Nine News, 2021).

The corpus analysis of media coverage surrounding Meghan Markle and Catherine Middleton’s first royal engagements reveals distinct patterns in how each woman is discursively constructed in public narratives. While both are introduced into the royal sphere through their romantic associations, the linguistic choices in the respective corpora point to divergent representations grounded in identity, values, and roles.

In the case of Kate, the most frequent co-occurring nouns included *William*, *Middleton*, *engagement*, *couple*, and *royal*, which clearly positioned her within the institution of monarchy and relational identity. Her persona is framed through her proximity to Prince William and other royal family members (Charles, Diana, Camilla). It emphasizes tradition, continuity, and duty. The language used to describe Kate showcases her as a figure seamlessly integrating into royal life. The presence of nouns *duchess* and *Cambridge* stresses her institutional role, while less frequent usage of words to describe emotions create a somewhat restrained portrait.

Conversely, the adjectives and vocabulary surrounding Meghan contribute to a more emotional and independent image. The articles highlight terms such as *personal*, *private*, and *public*. This reflects on ongoing conflicts regarding her privacy. Moreover, the emphasis on expressive verbs and adjectives associated with feelings (*loving*, *passionate*, *ambitious*) as well as references to her career achievements create an image of Meghan as a modern, independent woman. Her identity is less institutional and more autonomous. While Kate is meant to embody traditional values through service and duty, Meghan is framed as relatable and grounded, not able to fit the royal mold.

Overall, the corpus and pragmatic analysis present a similar tendency to CDA. Kate is represented as dutiful and completely adapted to royal expectations, whereas Meghan is positioned as emotionally open and career-focused woman, challenging the demands of the institution. These early discursive portraits set the tone for their ongoing media coverage, shaping public perceptions in ways that continue to influence their ongoing coverage in the media.

Another case in which the corpus analysis was of use concerned the performance of the pragmatic analysis. In order to create the foundation for this type of analysis, the articles covering the relationship announcement of both personalities were selected and two separate corpora created. After that, two wordlists of verbs were compiled and the most frequent verbs were later analyzed (Figures 23, 24).

verb (1,092 items | 10,457 total frequency)

Verb	Frequency ? ↓	Verb	Frequency ? ↓	Verb	Frequency ? ↓	Verb	Frequency ? ↓
1 be	2,277 ***	14 want	87 ***	27 date	47 ***	40 seem	37 ***
2 have	905 ***	15 announce	87 ***	28 find	47 ***	41 write	36 ***
3 say	225 ***	16 meet	83 ***	29 include	45 ***	42 reveal	32 ***
4 do	222 ***	17 think	74 ***	30 feel	45 ***	43 ask	32 ***
5 go	154 ***	18 share	71 ***	31 join	43 ***	44 put	30 ***
6 make	142 ***	19 give	69 ***	32 accord	42 ***	45 turn	30 ***
7 take	138 ***	20 come	65 ***	33 use	41 ***	46 decide	30 ***
8 get	133 ***	21 leave	64 ***	34 start	41 ***	47 appear	30 ***
9 know	115 ***	22 attend	60 ***	35 keep	41 ***	48 add	30 ***
10 become	108 ***	23 follow	57 ***	36 celebrate	39 ***	49 photograph	29 ***
11 tell	96 ***	24 wear	53 ***	37 work	38 ***	50 watch	28 ***
12 look	93 ***	25 spend	50 ***	38 enjoy	37 ***		
13 see	90 ***	26 marry	48 ***	39 move	37 ***		

Figure 23. The wordlist was generated based on articles covering the relationship announcement of Kate Middleton and Prince William

verb (1,092 items | 10,457 total frequency)

Verb	Frequency ? ↓	Verb	Frequency ? ↓	Verb	Frequency ? ↓	Verb	Frequency ? ↓
1 be	2,277 ***	14 want	87 ***	27 date	47 ***	40 seem	37 ***
2 have	905 ***	15 announce	87 ***	28 find	47 ***	41 write	36 ***
3 say	225 ***	16 meet	83 ***	29 include	45 ***	42 reveal	32 ***
4 do	222 ***	17 think	74 ***	30 feel	45 ***	43 ask	32 ***
5 go	154 ***	18 share	71 ***	31 join	43 ***	44 put	30 ***
6 make	142 ***	19 give	69 ***	32 accord	42 ***	45 turn	30 ***
7 take	138 ***	20 come	65 ***	33 use	41 ***	46 decide	30 ***
8 get	133 ***	21 leave	64 ***	34 start	41 ***	47 appear	30 ***
9 know	115 ***	22 attend	60 ***	35 keep	41 ***	48 add	30 ***
10 become	108 ***	23 follow	57 ***	36 celebrate	39 ***	49 photograph	29 ***
11 tell	96 ***	24 wear	53 ***	37 work	38 ***	50 watch	28 ***
12 look	93 ***	25 spend	50 ***	38 enjoy	37 ***		
13 see	90 ***	26 marry	48 ***	39 move	37 ***		

Figure 24. The wordlist was generated based on articles covering the relationship announcement of Meghan Markle and Prince Harry

An effective tool for detecting bias in the media is pragmatic analysis, particularly one of its branches, which is Speech Act Theory. This particular tool was applied to analyze and juxtapose the articles marking the beginning of romantic relationship between Catherine Middleton and Prince William and Meghan Markle and Prince Harry.

Each verb was analyzed according to five types of speech acts. According to John Austin, speech acts are divided into locutionary acts (the actual act of uttering sounds and words) and illocutionary acts (assigning meaning to these linguistic units) (Nordquist, 2024). It was the illocutionary acts that were analyzed in the course of this study, categorized according to John Searle's classification (Zaib, Mahmood, 2023).

During the research, it was found that assertives, declaratives, and directives in the articles about both figures showed no striking differences, which is explained by the function of such articles and the positioning of both women (both being in relationships with members of the royal family).

Thus, sentences containing these types of speech acts are mostly similar and provide general information about the couples' relationship history.

“As the pair spent more and more time together, the prince said they found commonalities between them.” (People, 2024) “We’re a couple,” Markle said in the October issue of Vanity Fair.” (Business Insider, 2017)

In both examples, the authors used the most common assertive *say* with the same purpose: to convey direct quotes of the speakers. Another example of an assertive used similarly is the verb *reveal*. However, it’s worth noting that in the articles about Meghan Markle and Prince Harry, this verb appeared far more frequently (67 times in articles about Meghan vs. 32 in those about Kate), which creates an impression that the couple was more eager to keep their relationship private.

“Kate Middleton's brother James reveals how the Princess of Wales told him she was marrying Prince William” (Tatler, 2024) “Prince Harry and Meghan Markle reveal details of how they met” (Hello, 2024)

The directive *announce* was frequently used in articles about both women, as these articles often aimed to inform the public about the couples' relationship status.

“William, 24, and 25-year-old Kate had been widely expected to announce their engagement and one bookmaker even stopped taking bets on a betrothal earlier this year.” (BBC, 2010)

“Buckingham Palace announced that Prince Harry and Meghan would spend the holidays with Meghan's mother.” (People, 2024)

Despite the similar way information is presented, some differences were still observed, mainly regarding commissives and expressives, especially in terms of how frequently they were used. This aspect can signal media bias and influence public perception of the two women.

The most frequently used commissives in articles about Catherine, Princess of Wales, included: *leave, date, marry, join, start, work*. In contrast, in articles about Meghan, Duchess of Sussex, the commonly used commissives were slightly different: *date, work, write, step, move, find*. This distinction may affect the reader and create a certain impression of each woman. Catherine appears as someone ready to become part of the royal family, whereas Meghan is portrayed as an independent individual with a vivid career background. This conclusion is supported by the following examples:

“Kate previously joined William on a family holiday back in 2005, but took great care not to be seen with the royals in front of any cameras.” (Cosmopolitan, 2025) — commissive join was used to express unity of Kate with her partner.

“I’m willing to learn quickly and work hard,” Middleton told reporter Tom Bradby during their official engagement interview.” (Vogue, 2022) — commissive work, in the meaning of fulfilling royal duties.

"We were very quietly dating for about six months before it became news, and I was working during that whole time." (USA Today, 2017) — Meghan Markle's quote using commissives date and work. Work here refers to her personal career endeavors.

"Us Weekly wrote that Prince Harry made a top-secret visit to Toronto in October, where Markle was filming her show, and claimed it was..." (Brides, 2024) — write used as a commissive, one of the most frequent in articles about Meghan, showing the excessive media intrusion into the couple's private life.

Expressives also significantly affect public perception of Meghan and Kate. The choice and frequency of their use differ and help build two contrasting images of the women. The most common expressives in articles about Kate include: *enjoy, celebrate, feel, help, delight*. On the other hand, expressives about Meghan were slightly different: *feel, love, celebrate, hope, enjoy*.

This choice of expressives portrays Catherine as someone content with her role and devoted to it. It creates an image of the "ideal princess" for the British public. However, the articles' authors emphasise neither her emotions nor romantic feelings.

In contrast to it, Meghan's portrayal is more romanticized and centered on love and feelings, rather than responsibilities associated with her new royal role.

"Kate, in turn, enjoyed spending time at all of the various family homesteads William brought her to." (E! News, 2024) — *enjoy*, the most frequently used expressive in articles about Catherine, indicating her joy from spending time with the royal family.

"At the time, it felt like a fantastical moment in a real-life fairy tale." (Vogue, 2022) — *feel*, commonly used in articles about Meghan, here expressing joy and the uniqueness of the moment she experienced as the future royal.

"Markle admits that she's dating the prince and says they're 'in love'." (Geo News, 2024) — *love*, indicating Meghan's openness about her strong romantic feelings.

"When you're talking to her, you feel like you're the only person on the planet. And it's just wonderful to see her so in love." (Vanity Fair, 2017) — *feel*, used to convey deep emotional affection and intimacy.

Despite the fact that only one part of speech (verbs) was chosen for analysis, it was still possible to identify media bias present from the very first articles about the Princess of Wales and the Duchess of Sussex. While the authors of these articles frequently used similar assertives, declaratives, and directives due to the shared purpose of reporting on relationships with royal figures, it is the difference in the frequency and nature of commissives and expressives that allows us to trace distinctions in the verbal portrayal of each woman. The choice of commissives and expressives in articles about Catherine creates the image of a person devoted to her duties, someone who finds happiness in serving her family and country. At the same time, romantic feelings are not emphasized in articles about the Princess of Wales. It contributes to the creation of a reserved, elevated image. The portrayal which is more commonly associated with political figures and public personalities. Catherine's verbal portrait presents her as a woman who finds joy and satisfaction in supporting her partner's family.

In contrast to it, the commissives and expressives used in articles about Meghan highlight the strength of her emotions and portray her as a down-to-earth, relatable individual. Someone whose romantic feelings are familiar to everyone. These articles place strong emphasis on Meghan's career, accomplishments, and professional background prior to her relationship, while references to duty and service to the royal family are rare.

Thus, corpus-based and pragmatic analysis of verbs reveals Catherine as a woman happy to embrace new responsibilities and become part of the royal family, and Meghan as an independent individual focused on her career and openly expressive of her deep feelings for her future husband. This image, constructed at the very beginning of their respective relationships, remains relevant and is reinforced by newer media coverage of both women.

3.4. Multimodal discourse analysis

Multimodality is a crucial notion, which shall be studied for the profound and comprehensive analysis of media sources, as, apart from lexical choices, media bias concerns images, headlines and even the layout of the article, which makes these aspects pivotal to be researched. To perform the analysis, the major tabloids of the United Kingdom and the USA were taken into consideration, in order to observe the notion at its fullest. The articles covering the engagement of Catherine and William and Meghan and Harry were selected to study the notion of multimodality. Even though each article includes the same information, the approach to its presentation was different in each specimen designated for the analysis. We have analyzed such tabloids and magazines as The Guardian, The USA Today, British Vogue, The Irish Independent, The Brides, People, The New York Times and Daily Mail.

The overview of the structure of the articles about the engagement of Meghan Markle is the following. The majority of the articles followed a particular structure, starting with the basic information about the engagement (details of the future wedding, how the couple felt, where it took place, etc.) accompanying it with the official engagement portrait of a couple or the slideshow of pictures (Figure 25, 26, 27).

Prince Harry Is Engaged to Meghan Markle!

The royal proposed to his American actress girlfriend—and she said yes!

By [Erin Hill](#) and [Simon Perry](#) | Published on November 27, 2017 05:12AM EST



Figure 25. People, 2017

Full transcript of Prince Harry and Meghan Markle's engagement interview

The two announced their engagement today.

By ABC News
November 27, 2017, 11:38 PM



The real-life fairy tale Meghan Markle shares her experiences with the royal family.

Prince Harry 'thrilled' after announcing he is to marry Meghan Markle

Couple appear in public for first time since announcing their engagement and give 20-minute television interview



Figure 26. The ABC News, 2017

Figure 27. The Guardian, 2017

Some sources, on the other hand, provided the readers with the image of the official statement of the palace, instead of the couple's photo (Figure 27).

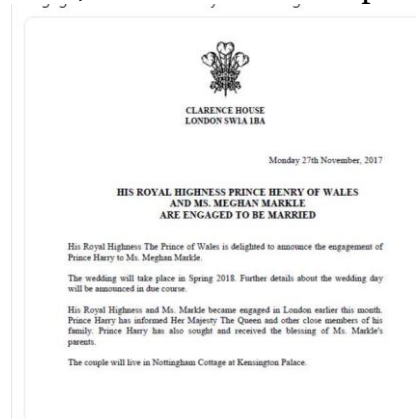


Figure 28. The New York Times, 2017

Vogue presented a detailed overview of the couple's engagement interview: "The couple, who were at ease with each other's company in front of television cameras, revealed that they have met with the Queen several times with Markle adding that she is a "remarkable woman (*Vogue*, 2017)". The overview was accompanied by the multiple screenshots from the interview, as well as couple's official engagement portraits (Figure 29). The full transcript of the interview was also provided by the ABC News, yet, it was not complemented with the pictures.



Figure 29. *Vogue*, 2017

The Brides focused on their dating history: “*The pair were set up on a blind date by a mutual friend and fashion designer, Misha Nonoo (The Brides, 2017)*” and presented it in the chronological order, using sub headlines to highlight every milestone of the couple’s relationship (Figure 30).

May 2016: A Blind Date

Not much is known about Prince Harry and Meghan's first date, but one thing is for certain. The pair were set up on a blind date by a mutual friend and fashion designer, Misha Nonoo. Matchmaker, matchmaker, make us a royal match next time, OK?

October 31, 2016: The Secret's Out

The day it all began—at least for the public. The *Sunday Express* reported that Prince Harry had been secretly dating the *Suits* star, and according to one source the royal was “happier than he’s been for many years.” That very day, Meghan posted an adorable banana couple hugging on her Instagram page, followed by a high-tea snap on November 1.

November 1, 2016: A Cryptic Quote

Canada's *Vancouver Sun* published a Q&A in which Meghan discussed her collection for retailer Reitmans. Her closing quote, while cryptic, sure said a lot. “So, my cup runneth over and I’m the luckiest girl in the world!” she shared.

Figure 30. *The Brides*, 2017

Most of the articles added the pictures taken during the official photoshoot to their articles. The majority of them included the close up of the Meghan’s hand with the engagement ring, and the information about it was presented below (Figure 31). The passage describing it was included in several right after the basic information. This is how it was presented by *The Daily Mail*: “*Meghan's engagement ring's two outer diamonds are from Diana's own collection with a central diamond from Botswana - where they went on safari in September - all set within a gold band*” (*The Daily Mail*, 2017). The ABC News presented the picture in the following way: “*Britain's Prince*

Harry holds hands with Meghan Markle wearing an engagement ring in the Sunken Garden of Kensington Palace, London” (ABC News, 2017).



Figure 31. The Daily Mail, 2017

However, a more significant part of the authors had decided to provide the detailed description of couples` family reactions to the news of the engagement, showcasing the family photos of the couple (Figure 32).



Figure 32. People, 2017

For instance, “*Prince William and Kate Middleton also sent their well-wishes in a statement: “We are very excited for Harry and Meghan. It has been wonderful getting to know Meghan and to see how happy she and Harry are together” (The Guardian, 2017).* The other section found in the main body of several articles focused on the Duchess of Sussex's background and former life, with the USA Today being the only outlet to place this information at the beginning “*Here comes the next royal bride: Meghan Markle, the first divorced, biracial, Jewish American actress to join the British royal family (the USA Today, 2017).* Three sources provided detailed accounts of the couple’s early conflict with the paparazzi. Notably, these sections were considerably

longer than others, with *The Guardian* devoting substantial attention to the issue: *“Markle, whose mother is African American and father is white, also described as “disheartening” and “discriminatory” some of the media coverage she received as Prince Harry’s girlfriend because it centred on her racial background”* (*The Guardian*, 2017). *People*, *The Brides* and *The USA today* had also attached some photos made by paparazzi at a certain period in the chronological order (Figure 33).



PHOTO: JAMES WHATLING/MEGA

Figure 33. *People*, 2017

The concluding sections of the articles published by *The Guardian* and *Vogue* featured comparisons between Meghan Markle and Princess Diana. “When asked how he thinks his mother would have liked his 36-year-old fiancée, the 33-year-old prince responded: *“Oh they’d be thick as thieves, without question, I think she would be over the moon, jumping up and down, you know so excited for me, but then, as I said, would have probably been best friends - best friends with Meghan”* (*Vogue*, 2017). Most authors chose to once again emphasize the bride-to-be’s background, offering glimpses into her past and including photos from her earlier life. *The New York Times*, in particular, referenced the Duchess’s previous career: *“She is best known for her roles in the television legal drama “Suits” and in the science-fiction series “Fringe,” and she appeared on the cover of Vanity Fair in September”* (*The New York Times*, 2017). It is important to highlight the difference in the length of articles and passages. Statistically, the passages covering the conflict, the bride’s background, and dating history were the most detailed, while the information about the engagement process and the upcoming wedding was relatively brief in all articles, except for those in *Vogue* and the *Irish Independent*. Most articles featured official portraits and close-up shots of the engagement ring, but some also included childhood photos, paparazzi pictures, and a copy of the official statement issued by the Palace when the news was announced.

When commenting on the announcement of Catherine's and William's engagement, it is crucial to note, that due to a significant time difference (Harry and Meghan's relationship began 7 years later), the coverage of the event was less ubiquitous and the number of multimedia resources, accompanying the articles is significantly smaller.

Most of the articles are compiled using the same structure: brief information on the event accompanied by the official portrait (Figure 34): *"The much-anticipated statement from Clarence House dropped into the press's inboxes on 16 November 2010: His Royal Highness Prince William of Wales was engaged to be married to his long-term girlfriend Miss Catherine Middleton!"* (Hello, 2023)

Prince William and Catherine Middleton's Official Engagement Portrait

Mario Testino captures the modern romance of Prince William and Miss Catherine Middleton, joining an illustrious lineup of *Vogue* photographers and artists who have captured the classical elegance and iconic style of a century of royal brides and fiancés

BY HAMISH BOWLES
December 12, 2010



Figure 34. Vogue, 2010

The copy of the official statement, however, was more widespread, and included in several resources (Figure 35). On the other hand, the close-up of the ring was not common, only BBC provided it how the readers, yet, the quality of the image was poor (Figure 36). And Daily Mail added the description of the ring underneath the official portrait: *"For the engagement ring, William presented Kate with a 12-carat oval blue Ceylon sapphire, surrounded by 14 solitaire diamonds, which had belonged to his late mother Princess Diana"* (Daily Mail, 2024).



Figure 35. Hello, 2023



Figure 36. BBC, 2010

Rather than focusing on the bride's background, most of the media resources described all the details on the upcoming wedding: *"No venue has been announced yet. For pomp itself, the ceremony is likely to fall between the extraordinary spectacle of Charles and Diana's 1981 wedding in St. Paul's Cathedral and Charles' subdued second marriage to Camilla, Duchess of Cornwall at Windsor Guildhall in 2005"* (Silive, 2010). Most of the media resources chose to emphasize this information using italics or the contrasting color of the text (Figure 37).

Following the marriage, the couple will live in north Wales, where Prince William will continue to serve with the Royal Air Force.

Figure 37. The Guardian, 2010

The traditional engagement interview was transcribed by The Guardian and The BBC, yet, video was not attached to the articles, which can be justified by the time period, when the news were published (Figure 38).

"We are both very, very happy," the prince said. He had proposed on holiday in Kenya in October, giving his fiancée his mother's engagement ring.

Miss Middleton said joining the Royal Family was a "daunting prospect".

"Hopefully I'll take it in my stride," she added.

The couple, both 28, met at St Andrews University, Fife.

William said using Diana's ring was "my way of making sure my mother didn't miss out on today and the excitement".

Prince Charles said he was "thrilled, obviously" about the engagement, joking: "They've been practising long enough."

Kate Middleton's parents Michael and Carole said they were "absolutely delighted" and "thrilled" by the news.

Figure 38. BBC, 2010

It also was not common to share multiple engagement portraits, in contrast to the engagement of Meghan and Harry, most media sources just shared the engagement portrait or the statement of the palace, without any additional images (Figure 39). Some also provided the screen shot from the official couple's interview (Figure 40).



Figure 39. Daily Mail, 2024



Prince William and Kate Middleton.

Figure 40. People, 2023

Some sources, however, commented on the reaction of couple's families to the occasion: "*The formal statement said William's father, Prince Charles, was "delighted"* (*The Guardian*, 2010). The Daily Mail was the only source to provide the engagement

portrait of Princess Diana and King Charles to juxtapose the two couples (Figure 41). The articles covering Meghan`s engagement, did not include the pictures of this kind.



Figure 41. The Daily Mail, 2010

Undoubtedly, the most prominent discrepancy in the coverage of both figures is the lack of information that may form biased treatment. The articles about Catherine`s engagement included no details on her former life, no scandals, no references to the pictures taken by paparazzi. At the same time, even the multimedia in the articles about Meghan were meant to make the readers emotional, including paparazzi pictures, photos from the childhood and the whole collection of engagement portraits. However, the chronological difference between the events could have also contributed the development of discrepancies in the multimodal discourse of both personalities.

CONCLUSIONS

The present thesis set out to examine the linguistic construction of media bias in the portrayals of Catherine, Princess of Wales, and Meghan, Duchess of Sussex, using four methodological approaches: Critical Discourse Analysis (CDA), Speech Act Theory, Corpus Linguistics, and Multimodal Discourse Analysis (MDA). The analysis was conducted on a dataset comprising 170 media articles and 200 visual elements (images, headlines, and layouts), with equal representation of the two figures. The material was sourced from major British and American media outlets including *BBC*, *The Guardian*, *The New York Times*, *USA Today*, *People*, *Vogue*, and *The Daily Mail*. Each article focused on major public events — engagements, weddings, and public appearances — and was selected to ensure comparability and methodological consistency.

Critical Discourse Analysis, based on Fairclough’s three-dimensional model, was performed on 70 articles dedicated specifically to the weddings of Catherine and Meghan. The analysis focused on discursive themes, lexical features, social norms, ideological framing, and intertextual references. In articles about Meghan Markle, emotionally charged language was used extensively. Examples include metaphors such as “*fairy-tale wedding*,” “*loverbirds*,” and idiomatic expressions like “*melt our hearts*”. Meghan was referred to by her surname alone — “Markle” — in 8 of the 30 analyzed articles, whereas Catherine was mostly identified by her title or full name. Meghan’s background as a former actress and a divorcee was mentioned repeatedly, with the terms “*divorcee*” and “*former Suits actress*”. What is more, there were at least five references to Princess Diana, they often involved a connection to jewelry, bouquets, or visual imagery. These mentions were meant to draw the distinct line between Meghan and her mother-in-law. In contrast to that, articles on Catherine emphasized traditions and royal titles. Most of the articles included detailed descriptions of attire, with references to designers, fabrics, and historical accessories. Moreover, the public significance of the wedding was consistently highlighted in the half of the articles (e.g. “*We’re all part of history. It’s just a wonderful experience*”), conversely, the articles about Meghan mostly emphasized the star-studded guest-list: “*The historic moment was witnessed by 600 guests—including Oprah Winfrey, George and Amal Clooney, and Elton John — a flag-waving crowd of around 100,000 outside, and millions of TV viewers around the world*”. The similar features included the prominent accent on the brides, in comparison with the grooms and the emphasis on the wedding-related vocabulary present in the vast majority of articles.

Corpus analysis was carried out using Sketch Engine and consisted of generating wordlists, conducting keyword analysis, and exploring lexical clustering through thesaurus functions. Corpus tools reinforced the findings of the Critical Discourse Analysis for both Meghan and Catherine. In Meghan’s dataset, Markle and Meghan topped the keyword list, while Prince and Harry ranked 13th and 23rd. Background-

related terms such as Doria (7th) and Suits (21st) were notable, as were wedding-specific terms like newlywed (9th), nuptials (20th), and bridesmaid (24th). The presence of celebrity names—Winfrey, Amal, Elton, Clooney, Beckham—emphasized the glamorous framing of the event. The noun wordlist confirmed frequent references to royal terminology (duke, duchess, queen) and to mother and Diana, with the word dress appearing only 39 times, indicating limited focus on Meghan's attire. The thesaurus for Meghan included both royal and personal references (carriage, actress, Chopra, guest), supporting earlier CDA observations. In contrast to it, Catherine's corpus highlighted ceremonial language, with tiara (4th), royal (10th), and palace (21st) among key terms. Fashion-related keywords such as McQueen (6th), lace (29th), and gown (36th) confirmed the emphasis on her appearance. However, Prince and William were mentioned more often than Kate or Middleton, suggesting a less central narrative role. The thesaurus for Kate included more family-related words (mother, father, family) and broader framing terms like Britain and world, indicating a national and global context to her portrayal.

A corpus-based analysis of media discourse on Meghan Markle and Catherine Middleton's early royal engagements identifies distinct linguistic patterns. References to Catherine frequently include institutional and relational nouns such as William, Middleton, engagement, and royal, indicating a lexical focus on her integration into the monarchy. The discourse in the articles about her is mostly formal with the limited usage of language. In contrast to it, Meghan's coverage features adjectives and verbs related to emotion and individuality, including personal, private, ambitious, and loving. Moreover, multiple references to her career are also included. This suggests a lexical emphasis on personal identity and emotional expression. The linguistic data reflect differing discursive constructions: Catherine is associated with institutional roles, while Meghan is characterized through affective and autonomous descriptors.

Speech Act Theory, following Searle's classification, was used to categorize verb usage into assertives, directives, commissives, expressives, and declaratives. The pragmatic analysis, based on Speech Act Theory and using Searle's classification, identified no major differences in the distribution of assertive, declarative, and directive speech acts across the two corpora. Verbs such as say and announce were commonly used in both, primarily to report direct speech or official statements. However, notable variation was observed in the frequency and type of commissive and expressive verbs. In articles about Catherine, frequently used commissives included join, marry, start, work, and leave. These verbs often described actions oriented toward relational or institutional alignment. Common expressives in the same corpus included enjoy, celebrate, feel, help, and delight, with usage often tied to events or public duties rather than emotional intimacy. In contrast, the Meghan corpus featured commissives such as date, work, write, step, move, and find, with a higher proportion relating to personal or professional autonomy. Expressives such as feel, love, celebrate, hope, and enjoy were more frequently used, often in emotionally charged or introspective contexts. These

findings indicate measurable linguistic differences in verb choice and speech act types, particularly in commissives and expressives, across the two corpora. These differences are observable at the level of frequency, lexical type, and pragmatic function.

Multimodal Discourse Analysis was applied to 200 visual materials. Articles about Meghan included a higher proportion of informal images, such as candid photos, paparazzi shots, and personal moments. Those about Catherine mostly included official portraits and the palace's statements, without any visual content which possibly could contribute the creation of controversy.

The findings across all four methods demonstrate measurable differences in media portrayal. Meghan's representation was more personalized, informal, and emotionally loaded. Apart from that, it includes repeated references to her American background and former lifestyle. Catherine's portrayal emphasized royalty, tradition, and institutional alignment, supported by formal language and visual presentation.

These results reflect consistent patterns within the dataset. However, the further framework may be developed. Firstly, the corpus could be expanded to include material beyond the engagement and wedding periods, especially coverage related to later events such as the coronation, the Oprah interview, or Prince Harry's memoir. A longitudinal or diachronic approach would allow the tracking of narrative shifts over time.

Secondly, a comparative study of tabloid versus broadsheet publications will be beneficial, as it could quantify the degree of bias. To perform such study, the articles shall be categorized and the frequency alongside with lexical units is to be compared. Similarly, ideological orientation (e.g., liberal versus conservative) could also be analyzed.

Thirdly, cross-cultural comparison presents a valuable opportunity. Non-English media representations of Catherine and Meghan — in French, German, Spanish, or Ukrainian outlets — could be included to assess how different sociocultural contexts influence discursive framing.

Fourth, audience reception could be taken into consideration. For instance, comment sections or social media responses to the same articles could be analyzed. Such research would showcase how readers interpret discourse. This would help assess whether the media framing matches public perception and how audience engagement varies depending on framing strategies.

Finally, intersectional analysis based on gender or race could be employed. Scholars could focus on sociolinguistics to perform this analysis. This may help showcase the level of interest in the portrayal of royal women in the media.

In summary, this research has proven that this kind of research requires multiple methods to be accurate enough. With a combination of corpus tools, CDA, pragmatic analysis, and multimodal interpretation, it has been documented how the discursive portraits of Catherine and Meghan are constructed and how media bias can be uncovered with the help of linguistics.

REFERENCES

1. *LancsBox X*. (n.d.). Lancsbox.lancs.ac.uk. <https://lancsbox.lancs.ac.uk>
2. A Multimodal Discourse Analysis of Pictures in ELT Textbooks: Modes of Communication in Focus. (2021). *Jordan Journal of Modern Languages and Literatures*, 13(4), 623–644. <https://doi.org/10.47012/jjml.13.4.2>
3. Piskorska, A. (2016). Perlocutionary effects and relevance theory. *Pragmatics & Beyond New Series*, 287–305. <https://doi.org/10.1075/pbns.268.11pis>
4. Lewandowska-Tomaszczyk, B., & Mammadov, A. (2022). *Analysing Media Discourse: Traditional and New*. Cambridge Scholars Publishing.
5. Bhatia, A. (2012). Critical Discourse Analysis: History and New Developments. *The Encyclopedia of Applied Linguistics*. <https://doi.org/10.1002/9781405198431.wbeal0277>
6. Brandt, S. (2004). Gunther Kress & Theo van Leeuwen, *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold, 2001. *Language in Society*, 33(01). <https://doi.org/10.1017/s0047404504221054>
7. Carpentier, N., De Cleen, B., & Van Brussel, L. (2019). Introduction: Communication and Discourse Theory, 1–32. <https://doi.org/10.2307/j.ctv36xvhx4.3>
8. Carvalho, A. (2008). Media(ted) discourse and society. *Journalism Studies*, 9(2), 161–177. <https://doi.org/10.1080/14616700701848162>
9. Chapelle, C. A. (2020). *The concise encyclopedia of applied linguistics*. Wiley Blackwell
10. Chen, E. L. Austin 1962 how to do things with words.
11. Chouliaraki, L., & Fairclough, N. (2021). *Discourse in Late Modernity*. Edinburgh University Press. <https://doi.org/10.1515/9780748610839>
12. *keyword* | *Sketch Engine*. (б.д.). *Sketch Engine*. <https://www.sketchengine.eu/glossary/keyword/>
13. Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
14. Fairclough, N. (1995). *Critical Discourse Analysis: the Critical Study of Language*. Routledge.
15. Fairclough, N. (2003). *Analysing Discourse : Textual Analysis for Social Research* / N. Fairclough..
16. Fendel, V. B. (2024). The I.Sicily Sketch Engine Corpus. *Journal of Open Humanities Data*, 10. <https://doi.org/10.5334/johd.258>
17. Fillmore, C. J. (2011). “Corpus linguistics” or “Computer-aided armchair linguistics.” *De Gruyter EBooks*. <https://doi.org/10.1515/9783110867275.35>
18. Firth, J. R. (1935). The technique of semantics. *Transactions of the Philological Society*, 34(1), 36–73. <https://doi.org/10.1111/j.1467-968x.1935.tb01254.x>

19. Forceville, C. J. (2010). The Routledge Handbook of Multimodal Analysis. *Journal of Pragmatics*, 42(9), 2604–2608. <https://doi.org/10.1016/j.pragma.2010.03.003>
20. Principles, Theories and Approaches to Critical Discourse Analysis – Discourse Studies. Discourse Studies. <https://discoursestudies.net/principles-theories-and-approaches-to-critical-discourse-analysis>
21. Freed, A. F. (2015). Institutional Discourse. *The International Encyclopedia of Language and Social Interaction*, 1–18. <https://doi.org/10.1002/9781118611463.wbielsi151>
22. González-Lloret, M. (2012). Pragmatics: Overview. *The Encyclopedia of Applied Linguistics*. <https://doi.org/10.1002/9781405198431.wbeal1338>
23. Green, G. M., Bach, K., & Harnish, R. M. (1983). Linguistic Communication and Speech Acts. *Language*, 59(3), 627. <https://doi.org/10.2307/413908>
24. Grice, H. P. (1975). *Logic and Conversation* (pp. 41–58). Academic Press. <https://www.ucl.ac.uk/lis/studypacks/Grice-Logic.pdf>
25. Gries, S. T. (2014). Quantitative corpus approaches to linguistic analysis: seven or eight levels of resolution and the lessons they teach us. *Cambridge University Press EBooks*, 29–47. <https://doi.org/10.1017/cbo9781139833882.005>
26. Han, C. (2015). How to Do Critical Discourse Analysis: A Multimodal Introduction. *Australian Journal of Linguistics*, 35(4), 415–418. <https://doi.org/10.1080/07268602.2015.1033673>
27. Harris, D. W., Fogal, D., & Moss, M. (2018). *Speech Acts*. In Oxford University Press eBooks. Oxford University Press. <https://doi.org/10.1093/oso/9780198738831.003.0001>
28. Herring, S. (n.d.). Discourse in Web 2.0: Familiar, Reconfigured, and Emergent. <https://homes.luddy.indiana.edu/herring/GURT.2011.prepub.pdf>
29. Nordquist, R. (2009). *What Is Corpus Linguistics?* ThoughtCo. <https://www.thoughtco.com/what-is-corpus-linguistics-1689936>
30. Huang, Y. (2016). Introduction. In Y. Huang (Ed.), *Oxford Handbooks Online*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199697960.013.33>
31. Hunston, S. (2002). *Corpora in Applied Linguistics*. <https://doi.org/10.1017/cbo9781139524773>
32. Laurence Anthony's Website. (2023). <http://www.laurenceanthony.net/software/antconc>
33. Janks, H. (1997). Critical Discourse Analysis as a Research Tool. *Discourse: Studies in the Cultural Politics of Education*, 18(3), 329–342. <https://doi.org/10.1080/0159630970180302>
34. Jewit, C. (2009): “Different approaches to Multimodality”, in C. Jewit (ed.) (2005): *The Routledge Handbook of Multimodal Analysis*. London: Routledge Taylor & Francis Group.

35. Jones, R. H. (2012). Multimodal Discourse Analysis. *The Encyclopedia of Applied Linguistics*. <https://doi.org/10.1002/9781405198431.wbeal0813>
36. Kazemian, B., & Hashemi, S. (2014). Critical Discourse Analysis of Barack Obama's 2012 Speeches: Views from Systemic Functional Linguistics and Rhetoric. *Theory and Practice in Language Studies*, 4(6). <https://doi.org/10.4304/tpls.4.6.1178-1187>
37. Kilgarriff, A. (n.d.). Thesauruses for natural language processing. *International Conference on Natural Language Processing and Knowledge Engineering, 2003. Proceedings. 2003*, 5–13. <https://doi.org/10.1109/nlpke.2003.1275859>
38. Kilgarriff, A., & Iztok Kosem. (2012). Corpus tools for lexicographers. *Oxford University Press EBooks*, 31–56. <https://doi.org/10.1093/acprof:oso/9780199654864.003.0003>
39. Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1(1), 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
40. Kilgarriff, A., Marcowicz, F., Smith, S., & Thomas, J. (2015). Corpora and Language Learning with the Sketch Engine and SKELL. *Revue Française de Linguistique Appliquée*, XX(1), 61. <https://doi.org/10.3917/rfla.201.0061>
41. Kim, Y. (2007). Paul Baker: Using Corpora in Discourse Analysis. *Continuum, Applied Linguistics*, 28(2), 327–330. <https://doi.org/10.1093/applin/amm006>
42. Kress, G., & Selander, S. (2012). Multimodal design, learning and cultures of recognition. *The Internet and Higher Education*, 15(4), 265–268. <https://doi.org/10.1016/j.iheduc.2011.12.003>
43. Kuebler, S., & Zinsmeister, H. (2014). *Corpus Linguistics and Linguistically Annotated Corpora*. Bloomsbury Publishing
44. Lali Tavadze, Ivdit Diasamidze, Natia Katamadze, & Davitadze, L. (2024). Modern tendencies in media discourse. *International Journal of Innovative Technologies in Social Science*, 1(41). https://doi.org/10.31435/rsglobal_ijitss/30032024/8123
45. Laurence Anthony. (2013). A critical look at software tools in corpus linguistics. *Linguistic Research*, 30(2), 141–161. <https://doi.org/10.17250/khisli.30.2.201308.001>
46. Leech, G. N. (2012). *Change in contemporary English : a grammatical study*. Cambridge University Press
47. Lindquist, H. (2014). *Corpus linguistics and the description of English*. Edinburgh University Press
48. Litosseliti, L. (2010). *Research methods in linguistics*. Bloomsbury Academic
49. Litosseliti, L. (2024). *Research Methods in Linguistics*. Bloomsbury Publishing
50. Luca, I.-S. (2020). A Multimodal Discourse Analysis in Media. *Romanian Journal of English Studies*, 17(1), 74–80. <https://doi.org/10.1515/rjes-2020-0009>

51. Luca, I.-S. (2020). A Multimodal Discourse Analysis in Media. *Romanian Journal of English Studies*, 17(1), 74–80. <https://doi.org/10.1515/rjes-2020-0009>
52. Maachi, H., & Khamar, H. (2025). The Contribution of Corpus Linguistics and Natural Language Processing Tools in the Development of School Lexicography. *Communications in Computer and Information Science*, 41–55. https://doi.org/10.1007/978-3-031-80438-0_4
53. Martín, J., Tutor, I., & Cendón, B. (2015). Ideology through Musical Discourse: A Critical Discourse Analysis on Three of the Most Controversial Rock/Metal Songs during the 1980s. https://uvadoc.uva.es/bitstream/handle/10324/18919/TFG_F_2016_46.pdf
54. McEnery Tony McEnery. (2019). *Corpus Linguistics*. Edinburgh University Press
55. McEnery, T., & Hardie, A. (2011). *Corpus Linguistics*. <https://doi.org/10.1017/cbo9780511981395>
56. McEnery, T., & Hardie, A. (2012). *Corpus Linguistics*. Cambridge University Press
57. Nordquist, R. (2024). Speech-Act Theory: Definition and Examples. ThoughtCo. <https://www.thoughtco.com/speech-act-theory-1691986>
58. O'keeffe, A. (2006). *Investigating Media Discourse*. Routledge.
59. Olejniczak, J. (2018). Using corpora to aid qualitative text analysis. *doaj (DOAJ: Directory of Open Access Journals)*. <https://doi.org/10.15503/jecs20182.154.164>
60. Pearce, K. J. (n.d.). Media and Mass Communication Theories. *Encyclopedia of Communication Theory*. <https://doi.org/10.4135/9781412959384.n231>
61. Rafiah, S. (2018). Understanding CDA: Histories, Remembering, And Futures. *Script Journal: Journal of Linguistic and English Teaching*, 3(2), 175. <https://doi.org/10.24903/sj.v3i2.170>
62. Rahimi, F., & Riasati, M. (2011). Critical Discourse Analysis: Scrutinizing Ideologically-Driven Discourses. *International Journal of Humanities and Social Science*, 1(16). https://www.ijhssnet.com/journals/Vol_1_No_16_November_2011/13.pdf
63. Rheindorf, M., & Springerlink (Online Service. (2019). *Revisiting the Toolbox of Discourse Studies : New Trajectories in Methodology, Open Data, and Visualization*. Springer International Publishing
64. Ross, R. (1970). On Declarative Sentences. In R. A. Jacobs, & P. S. Rosenbaum (Eds.), *Readings in English Transformational Grammar* (pp. 222-272). Xerox College Publishing.
65. Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4), 696–735. <https://doi.org/10.2307/412243>
66. Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*, 53(2), 361. <https://doi.org/10.2307/413107>

67. Schollon, R. (2007). *Analyzing Public Discourse*. Routledge. <https://doi.org/10.4324/9780203375211>
68. Schröder, K. C. (1998). Book reviews : Norman Fairclough, *Media Discourse*. London: Edward Arnold, 1995. 214 pp. ISBN 0-340-63222-4 (hbk); ISBN 0-340-58889-6 (pbk. *European Journal of Cultural Studies*, 1(2), 285–290. <https://doi.org/10.1177/136754949800100209>
69. Searle, J. R. (1969). *Speech Acts*. <https://doi.org/10.1017/cbo9781139173438>
70. Searle, J. R., & Vanderveken, D. (1985). *Foundations of illocutionary logic*. Cambridge University Press.
71. Sriwimon, L., & Zilli, P. J. (2017). Applying Critical Discourse Analysis as a Conceptual Framework for Investigating Gender Stereotypes in Political Media Discourse. *Kasetsart Journal of Social Sciences*, 38(2), 136–142. <https://doi.org/10.1016/j.kjss.2016.04.004>
72. Stubbs, M. (1997). Whorf’s Children: Critical Comments on Critical Discourse Analysis. In A. Ryan, & A. Wray (Eds.), *Evolving Models of Language* (pp. 100–116). *Multilingual Matters*.
73. Swales, J. M. (2014). 1990. Genre analysis: English in academic and research settings. Cambridge: Cambridge University Press, selected 45–47, 52–60. *The Discourse Studies Reader*, 306–316. <https://doi.org/10.1075/z.184.513swa>
74. Tan, S., O’Halloran, K., & Wignell, P. (2020). Multimodality. *The Cambridge Handbook of Discourse Studies*, 263–281. <https://doi.org/10.1017/9781108348195.013>
75. Thuy, T. T. H. (2017). *READING IMAGES - THE GRAMMAR OF VISUAL DESIGN*. *VNU Journal of Foreign Studies*, 33(6). <https://doi.org/10.25073/2525-2445/vnufs.4217>
76. Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. J. Benjamins.
77. Tognini-Bonelli, E. (2002). Corpus Linguistics at Work. *Computational Linguistics*, 28(4), 583–583. <https://doi.org/10.1162/coli.2002.28.4.583a/>
78. Brezina, V. (2018). *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge University Press.
79. van Dijk, T. A. (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>
80. van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. Sage Publications, Inc.
81. van Dijk, T. A. (2015). Critical discourse analysis. *The Handbook of Discourse Analysis*, 2(1), 466–485. <https://doi.org/10.1002/9781118584194.ch22>
82. van Dijk, T. A. (2015). Critical discourse analysis. *The Handbook of Discourse Analysis*, 2(1), 466–485. <https://doi.org/10.1002/9781118584194.ch22>

83. Vaughan, E., & O’Keeffe, A. (2015). Corpus Analysis. *The International Encyclopedia of Language and Social Interaction*, 1–17. <https://doi.org/10.1002/9781118611463.wbielsi168>
84. What Are the Different Types of Discourse? (with pictures). (n.d.). Language Humanities. <https://www.languagehumanities.org/what-are-the-different-types-of-discourse.htm>
85. White, P. (1977). Death, Disruption and the Moral Order: the narrative impulse in mass-media “hard news” reporting [Review of Death, Disruption and the Moral Order: the narrative impulse in mass-media “hard news” reporting]. *Genres and Institutions: Social Processes in the Workplace and School*, 101–133. [https://prwhite.info/prwhite,%201997,%20Death,%20Disruption%20and%20the%20Moral%20Order%20\(in%20news\).pdf](https://prwhite.info/prwhite,%201997,%20Death,%20Disruption%20and%20the%20Moral%20Order%20(in%20news).pdf)
86. Wikipedia Contributors. (2025, April 30). Sketch Engine. Wikipedia; Wikimedia Foundation
87. Wodak, R. (2019). The Trajectory of Far-Right Populism – A Discourse-Analytical Perspective. Routledge EBooks, 21–37. <https://doi.org/10.4324/9781351104043-2>
88. Wodak, R., & Busch, B. (2004). Approaches to Media Texts. *The SAGE Handbook of Media Studies*, 105–122. <https://doi.org/10.4135/9781412976077.n6>
89. Wodak, R., & Meyer, M. (2009). Critical Discourse Analysis: History, Agenda, Theory, and Methodology. In R. Wodak, & M. Meyer (Eds.), *Methods for Critical Discourse Analysis* (pp. 1, 33). London: Sage.
90. Wodak, R., & Meyer, M. (2015). *Methods of critical discourse studies*. Sage.
91. Wolfram, W. (2021). Sociolinguistic variation and the public interest. *Cadernos de Linguística*, 2(1), e357. <https://doi.org/10.25189/2675-4916.2021.v2.n1.id357>
92. Wordlists - word frequency lists | Sketch Engine. (2019, May 14). Sketch Engine. <https://www.sketchengine.eu/guide/wordlist-frequency-lists/#toggle-id-1>
93. Wright, J. (2015). *International Encyclopedia of the Social & Behavioural Sciences* (2nd ed., pp. 238–245). Elsevier
94. Zaib, Z., Shah, S. K., & Ilyas Mahmood, M. (3 C.E.). A Comparative Analysis of Searle’s Speech Act Theory and Cohen’s Model: An Exploration of the Social Contexts in Which Explicit and Implicit Speech Acts Are Used [Review]. *Journal of Development and Social Sciences*, 4(3), 1224–1235. [https://doi.org/10.47205/jdss.2023\(4-III\)114](https://doi.org/10.47205/jdss.2023(4-III)114)