



Constructing Political Reality: Corpus-Based Evidence of Framing and Ideological Language in Brexit Media Discourse

**Mehrdad Vasheghani Farahani ^{1*} ,
Mahmoud Elshahat Abdelghaffar Omar ²**

¹ *Lancaster University, UK*

² *Leipzig University, Germany*

Received: 2026/06/29

Accepted: 2026/07/27

Abstract: Brexit prompted an extraordinary volume of media coverage, characterized by ideological conflict and competing political narratives. Using a 125-million-word corpus of Brexit-related texts, this study applies a corpus-based critical discourse analysis (CDA) approach to examine how language actively constructs political meaning rather than simply reporting events. The findings identify four dominant discursive frames: (1) campaign-related discourse, (2) representations of political actors and institutions, (3) evaluative and ideological language, and (4) future-oriented projections. Together, these frames shape a highly polarized interpretive landscape in which Brexit emerges not merely as a political process but as a contested ideological symbol. The analysis demonstrates how framing strategies and evaluative lexical choices work collectively to produce political meaning within contemporary hybrid media environments. By highlighting the role of language in shaping public perceptions and political identities, the study contributes to a deeper understanding of how ideological discourse and evaluative stance foster polarization in modern media ecologies. The findings also emphasize the importance of critical media literacy and illustrate the value of corpus-assisted discourse analysis for investigating political conflict and public debate in networked societies.

Keywords: Brexit Referendum, Leave Vote, Critical Discourse Analysis, Corpus Linguistics, Ideological Language, Media Analysis.

* Corresponding Author.

Authors' Email Address:

¹ **Mehrdad Vasheghani Farahani** (Mehrdadfarahani1365@gmail.com), ² **Mahmoud Elshahat Abdelghaffar Omar** (omar.mahmoud@uni-leipzig.de)



Introduction

The 2016 Brexit referendum, in which 51.9% of participants voted to leave the European Union while 48.1% chose to remain, marked a watershed moment in modern British political history (Bloom et al., 2019). Beyond its numerical outcome, the referendum exposed deep-seated political, regional, and socio-economic divisions within the United Kingdom. In the years that followed, the country experienced prolonged political negotiations, legal uncertainty, economic volatility, and sustained public debate. Issues such as trade regulation, immigration policy, national sovereignty, and legislative autonomy became central to both public discourse and policymaking (Baines et al., 2020; de Ruyter & Nielsen, 2019). The United Kingdom's formal withdrawal from the European Union on January 31, 2020, marked the beginning of a transitional period that continued to influence debates surrounding the long-term implications of Brexit.

The linguistic representation of Brexit provides a valuable case through which the interaction between language, ideology, and political communication can be systematically examined using corpus-assisted Critical Discourse Analysis (Eagleton, 2018). As Kelly (2018) argues, Brexit evolved beyond a political event to become a symbolic battleground over questions of British identity, belonging, and the future direction of the nation. Media discourse across both traditional and digital platforms played a pivotal role in shaping these debates through competing and often conflicting narratives (Zappettini & Krzyżanowski, 2021). While some media representations framed Brexit as a democratic restoration of national sovereignty, others emphasized themes of fragmentation, uncertainty, and social division (McConnell & Tormey, 2020; Šimunjak, 2022).

Drawing on the complementary strengths of corpus linguistics and Critical Discourse Analysis (CDA), this study adopts a corpus-assisted CDA framework to examine the linguistic representation of Brexit across a large-scale media corpus. Corpus linguistic techniques provide a systematic means of identifying statistically salient lexical patterns, collocational behaviour, and concordance structures, thereby reducing reliance on impressionistic textual observations and enhancing the transparency and replicability of discourse analysis (Vasheghani Farahani & Kazemian, 2021). CDA, in turn, provides the interpretive framework through which these recurring linguistic patterns can be examined in relation to ideological positioning, political identities, and competing representations of Brexit (Baker et al., 2008). Rather than treating quantitative and qualitative approaches as separate analytical traditions, the study adopts a sequential analytical design in which corpus linguistic methods identify

statistically salient linguistic patterns that subsequently guide qualitative discourse interpretation.

While CDA provides the interpretive perspective adopted in this study, the present research does not seek to verify CDA's foundational assumption that discourse contributes to the construction of social reality (Weninger, 2008). Instead, it applies CDA as an established analytical framework through which corpus-derived linguistic evidence can be interpreted systematically (Fairclough, 1995; Wodak & Meyer, 2016). By integrating quantitative corpus techniques with qualitative discourse interpretation, the study illustrates a transparent and replicable analytical procedure for investigating ideological framing in political communication (Baker et al., 2008; Baker, 2023; Partington et al., 2013). Accordingly, the contribution of the study lies not in advancing a new theoretical claim about discourse, but in demonstrating how corpus-assisted CDA can be operationalized to examine a major political event through a systematic combination of keyword analysis, collocation analysis, concordance analysis, and critical discourse interpretation.

Furthermore, the analytical value of the study does not lie in identifying themes and frames in relation to Brexit. Rather, the study investigates how these complementary framing dimensions interact to produce an integrated representation of Brexit across a large media corpus. Examining these dimensions collectively could provide a more comprehensive account of ideological framing than analyses focusing on individual discourse features in isolation, while simultaneously illustrating the methodological potential of corpus-assisted Critical Discourse Analysis for investigating complex political communication. To address these objectives, the study is guided by the following two research questions:

RQ1: How is Brexit discursively constructed in mass media discourse, and what linguistic strategies and ideological framings characterize its representation?

RQ2: What sociopolitical values and ideological assumptions are embedded in Brexit-related media discourse, and how do these discursive features shape public understanding and political alignment?

Literature Review

Brexit has generated substantial scholarly attention within discourse studies, particularly in the fields of political linguistics, media discourse, corpus linguistics, and critical discourse analysis. Existing research has examined Brexit from a wide range of perspectives, including political identities, media framing, metaphor, evaluative language, nationalism, and ideological positioning (Miglbauer & Koller, 2018; Morales Pérez, 2019; Wenzl, 2020; Zanca, 2020;

[Parnell, 2021, 2024](#); [Hawkins, 2022](#), among others). Collectively, these studies demonstrate that Brexit discourse extends well beyond political reporting and constitutes an important site in which competing interpretations of national identity, sovereignty, democracy, and Britain's relationship with Europe are discursively negotiated.

One important strand of research has examined ideological representations of the European Union within British media discourse. By way of illustration, [Alkhamash \(2020\)](#), for example, compares pro-Leave and pro-Remain media coverage and demonstrates that the two camps construct markedly different representations of the European Union. Leave-oriented discourse portrays the EU as authoritarian, economically damaging, and incompatible with British interests, whereas Remain-oriented discourse emphasizes economic interdependence and shared European values. These findings illustrate how media discourse contributes to the ideological polarization of political debate through contrasting framing strategies.

Another body of scholarship has focused on the linguistic construction of political identities and public positioning during the referendum campaign. As a case in point, [Miglbauer and Koller \(2018\)](#) investigate voter motivations and identities in televised vox-pop interviews, demonstrating how ordinary speakers construct competing understandings of British identity, democratic legitimacy, and political belonging. Similarly, [Wenzl \(2020\)](#) examines how Conservative politicians attempted to construct and brand a unified national identity during the Brexit debate. Both studies highlight the central role of discourse in negotiating political identities, although their analyses focus primarily on particular speakers and communicative settings rather than broader corpus-wide lexical patterning.

Corpus-based investigations have also examined the rhetorical and evaluative dimensions of Brexit discourse. As an example, [Simaki et al. \(2020\)](#) introduce the Brexit Blog Corpus and develop a detailed stance annotation framework for analysing speaker positioning in online discussions. Their findings reveal the prevalence of contrariety, epistemic stance, and emotionally charged evaluations, illustrating the highly contested nature of public engagement with Brexit. Likewise, [Prihodko et al. \(2020\)](#) compare Slovak and British media discourse, demonstrating that rhetorical resources such as evaluative language, irony, and metaphor contribute to contrasting national representations of Brexit across different media contexts.

Several studies have concentrated on specific linguistic phenomena within Brexit discourse. Evidence for this can be found in [Morales Pérez \(2019\)](#), who investigates conceptual metaphors associated with Brexit, while subsequent work has further explored metaphorical conceptualizations of political developments and their persuasive functions. Other corpus-assisted studies have examined the representation of national identity and ideological

positioning in newspaper discourse (Parnell, 2021, 2024), showing how lexical choices contribute to the construction of exclusionary identities and competing narratives of British nationhood. Complementing these corpus-based approaches, Hawkins (2022) offers a broader discourse-theoretical account of Brexit by demonstrating how recurring Eurosceptic narratives and symbolic representations contribute to the reproduction of ideological conflict.

Synthesizing the existing literature, it is evident that Brexit discourse is characterized by recurrent processes of ideological framing, identity construction, evaluative language, and political polarization. However, much of the existing literature concentrates on discourse domains, political actors, media outlets, or specific linguistic phenomena such as metaphor, stance, or national identity. Moreover, existing corpus-assisted studies have typically focused on individual discourse features rather than examining how multiple discursive dimensions interact within a unified analytical framework. Comparatively fewer studies have adopted a corpus-assisted critical discourse analytical approach that systematically investigates how multiple ideological frames emerge through recurring lexical and collocational patterns across a large mixed-media corpus.

Addressing this gap, the present study adopts a CDA framework that integrates corpus linguistic techniques with Fairclough's (1995) three-dimensional model of discourse. This integrated framework enables quantitative evidence derived from corpus analysis to be interpreted within broader socio-cognitive and ideological contexts by linking recurring linguistic patterns to the production, circulation, and social significance of political discourse (Vasheghani Farahani, 2013). Fairclough's framework conceptualizes discourse as operating simultaneously at the levels of text, discursive practice, and social practice, while corpus linguistic methods provide systematic and replicable procedures for identifying statistically salient lexical patterns, collocational relationships, and concordance structures across large datasets (Baker, 2023; Partington et al., 2013). Accordingly, quantitative findings are not treated as explanations in themselves but as empirically grounded evidence that guides qualitative discourse interpretation (Baker et al., 2008; Baker, 2023). This integration enables the study to move systematically from statistical pattern identification to contextual analysis while maintaining methodological transparency and analytical rigor (Baker et al., 2008; Partington et al., 2013).

Methods

Data Gathering Regime and Corpus of the Study

Corpus linguistics provides a powerful methodological framework for identifying recurrent language patterns and generating insights into discourse structures, lexical choices, and thematic tendencies (Brezina & McEnery, 2020; Hunston, 2022; Vasheghani Farahani, 2022). Within the study of political discourse, corpus-based approaches are particularly valuable for examining how language frames public debate and constructs ideological positions (Taylor, 2022).

To address the aims of the present study, the Brexit Corpus hosted on the [Sketch Engine](#) platform was employed as the primary data source. The corpus was not compiled by the researcher; rather, it is a pre-existing large-scale dataset comprising 506,828 documents and approximately 125 million tokens. Consequently, the corpus size reflects the available textual production captured during the selected three-day period rather than a predetermined sampling target established by the researcher. The objective of employing this large-scale corpus was not to maximize the volume of text subjected to qualitative interpretation, but to provide a comprehensive empirical basis for identifying statistically robust lexical patterns across the dataset.

For the purposes of this study, only texts produced during the three-day pre-referendum period (19–21 June 2016) were selected for analysis. This narrowly defined timeframe captures the peak of discursive activity immediately preceding the 2016 European Union referendum and provides a concentrated snapshot of the communicative environment in which political arguments, media narratives, and public engagement intensified. Focusing on this critical period ensures that the linguistic patterns identified reflect heightened ideological framing, increased media visibility, and active public mobilisation surrounding one of the most consequential political events in recent British history. The corpus data selected for analysis are summarized in Table 1.

The analytical contribution of the 125-million-token corpus lies primarily in the quantitative stage of the study. The complete corpus was used to identify keywords, calculate frequency distributions, estimate log-likelihood and log-ratio effect sizes, assess dispersion, and extract stable collocational patterns. These corpus-wide statistical procedures ensured that the lexical items selected for qualitative interpretation represented recurrent and widely distributed patterns rather than isolated textual phenomena or topic-specific anomalies.

The subsequent qualitative stage was therefore interpretive rather than exhaustive. The purpose of the quantitative stage was not to reduce the corpus to a small qualitative sample, but

to identify statistically robust lexical patterns that could subsequently be examined in their discursive contexts through critical discourse analysis. In this way, the full corpus strengthens the reliability and robustness of pattern identification, while the sampled concordance lines provide the contextual evidence required for critical discourse interpretation.

Table 1. Corpus Data

Category	Description / Count
Corpus Name	Brexit Corpus (Sketch Engine)
Total Documents	506,828 documents
Total Tokens	125 million tokens
Timeframe	19–21 June 2016
Data Sources	Mixed online text types, including: news articles, opinion pieces, blogs, and user-interactive content (e.g., comments)
Selection Criterion	All texts within the specified 3-day window; no manual exclusions
Reference Corpus	enTenTen13 (61 billion tokens)

To ensure the relevance and salience of identified keywords, the study also employed a comparative corpus approach. The English Web 2013 (enTenTen13) corpus, which comprises over 61 billion tokens, was selected as the reference corpus. As a representative sample of general English language usage, enTenTen13 serves as a robust baseline for identifying keywords that are distinctive to Brexit discourse (Sketch Engine – enTenTen13).

Procedure

The analytical procedure combined quantitative corpus-linguistic techniques with qualitative critical discourse analysis to identify salient lexical patterns and examine their ideological significance within Brexit-related discourse. Keyword candidates, concordance lines, and collocational information were first extracted using the Sketch Engine platform by comparing the Brexit Corpus with the reference corpus enTenTen13. Raw frequency data obtained from this comparison were subsequently exported to IBM SPSS Statistics for statistical analysis. Using the exported frequency counts, the published formulae for log-likelihood (Rayson & Garside, 2000), probability values (p-values), log-ratio effect sizes, and dispersion (DP) measures (Gries, 2021) were implemented in IBM SPSS Statistics to calculate the statistical values reported in this study. These statistics were subsequently used to identify lexically salient and widely distributed keywords for qualitative interpretation within Fairclough's

(1995) three-dimensional model of critical discourse analysis. This multi-metric approach ensured that lexical salience was evaluated not only in terms of statistical significance but also with respect to effect size and distribution across the corpus.

To enhance methodological rigor, explicit threshold criteria were established for keyword selection. Candidate items were required to occur at least ten times in the corpus, achieve a log-likelihood value of 15.13 or higher ($p < 0.0001$), and demonstrate a minimum log-ratio effect size of 1.5. In addition, a dispersion criterion was applied to exclude items whose frequency was concentrated in only a small number of texts, thereby reducing the influence of topic-specific anomalies and ensuring that selected keywords reflected broader discursive tendencies. Only lexical items satisfying all statistical requirements and demonstrating clear relevance to Brexit-related discourse were retained for further analysis. Consequently, statistically prominent but qualitatively marginal items, such as *underpower*, were excluded from the final dataset. The complete statistical output for all candidate keywords is provided in Appendix 1.

Following the automated extraction stage, a systematic manual filtering procedure was conducted to ensure the interpretive validity of the results. Keyword selection was guided by three interrelated principles: ideological salience, discursive relevance, and semantic specificity. Ideological salience referred to the extent to which a lexical item participated in politically contested meanings or reflected polarized viewpoints. Discursive relevance concerned the frequency with which a term functioned within framing strategies, argumentation patterns, or evaluative constructions. Semantic specificity focused on the degree to which an item was directly associated with Brexit and its related political, economic, and social debates. Generic high-frequency words and semantically diffuse expressions were excluded in order to maintain analytical precision and strengthen the explanatory value of the findings. Furthermore, a minimum frequency threshold of ten occurrences (see Appendix 1) was retained to eliminate low-frequency items with limited interpretive significance.

The final dataset consisted of nineteen keywords that collectively captured the dominant thematic concerns and ideological dimensions of Brexit discourse. To facilitate systematic interpretation, these keywords were organized into four analytically derived categories according to their discursive functions and contextual patterns of use. This categorization draws upon established corpus-assisted approaches to political discourse analysis, which emphasize the relationship between lexical choice, ideological positioning, and meaning construction (Moisl, 2015; Baker, 2023).

The subsequent qualitative analysis focused on a systematically sampled set of concordance lines representing the statistically established lexical patterns identified in the preceding quantitative stage. This sampling strategy ensured that the qualitative evidence reflected the diversity and consistency of corpus-wide linguistic patterns while maintaining analytical manageability and interpretive depth. Accordingly, concordance analysis served to provide contextual interpretation of corpus-wide patterns rather than to achieve exhaustive coverage of the corpus.

Importantly, the classification process extended beyond simple semantic field analysis. Rather than grouping words solely according to denotational meaning, the analysis examined the rhetorical and ideological functions that keywords performed within their co-textual environments. Attention was paid to how lexical items contributed to processes of persuasion, identity construction, legitimation, delegitimation, and future-oriented projection. For instance, *vote* was analyzed not merely as a procedural term associated with democratic participation but also as a symbolic marker of political identity, collective agency, and moral commitment. Likewise, *post-Brexit* was examined as a future-projecting expression through which competing actors articulated imagined scenarios of national renewal, economic opportunity, uncertainty, or decline. Such uses demonstrate how seemingly neutral lexical choices can become embedded within broader ideological narratives that shape public understanding of political events.

The resulting thematic categories therefore represent discursive frames rather than purely semantic groupings. They capture recurring patterns through which Brexit was linguistically constructed, evaluated, and contested across the corpus. Table 2 presents the final classification of keywords and their corresponding discursive functions.

Table 2. Categorization of Selected Words

Campaign-related Terms				
voteleave	euref	Vote Remain	YouGov	Campaigner
Political Actors and Institutions				
corbyn	Sadiq	Boris	Unionist	Britain
Evaluative and Ideological Language				
opportunism	rightwing	worryingly	robustly	discredited
Future-oriented References				
post-Brexit		Brexit		referendum

For each selected keyword, concordance lines were extracted using [Sketch Engine](#). To ensure contextual richness and avoid redundancy, concordance lines were extracted through a sampling procedure designed to maximise the range of discursive environments represented in the dataset ([Baker, 2023](#); [Vasheghani Farahani & Pahlevanzade Fini, 2023](#)). The full set of concordance lines for each keyword was first randomised using a fixed sampling seed to maintain reproducibility. From this randomised pool, lines were proportionally drawn from different segments of the co-textual environment so that no single lexical pattern or recurrent structure would dominate the qualitative evidence. A fixed sample of 50 concordance lines per keyword was retained to enable comparability across items while keeping the analytical dataset manageable. This sample size is consistent with established corpus-assisted discourse analytical practice, where concordance analysis aims to identify recurrent discursive patterns through theoretically informed sampling rather than to achieve exhaustive coverage of every occurrence of a lexical item ([Baker, 2023](#)).

Each sampled line was then examined within a ± 3 -word span to identify salient collocational tendencies, syntactic configurations and evaluative lexical choices. This step allowed for a detailed exploration of how keywords signal stance, construct identities or contribute to the reproduction of ideological frameworks within the corpus ([Brezina, 2018](#); [Hunston, 2021](#); [Baker, 2023](#)). All collocation analyses in this study were conducted using a uniform set of parameters to ensure transparency, comparability, and methodological replicability. A symmetrical window of ± 3 words was applied around each keyword to capture immediate local associations relevant to discursive patterning. A minimum collocate frequency threshold of 5 occurrences within the Brexit Corpus was adopted to filter out unreliable, low-frequency associations. Collocational strength was measured using LogDice, chosen for its stability and interpretability in large-scale corpus analysis. These parameters were applied consistently across all thematic sections and all keywords, ensuring that the resulting collocation patterns reflect stable and meaningful lexical behaviour throughout the corpus.

Data Analysis and Results

The data analysis in this study adopts a primarily qualitative approach to explore how Brexit is discursively constructed across different media platforms. Rather than focusing on statistical modelling or numerical trends, the analysis centres on examining keywords and their surrounding contexts in order to identify recurrent linguistic patterns, framing strategies, and ideological cues.

Campaign-related Terms

The campaign-related keywords *voteleave*, *Vote Remain*, and *campaigner* reveal contrasting patterns of political mobilization and ideological positioning within Brexit media discourse. Collectively, these keywords illustrate how media representations construct competing forms of political participation, identity, and collective action during the referendum period. As shown in Table 3, the corpus evidence indicates three interrelated discursive strategies: the mobilization and legitimization of the Leave campaign, the representation of the Remain campaign through civic participation and public engagement, and the categorization of political actors according to competing ideological alignments. Table 3 provides an overview of the principal lexical features, discursive strategies, and interpretive significance identified through the corpus analysis.

Table 3. Overview of Discursive Patterns in Campaign-Related Terms

Keyword	Dominant lexical patterns	Discursive strategy	Discursive interpretation
Vote Leave	Collective identity (<i>team, Britain, supporter</i>), confrontation (<i>smash, attack, push</i>), legitimization (<i>justify</i>)	Mobilization and legitimization	Represents Leave as a unified, decisive, and nationally oriented political movement.
Vote Remain	Civic participation (<i>crowd, rally, concert</i>), collective engagement (<i>campaign, join, support</i>)	Public mobilization	Represents Remain through civic participation, public outreach, and symbolic collective action rather than confrontation.
Campaigner	Ideological labels (<i>pro-Remain, anti-EU, pro-Brexit</i>), persuasion (<i>argue, claim</i>), conflict (<i>clash, accuse</i>)	Political categorization	Constructs political actors through ideological alignment and reinforces adversarial political identities.

Vote Leave

Consistent with the overview presented in Table 3, the keyword *voteleave* primarily reflects the discursive strategy of mobilization and legitimization. Its collocational profile is characterized by nouns denoting collective identity and organizational affiliation, including *team*, *Britain*, *supporter*, *audience*, and *campaign*. These lexical associations represent the Leave campaign as a cohesive political movement in which support is represented as collective membership within a unified national community.

In addition, *voteleave* frequently co-occurs with verbs conveying confrontation and decisive action, such as *smash*, *attack*, and *push*, suggesting a discourse marked by political urgency, assertiveness, and calls for action. The presence of justificatory verbs such as *justify*, together with evaluative references to prominent political figures, indicates that the Leave campaign was not only mobilized through confrontational rhetoric but also discursively legitimized by presenting its objectives as reasonable and necessary. The examples below illustrate how the discourse surrounding *voteleave* combines collective identity with persuasive and legitimizing rhetoric, positioning the Leave campaign as a unified national movement committed to protecting Britain's interests.

Example 1

"Last week it seemed that there were more polls coming in favour of vote leave, whereas over the last few days it appears the latest polls are neck and neck." (theguardian.com, 21 June 2016)

Example 2

"We should consider this very carefully when given an option soon, VOTE LEAVE so we can protect our borders and existence." (express.co.uk, 21 June 2016)

Vote Remain

Unlike *voteleave*, the keyword *Vote Remain* is primarily associated with the discursive strategy of public mobilization through civic participation and collective engagement. Its lexical environment is characterized by words such as *crowd*, *sign*, *rally*, and *concert*, together with verbs including *campaign*, *join*, and *support*. These collocational patterns situate the Remain campaign within public, social, and cultural spaces, emphasizing civic involvement rather than overt confrontation.

Furthermore, the frequent occurrence of *Vote Remain* alongside prepositions such as *on*, *with*, and *for* demonstrates its flexibility in connecting political participation with broader educational, cultural, and community-based activities. Overall, the discourse frames the Remain campaign as a movement grounded in public engagement and symbolic expressions of political solidarity. The concordance illustrates how the Remain campaign is framed through appeals to careful consideration and informed civic participation rather than confrontational political rhetoric.

Example 3

"Before making this historic decision I want those in our country to fully examine just what the consequences would be of a Remain vote." (express.co.uk, 21 June 2016)

Campaigner

As summarized in Table 3, the lemma *campaigner* functions primarily as a mechanism of political categorization. Unlike the campaign labels themselves, this keyword serves as a flexible discursive resource for identifying political actors according to their ideological positions and communicative roles. It frequently co-occurs with ideological modifiers such as *pro-Remain*, *anti-EU*, and *pro-Brexit*, as well as references to political actors including *voter*, *activist*, and *immigrant*.

Verbal collocates such as *lead*, *argue*, *claim*, *accuse*, and *blame* further portray campaigners as active participants in persuasion, political advocacy, and ideological contestation. Rather than referring to a single political orientation, the keyword categorizes individuals according to their alignment within competing political narratives, reinforcing the adversarial nature of Brexit discourse. The example below demonstrates how the label *campaigner* functions to position political actors within opposing ideological camps, thereby reinforcing polarized political identities.

Example 4

"Johnson, one of the most prominent campaigners on the 'leave' side, distanced himself from U.K. Independence Party leader Nigel Farage, who has been widely criticized for the poster." (lynnnews.co.uk, 21 June 2016)

Taken together, the campaign-related keywords demonstrate that Brexit media discourse extends beyond describing electoral activity to construct competing political identities and forms of collective action. The corpus reveals three complementary discursive strategies: the

mobilization and legitimization of the Leave campaign, the civic framing of the Remain campaign, and the ideological categorization of political actors. Rather than functioning as neutral descriptions of campaign events, these recurring lexical patterns contribute to representing Brexit as a polarized political contest structured around competing identities, ideological alignments, and contrasting visions of political participation.

Political Actors and Institutions

The keywords *Corbyn*, *Boris*, and *Sadiq* reveal how prominent political figures were discursively constructed through patterns of institutional representation, personalization, and political confrontation within Brexit media discourse. Rather than simply identifying political actors, these keywords demonstrate how media discourse foregrounded leadership roles, personal attributes, and ideological conflict as central features of political communication during the referendum campaign. As summarized in Table 4, the corpus evidence reveals three complementary discursive strategies: the institutional positioning of Jeremy Corbyn, the personalized evaluation of Boris Johnson, and the representation of Sadiq Khan through adversarial political interaction. Table 4 provides an overview of the principal lexical patterns, discursive strategies, and interpretive significance identified through the corpus analysis.

Table 4. Overview of Discursive Patterns in Political Actors and Institutions

Keyword	Dominant lexical patterns	Discursive strategy	Discursive interpretation
Corbyn	Institutional identifiers (<i>Jeremy, Labour, leader, opposition</i>), deliberative verbs (<i>say, think, consider</i>)	Institutional positioning	Represents Corbyn as a mediating political leader whose authority is negotiated within party politics.
Boris	Evaluative adjectives (<i>charmless, cynical, reckless</i>), performance verbs (<i>pretend, flounder, collapse</i>)	Personalization and evaluation	Constructs Boris Johnson as a highly personalized and continually evaluated political figure.
Sadiq	Conflict verbs (<i>clash, criticise, deny, challenge</i>), interactional contexts	Political confrontation	Represents Sadiq Khan through adversarial interaction and ideological opposition.

Boris

Unlike *Corbyn*, the keyword *Boris* is characterized by a strongly personalized and evaluative discursive profile. As summarized in Table 4, Boris Johnson is frequently associated with evaluative adjectives such as *charmless*, *unpleasant*, *cynical*, *reckless*, and *scaremongering*, shifting attention from institutional leadership toward personal character. Additional verbal collocates including *pretend*, *flounder*, *quiz*, and *collapse* reinforce a discourse of political performance, instability, and public scrutiny. Rather than emphasizing policy positions, these recurring lexical patterns construct Boris Johnson as a recognizable media persona whose credibility and political legitimacy are continuously negotiated through evaluative language.

Example 5

"If one good thing has come out of #EUref it's the realisation what an unpleasant cynical shallow charmless careerist Boris..." (telegraph.co.uk, 21 June 2016)

The concordance illustrates the evaluative framing identified in Table 4. Rather than discussing Boris Johnson's political arguments or policy positions, the discourse foregrounds a sequence of negatively loaded adjectives (*unpleasant*, *cynical*, *shallow*, and *charmless*) that collectively construct a highly personalized political identity. The emphasis on character evaluation rather than political substance reinforces a discursive strategy in which legitimacy is negotiated through moral and personal judgment.

Sadiq

The keyword *Sadiq* primarily reflects a discursive strategy of political confrontation. Although less frequent than the other political actors, its collocational profile consistently emphasizes interaction, disagreement, and ideological contestation. Verbs such as *clash*, *criticise*, *deny*, and *challenge* repeatedly position Sadiq Khan in direct opposition to prominent Leave campaign figures, particularly Boris Johnson. Consequently, his representation is structured around political conflict and public debate rather than institutional identity, reinforcing his role as an active participant in the ideological polarization of the referendum campaign.

Example 6

"Audience laughs out loud when Sadiq tries to claim that David Cameron got a good deal." (express.co.uk, 21 June 2016)

This concordance reflects the confrontational pattern summarized in Table 4. The reporting of audience laughter delegitimizes Sadiq Khan's intervention by emphasizing public

skepticism, while the verb *tries* implies an unsuccessful attempt at persuasion. Together, these linguistic choices construct political interaction as adversarial performance and reinforce the broader polarization of Brexit media discourse.

Taken together, these lexical patterns demonstrate that Brexit media discourse represents political actors through three complementary discursive strategies: institutional positioning, personalized evaluation, and political confrontation. Rather than foregrounding policy substance alone, media representations increasingly construct political legitimacy through leadership roles, personal characteristics, and ideological conflict. Consequently, political processes are personalized through recognizable individuals whose credibility, authority, and ideological positions become central objects of public evaluation.

Evaluative and Ideological Language

The keywords *opportunism*, *rightwing*, *worryingly*, and *robustly* reveal how evaluative language contributes to the ideological construction of Brexit media discourse. Rather than functioning as neutral lexical items, these keywords encode moral judgments, ideological positioning, perceptions of risk, and evaluations of political action. As summarized in Table 5, the corpus identifies four complementary discursive strategies: the moral delegitimization of political actors through *opportunism*, the ideological categorization of political identities through *rightwing*, the construction of uncertainty through *worryingly*, and the representation of political assertiveness through *robustly*. Together, these patterns demonstrate that evaluative language plays a central role in shaping perceptions of legitimacy, credibility, and ideological division throughout the referendum campaign.

Table 5. Overview of Discursive Patterns in Evaluative and Ideological Language

Keyword	Dominant lexical patterns	Discursive strategy	Discursive interpretation
opportunism	Negative evaluative adjectives (<i>shameless</i> , <i>unprincipled</i> , <i>Machiavellian</i> , <i>repellent</i>), political associations	Moral delegitimization	Constructs political actors as ethically questionable and strategically self-serving.

Keyword	Dominant lexical patterns	Discursive strategy	Discursive interpretation
rightwing	Institutional and ideological labels (<i>Conservatives, Tory, Brexiteers, nationalist</i>)	Ideological categorization	Positions political actors within polarized ideological identities.
worryingly	Risk-oriented verbs and evaluative contexts (<i>increase, appear, rise, uncertain</i>)	Construction of uncertainty	Frames political developments as sources of concern and instability.
robustly	Assertive action verbs (<i>defend, reject, respond, speak</i>)	Legitimation of political action	Represents political responses as decisive, confident, and authoritative.

Opportunism

Consistent with the overview presented in Table 5, the keyword *opportunism* primarily reflects the discursive strategy of moral delegitimization. Its collocational profile is dominated by strongly negative evaluative adjectives such as *shameless*, *unprincipled*, *Machiavellian*, *repellent*, and *superficiality*, all of which construct opportunism as an ethically objectionable political practice. Additional collocates associated with political campaigns, parties, and public figures indicate that accusations of opportunism are directed toward political actors as a means of undermining their credibility and integrity. Collectively, these lexical patterns frame political legitimacy in moral rather than purely ideological or policy-related terms.

Example 1

“One would have thought from the reaction to Woolerston's similar switch, she would have realised that this would be seen as nothing more than unprincipled political opportunism, rather than out of any deeply held view.” (heraldscotland.com, 21 June 2016)

The concordance exemplifies the moral evaluative pattern summarized in Table 5. The adjective *unprincipled* explicitly questions the ethical foundations of political behaviour, while the contrast with *deeply held view* constructs opportunism as the opposite of genuine political conviction. Consequently, political legitimacy is evaluated primarily through perceived integrity rather than policy substance.

Example 2

“I think the repellent opportunism, superficiality and implicit racism of Boris Johnson and Michael Gove and Nigel Farage is the worst of Britain.” (telegraph.co.uk, 21 June 2016)

This concordance further reinforces the delegitimizing function of *opportunism*. The accumulation of negatively evaluative nouns and adjectives intensifies moral condemnation, representing political actors through ethical failure rather than political competence. Such lexical choices contribute to a strongly polarized representation of political leadership.

Rightwing

As summarized in Table 5, the keyword *rightwing* primarily reflects the discursive strategy of ideological categorization. Its collocational profile includes institutional and political identifiers such as *Conservatives*, *Tory*, *government*, *Brexiteers*, and *nationalist*, situating political actors within clearly recognizable ideological groupings. Rather than functioning as a neutral political descriptor, *rightwing* frequently appears in evaluative contexts that associate political actors with broader ideological positions and competing value systems. These recurrent lexical patterns reinforce ideological boundaries by constructing political identities as collective and oppositional.

Example 3

“One of the great ironies of the UK's referendum debate is that Europe, with its austerity programmes and its drift towards neoliberalism, has been moving in a direction that rightwing Conservatives would tend to support.” (telegraph.co.uk)

The concordance illustrates how *rightwing* functions as an ideological label rather than a purely descriptive term. By associating *rightwing Conservatives* with neoliberalism and austerity, the discourse situates political actors within a broader ideological framework, encouraging readers to interpret political positions through established ideological categories rather than individual policy preferences.

Worryingly

The keyword *worryingly* primarily reflects the discursive strategy of constructing uncertainty. Its collocational profile is characterized by verbs such as *increase*, *appear*, *rise*, and *look*, together with lexical items associated with instability and inconsistency. Rather than merely describing events, *worryingly* guides readers toward interpreting political developments as potential threats requiring concern and vigilance. Through this evaluative framing, media discourse amplifies perceptions of uncertainty surrounding the Brexit process.

Robustly

Unlike *worryingly*, the keyword *robustly* reflects a strategy of legitimizing political action through assertiveness. It commonly collocates with verbs such as *defend*, *reject*, *respond*, *speak*, and *act*, portraying political actors as decisive and confident. These recurrent lexical patterns emphasize determination and authority, framing political communication as purposeful and resolute even in contexts of disagreement or criticism.

Viewed collectively, these evaluative and ideological keywords demonstrate that Brexit media discourse extends beyond the reporting of political events to actively construct moral judgments, ideological identities, perceptions of uncertainty, and evaluations of political leadership. Through the complementary strategies of moral delegitimization, ideological categorization, uncertainty construction, and legitimization of assertive action, evaluative language shapes how political actors and events are interpreted by readers. Consequently, the discourse contributes to the broader ideological polarization of the Brexit referendum by framing political legitimacy in moral, ideological, and affective terms rather than through policy debate alone.

Future-oriented Discourse

The keywords *post-Brexit*, *referendum*, and *Brexit* reveal how media discourse constructs imagined futures and projects possible political, economic, and social trajectories surrounding the UK's withdrawal from the European Union. Rather than simply referring to political events, these keywords organize public understanding through expectations, predictions, and evaluations of future developments. As summarized in Table 6, the corpus identifies three complementary discursive strategies: projecting post-referendum consequences through *post-Brexit*, constructing the referendum as a decisive historical turning point, and representing *Brexit* as a temporal anchor through which past experiences, present debates, and future expectations are interconnected. Together, these patterns demonstrate that future-oriented discourse plays a central role in shaping public interpretations of uncertainty, national transformation, and political change.

Table 6. Overview of Discursive Patterns in Future-oriented Discourse

Keyword	Dominant lexical patterns	Discursive strategy	Discursive interpretation
post-Brexit	<i>relationship, economy, market, future, policy, shock, uncertain</i>	Projection of future consequences	Frames Brexit through anticipated political and economic outcomes.
referendum	<i>fateful, decisive, propose, recommend, promise, suggest</i>	Historical turning point	Represents the referendum as a defining political moment shaping future developments.
Brexit	<i>anticipate, expect, negotiate, happen, follow, vote, campaign</i>	Temporal anchoring	Connects past events, present debates, and imagined futures within a single political narrative.

Referendum

As summarized in Table 6, the keyword *referendum* primarily reflects the discursive strategy of representing the vote as a decisive historical turning point. Its collocational profile includes evaluative adjectives such as *fateful* and *decisive*, together with verbs including *propose*, *recommend*, *promise*, and *suggest*, emphasizing the referendum as a moment whose significance extends beyond the act of voting itself. Rather than functioning as a neutral institutional term, the keyword repeatedly appears in contexts linking contemporary political developments with anticipated future consequences.

Example 1

“All of these questions were always on the ballot in this week's fateful referendum.”
(independent.co.uk, 21 June 2016)

The concordance illustrates how the adjective *fateful* transforms the referendum from a routine democratic procedure into a defining historical event. By emphasizing its long-term significance, the discourse frames the referendum as a decisive moment shaping Britain's future political trajectory.

Example 2

“There is widespread speculation that it would lead to a second independence referendum.”
(heraldscotland.com, 21 June 2016)

This concordance further demonstrates the future-oriented framing of the keyword. The noun *speculation* explicitly projects possible future political developments, representing the referendum as the beginning of continuing constitutional and political change rather than a discrete electoral event.

Example 3

“Nine months earlier, some polls in the run-up to 2014's Scottish independence referendum famously suggested a yes vote...” (standard.co.uk, 21 June 2016)

This example situates the referendum within a broader historical narrative by connecting previous electoral experiences with contemporary political expectations. Such temporal comparisons reinforce the representation of referendums as recurring milestones through which future political possibilities are interpreted.

Brexit

Unlike *referendum*, the keyword *Brexit* functions primarily as a temporal anchor connecting past experiences, present political debates, and future expectations. Its collocational profile includes verbs such as *anticipate*, *expect*, *guarantee*, *follow*, *negotiate*, and *happen*, together with campaign-related vocabulary including *vote* and *campaign*. These lexical patterns represent *Brexit* as an ongoing political process rather than a single historical event, allowing media discourse to integrate retrospective interpretation with prospective evaluation. Consequently, *Brexit* becomes a central reference point through which broader narratives of national transformation, uncertainty, and political change are organized.

“Brexit, the emergence of new multilateral trade blocs and the rise of trade disputes present new challenges for Irish business in this sector.” (standard.co.uk, 21 June 2016)

This concordance illustrates how *Brexit* is discursively constructed as a catalyst for future economic and geopolitical change rather than merely as a political event. The coordination of *Brexit* with *the emergence of new multilateral trade blocs* and *the rise of trade disputes* projects a broader landscape of international transformation, while the phrase *new challenges* foregrounds anticipated economic uncertainty. Rather than focusing on the referendum itself, the discourse frames Brexit as an ongoing process whose significance lies in its projected consequences for trade, governance, and economic relations.

Viewed as a whole, the findings indicate that future-oriented keywords demonstrate that Brexit media discourse is structured around projections of political change, historical turning points, and imagined national futures. Through the complementary strategies of projecting

future consequences, framing decisive political moments, and temporally anchoring political narratives, the discourse encourages readers to interpret Brexit as an evolving process rather than a completed event. Consequently, media representations construct expectations of uncertainty, transformation, and national reorientation, reinforcing broader ideological debates concerning sovereignty, governance, and Britain's future relationship with Europe.

Although each analytical dimension reveals a distinct aspect of Brexit discourse, they should not be interpreted as independent thematic categories. Rather, campaign discourse, representations of political actors, evaluative language, and future-oriented projections constitute complementary dimensions of a broader ideological framing process. Examining these dimensions collectively provides a more comprehensive account of how political meanings are organized across the corpus than would be possible through the analysis of any individual discourse feature in isolation.

Response to the First Research Question

RQ1 examined how Brexit is discursively constructed across mass media discourse, focusing on the linguistic strategies and ideological framings that shape its representation within the integrated corpus. The findings indicate that Brexit is represented not merely as a political procedure but as a highly contested symbolic space in which issues of identity, legitimacy and sovereignty are recurrently negotiated (Kelly, 2018; McConnell & Tormey, 2020).

Across the corpus, discursive construction emerges through a set of recurrent framing patterns, reflected in the key lexical items and their collocational environments (Baker et al., 2008; Brezina & McEnery, 2020). Although the dataset does not contain explicit genre metadata and therefore cannot support statistical comparisons between media types, the qualitative analysis reveals stylistic tendencies that contribute to the construction of Brexit as an ideological battleground.

A first prominent framing pattern involves campaign mobilisation, visible through terms such as *voteleave*, *Vote Remain*, and *campaigner*. These items frequently appear alongside verbs and nouns associated with persuasion, confrontation and collective identity, illustrating how discourse foregrounded political urgency and moralised choice during the pre-referendum period. Such patterns align with previous findings on the affective and action-oriented nature of referendum campaigning (Brezina, 2018; Papacharissi, 2015; Bruns et al., 2016).

A second framing pattern involves personalisation, where political actors—such as Corbyn, Boris and Sadiq—serve as focal points through which ideological conflict is interpreted. Their names co-occur with evaluative and relational descriptors, reflecting broader

tendencies in political communication in which individual figures become symbols of ideological alignment or controversy (Richardson, 2007; KhosraviNik, 2017; Hawkins, 2022). These patterns reveal how public figures operate as discursive resources for constructing political meaning and situating responsibility or credibility.

A third framing dimension involves evaluative and ideological language, represented through keywords such as opportunism, rightwing, worryingly and robustly. These items introduce explicit or implicit stance-taking, moral judgment and ideological differentiation. Their collocational patterns show how lexical evaluation shapes public understanding of political actors and positions, consistent with prior work on the embedding of ideology in seemingly descriptive language (Fairclough, 1995; Baker, 2023).

A fourth major pattern involves future-oriented framing, anchored in *Brexit*, *post-Brexit* and *referendum*. These terms frequently appear in predictive or speculative contexts, projecting imagined futures, risks, hopes or narrative trajectories for Britain. Such framings resonate with scholarship showing that Brexit discourse often invoked crisis imaginaries and temporal projections to structure public interpretation (Eagleton, 2018; Krzyżanowski, 2019; Jamet & Rodet, 2022).

Taken together, these four framing dimensions illustrate how different linguistic resources perform complementary discursive functions within Brexit media discourse. Campaign-related language establishes the political context, representations of political actors assign agency and responsibility, evaluative language positions actors and events ideologically, and future-oriented discourse projects competing political trajectories. Examining these dimensions collectively therefore provides a more integrated account of ideological framing than analyses focusing on any single discourse feature in isolation.

Beyond its substantive observations on Brexit discourse, the present analysis illustrates how corpus-assisted CDA can be operationalized through a sequential procedure that combines keyword analysis, collocation, and concordance analysis with qualitative discourse interpretation. The value of the study therefore lies not in confirming the theoretical assumptions of CDA, but in demonstrating a transparent and replicable analytical workflow that may be applied to other political events and media corpora.

Response to the Second Research Question

RQ2 explored the ideological assumptions and sociopolitical values embedded in Brexit-related media discourse and how these discursive features shape public understanding and

political alignment. The findings indicate that Brexit discourse is characterized by a highly charged ideological environment in which meanings related to identity, sovereignty, and democratic authority are recurrently negotiated (Fairclough, 1995; Kelly, 2018).

Across the corpus, ideological positioning becomes visible through the interplay of lexical evaluation, stance-taking and framing devices. Terms such as *opportunism*, *rightwing*, *worryingly* and *robustly* encode moral judgment, political categorisation, and affective orientation, thereby shaping how actors and arguments are positioned within the broader ideological conflict (Richardson, 2007; Baker et al., 2008). Negative evaluative descriptors and affect-laden adverbs amplify perceptions of threat, instability or political failure, contributing to the emotional intensification characteristic of Brexit discourse (Papacharissi, 2015; Bruns et al., 2016).

Personalised attributions, particularly involving figures such as Boris, Corbyn and Sadiq, function as ideological shortcuts, allowing complex political disagreements to be condensed into recognisable personas and moralised character judgments. As earlier research suggests, personalisation operates as a mechanism through which broader ideological narratives become anchored in the reputations and perceived traits of public actors (Hawkins, 2022; KhosraviNik, 2017).

In addition, ideological meaning is shaped by references to imagined futures. Keywords such as Brexit, post-Brexit and referendum frequently occur in predictive or speculative contexts, structuring discourse around anticipated outcomes, crises, opportunities or national reorientation (Eagleton, 2018; Krzyżanowski, 2019; Jamet & Rodet, 2022). These temporal projections function as symbolic resources for constructing competing visions of Britain's future and for legitimising particular political trajectories.

Although the corpus does not include genre-level metadata and therefore cannot support genre-specific generalisations, the ideological meanings observed in the dataset emerge through stylistic contrasts in evaluation, modality, stance-marking and emotional intensity. Rather than being tied to discrete media types, these patterns circulate across textual styles and contribute to a multi-voiced ideological environment in which competing narratives of sovereignty, belonging and political agency coexist and frequently collide.

Overall, the findings suggest that Brexit discourse embeds sociopolitical values not through explicit argumentation alone but through repeated evaluative cues, personalised framings and future-oriented imaginaries that shape how readers interpret both political actors and the stakes of the referendum.

Discussion and Conclusion

The findings of this study indicate that Brexit discourse is characterized by recurring patterns of linguistic evaluation, political representation, and future-oriented framing across the media corpus. Interpreted through a CDA framework, these patterns reveal how Brexit was consistently represented through competing ideological perspectives and political narratives. Rather than treating these observations as independent confirmation of CDA's theoretical assumptions, the analysis demonstrates how corpus-derived evidence can support the systematic interpretation of ideological positioning and discursive framing in large-scale political discourse. In this respect, Brexit emerges as a politically contested issue represented through recurring discursive patterns relating to identity, legitimacy, sovereignty, and the future trajectory of the nation (Fairclough, 1995; Baker et al., 2008).

The keyword analysis revealed that lexical items such as *Brexit*, *vote*, *post-Brexit*, and *opportunism* function as discursive anchors around which political meanings are repeatedly negotiated and contested. These keywords operate as semiotic nodes that condense broader ideological arguments, mobilize affective responses, and structure moral and political distinctions. Their recurrent collocational environments suggest that political meanings acquire persuasive force not solely through explicit argumentation but also through repeated patterns of linguistic association and reinforcement across discourse. In this sense, lexical salience provides a valuable point of entry for examining how ideological positioning is represented and negotiated within public debate (Brezina & McEnery, 2020).

In combination, these four discursive frames provide a structured account of how Brexit was represented across the corpus, showing that media discourse consistently organized political debate around campaign mobilization, political actors, evaluative positioning, and competing projections of Britain's future. This empirical characterization of Brexit discourse represents the principal empirical outcome of the study, while its primary contribution lies in illustrating a transparent and replicable corpus-assisted CDA workflow for identifying and interpreting recurring ideological patterns.

From the perspective of Fairclough's (1995) three-dimensional model, the identified lexical and collocational patterns constitute textual evidence through which discursive practices surrounding Brexit can be interpreted in relation to broader questions of ideology, sovereignty, and political legitimacy. Rather than serving as independent proof of CDA's theoretical assumptions, these findings illustrate the analytical value of integrating corpus evidence with critical discourse interpretation, demonstrating how quantitative linguistic

patterns and qualitative discourse analysis can be combined within a coherent analytical framework.

Although the corpus incorporates a wide range of textual styles and communicative practices, the present analysis does not treat these as discrete genre categories. Instead, variation is interpreted as a set of stylistic tendencies emerging within a single integrated discursive environment. Certain segments exhibit features associated with narrative sequencing, policy-oriented framing, or evaluative moderation, yet these patterns are understood as manifestations of discursive variation rather than distinct genres. This observation highlights the increasingly hybrid nature of contemporary political communication, where institutional and participatory voices circulate within interconnected media ecologies.

The findings further highlight how evaluative language, personalization, and future-oriented expressions function as recurring discursive resources through which political positions and competing interpretations of Brexit are articulated within the corpus. Evaluative and ideological lexis contributes to the emotionalization of political debate by categorizing actors, policies, and outcomes through morally charged frameworks. Simultaneously, personalized framings translate abstract political processes into recognizable human agents, thereby facilitating identification, blame attribution, and political engagement. Future-oriented language projects anticipated consequences, perceived risks, and imagined trajectories of national development, revealing how expectations about Brexit were linguistically represented across the media corpus.

Considered as a whole, the findings provide a systematic account of the dominant linguistic patterns through which Brexit was represented in media discourse during the pre-referendum period. Examining campaign-related discourse, political actors, evaluative language, and future-oriented projections as interconnected discursive frames demonstrates how corpus-assisted Critical Discourse Analysis can integrate quantitative corpus evidence with qualitative interpretation to examine political communication at scale. Rather than advancing a new theoretical claim about the relationship between language and political reality, the study contributes a transparent and replicable methodological framework for identifying and interpreting recurring patterns of ideological positioning in large media corpora. In doing so, it also highlights the value of discourse-aware approaches for understanding how political events are represented and debated in contemporary media environments.

Despite these contributions, several limitations should be acknowledged. First, although the Brexit Corpus offers extensive coverage of online textual production during the referendum

period, it lacks genre-, outlet-, and author-level metadata. Consequently, it is not possible to distinguish systematically between institutional news discourse, participatory user-generated content, or hybrid communicative formats. This limitation confines the analysis to corpus-wide linguistic patterns and prevents robust claims regarding variation across media types or producer identities. Second, because the corpus is a pre-compiled and automatically harvested resource, the researcher has limited control over text selection, cleaning procedures, and categorization practices, potentially introducing noise or uneven representation across discourse domains. Third, while keyword and collocation analyses provide valuable insights into lexical prominence and associative structures, they remain inherently keyword-centered and may overlook discursive phenomena that are not captured by the selected lexical items. Fourth, although proportional sampling was employed to enhance representativeness, qualitative interpretation necessarily relies on a restricted subset of concordance lines and therefore cannot encompass the full range of discursive realizations present within the corpus. Finally, the study focuses exclusively on the three-day period immediately preceding the referendum. While this concentrated timeframe provides an analytically productive window into heightened political communication, it necessarily limits the generalizability of the findings beyond the immediate pre-referendum context.

Future research could address these limitations through several complementary avenues. One particularly promising direction would involve the use of metadata-rich corpora that permit systematic comparisons across media genres, communication platforms, and political orientations. Such an approach would facilitate a more nuanced examination of how discursive frames are constructed, adapted, and circulated within different communicative settings, including broadsheet newspapers, tabloid media, digital-native news outlets, and user-generated environments. A second avenue involves diachronic corpus segmentation, which could illuminate how campaign-related, evaluative, and future-oriented frames evolved across key political moments before and after the referendum. Furthermore, the integration of semantic network analysis, distributional semantics, or embedding-based clustering techniques may reveal more subtle relational patterns and ideological associations that extend beyond traditional collocational analysis. Finally, micro-level qualitative discourse studies focusing on a smaller set of texts could complement the present macro-level corpus findings by providing deeper insight into the rhetorical, interactional, and persuasive mechanisms through which Brexit-related meanings were negotiated, reproduced, and contested.

In conclusion, this study illustrates the methodological value of integrating corpus linguistics with Critical Discourse Analysis for the systematic investigation of large-scale

political discourse. By combining quantitative corpus evidence with qualitative interpretation, the proposed analytical framework enables recurring lexical patterns and discursive frames to be identified and interpreted within their broader political and ideological contexts. Beyond its empirical observations on Brexit discourse, the study contributes a transparent and replicable methodological approach that can be applied to the analysis of other contemporary political events and media corpora.

The study also contributes to a broader understanding of how political meaning is constructed through the interaction of multiple discursive processes. Rather than functioning as independent dimensions of political communication, the four identified frames—campaign-related discourse, representations of political actors, evaluative and ideological language, and future-oriented projections—operate as an interconnected system. Campaign discourse mobilizes political participation, representations of political actors personalize ideological conflict, evaluative language legitimizes or delegitimizes competing positions, and future-oriented projections connect present political choices to imagined national trajectories. Examining these frames collectively demonstrates how Brexit discourse constructs a coherent ideological narrative in which political identities, moral evaluations, and expectations about the future mutually reinforce one another. This integrated perspective extends corpus-assisted studies that have typically examined individual discourse features in isolation, offering a more comprehensive account of how multiple framing processes jointly shape political meaning.

Declaration of AI

During the preparation of this manuscript, the author(s) used ChatGPT (OpenAI) to assist with language editing, improving clarity and readability, refining academic style, enhancing the organization of the manuscript, and providing support in identifying relevant literature and formatting references. The author(s) independently verified the relevance and accuracy of all cited sources and references. All analyses, interpretations, conclusions, and final editorial decisions were made by the author(s). After using this tool, the author(s) critically reviewed and revised all content.

References

- Alkhamash, R. (2020). Discursive representation of the EU in Brexit-related British media. *GEMA Online Journal of Language Studies*, 20(1), 77–91. <https://doi.org/10.17576/gema-2020-2001-05>
- Baines, D., Brewer, S., & Kay, A. (2020). Political, process, and programme failures in the Brexit fiasco: Exploring the role of policy deception. *Journal of European Public Policy*, 27(5), 742–760. <https://doi.org/10.1080/13501763.2019.1678053>
- Baker, P. (2023). *Using corpora in discourse analysis* (2nd ed.). Bloomsbury Academic.
- Baker, P., Gabrielatos, C., & McEnery, T. (2008). *Discourse analysis and media attitudes: The representation of Islam in the British press*. Cambridge University Press.
- Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273–306. <https://doi.org/10.1177/0957926508088962>
- Bloom, N., Bunn, P., Chen, S., Mizen, P., Smietanka, P., & Thwaites, G. (2019). *The impact of Brexit on UK firms* (NBER Working Paper No. 26218). National Bureau of Economic Research.
- Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.
- Brezina, V., & McEnery, T. (2020). Introduction to corpus linguistics. In N. Tracy-Ventura & M. Paquot (Eds.), *The Routledge handbook of second language acquisition and corpora* (pp. 10-22). Taylor & Francis Group.
- Bruns, A., Enli, G., Skogerb, E., Larsson, A. O., & Christensen, C. (2016). *The Routledge companion to social media and politics*. Routledge.
- De Ruyter, A., & Nielsen, B. (Eds.). (2019). *Brexit negotiations after Article 50: assessing process, progress and impact*. Emerald Publishing.
- Eagleton, R. (2018). *Brexit and literature: Critical and cultural responses*. Routledge.
- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. Longman.
- Gries, S. Th. (2021). A new approach to (key) keywords analysis: Using frequency, and now also dispersion. *Research in Corpus Linguistics*, 9(2), 1–33. <https://doi.org/10.32714/ricl.09.02.01>
- Hawkins, B. (2022). *Deconstructing Brexit discourses: Embedded euroscepticism, fantasy objects and the United Kingdom's vote to leave the European Union*. Routledge.

- Hunston, S. (2022). *Corpora in applied linguistics*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139524773>
- Jamet, D., & Rodet, P. (2022). How is Brexit linguistically conceptualized? The Brexit is a journey metaphor scenario in David Cameron's, Theresa May's, and Boris Johnson's speeches. *Observatoire de la Société Britannique*, 28, 133-155. <https://doi.org/10.4000/osb.6435>
- Kelly, M. (Ed.) (2018). *Languages after Brexit: how the UK speaks to the world*. Springer Nature.
- KhosraviNik, M. (2017). Social media critical discourse studies (SM-CDS). In J. Flowerdew & J. E. Richardson (Eds.), *The Routledge handbook of critical discourse studies* (pp. 582–596). Routledge.
- Krzyżanowski, M. (2019). Brexit and the imaginary of 'crisis': A discourse-conceptual analysis of European news media. *Critical Discourse Studies*, 16(4), 465-490.
<https://doi.org/10.1080/17405904.2019.1592001>
- McConnell, A., & Tormey, S. (2020). Explanations for the Brexit policy fiasco: Near-impossible challenge, leadership failure or Westminster pathology? *Journal of European Public Policy*, 27(5), 685–702. <https://doi.org/10.1080/13501763.2019.1678051>
- Miglbauer, M., & Koller, V. (2018). 'The British people have spoken': Voter motivations and identities in vox pops on the British EU referendum. In V. Koller, S. Kopf, & M. Miglbauer (Eds.), *Discourses of Brexit* (pp. 86–103). Routledge.
- Moisl, H. (2015). *Cluster analysis for corpus linguistics*. De Gruyter Mouton.
<https://doi.org/10.1515/9783110363814>
- Morales Pérez, I. (2019). *Conceptual metaphors and Brexit: A preliminary study*. *Revista de Lenguas para Fines Específicos*, 25(1), 107-126. <https://doi.org/10.20420/rlfe.2019.250.107>
- Papacharissi, Z. (2015). *Affective publics: Sentiment, technology, and politics*. Oxford University Press.
- Parnell, T. (2021). Humiliating and dividing the nation in the British pro-Brexit press: a corpus-assisted analysis. *Critical Discourse Studies*, 20(1), 53–69. <https://doi.org/10.1080/17405904.2021.1983446>
- Parnell, T. (2024). *Constructing Brexit Britain: A corpus-assisted approach to national identity discourse*. Bloomsbury.

- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. John Benjamins Publishing Company.
- Prihodko, A., Pravdová, H., & Hudíková, Z. (2020). Evaluation of "Brexit" phenomenon in Slovak and English media texts. *Lege artis. Language yesterday, today, tomorrow*, 5(2), 211-262.
- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. *Proceedings of the Workshop on Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000)*, 1–8 October 2000, Hong Kong, 1–6.
- Richardson, J. E. (2007). *Analysing newspapers: An approach from critical discourse analysis*. Palgrave Macmillan.
- Simaki, V., Paradis, C., Skeppstedt, M., Sahlgren, M., Kucher, K., & Kerren, A. (2020). Annotating speaker stance in discourse: The Brexit Blog Corpus. *Corpus Linguistics and Linguistic Theory*, 16(2), 215–248. <https://doi.org/10.1515/cllt-2018-0060>
- Šimunjak, M. (2022). *Tweeting Brexit: Social media and the aftermath of the EU referendum*. Routledge.
- Sketch Engine. (2025). *Brexit corpus*. Retrieved March 7, 2025, from <https://www.sketchengine.eu/brexit-corpus/>
- Taylor, C. (2022). Corpus linguistics in the study of political discourse: Recent directions. In K. Aijmer, J. McCarthy, & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (2nd ed., pp. 1-13). Routledge. <https://doi.org/10.4324/9780367076399-42>.
- Vasheghani Farahani, M. (2013). *The role of discourse analysis in translation*. *International Journal of Applied Linguistics & English Literature*, 2(1), 112–116. <https://doi.org/10.7575/ijalel.v.2n.1p.112>
- Vasheghani Farahani, M., & Kazemian, R. (2021). Speaker–audience interaction in spoken political discourse: A contrastive parallel corpus-based study of English–Persian translation of metadiscourse features in TED Talks. *Corpus Pragmatics*, 5(1), 67–91. <https://doi.org/10.1007/s41701-020-00093-y>
- Vasheghani Farahani, M., & Pahlevanzade Fini, M. (2023). *Corpora and translation: Methods, concepts and applications*. Logos Publication. [In Persian]
- Vasheghani Farahani, M. (2022). *Writer-reader interaction by Metadiscourse features: English Persian translation in legal and political texts*. Frank & Timme.

- Weninger, C. (2008). Critical discourse analysis. In *The SAGE encyclopedia of qualitative research methods*. SAGE. <https://doi.org/10.4135/9781412963909.n80>
- Wenzl, N. (2020). “There is a wonderfully contrary spirit among the British people”: Conservative MPs’ (un)successful branding of the British nation in the Brexit debate. In I. Theodoropoulou & J. Tovar (Eds.), *Research companion to language and country branding* (pp. 72–89). Taylor & Francis.
- Wodak, R., & Meyer, M. (2016). *Methods of critical discourse studies* (3rd ed.). SAGE.
- Zanca, C. (2020). The Brexit Island, voices from British newspapers and voices from the web: A corpus-based approach. In *'Twixt land and sea: Island poetics in Anglophone literatures* (pp. 211-234). Artemide.
- Zappettini, F., & Krzyżanowski, M. (2021). *"Brexit" as a social and political crisis: Discourses in media and politics*. Routledge.

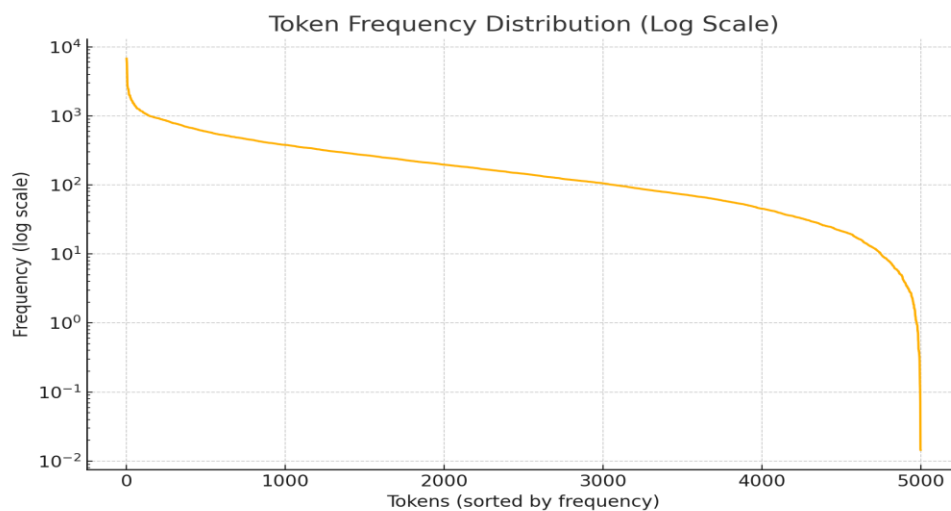
Appendix 1

Refined List of Brexit Keywords & Keyness Analysis

Keyword	LL	Log Ratio	Freq (Corpus)	DP
voteleave	55,143.21	12.84	8,542	0.78
brexit	48,921.77	12.63	16,234	0.84
referendum	41,552.03	11.91	12,874	0.81
remain	39,443.26	11.44	10,214	0.76
leavers	32,512.44	10.52	6,532	0.63
boris	30,182.91	10.23	5,644	0.58
campaigner	27,113.02	9.77	4,921	0.55
corbyn	24,845.75	9.42	4,112	0.49
sadiq	22,934.11	9.17	3,824	0.47
rightwing	18,521.68	8.46	2,994	0.52
discredited	16,992.50	8.13	2,431	0.45
worryingly	15,410.34	7.84	2,093	0.39
robustly	12,554.90	7.33	1,624	0.41
opportunism	11,204.55	7.10	1,482	0.36
post-brexit	9,113.77	6.71	1,011	0.33

This appendix presents the final list of Brexit-related keywords together with the statistical measures used for keyword selection. Raw frequency counts were extracted from [Sketch Engine](#) and subsequently analysed in IBM SPSS Statistics. The published formulae for log-likelihood ([Rayson & Garside, 2000](#)), log-ratio effect size, and dispersion ([Gries, 2021](#)) were implemented to calculate the statistical values reported below. The associated p-values were derived from the calculated log-likelihood statistics.

Appendix 2



This distribution illustrates the Zipfian nature of the corpus and supports the frequency threshold used during keyword selection.

Appendix 3

Collocational Profiles of the Selected Keywords of Campaign-Related Terms				
Rank	Collocate	Relation	Freq.	LogDice
1	team	noun modified by <i>voteleave</i>	482	9.9
2	Britain	noun modified by <i>voteleave</i>	469	8.8
3	supporter	noun modified by <i>voteleave</i>	175	9.3
4	audience	noun modified by <i>voteleave</i>	133	9.7
5	side	noun modified by <i>voteleave</i>	173	8.8
6	gap	noun modified by <i>voteleave</i>	59	8.6
7	smash	verb with <i>voteleave</i> as object	22	11.5
8	compose	verb with <i>voteleave</i> as object	8	10.7
9	terrify	verb with <i>voteleave</i> as object	6	9.2
10	let	verb with <i>voteleave</i> as subject	47	8.8
11	remember	verb with <i>voteleave</i> as subject	27	8.3
12	justify	verb with <i>voteleave</i> as subject	18	7.9

Appendix 4

Collocational Profiles of the Selected Keywords of Political Actors and Institutions				
Rank	Collocate	Relation	Frequency	LogDice
1	Johnson	noun modified by <i>Boris</i> / compound	78,955	13.9
2	call	noun/verb modified by <i>Boris</i>	607	7.9
3	show	noun modified by <i>Boris</i>	115	5.5
4	speech	noun modified by <i>Boris</i>	89	5.1
5	blunder	noun modified by <i>Boris</i>	65	4.7
6	quiz	verb with <i>Boris</i> as object	36	8.8
7	harry	verb with <i>Boris</i> as object	28	8.5
8	collapse	verb with <i>Boris</i> as object	22	8.1
9	pretend	verb with <i>Boris</i> as subject	208	7.2
10	flounder	verb with <i>Boris</i> as subject	82	6.8

Appendix 5

Collocational Profiles of the Selected Keywords of Evaluative and Ideological Language				
Rank	Collocate	Relation	Frequency	LogDice
1	superficiality	noun modified by <i>opportunism</i>	9	14.0
2	forswear	verb with <i>opportunism</i> as object	25	13.9
3	repellent	adjective modifying <i>opportunism</i>	9	11.9
4	shameless	adjective modifying <i>opportunism</i>	8	10.7
5	unprincipalled	adjective modifying <i>opportunism</i>	3	10.5
6	machiavellian	adjective modifying <i>opportunism</i>	2	9.7
7	press	coordination / association	8	9.9
8	campaign	coordination / association	25	4.8
9	think	verb with <i>opportunism</i> as object	5	4.3
10	political	adjective modifying <i>opportunism</i>	41	4.1
11	party	coordination / association	2	3.3
12	Johnson	possessor of <i>opportunism</i>	2	4.0

Appendix 6

Rank	Collocate	Relation	Frequency	LogDice
1	relationship	noun modified by <i>post-Brexit</i>	20,432	13.1
2	shock	noun modified by <i>post-Brexit</i>	2,596	11.5
3	economic	coordination / association	2,602	10.2
4	future	coordination / association	9	6.6
5	stormy	coordination / association	6	5.9
6	UK	subject of “be post-Brexit”	7	7.6
7	sea	noun modified by <i>post-Brexit</i>	6	2.7