

Brexit references in UK financial disclosure discourse (2015-2017): a corpus study

Vasiliki Simaki

Abstract. Major political events become a dominant topic of the discussions in the countries involved, and various actors, often other than political parties, take a stance towards them. The withdrawal of the UK from the European Union as a result of the 2016 UK referendum was one of those, and a quite polarised atmosphere from supporters of both sides was observed in the UK and the rest of Europe since the referendum was announced in 2015. Based on the impact of the Brexit on the economy, this study focused into the financial domain of the UK, and the ways that UK financial services referred towards it in their financial disclosures before, during and just after the referendum. For that purpose, I collected the 2015, 2016 and 2017 annual reports from five UK-based financial companies (Barclays, HSBC, Lloyds, Royal Bank of Scotland and Santander UK), and I used thematic keywords to identify all the texts referring to the 2016 UK referendum, its outcome, and the response of the financial services to the exiting process of the UK from the European Union. This set of texts composed the Brexit-related data set, in which different analytical tasks were performed. I explored the context in which the thematic keywords are found, compared statistically the three yearly subsets of the corpus, and searched the significant words of the subset in terms of their keyness strength. This case study revealed that the discourse around a major political event such as Brexit cannot be considered as neutral or objective, and that financial companies clearly expressed negative opinion regarding the 2016 referendum and the UK's decision to exit the European Union.

1. Introduction

Major political events often become focal points of discussion, attracting the attention from various stakeholders and organisations. Influential figures, business leaders and financial actors, when engaging in these discussions (directly or indirectly) can shape public opinion and drive market and/or societal changes. The exit of the UK from the European Union, as a result of the 2016 UK referendum, has been a dominant topic of the political discussions in the UK and all over Europe as a potential scenario first, and then as a signed treaty and reality. Since the referendum was announced in 2015, a highly polarised atmosphere from supporters of both sides discussing passionately the terms and processes was observed. The Brexit withdrawal agreement treaty was signed on January 2020, but even before that people in various fields in

politics, economics, market factors, and experts in geopolitics and sociology started talking about the consequences of the event to the UK and the rest of Europe. In an early report from 2018, Oliver Wyman and Clifford Change¹ analysed the Brexit-estimated costs from trade barriers to UK and EU27 firms. They stated that a particularly significant impact (70% of the estimated costs) of the UK's EU exit would be incurred by five sectors in the UK: financial services, automotive industries, agriculture, food and drink, consumer goods, and chemical and plastic industries. This finding inspired this study to focus on one of these domains, and more specifically on financial services, in order to identify the domain's stance towards Brexit in their financial disclosures of annual reports. This case study aims to explore how influential and powerful voices, as for instance the voice of financial services, respond to a major event with discourses that potentially shape the public opinion.

The 2016 UK referendum has been studied from a wide variety of angles and by researchers from different disciplines, such as political and social sciences (e.g., Andreouli et al. 2020, Evans et al. 2017) and economics (e.g., Bulmer et al. 2018, Politt 2022). The various discourses of and communicative strategies regarding Brexit have also been analysed extensively within the Critical Discourse Analysis (CDA) framework. For instance, Zappettini (2019) explores the discourses from the official 'leave' and 'remain' campaigns during the Brexit referendum, and the study focuses on how trade and immigration were framed in these campaigns to legitimise a new 'toxic (inter)national' logic. The author's approach highlighted the main strategies and arguments used by both sides and he argues that the campaigns' rhetoric promoted the idea that by leaving the EU, Britain could 'take back control' and pursue policies that exclude 'outsiders' (Zappettini 2019). In another study, Krzyżanowski (2019) explores how Brexit was framed as both an imaginary and real crisis in European news media. It examines media discourse in Austria, Germany, Poland, and Sweden, showing how past crises were recontextualized to construct Brexit as a significant event. Metaphors, and more specifically the use of marriage and divorce metaphors in Brexit discourse, have also been studied, and Đurović and Silaški (2018) highlighted how these metaphors simplify and clarify the complex relationship between Britain and the EU, using vivid images like 'rocky marriage' and 'messy divorce'. Finally, in their edited volume Koller et al. (2019) examine the discourses surrounding the Brexit referendum. By using a variety of data sources, including political speeches (e.g., Cap 2019) and social media (e.g., Zappavigna 2019), and covering topics such

¹ Oliver Wyman and Clifford Change report (2018): <http://www.oliverwyman.com/our-expertise/insights/2018/mar/red-tape-cost-brexit.html>

as national identity, immigration and the role of media, the ways that different narratives influenced the outcome of the referendum are analysed, and the subsequent discourses that emerged are discussed.

Financial disclosure is the text genre that comprises financial reporting documents of different types, from written texts such as annual reports and earnings announcements, to transcribed oral texts such as earning calls and presentations. Despite the differences in structure, content and context, these different texts share the same purpose, which is the publication of the company's financial status to an expert audience and to society. Corporate reporting has become a communicative activity for a firm, not only in terms of fulfilling its legal obligation to the reporting authorities, but also to demonstrate a positive image to people involved in the process (shareholders, employees, investors, management), and to the general public. Therefore, the study and analysis of the specific text genre is important from a language perspective as well, to identify the linguistic ways in which corporations choose to present their financial data, their strategy, and their goals for the future, not only for the understanding and decoding of the narrative, but also for educational and training purposes.

Annual reports are defined as a sub-genre of organisational communication, and it is the most important tool that companies use to communicate financial information (Bhatia 2004; Courtis 1998). An annual report is a complex multimodal document, consisting of different narratives and sections, aiming not only to the restricted audience of the corporate community and controlling authorities, but also to a larger societal audience, to the media, and to potential customers and investors. It is a complex genre with a dual informational-promotional function: transmit factual data and incorporate a considerable amount of promotional rhetoric to inspire confidence in stakeholders (De Groot et al 2006; Bhatia 2010). Malavasi (2005) has also suggested that annual reports could be studied as a promotional genre, based on the presence and use of evaluative language in this type of documents.

For this study, five large UK-based financial companies were selected: Barclays PLC, HSBC Holdings PLC, Lloyds Banking Group, Royal Bank of Scotland and Santander UK PLC. The annual reports of these companies for the economic years 2015, 2016 and 2017 were collected, and the content related to Brexit was extracted and analysed. My purpose was to observe how Brexit, as an example of a major political event, was discussed by influential institutions, i.e. the financial companies. The sections of the Brexit-related references are located in the reports, similarities and differences of these references in the relevant data from 2015, 2016 and 2017 are observed, and the ways in which financial companies present the

concept are analysed. A statistical analysis of the Brexit-related data set is performed, and keywords are extracted after a comparison of my data to a general reference corpus.

The article is organised as follows. Section 2 describes the background work of my study. In Section 3, the methodology of this study and the corpus are presented. Section 4 presents the various analytical tasks performed on the Brexit-related data set. Finally, Section 5 concludes this article.

2. Corpus methods for the analysis of financial disclosures

The analysis of financial reporting has attracted significant research interest in linguistics, with researchers exploring various aspects such as genre-specific characteristics, thematic elements, and communicative strategies with the use of various tools (e.g., El-Haj et al. 2019). For this paper, studies that followed a corpus-based methodology for the analysis of annual reports were on focus. Bhatia (2008) identified two distinct types of discourse within annual reports. Most of the annual reports could be characterized as accounting discourse, based on the financial data of the company's past performance. However, there was also public relations discourse, particularly in the narrative sections of the report and the Chairman's letter, which aimed to promote a positive image of the company and to inspire confidence among stakeholders.

Annual reports have changed significantly over the years in terms of length (which has increased) and focus. The focus of recent annual reports has shifted to providing more promotional information about the company's activities and achievements, than financial information that was the company's focus in the past (Beattie et al. 2008). The standardisation of these documents to the specific requirements set by the reporting authorities (i.e. European Accounting Standards) resulted to the phenomenon that companies are increasingly reliant on the non-financial elements in the report as a means of distinguishing themselves from their competitors. A corpus approach proved an invaluable tool in facilitating the analysis of a large body of individual text realizations and in uncovering something about those underlying text-genres (De Groot et al 2006). In a genre analysis focusing on corporate disclosure practices in Hong Kong, Bhatia (2008) claimed that in order to have a good understanding of corporate disclosure documents, a comprehensive and evidence-based awareness of the motives, patterns and rules of this specific text type needs to be developed.

In her study, Li (2008) measured the readability of 50,000 company-years (US companies, 10-K filing²), in terms of the syntactic textual features and the report's length. She found that companies with lower earnings tended to file annual reports that are more difficult to read. On the other hand, an increase in earnings from the previous year also resulted in annual reports that were easier to read compared with the previous year's reports, and vice-versa. In an early study using readability measures (passive voice metrics, sentence and word lengths) for the analysis of *Management's Discussion and Analysis* section (MD&A) of the annual reports, it was shown that this section needs to be easier to read, and more comprehensible, or as the authors summarised "less impressive and more expressive" (Schroeder and Gibson 1990). In another study, Curtis and Hasan (2002) studied for the first time the readability of a corpus of annual reports in Hong Kong and Malaysia (English, Chinese, Malay). They showed that the non-English versions of the ARs were written with easier reading ease than the English versions (based on the readability metrics Flesch, Yang and Yunus for the English, Chinese and Malay data respectively). Their findings proved also that the texts in English in the Malaysian reports were easier to read than the English passages in the Hong Kong reports.

In a large corpus of annual reports (over 15,000 non-US companies from 42 countries over the period 1998-2011), Lang and Stice-Lawrence (2015) focused on the length of the disclosure, the presence of boilerplate, the comparability with US and non-US companies, and the complexity. They found that textual attributes are associated with regulation and incentives for more transparent disclosures, and they are correlated with economic outcomes such as liquidity, institutional ownership, and analyst following. Rutherford (2005) performed a stylistic analysis of the *Operating and Financial Report* section of the annual report (OFR, replaced by the Strategic Plan section since 2014), by using genre theory and corpus linguistics methods. The *Polyanna Effect*³ observed in this section of the report, shows that writers keep using positive words despite the reporting authority's instructions towards the use of neutral language. The study proved that this phenomenon was stronger in poorly performing companies, based on the stylistic analysis of the document. This article examined the lexical choices made by the writers of the OFR, and analysed word frequencies, charged words and the OFR's content. The authors identified a specific rhetorical frame within which such narratives that needs to be produced.

² <https://www.investopedia.com/terms/1/10-k.asp>

³ Or else *positivity bias*, it is the tendency for people to remember pleasant items more accurately than unpleasant ones.

Wang et al. (2012) used a corpus consisting of 120 pieces of 2011 annual reports (more than 8 million words), and nine sub-corpora to investigate lexical features in annual reports in an attempt to better understanding this text type. In their study, they analysed the lexical richness, the word length and the keywords, the first personal pronouns, the hedges and the evaluative words. Their results showed that the financial aspect was mostly concerned in these reports. Words and sentence lengths were estimated in the different sections of the ARs, and the phenomenon of the increased use of positively charged words in comparison to the negative ones was also confirmed in this study. In another study, Garzone (2005) focused on concessive connectives (e.g., *yet, although, however*) that countered a conceded weak proposition with a stronger one. These forms were found to be characteristic of executives' annual letters, and seemed to have primarily rhetorical and promotional purposes.

In other studies, researchers focus their interest on specific sections of the annual report, such as the *Chairman's statement*. Chairman's statements were studied by Hildebrandt and Snyder (1981) in order to find the impacts of corporate performance on language used in these narratives. Results showed that the linguistic choices of the writers were predominantly positive no matter what financial status a company had (the so-called Polyanna Effect). Their results were confirmed by Thomas (1997), who analysed the Chairman's statements of the same company's annual reports in five years and found that more words were used to present positive information. In another study, Jacobson (1988) researched on evaluative words in Chairman's statements and concluded that more positive words were used instead of negative words in describing bad news which influenced readers' judgments. De Groot et al. (2006) performed a genre-based study of the textual and pictorial themes in a corpus of Dutch-English and British-English CEO's and Chairman's statements. They observed significant thematic⁴ differences, which can be attributed to the different communicative and historical differences, and to the current affairs in a particular business community.

The *CEO's letter/statement* is also a section of the annual report that was studied from a corpus perspective, due to its significantly different communicative purpose and discourse type compared to the rest of the report. In his work, Hyland (1998) studied the rhetorical text features that the writers of the letters use in order to create a positive image of an effective corporation. He focused on the detection of markers of metadiscourse in CEO's and directors' letters extracted from a corpus containing 137 annual reports in English published in Hong Kong

⁴ The themes here are defined as clusters of words that may have incoherent meanings or connotations outside the scope of a particular text but that are conceptually related within the boundaries of that text.

between 1992 and 1994. Based on the literature on modality and metadiscourse, he created an inventory of textual (logical connectives, sequencers, frame markers, endophoric markers, code glosses) and interpersonal (hedges, emphatics, attributors, attitude markers, relational markers) markers in order to recover the writers' intention by explicitly establishing preferred interpretations of propositional meanings, and to reveal the writers' attitude towards both the propositional information and the readers accordingly. Both CEOs' letters and directors' reports employed more textual than interpersonal forms, and despite that CEOs' letters contained about 2.5 times more metadiscourse markers per 100 words and included six times more interpersonal metadiscourse markers, there were considerable similarities of usage within the two document types. In a later study, taking inspiration from Aristotelian rhetoric, he used the notion of logos (the appeal to rationality) to interpret these devices in letters written by CEOs to shareholders. He showed how this document was used strategically by top executives to support claims and draw conclusions, and how logos was employed to present arguments in a more persuasive way (Hyland 2005). In a comparative study, Conaway and Wardrope (2010) analysed thematically 30 letters from US companies and 24 from Latin American companies. They used a content analysis software to detect thematic⁵, stylistic and readability features, and they analysed the extent to which reoccurring words and relationships appeared in the letters. They concluded that the purpose of CEO's letters may be more strategic and rhetorical than informative, and they serve to communicate more than factual information to their constituencies.

Finally, an important narrative of the annual report is the *Letter to the shareholders* section, and according to the Hyland (1998) and Crombie and Samujh (1999) it is claimed to be the most widely read one. Dragsted (2014) presents a genre- and corpus-based case study of letters to shareholders in annual reports before, during and after the financial crisis, in which changes in themes and linguistic strategies in letters to shareholder from Danske Bank annual reports were detected and analysed. It was shown that the readability and word count of the letters changed over the years: in periods of crisis the readability score and the word count were increased, while in more stable periods these indicators were lower. A qualitative interpretation of the data also highlighted the change of the communicative strategy of the company towards relation-building.

⁵ Corporate governance, Customer relations, Environmental (external economic, political, natural forces), Financial reporting, Infrastructure and expansion, Leadership, Social responsibility, Vision, mission, and outlook

3. Methodology and corpus description

In this section, the methodology and the corpus that was used in this study are presented.

3.1 Methodology

In this study, my goal was to describe and analyse how the selected UK financial services refer to the Brexit in their annual reports. For that purpose, the companies' annual reports from the years 2015, 2016 and 2017 were selected, as different references to Brexit were expected in discourse due to the changes in the political agenda, but also the actual developments. In 2015, the referendum was announced⁶, and the first reactions from different economic and political key factors were monitored. In 2016, both sides (*remain in* and *leave the EU*) campaigned for their purpose, and the political scene as well as the society was quite polarised. This was also the year that the referendum took place (23 June 2016), the Brexit result became a fact, and the exit process started to be discussed. Finally, 2017 was the year when, although the deadline of the UK exit from the EU was a *fait accompli*, the rules and the transition procedures were still debated, and strong reactions from both *pro* and *con* sides were recorded.

In this set of documents, I searched for the text chunks in which Brexit was discussed, and I created a new set of texts: the *Brexit-related data set*. This content was collected manually, by browsing the extracted annual reports for all the key thematic terms related to Brexit such as *Brexit*, *exit*, *referendum*, and all the different sequences of *leave the EU* (see Section 3.2). The context (words, sentences or paragraphs) around these terms was manually evaluated regarding its relevance to the Brexit topic and extracted. The differences in the frequency of these terms among the yearly subsets are discussed in more detail in Section 4. This variance led to the comparison of the yearly subsets and the analysis of the context around the Brexit terms, in order to identify common patterns and differences between the 2015, 2016 and 2017 sets. Statistical tasks and a keyword analysis of the corpus was then performed.

⁶ <https://www.consilium.europa.eu/en/policies/eu-relations-with-the-united-kingdom/the-eu-uk-withdrawal-agreement/2016-uk-settlement-process-timeline/>

3.1 Corpus description

For the purposes of this study, I extracted the 2015, 2016 and 2017 annual reports of five UK-based financial companies⁷ (Barclays PLC⁸, HSBC UK⁹, Lloyds Banking Group¹⁰, Royal Bank of Scotland¹¹ and Santander UK PLC¹²). These documents were publicly available in .pdf format from the companies’ website. The CFIE-FRSE Web Tool¹³ was then used to convert the .pdf files to editable .txt files. The overall size of this set of data is about 3.2 million words (3,151,873 words). In Table 1, the detailed size of the set is presented, for each financial company and yearly set.

Table 1. The size of the annual reports for each financial company per year and in total.

ARs	Barclays	HSBC	Lloyds	RBS	Santander UK	Total
2015	85,842	288,150	215,678	245,769	221,877	1,057,316
2016	247,058	185,697	205,343	266,955	171,629	1,076,682
2017	221,197	185,739	198,406	251,592	162,941	1,019,875
Total:	554,097	659,586	617,427	764,316	556,447	3,151,873

From this data set, I manually extracted all the text chunks related to the 2016 UK referendum. In each annual report (15 documents in total), I searched for specific terms related to my topic of interest, i.e. *Brexit*, *referendum* (only when it is about the 2016 UK referendum), *exit* (only when it is referred to the UK’s exit from the EU) and all the different sequences of *leav* the EU** (e.g., *leaving the EU*, *leave the European Union*, etc.), and extracted these terms with their surrounding context. In Table 2, the occurrences of these Brexit terms in the corpus are presented:

⁷ For this study, the selection of RBS over Bank of Scotland was based on the fact that Bank of Scotland is part of the Lloyds Banking Group. Any political affiliations of each one of the financial services selected are not discussed or taken into consideration in this study but this can be a future direction of my research.

⁸ <https://www.barclays.co.uk/>

⁹ <https://www.hsbc.co.uk/>

¹⁰ <https://www.lloydsbankinggroup.com/>

¹¹ <https://www.rbs.co.uk/>

¹² <https://www.santander.co.uk/>

¹³ CFIE-FRSE Web Tool: <https://cfie.lancaster.ac.uk:8443/>

Table 2. The occurrence of the Brexit terms in the 2015, 2016 and 2017 annual reports.

Keyword	2015 ARs	2016 ARs	2017 ARs	Total
Brexit	1	1	58	60
exit	15	44	50	109
referendum	28	129	23	178
leav* the EU*	4	59	18	81

Table 2 shows the frequency of the Brexit terms in the set of the 2015, 2016 and 2017 annual reports. I selected the size of the text extracted in each case, based on the text's topic and relevance to Brexit. More specifically, the sentence containing the Brexit term was extracted and all surrounding Brexit-related context. The text excerpts extracted vary from one sentence (in which the Brexit term is present) to a bigger chunk of text, one or more paragraphs. The manual selection of this subset resulted to a relatively small but accurate data set, where only in a few texts Brexit is not the main topic, and it is just listed among other political factors¹⁴. This data set contains 28,645 words in total. In Table 3, more detailed information about the Brexit-related data set can be found.

Table 3. The size of the Brexit-related data set.

Financial company	2015	2016	2017	Total
Barclays PLC	258	5,510	2,982	8,750
HSBC UK	216	1,406	1,102	2,724
Lloyds Banking Group	645	2,406	2,022	5,073
Royal Bank of Scotland	835	4,658	4,360	9,844
Santander UK PLC	316	1,259	679	2,254
Total:	2,270	15,239	11,136	28,645

Table 3 shows the distribution of the Brexit-related text among the different financial companies for the years 2015, 2016 and 2017. I observed that the references related to Brexit in the 2015 subset were not as many as the ones in the 2016 and 2017 subsets, but this seemed normal as the referendum was just announced. As a result, not many reactions were recorded at the time of the reports' publication. The largest subset of the Brexit-related data set is the 2016 one, as the highest number of references to Brexit (in terms of number of words) was

¹⁴ An example of such text chunk from the data set follows: "Changes to the approach in 2016 (4) Changes in the macro environment I considered other macro factors to determine if changes in the approach were required, for example the impact of the United Kingdom's decision to leave the European Union, the devaluation of the Mexican Peso and changes in the credit environment. I reported to the GAC in December that I did not believe that these changed my original risk assessment." (Corporate governance: Report of the independent auditors to the members of HSBC UK, HSBC Holdings PLC, Annual Report 2016)

observed in all financial companies' 2016 reports. This is something that can be also considered as expected, as 2016 was a quite full year during which a polarised pre-referendum campaign from both sides was held, and negative reactions by market factors were followed the referendum's outcome. The size of the 2017 subset is smaller than 2016, but this can be justified due to the uncertainty and the lack of specific information about the Brexit process, as well as the political instability and the market's pending (and/or defensive) position regarding the developments. Important differences in the Brexit-related text size among the different financial companies were also observed: HSBC and Santander appear to be quite sparing (less than 3,000 words) with their references to the referendum, its result and consequences, while Barclays and Royal Bank of Scotland refer the most to it (about 9,000-10,000 words). Finally, Lloyds seems to be somewhere in the middle with about 5,000 words of Brexit-related data.

4. Brexit-related data set analysis

In this section, the analytical tasks conducted during this study are presented. First, the distribution of the Brexit-related data in the different sections of the annual reports is presented (Section 4.1), and then, I continue with a collocation analysis of these texts (Section 4.2), their statistical and keyword analysis (Section 4.3).

4.1 Distribution of the Brexit-related data set in the annual report

Each text chunk of the Brexit-related data set is annotated with the annual report's section in which the text is located, such as *Chairman's statement*, *Risk management*, etc. In order to compare the section/s where the different financial companies discuss Brexit, I tried to generalise and align the different annual report's section names that each company uses according to their purpose and information provided. This has been a challenging task, since not every financial company follows a specific template with given sections for their reports, and only few sections, such as the *Chairman's statement*, are clearly common among the companies' reports. Contrary to the USA financial reporting standards that are required by the U.S. Securities and Exchange Commission (SEC), widely known as *Form 10-K*¹⁵, the EU and UK financial regulators (the European Commission and the Financial Reporting Council–FRC respectively) do not provide a strict template for the financial reporting documents but only instructions and rules about the required content to be included in the report. In some cases, for

¹⁵ <https://www.sec.gov/files/form10-k.pdf>

instance, the *Risk overview* was part of the *Corporate Governance* section, or an independent section in the report, so I decided to list it as a separate section. The order of these sections is not the same in all reports, and apart from basic sections such as the *Chairman's statement* that appears at the beginning, and the *Forward looking statements* that appear at the end, other sections can be placed anywhere in the report. In Table 4, the sections of the annual reports where the Brexit-related data is found are presented.

Table 4. The sections of the annual reports that Brexit-related data was found in the 2015, 2016 and 2017 subsets for the five financial companies.

Financial company	Barclays			HSBC			Lloyds			Royal Bank of Scotland			Santander UK		
	2015	2016	2017	2015	2016	2017	2015	2016	2017	2015	2016	2017	2015	2016	2017
Chairman's statement	✓	✓	✓		✓						✓	✓			
CEO's review		✓	✓				✓	✓	✓		✓	✓			
Strategic report	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓			
Corporate Governance		✓	✓			✓			✓		✓		✓	✓	✓
CG: Directors' report		✓	✓	✓	✓			✓	✓					✓	
Risk review	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
Financial review		✓	✓		✓	✓					✓	✓			
Independent Auditors' report		✓			✓	✓					✓	✓			
Other information/ Shareholder information/ Forward looking statements		✓	✓		✓		✓	✓	✓	✓	✓	✓	✓	✓	

Table 4 shows the sections of the annual reports where the Brexit-related data was found in each document. It is observed that in 14 out of the 15 documents, text about Brexit is located in the *Risk review* section of the annual report. This is an important indication showing that all

financial companies acknowledged Brexit as a risk factor for their growth, and its consequences to the company, the market and the overall economy needed to be managed. Brexit-related data is also found in the last section (*Other information/Shareholder information/Forward looking statements*) of 11 annual reports, and in the *Strategic report* of ten reports. In eight documents, I identified Brexit-related text in the *Corporate governance*, and in seven reports in the *Corporate governance* subsection *Directors' report*. Due to its small size, the Brexit-related data of the 2015 reports is not spread in more than four sections (for the case of the 2015 Lloyds annual report), and in most 2015 documents this content is found in one and up to three different sections. The Brexit-related content of the 2016 and 2017 reports is spread in at least five sections of the document for four out of five companies (Barclays, RBS, HSBC and Lloyds). Barclays' references to Brexit are found in nine sections of the 2016 annual report, and eight of the 2017 report. The Royal Bank of Scotland mentions Brexit in eight sections in the 2016 report, and six in the 2017 document. HSBC discusses this topic in seven sections of its 2016 report, and five of the 2017 report. Lloyds's references to Brexit can be found in five of the 2016 report, and six of the 2017 report. Finally, Santander's references to Brexit can be found in four sections in the 2016 report and only two sections in the 2017 report. This spread of the Brexit-related content into the different sections of the annual reports shows that the referendum and UK's European exit was a very relevant and key factor to various aspects of the financial companies' life. In the next section, the various analytical tasks in the Brexit-related data set are described.

4.2 Context analysis of the Brexit terms

In this section, an analysis of the context of the Brexit terms in the Brexit-related data set is performed. For that purpose, the texts were browsed and a collocate analysis was followed. The collocation analysis (McEnery & Hardie 2011) will point to recurrent, *preferred* and salient lexical choices that are associated with the Brexit terms. As shown in Table 2, the different Brexit terms in the corpus were not used in a balanced or alternating way in each of the three yearly subsets, but some of them were more preferred than others depending on the year. This section is further divided in four subsections, each of them presenting the analysis of the Brexit terms *Brexit*, *exit*, *referendum* and *leav* the EU**.

4.2.1 The analysis of the context of the **Brexit** term

In the 2015 and 2016 annual reports, the most politically charged term, *Brexit*, was almost non-existent with only one occurrence in each subset, as companies in these reports possibly avoided referring to the term as the potential outcome of the referendum. I can only assume that companies tended to be reluctant in using this term, and one of the reasons may also be the fact that this term was considered more as the short version of the *British exit* than an established and acceptable term, which was not used yet in official documents. In the 2017 annual reports, this changes significantly, and the *Brexit* term occurred 58 times being the most frequent *Brexit* term. In (1)-(4), some examples extracted from the 2017 subset are presented in order to show some of the context in which this term occurred.

- 1) We are also creating a new UK ring-fenced bank from scratch, resolving and continuing to resolve multiple large legacy conduct matters, as well as preparing the Group for **Brexit**. (Barclays, 2017)
- 2) We continued with our **Brexit** preparations to ensure that Barclays can preserve our access to the EU markets for our customers and clients. (Barclays, 2017)
- 3) Additional structural changes to the Group's operations will also be required as a result of **Brexit**. (RBS, 2017)

In (1)-(3), *Brexit* was a fact for the financial companies, and their concern was to be prepared for the *Brexit* transition and to foresee any consequences this may cause to their company and clients. Besides the uncertainty that was expressed in the 2017 reports about the timeline and the process that was to be followed, the companies tended to refer to themselves more as a strong, powerful and solid *Group* that can manage any *Brexit*-related risk, by making the best effort not to disturb the company's and clients' growth.

Then, the most frequent lexical items that co-occur with the *Brexit* terms were extracted. The AntConc concordancer (Anthony 2024) was used to search for collocates occurring in a five-window span around each of the terms, excluding from the results any stop-words (commonly used words like *the, is, and*). For the *Brexit* term, I only extracted the collocates of the 2017 subset, since it was not a recurrent term in the 2015 and 2016 subsets, as mentioned above. In Table 5, the 10 most frequent collocates in a five-window span of the 2017 subset that were found around the *Brexit* term are presented.

Table 5. The 10 most frequent collocates of the *Brexit* term in the 2017 subset.

2017 subset
result
developments
Group
uncertainty
transformation
required
including
restructuring
political
impact

As shown in Table 5, *result* was the most frequent collocate that refers to the Brexit result. The *developments* collocate refers to the political developments towards the Brexit process, and the *transformation* and *restructuring* collocates refer to the various activities that the companies need to plan as part of their risk management strategy towards Brexit. *Group* is also a frequent collocate that refers to the financial company as a strong Group, as discussed above. *Uncertainty* was also a frequent collocate which has a negative meaning, and it is directly connected to the Brexit process and transition (it was not discussed as a potential scenario anymore), and the uncertainty that this process created to the company, market, clients, growth, etc. *Political* collocates in most cases with *developments*, *risk/s* and *uncertainty*, while *impact* was mostly used as a verb in order to describe how and under what terms the Brexit or the Brexit-related actions were going to impact the companies, the market, the clients, etc. Finally, *required* and *including* were collocates with a more neutral meaning and use in the text.

4.2.2 The analysis of the context of the exit term

The second Brexit term searched in all three Brexit-related data set subsets was the term *exit* (strictly within the context of the *UK exit from the EU*). This item was used more frequently than the *Brexit* term in the 2015 and 2016 subset (15 and 44 times respectively). In (4)-(5), I present examples of the *exit* term found in the 2015 subset.

- 4) In the UK, the referendum on EU membership gives rise to some political uncertainty and raises the possibility of a disruptive and uncertain **exit** from the EU, with attendant consequences for investment and confidence. (Barclays, 2015)

- 5) An **exit** could have a significant impact on UK, European and global macroeconomic conditions, as well as substantial political ramifications. (HSBC, 2015)

As shown in (4)-(5), even though the exit of the UK from the EU was not a fact yet, already the financial companies discussed the significant impact of the Brexit, and the instability, uncertainty and negative impact that this event was going to bring to the companies, market, and political developments. Without suggesting any measures to mitigate the negative impact of Brexit, companies tended to discuss in an abstract way the implications of the exit to the political and economic scene, as well as the lack of information by the UK government on how to handle such a development. In (6)-(8), some examples of the exit term in the 2016 subset are presented.

- 6) Similarly, the impact of the planned **exit** of the UK from the EU could potentially have an impact on our ability to hire and retain key employees. (Barclays, 2016)
- 7) A way forward to capture the opportunity and mitigate the risk of the UK's **exit** from the EU needs to be found, depending on the final international agreement. (Barclays, 2016)
- 8) The result of the referendum means that the long-term nature of the UK's relationship with the EU is unclear and there is uncertainty as to the nature and timing of any agreement with the EU on the terms of **exit**. (Barclays, 2016)

In these examples, scepticism, uncertainty, and the companies' efforts to assess the referendum outcome are observed, as well as their first thoughts on how to mitigate the risks and potential impact of the UK's EU exit to their growth. To some extent, in the 2016 narratives, the UK exit was still discussed as a hypothetical and long-term scenario since the referendum was recent and its outcome was still not well processed. As a result, the companies, the market and the clients seemed to be benumbed to the result and the developments. In contrast, this is not the case for the following year (2017), when the referendum result was an actual fact, and further actions were to be made to make it happen. In the 2017 subset the *exit* term appeared 50 times, and in (9)-(11) some examples are provided.

- 9) We are making comprehensive plans for the UK's planned **exit** from the EU and we believe we will provide an uninterrupted service to our clients, consumers and other stakeholders during and after the transition. (Barclays, 2017)

- 10) An uncertain UK and global economic outlook and uncertainty relating to EU **exit** negotiations have the ability to impact the Commercial Banking portfolios. (Lloyds, 2017)
- 11) We hold regular meetings with UK authorities to discuss strategic contingency plans covering a wide range of scenarios relating to the UK’s **exit** from the EU. (HSBC, 2017)

In Examples (9)-(11), the companies’ narrative towards the UK exit from the EU did not change much from the 2016 content, with the potential impact and the uncertainty that the Brexit result created still being discussed. The timeline of the developments as well as possible actions by the UK-EU authorities and the companies’ reaction to those were described, but in an abstract and theoretical way. In Table 6, I present the 10 most frequent collocates in a five-window span to the *exit* term for the three yearly subsets.

Table 6. The 10 most frequent collocates of the *exit* term in the 2015, 2016 and 2017 subsets.

2015 ARs	2016 ARs	2017 ARs
UK	EU	EU
EU	UK	UK
Union	<i>risks</i>	negotiations
European	<i>referendum</i>	<i>impacts</i>
<i>uncertainty</i>	outcome	<i>risks</i>
supporting	<i>possible</i>	<i>potential</i>
relating	<i>impact</i>	<i>impact</i>
<i>referendum</i>	Group	countries
<i>potential</i>	scenarios	<i>uncertainty</i>
occurs	<i>risk</i>	triggered

In Table 6, the common words in all three subsets are highlighted, and the common words in at least two subsets are in italics. The *exit* term collocated more frequently, as expected, with EU and UK (*UK exit from the EU* or *EU exit*), and this is a common characteristic in all three yearly subsets. The 2015 collocates *Union* and *European* were also frequent as the full version of the *EU*. The words *uncertainty* and *referendum* that are directly related to the UK exit from the EU were also frequent and co-occurred in the 2017 (the *uncertainty* one) and the 2016 subsets (the *referendum* one), as well as the word *potential*. The words *possible* and *potential* are synonyms, so I considered them as a common collocate used around *exit* in all three yearly narratives, referring to possible/potential risks, changes, impacts, etc. of the UK exit from the EU. *Supporting*, *relating* and *occurs* are more neutral words. In the same context with the *exit*

term, the words *risk-s*, and *impact-s* (as noun) occurred frequently in 2016 and 2017 subsets. In the 2016 subset, the remaining collocates that are common with the 2017 ones are *outcome*, *Group* and *scenarios*. The words *outcome* and *scenarios* were used in a non-charged, more descriptive way to refer to the terms of the UK exit and the scenarios about the timing and the developments. The word *Group* had the same use as mentioned in previous cases.

In the 2017 narrative, the unique collocates are *negotiations*, *countries*, and *triggered*. The *negotiations* collocate was used frequently in a way to pinpoint the difficulty of this process, and to highlight the uncertainty of such action (e.g., *While elections across the EU during 2017 have temporarily stemmed a populist tide, political uncertainty remains high in the UK as **negotiations** progress towards an exit from the EU (see 'Process of UK withdrawal from the European Union' in Areas of special interest on page 66). (HSBC, 2017)*). In other cases, it was just used neutrally in a brief sentence or a headline to refer to the negotiation process or terms of the UK exit from the EU, and then, more details about this are given in the next few sentences, frequently in a charged and not very optimistic way (e.g., *EU exit KEY ISSUES **Negotiations** are on-going to determine the terms of the UK's exit from the EU. The uncertainty regarding the timing and the process itself could affect the outlook for both the UK and global economy. (Lloyds, 2017)*). The *countries* collocate refers to the European countries as a group or individuals, and it is used in the context of describing whether the UK's relationship with these countries will change after Brexit or not. Finally, the *triggered* collocation is used when the UK's EU exit process is discussed (e.g., *On 29 March 2017 the UK Government **triggered** the exit process contemplated under Article 50 of the Treaty on European Union. (RBS, 2017)*).

4.2.3. The analysis of the context of the referendum term

The third term in the Brexit-related data set was the term *referendum* (only within the context of the 2016 UK referendum). This term was the most frequently used in the 2016 subset with 129 occurrences, while it was less frequent in the 2015 and 2017 subsets with 28 and 23 occurrences respectively. In (12)-(17), two examples of sentences with the *referendum* term from each yearly subset are presented:

- 12) We continue to deal with a range of uncertainties in the external environment, including those caused by the **referendum** on the UK's continuing membership of the European Union. (RBS, 2015)

- 13) We are closely monitoring political developments in the lead up to the **referendum**, assessing potential risks, and planning mitigating actions where appropriate. (Santander UK, 2015)
- 14) This uncertainty is compounded by the UK's decision to leave the EU following the outcome of the EU **Referendum** which may result in further changes to the prudential and regulatory framework applicable to the Group. (RBS, 2016)
- 15) A meeting was held specifically to consider the risk considerations arising from the outcome of the EU **Referendum** result, further details of which can be found on page 72. (Barclays, 2016)
- 16) The Group is subject to political risks, including economic, regulatory and political uncertainty arising from the **referendum** on the UK's membership of the European Union which could adversely impact the Group's business, results of operations, financial condition and prospects. (RBS, 2017)
- 17) Political risks continue to evolve with the UK's vote to leave in the EU **referendum** creating significant economic, political and regulatory uncertainty. (RBS, 2017)

As shown in Examples (12)-(13), the narrative remained the same as observed with the other Brexit terms. The companies showed a wait-and-see attitude, but they highlighted the uncertainty and the risks of a potential Brexit referendum outcome. The changes, the risks, the uncertainty and the potential consequences of the referendum outcome were discussed in the sentences where the *referendum* term occurred in the 2016 subset (Examples (14)-(15)) in a more systematic way, since the outcome of the UK vote in the 2016 referendum was then a fact. The high frequency of the term in this subset is not surprising, as that was the year that the referendum took place and the event as well as the result was widely discussed. Finally, in the examples from the 2017 subset ((16)-(17)), the narrative did not change much. Only in a few sentences, in which more detailed information about the consequences of the Brexit was provided, a different narrative is observed with companies' actions and changes in financial data described. In Table 7, the ten most frequent collocates of the *referendum* term in the 2015, 2016 and 2017 subsets are presented.

Table 7. The 10 most frequent collocates of the *referendum* term in the 2015, 2016 and 2017 subsets.

2015 ARs	2016 ARs	2017 ARs
<i>UK</i>	EU	EU
EU	<i>following</i>	<i>UK</i>
<i>membership</i>	<i>membership</i>	outcome
<i>risks</i>	<i>result</i>	political
outcome	outcome	<i>June</i>
lead	<i>June</i>	<i>following</i>
hold	impact	<i>result</i>
rise	post	uncertainty
potential	after	Scottish
period	<i>risk</i>	monitored

In Table 7, the common items between the three yearly subsets are highlighted, and the common collocates in at least two yearly subsets are in italics. The *referendum* term collocated frequently as expected with EU, and this is a common word in all three subsets. The *outcome* collocate is also common in all three subsets, and it refers to the result of the referendum. The *UK* was the most frequent collocate to the *referendum* term in the 2015 subset, and it was also a frequent in the 2017 subset. The words *membership* and *risk-s* co-occurred with the *referendum* term in the 2015 and 2016 subsets. The first refers to the UK's membership in the European Union, and the second to the potential risk-s of the UK's EU exit. The *lead*, *hold*, *rise*, and *period* collocates have a neutral meaning, and the word *potential* refers to the impact, risk-s, outcome and consequences of the referendum.

In the same context with the *referendum* term the words *following*, *result* and *June* co-occurred frequently in the 2016 and 2017 subsets. The *following* collocate was mostly used when actions or impacts of the referendum outcome were about to be presented (e.g., *I am very proud in particular of how, **following** the EU referendum last June, Barclays continued to be a constructive partner to our customers and clients, and to the Government, as we dealt with the initial impact of that decision.* (Barclays, 2016)). The *result* collocate had the same use as the *outcome* one, and they were both used in an alternating way, though the word *result* was more frequent in both 2016 and 2017 subsets. *Result* was mostly used in the sequence *referendum result*, while *outcome* was mostly used in the sequence *outcome of the (EU) referendum*. The *June* collocate refers to June 2016, when the UK referendum took place (23rd of June 2016). In the 2016 subset, the *impact* collocate was used in the same way that was described for the other terms, and the *post* collocate was used to mark the time period after the

referendum (e.g., *Business and consumer confidence did fall immediately **post** referendum, but most of this has now been recovered and consumers' retail spending growth actually accelerated in the months after the referendum.* (Lloyds, 2016)). The same applies to the *after* collocate.

In the 2017 subset, the *political* collocate was used to describe the impact of the referendum outcome to the political developments in the UK, and it co-occurs in most cases with *uncertainty* and *instability* (e.g., *The outlook for the UK and global economy remains uncertain due to a number of factors including: the UK's vote to leave in the EU referendum, wider **political** instability, an extended period of low interest rates, high debt levels and delays in normalising monetary policy.*(RBS, 2017)). The *uncertainty* collocate was used in the same way and context as described in the previous cases, the *Scottish* one refers to the Scottish referendum as a scenario aligned to the UK referendum that was only discussed in the RBS' report, and finally, *monitored* was used in the context of highlighting the impact, the uncertainty, and the risk of the referendum outcome.

4.2.4. The analysis of the context of the *leav the EU* term**

The fourth Brexit term, or, more precisely, term-sequence, I searched in the corpus is the *leav** the EU* term. I searched in the three yearly subsets all possible sequences such as *leave the EU*, *leave the European Union*, *leave the Eurozone*, *leave the euro*, *leaving the EU*, and I hit only four occurrences in the 2015 subset, 59 in the 2016 and 18 in the 2017. In (18)-(22), examples of those occurrences are presented.

- 18) Following the referendum in June 2016, in the event that there is a vote in favour of **leaving the EU**, a period of negotiation is likely, widely anticipated to be around two years, with unpredictable implications on market conditions. (Barclays, 2015)
- 19) In light of these potential developments as well as the impact of the UK's decision to **leave the EU** following the result of the EU Referendum, there remains uncertainty as to the rules which may apply to the Group going forward. (RBS, 2016)
- 20) Finally, there were the challenges presented by emerging economic and political risks, notably those associated with the EU Referendum and the subsequent vote by the UK to **leave the EU**. (Barclays, 2016)
- 21) This had been prompted by the rising level of personal debt in the UK and concerns of weaker growth and higher inflation resulting from the country's vote to **leave the European Union**. (Barclays, 2017)

- 22) In addition, it is possible (although of low likelihood) that a disorderly termination of the Article 50 process could occur, resulting in the UK **leaving the EU** before 29 March 2019. (RBS, 2017)

The *leav* the EU** term was not frequent in the 2015 subset (only four cases), and as a result, I did not have enough data to observe any differences in the companies' narrative. In the 2016 subset, the term was quite frequent (59 hits in the data), but no differences were observed in the narrative in comparison to the other Brexit terms. This sequence was considered as the most descriptive term of the 2016 UK referendum outcome. The choice of the verb *leave* in this sequence is interesting, from a semantics point of view, as other verbs such as *quit* and/or *withdraw* could be used instead in a more literal approach to the Brexit process, but *leave* is the most frequent one in my corpus, and it gives a more sentimental and polarised approach to the UK's vote to exit the EU. In the 2017 subset, the *leav* the EU** term was less frequent with only 18 hits in the data. In these texts, the narrative did not change significantly, but in some of the cases more information about the process was added. Since the timeline and the process of the Brexit was still an abstract concept (the exit's completion initial date was on the 29th of May 2019, but then extensions were given¹⁶), the companies highlighted the potential risks and fears as inevitable events, and emphasised with negative words such as *weak*, *concern*, *uncertain*. In Table 8, the ten most frequent collocates of the *leav* the EU** term in a five-window span in the 2016 and 2017 subsets are presented.

Table 8. The 10 most frequent collocates of the *leav* the EU** term in the 2016 and 2017 subsets.

2016 ARs	2017 ARs
EU	EU
UK	UK
vote	decision
decision	resulting
following	March
result	following
outcome	before
June	uncertainty
favour	terms
expected	significant

¹⁶ <https://www.consilium.europa.eu/en/policies/eu-relations-with-the-united-kingdom/the-eu-uk-withdrawal-agreement/timeline-eu-uk-withdrawal-agreemen>

In Table 8, the common items between the 2016 and 2017 subsets are highlighted. Four out of ten collocates are common in both sets. The first two collocates *UK* and *EU* were expected to be the most frequent ones in both subsets. The *decision* collocate was also to be expected, since the sequence *decision to leave the EU** occurs 12 times in the 2016 subset and six times in the 2017 subset. The *following* collocate was mostly used to connect the surrounding atmosphere of uncertainty to the UK's decision to leave the EU (e.g., *In the UK, following the decision to leave the EU, concerns persist about the upcoming exit negotiations and the ultimate nature of the EU-UK relationship.* (HSBC, 2016)). In the 2016 subset, the *vote* collocate was mostly used as a noun in the sequence *(UK's) vote to leave the EU**, and it functioned as a synonym to the *referendum outcome*. The *result* and *outcome* collocate refer to the referendum's result, and the *June* collocate refers to June 23rd, which was the exact date that the referendum took place. The *favour* collocate was used in only two cases where the *vote in favour of leaving the EU* was discussed. Finally, the *expected* collocate refers to the expected decisions, changes and impact of the UK's decision to leave the EU.

In the 2017 subset, the *resulting* collocate was used mostly after the *leav* the EU** term in a way to highlight the consequence of that decision. The *March* collocate refers to the 29th March 2019, the date when the UK's exit process from the EU should have been completed. *Before* and *terms* are neutral words used in this context to define until when all actions should have been completed and the terms of the Brexit process. The *uncertainty* collocate was used in the same way as in the previous cases, and finally, the *significant* collocate co-occurs with the impact, risks, uncertainty and volatility that leaving the EU would cause. In the next section, statistical and keyword analysis of the Brexit-related data set is performed.

4.3. Statistical and keyness analysis of the Brexit-related data set

The different frequency of the Brexit terms, as well as the slightly varied narratives in each of the three yearly subsets, led me to an overall comparison of the three yearly sets, where differences and similarities were highlighted. For that purpose, a statistical analysis of the lexical items of the Brexit-related data set was performed. The AntConc concordancer (Anthony 2024) was used to generate the dataset's word list, and the 100 most frequent lexical items were extracted. The observed and the expected values of these items in each yearly subset were calculated, and a Chi-square (χ^2) test was performed. The purpose of this test was to determine the population distribution of each subset in the Brexit-related data set. To achieve

this goal, a null hypothesis was formulated, which states that either there is no specific preference in the given categories, or that the percentages preferred by the variables are not different from the percentages of other populations. The null hypothesis was evaluated by comparing the expected values of the variables in each subset to the observed ones. The null hypothesis was rejected for 32 variables that had a p-value <0.05. In Table 9, the results of the test are presented (the actual values for the 32 constructions in each subset can be found in Appendix 1).

Table 9. The statistically significant words in alphabetical order with their % of occurrences across each yearly subset, and their Chi-square p-value.

Words	2015	2016	2017	p-value
ability	0.31%	0.10%	0.15%	0.048
any	0.57%	0.24%	0.20%	0.005
as	1.01%	0.78%	1.21%	0.002
Barclays	0.04%	0.41%	0.27%	0.006
Brexit	0.04%	0.01%	0.52%	0.000
could	0.62%	0.21%	0.34%	0.002
developments	0.44%	0.18%	0.22%	0.049
effect	0.22%	0.07%	0.04%	0.019
European	0.62%	0.28%	0.40%	0.019
exit	0.66%	0.26%	0.40%	0.006
following	0.13%	0.31%	0.15%	0.021
for	0.35%	0.85%	0.86%	0.040
geopolitical	0.26%	0.04%	0.12%	0.001
government	0.31%	0.06%	0.09%	0.001
Group	0.88%	0.81%	1.12%	0.036
leave	0.09%	0.28%	0.13%	0.015
membership	0.75%	0.16%	0.05%	0.000
on	1.72%	1.28%	0.91%	0.001
or	1.41%	0.56%	0.59%	0.000
ratings	0.35%	0.10%	0.08%	0.002
RBS	0.48%	0.12%	0.13%	0.000
referendum	1.28%	0.85%	0.22%	0.000
regulatory	0.44%	0.28%	0.51%	0.011
risk	0.70%	0.80%	0.54%	0.041
risks	0.70%	0.36%	0.47%	0.049
structural	0.04%	0.14%	0.24%	0.040
UK	2.16%	1.38%	1.61%	0.015
Union	0.48%	0.12%	0.19%	0.001
vote	0.26%	0.20%	0.03%	0.000
was	0.00%	0.28%	0.09%	0.000
with	0.13%	0.51%	0.73%	0.001
would	0.26%	0.09%	0.07%	0.021

As shown in Table 9, 32 lexical items were statistically significant and had a p-value lower than 0.05. To establish that there is a significant difference in the frequency of these words among the three yearly subsets and to calculate the familywise error rate, a Bonferroni correction post-hoc test was further applied, setting a lower significance level (<99% of confidence). The p-value equal to 0.0001 was calculated by dividing the familywise error rate (0.05) by the number of tests (96 in my case) and the critical value was equal to ± 3.469 . In Table 10, the adjusted residuals of the Bonferroni correction are presented.

Table 10. The adjusted residuals of the Bonferroni correction post-hoc test for the 32 words.

Words	2015	2016	2017
ability	2.244	-1.674	0.471
any	3.186	-0.308	-1.450
as	0.234	-3.432	3.383
Barclays	-2.467	2.690	-1.387
Brexit	-1.796	-8.008	9.193
could	2.971	-2.778	1.198
developments	2.394	-1.270	-0.027
effect	2.696	-0.075	-1.417
European	2.213	-2.344	1.173
exit	2.667	-2.556	1.139
following	-1.046	2.784	-2.270
for	-2.548	0.665	0.731
geopolitical	2.977	-2.927	1.347
government	3.588	-1.900	-0.043
Group	-0.299	-2.346	2.567
leave	-1.318	2.870	-2.207
membership	7.175	-0.294	-3.675
on	2.534	1.848	-3.296
or	4.804	-1.836	-0.782
ratings	3.578	-0.717	-1.248
RBS	4.106	-1.477	-0.763
referendum	3.980	4.703	-7.020
regulatory	0.454	-2.971	2.790
risk	0.082	2.382	-2.483
risks	2.092	-1.890	0.775
structural	-1.526	-1.452	2.332
UK	2.530	-2.173	0.822
Union	3.610	-2.284	0.337
vote	1.726	2.971	-3.998
was	-2.117	3.988	-2.909
with	-2.870	-1.292	2.913
would	2.752	-0.526	-0.986

In Table 10, the results of the Bonferroni correction are presented for the 32 lexical items that were statistically significant after the Chi-square test. The frequencies of occurrence of nine words differed significantly between the three subsets. The cases where the statistically significant values were greater than the expected ones (for values higher than 3.469) were marked with green, and the cases where the statistically significant observed values were smaller than the expected ones (for values lower than -3.469) with red. The words *government*, *membership*, *or*, *ratings*, *RBS*, *referendum*, and *Union* proved to be significant for the 2015 subset. *Government* refers to the UK government, *membership* is either the *UK's membership of the EU* or the *EU membership*, or, it is just a frequent neutral word in this subset. *Ratings* is used when the *credit ratings of the UK government* is discussed, *RBS* is the acronym of the Royal Bank of Scotland, and *referendum* refers to the 2016 UK referendum. The words *referendum* and *was* are statistically significant words for the 2016 subset, with the first referring to the 2016 UK referendum and the second being a frequent function word. The word *Brexit* is the only significant word for the 2017 subset. The low number of the statistically significant words shows that there is not an important difference among the three yearly subsets, and especially between the 2016 and 2017 Brexit-related data.

The statistical comparison of the 2015, 2016 and 2017 subsets, which confirmed the similarity in the narrative between these reports, followed a keyness analysis (e.g., Baker 2004, Gavrielatos 2018) to identify the most important words of the Brexit-related data set. For that purpose, the keywords function of the AntConc corpus tool that calculates the words' keyness strength in terms of the values' log-likelihood was employed, and the British National Corpus (BNC)¹⁷ was used as a reference corpus. I compared the Brexit-related data set's wordlist to the BNC's wordlist, as well as each yearly subset's wordlists to the BNC's wordlist. In Table 11, the list with the 20 highest ranked words, according to their keyness score, are presented excluding the stop-words.

¹⁷ <http://www.natcorp.ox.ac.uk/>

Table 11. The 20 highest ranked keywords of the Brexit-related data set, and separately of the 2015, 2016 and 2017 subsets.

Brexit-related data set	2015	2016	2017
EU	<i>EU</i>	<i>EU</i>	<i>EU</i>
UK	<i>UK</i>	<i>UK</i>	<i>UK</i>
referendum	<i>referendum</i>	<i>referendum</i>	Brexit
impact	<i>RBS</i>	<i>risk</i>	<i>Group</i>
risk	<i>exit</i>	<i>impact</i>	<i>regulatory</i>
Group	<i>risks</i>	<u>Barclays</u>	<i>risks</i>
regulatory	membership	<i>Group</i>	<i>exit</i>
risks	<i>impact</i>	<i>risks</i>	<i>impact</i>
Brexit	<u>geopolitical</u>	<u>volatility</u>	<i>risk</i>
Barclays	<i>regulatory</i>	<i>regulatory</i>	<i>uncertainty</i>
exit	<i>risk</i>	<i>exit</i>	<u>Barclays</u>
uncertainty	<i>uncertainty</i>	<u>potential</u>	<i>including</i>
including	<i>including</i>	<i>including</i>	<i>referendum</i>
volatility	<i>Group</i>	<i>uncertainty</i>	<u>financial</u>
RBS	ratings	<i>credit</i>	<u>geopolitical</u>
potential	developments	economic	structural
credit	<u>European</u>	markets	<u>volatility</u>
financial	changes	Eurozone	<i>RBS</i>
economic	<i>credit</i>	<i>RBS</i>	<u>European</u>
global	<u>potential</u>	<u>financial</u>	<i>credit</i>

Table 11 shows the 20 highest ranked keywords in a descending, in terms of keyness strength, order of the Brexit-related data set, and of the yearly subsets. In this table, the words that are common in all three yearly subsets are italicised, the words that are common in two subsets are underlined, and the words that are unique are emphasised (highlighted with bold). In some cases, e.g., *referendum*, *membership* and *Brexit*, the keywords extraction process concompanies the results of the statistical analysis. In general, few differences and unique keywords are observed among the three lists. In Table 12, the Brexit-related data set's keywords according to their use in the text are grouped into three categories: (i) keywords that are terms related to Brexit, (ii) keywords that are charged words used when the Brexit process is discussed and describe the actions to be taken (either by the companies or the UK government) and the surrounding atmosphere, and, (iii) keywords that are used in the Brexit-related data set, but are more general and neutral:

Table 12. The overall Brexit-related data set's keywords categorised in terms of their use in the text.

Brexit terms	Charged words	Neutral words
EU	risk/risks	Group
UK	impact	Barclays
referendum	regulatory	including
Brexit	uncertainty	RBS
exit	volatility	potential
		credit
		financial
		economic
		global

In the first column of Table 12, the keywords that are Brexit terms are presented. In the second column, the keywords that are mostly used to describe the surrounding atmosphere towards Brexit are presented. As discussed in previous sections the words *risk/s*, *impact*, and *uncertainty* are mostly used to describe in a negative way the companies' reaction and impression about the 2016 UK referendum. A new word, *volatility*, is added to these words without changing the spirit, and it emphasises the instability that the referendum's outcome has created to the companies, the clients, and the market (e.g., *The longer term effects of Brexit on the Group's operating environment are difficult to predict, and subject to wider global macro-economic trends and events, but may significantly impact the Group and its customers and counterparties who are themselves dependent on trading with the EU or personnel from the EU and may result in periodic financial volatility and slower economic growth, in the UK in particular, but also in Republic of Ireland, Europe and potentially the global economy.* (RBS, 2017)). In this category, I also included the adjectives *regulatory*, as in that context, it refers to the transformations, reforms and changes that need to be done throughout the Brexit process. In the third column, the words that are used in a neutral way within the context are listed. In this group words like *Group*, *Barclays*, *RBS*, or the adjectives that refer to many cases to the Brexit actions and consequences, but not always, such as *potential*, *financial*, *economic*, *global*, and other neutral words such as *including* and *credit* are included.

Before this analysis, I expected to observe a strict, neutral, official and objective narrative towards politico-economic developments like Brexit. The abstract transition terms and the difficulties that the vague frame towards the pre- and post-referendum were surely expected to be part of the narrative from the financial companies, but instead of a cool, calm and collected attitude regarding the whole process and results, an obviously negative and polarised position

was observed. This is an overall conclusion for all companies' narrative towards Brexit, and I assume that a charged discourse from their side contributed to the overall negative atmosphere rather than attenuating the public reactions. The lack of a specific frame and instructions regarding the Brexit process from the part of the UK government was an argument to the companies' narrative, as they did not have significant information to assess the actual cost and consequences from Brexit for their own business and the impact to their clients and market. The pound devaluation that was observed since November 2016 (e.g., Plakandaras et al. 2017) in combination to the slow growth rate of the UK¹⁸ were among the factors that supported the companies' negative position. I cannot be assertive that financial companies tend to take such strong position towards other politico-economic developments, but my study shows that they took a relatively strong opinion towards Brexit, and their narrative may have influenced their clients, market factors, media sources and the public opinion.

5. Conclusion

Various stakeholders often take a stance towards a political event and engage in discourses that can influence not only the market but also the society. A corpus approach can be followed to study these discourses and identify the actors' positioning towards the given events. In this paper, I selected a major event (i.e., Brexit) and the discourse of financial services (i.e., annual reports) towards this event, and presented a corpus-based and statistical methodology to analyse their positioning. The impact of Brexit to the financial services led me to the compilation of a small corpus of manually selected content related to Brexit from the 2015, 2016 and 2017 annual reports of five UK-based financial companies. This data set was analysed, and the three yearly subsets were compared in order to examine whether there are differences in the companies' narrative regarding the 2016 UK referendum. The results of the statistical analysis showed that there are not many important differences in the way that financial companies refer to the Brexit in 2015, 2016 and 2017, regardless the different conditions and developments during this period. I observed differences in the use of Brexit terms such as *Brexit* and *referendum*, and minor differences in the three yearly subsets that do not change significantly the overall narrative towards the 2016 UK referendum and the Brexit process. The financial companies started their narrative in 2015 with the uncertainty that the referendum announcement and potential EU exit brought to the UK. They continued in 2016

¹⁸ <https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=GB>

by arguing more systematically about the impact, uncertainty and risks to the UK economy and growth rates of the referendum outcome, and finally, in 2017, they kept denoting the negative impacts of the Brexit as a fact.

I extracted the Brexit-related data set's keywords by using the BNC as a reference corpus, and apart from Brexit terms and neutral words that appear in the keywords list, I detected few charged words such as *risk*, *uncertainty*, *volatility* and *impact*. The presence of these negatively polarised words in the reports gave me the impression that financial services tended to position themselves in a generic and negative way towards the UK's exit decision from the EU. Although financial companies affirmed their ability to deal with any Brexit scenario, and they revealed to their shareholders and clients that special teams are sorting out the company's plan about the management of the Brexit-related risk and impact, they kept using a negatively charged language, maintaining the uncertainty and negative atmosphere around.

References

- Andreouli, E., Greenland, K., & Figgou, L. (2020). Lay discourses about Brexit and prejudice: "Ideological creativity" and its limits in Brexit debates. *European Journal of Social Psychology*, 50(2), 309-322. <https://doi.org/10.1002/ejsp.2625>
- Anthony, L. 2024. AntConc (Version 4.3.1) [Computer Software], Waseda University, Tokyo, Japan. Available at <http://www.laurenceanthony.net/software>
- Baker, P. (2004). Querying keywords: Questions of difference, frequency, and sense in keywords analysis. *Journal of English linguistics*, 32(4), 346-359. <https://doi.org/10.1177/0075424204269894>
- Beattie, V., Dhanani, A., & Jones, M. J. (2008). Investigating presentational change in UK annual reports: A longitudinal perspective. *The Journal of Business Communication* (1973), 45(2), 181-222. <https://doi.org/10.1177/0021943607313993>
- Bhatia, V. K. (2010). Interdiscursivity in professional communication. *Discourse & communication*, 4(1), 32-50. <https://doi.org/10.1177/1750481309351208>
- Bhatia, V. K. (2008). Genre analysis, ESP and professional practice. *English for specific purposes*, 27(2), 161-174. <https://doi.org/10.1016/j.esp.2007.07.005>
- Bhatia, V. (2004). *Worlds of written discourse: A genre-based view*. A&C Black.
- Bulmer, S., & Quaglia, L. (2018). The politics and economics of Brexit. *Journal of European public policy*, 25(8), 1089-1098. <https://doi.org/10.1080/13501763.2018.1467957>

- Cap, P. (2019). 'Britain is full to bursting point!': Immigration themes in the Brexit discourse of the UK Independence Party. In *Discourses of brexit* (pp. 69-85). Routledge.
<https://doi.org/10.4324/9781351041867-5>
- Conaway, R. N., & Wardrope, W. J. (2010). Do their words really matter? Thematic analysis of US and Latin American CEO letters. *The Journal of Business Communication* (1973), 47(2), 141-168. <https://doi.org/10.1177/0021943610364523>
- Courtis, J. K., & Hassan, S. (2002). Reading ease of bilingual annual reports. *The Journal of Business Communication* (1973), 39(4), 394-413.
<https://doi.org/10.1177/002194360203900401>
- Courtis, J. K. (1998). Annual report readability variability: tests of the obfuscation hypothesis. *Accounting, Auditing & Accountability Journal*, 11(4), 459-472.
<https://doi.org/10.1108/09513579810231457>
- Crombie, W. and Samujh, H. (1999): Negative Messages as Strategic Communication: A Case Study of a New Zealand Company's Annual Executive Letter. *Journal of Business Communication*, 36(3): 229-246. <https://doi.org/10.1177/002194369903600301>
- Dragsted, B. (2014). A case study of letters to shareholders in annual reports before, during and after the financial crisis. *LSP Journal-Language for special purposes, professional communication, knowledge management and cognition*, 5(2).
- Garzone, G. (2005). Pragmatic and discursual features of annual executive letters: observations on the rhetorical and evaluative function of concessive constructions. In *Cross-cultural encounters: Linguistic perspectives* (pp. 130-143). Officina Edizioni.
- Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus Approaches to Discourse: A Critical Review* (pp. 225–258). Routledge, 10, 9781315179346-11. <https://doi.org/10.4324/9781315179346-11>
- De Groot, E. B., Korzilius, H., Nickerson, C., & Gerritsen, M. (2006). A corpus analysis of text themes and photographic themes in managerial forewords of Dutch-English and British annual general reports. *IEEE Transactions on Professional Communication*, 49(3), 217-235. <https://doi.org/10.1109/TPC.2006.880755>
- Đurović, T., & Silaški, N. (2018). The end of a long and fraught marriage: Metaphorical images structuring the Brexit discourse. *Metaphor and the Social World*, 8(1), 25-39.
<https://doi.org/10.1075/msw.17010.dur>
- El-Haj, M., Rayson, P., Walker, M., Young, S., & **Simaki, V.** (2019). In Search of Meaning: Lessons, Resources and Next Steps for Computational Analysis of Financial Discourse.

Journal of Business Finance and Accounting, 46, 265-306.

<https://doi.org/10.1111/jbfa.12378>

Evans, G., & Menon, A. (2017). *Brexit and British politics*. John Wiley & Sons.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. A&C Black.

Hyland, K. (1998). Exploring corporate rhetoric: Metadiscourse in the CEO's letter. *The Journal of Business Communication* (1973), 35(2), 224-244.

<https://doi.org/10.1177/002194369803500203>

Jacobson, G. (1988). How valuable is the annual report?. *Management Review*, 77(10), 51.

Koller, V., Kopf, S., & Miglbauer, M. (Eds.). (2019). *Discourses of Brexit*. London: Routledge. <https://doi.org/10.4324/9781351041867>

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>

Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, 60(2-3), 110-135. <https://doi.org/10.1016/j.jacceco.2015.09.002>

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3), 221-247.

<https://doi.org/10.1016/j.jacceco.2008.02.003>

Malavasi, D. (2005). Banks' annual reports: an analysis of the linguistic means used to express evaluation. In *Proceedings of The Association for Business Communication 7th European Convention*.

McEnery, T., & Hardie, A. (2011). *Corpus linguistics: Method, theory and practice*.

Cambridge University Press. <https://doi.org/10.1017/CBO9780511981395>

Pollitt, M. G. (2022). The further economic consequences of Brexit: energy. *Oxford Review of Economic Policy*, 38(1), 165-178. <https://doi.org/10.1093/oxrep/grab044>

Plakandaras, V., Gupta, R., & Wohar, M. E. (2017). The depreciation of the pound post-Brexit: Could it have been predicted?. *Finance research letters*, 21, 206-213.

<https://doi.org/10.1016/j.fl.2016.12.003>

Rutherford, B. A. (2005). Genre analysis of corporate annual report narratives: A corpus linguistics-based approach. *The Journal of Business Communication* (1973), 42(4), 349-378. <https://doi.org/10.1177/0021943605279244>

Schroeder, N., & Gibson, C. (1990). Readability of management's discussion and analysis. *Accounting Horizons*, 4(4), 78.

- Thomas, J. (1997). Discourse in the marketplace: The making of meaning in annual reports. *The Journal of Business Communication* (1973), 34(1), 47-66.
<https://doi.org/10.1177/002194369703400103>
- Wang, H., Li, L., & Cao, J. (2012). Lexical features in corporate annual reports: a corpus-based study. *European Journal of Business and Social Sciences*, 1(9), 55-71.
- Zappavigna, M. (2019). Ambient affiliation and# brexit: Negotiating values about experts through censure and ridicule. In *Discourses of Brexit* (pp. 48-68). Routledge.
<https://doi.org/10.4324/9781351041867-4>
- Zappettini, F. (2021). The Brexit referendum: How trade and immigration in the discourses of the official campaigns have legitimised a toxic (inter) national logic. In *"Brexit" as a Social and Political Crisis* (pp. 23-39). Routledge.
<https://doi.org/10.4324/9781003166832-3>