Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327
مجلة اللغة الانكليزية في العالم العربي

Arab World English Journal (AWEJ) Volume 9. Number 4. December 2018 Pp. 19- 38
DOI: https://dx.doi.org/10.24093/awej/vol9no4.2

A Corpus-assisted Critical Discourse Analysis of the Discursive Representation of Immigration in the EU Referendum Debate

Ebtisam Saleh Aluthman
Department of Applied Linguistics, College of Languages
Princess Nourah bint Abdulrahman University
Saudi Arabia

Abstract

This paper presents a critical account of the representation of immigration in the Brexit corpus—a collective corpus of 108,452,923 words compiled mostly from blogs, tweets, and daily news related to Brexit debate. The study follows the methodological synergy approach proposed by Baker et al. (2008), a heuristic methodological approach that combines methods of discourse analysis and corpus-assisted statistical tools including keyword, collocation, and concordance analysis. Drawing on this methodological synergy approach, the investigation yields significant findings contextualized within the socio-economic-political context of the European Union (EU) leave referendum to trace how the issue of immigration is represented in the discourses of the Remain and Leave campaigns. The frequency results show that immigration is one of the most salient topics in the Brexit corpus. Concordance analysis of the word *immigrants* and collocation investigation of the word *immigration* reveal opposing attitudes toward immigration in the EU referendum debate. The analysis uncovers negative attitudes toward the uncontrolled flow of immigrants from other EU countries and public concerns about immigrants' negative impacts on wages, education, and health services. Other findings reveal positive attitudes toward immigrants emphasizing their positive contributions to the UK economy. The study concludes with an argument of the significant association between the political and socio-economic ideologies of a particular society and the language communicated in its media.

Keywords: A corpus-assisted discourse analysis, Brexit, discursive representation, immigration

Cite as: Aluthman, E. S. (2018). A Corpus-assisted Critical Discourse Analysis of the Discursive Representation of Immigration in the EU Referendum Debate. *Arab World English Journal, 9* (4), 19- 38. DOI: https://dx.doi.org/10.24093/awej/vol9no4.2

Preliminaries: A Historical Account of the Brexit

The term *Brexit*—a blended word of *Britain* and *exit*—came into existence to refer to the UK's official withdrawal from the EU. The EU was formed in the 1950s by European countries that desired to build stronger ties after World War II. In 1952, Belgium, France, Italy, Luxenberg, Netherland, and West Germany signed the Treaty of Rome, forming the European Economic Community (EEC). In 1973, the EEC welcomed three new members, the UK, Denmark, and Ireland. A couple of years after joining the EEC, the UK held its first national EEC in/out referendum, and 67.2% of voters chose to stay in the EEC.

On November 1, 1993, the EEC became known as the EU, reflecting the development of the European association from an economic union into a political union integrating 28 members and enacting its own laws and reforms within a flourishing political and economic state. The Schengen visa was introduced in 1995 permitting free movement among EU countries .Even so, the UK did not open its borders entirely like other EU countries. Also, the UK chose to keep the British pound as its official currency rather than approving the Euro (see Hobolt, 2016; Peers, 2016).

In 2009, the EU adopted Article 50 under the Lisbon Treaty, establishing an official mechanism for any EU country leaving the EU. Around this time, many major issues started to incite national calls for the UK to leave the EU. The EU witnessed a great economic collapse in Greece, requiring efforts by all parties to stabilize Greece's economy. Additionally, the EU experienced floods of immigrants whose preferred destinations were the UK, France, and Germany. Although personally rejecting calls to leave the EU in 2012, Prime Minister David Cameron announced a national in/out referendum under pressure from many of the members of Parliament and the Independence Party. On June 23, 2016, 17.4 million UK citizens voted, and 51.9% of them chose to leave the EU. Cameron resigned immediately. The value of the British pound fell 15% lower than the United States dollar and the UK and the EU began negotiating their divorce bill (see Hobolt, 2016; Peers, 2016).

The term *Brexit* was analyzed semantically by Fontaine (2017), implementing a systemic functional linguistic approach with a corpus of 1,641,903 words. Buckledee (2018) explores the ways the Leave campaign overcame the Remain party from a linguistic perspective. He investigates the ways language affects the political process as voters are exploited through persuasive and emotive linguistic strategies using influential metaphors and inspiring tones (Buckledee, 2018). Although Brexit has been a major cause of concern in the UK and EU socio-economic-political scenes, the field of linguistics lacks a discursive linguistic analysis of the massive political discourse that has come into existence since the first inclusion of the term *Brexit* in the Oxford English Dictionary (2012). Among the most provocative issues in the Brexit debate is immigration. A main claim of the Leave campaign is that leaving the EU will allow the UK to better regulate the movement of immigrants from EU countries. The aim of the present study is to investigate the discursive patterns of the representations of immigration in the discourses of the Leave and Remain campaigns utilizing corpus-assisted discourse analysis (CADA) tools. The availability of the two representative sub-corpora, Opinion on Brexit–agreement and Opinion on Brexit—disagreement, with the same amounts of words and tokens (see Appendix A) is important for yielding reliable findings (Baker, 2011). An account of the CADA is given in the following section along with its applications in a variety of contexts.

A Corpus-Assisted Critical Discourse Analysis Approach

The use of corpus linguistics (CL) dates back to the 1950s when compiling an electronic, readable format of a particular language was a novel innovation used by a few linguists with great enthusiasm. A few decades later, CL had been integrated into discourse analysis studies. Sinclair (2004) describes CL and discourse analysis as "the twin pillars of language research. … They both encourage the formulation of radically new hypotheses [and] the dimensions of patterns that they deal with are, on the whole, larger than linguistics is accustomed to" (p. 11). Only a few years ago, the combination of the qualitative methods of discourse analysis and quantitative statistical tools in CL permitted significant developments in the field of discourse analysis. The integration of these two disciplines resulted in the emergence of what was called corpus-assisted discourse studies (CADS) (Partington, 2004, 2006). The integration of CL and Critical Discourse Analysis (CDA), which views discourse as "a form of social practice" (Fairclough & Wodak, 1997, p. 258), within this developing body of research has been come to be known as Corpus-assisted Critical Discourse Analysis (CACDA). In a very significant study that initiated this synergy approach, CACDA was taken up by Baker et al. (2008), who investigate the patterns of discourses on refugees and immigrants in a data set of 140 million words in the UK press. Due to its multidisciplinary nature, CACDA research is best considered to be a multidimensional endeavor that can be understood only within the context of its multidisciplinary areas. The most important area of them is CDA.

CDA has rapidly expanded in the 20th century as a theoretical framework with three main approaches: Van Dijk's (1996, 1997, 1998, 2001) cognitive-discourse framework, Wodak's (2001) historical framework, and Fairclough's (1995) social framework. According to Wodak (2001), CDA is primarily concerned with investigating the associations among language, power, and ideology. Similarly, Van Dijk (1997, 1998) proposes that CDA is best used to uncover discursive representations of inequalities, social prejudices, and discriminatory social practices and how these discursive representations are constituted, maintained, and adapted within their social and cultural contexts. From the perspective of methodology, McKay (2009) suggests that CDA applies tools and methods adapted from a variety of disciplines, such as text linguistics, rhetoric, sociolinguistics, and semantics, to uncover how language, power, and ideology are related to each other and how these dependent relations are represented through texts.

The application of CDA has been proved to be advantageous in investigating discourse in a diversity of contexts, including political issues (e.g., Fairclough, 2001, 2003b; Kress, 1994), discrimination and racism representation (e.g., Van Dijk, 1996, 1998, 2000), and media (e.g., Fairclough,1995b). However, the adequacy of CDA as an analytical approach has often been subjected to critical criticisms (Baker et al., 2008; Koller & Mauntner, 2004; Stubbs, 1997; Widdowson, 2000, 2001, 2004). Among issues that have provoked debate within this body of criticism is the subjective selection of texts to be analyzed. Koller & Mauntner (2004) state that the danger is that the texts selected within the CDA framework attract the researcher's attention but are not representative of the discourse addressed. CDA researchers are accused of "cherry picking" texts that best prove their positions (Widdowson, 2000, 2001, 2004). Another major criticism of CDA is stated by Stubbs (1997), who criticizes CDA researchers for focusing on small-scale based studies and overlooking significant language forms and patterns occurring within millions of running texts.

These criticisms have spurred the integration of CL tools into CADS (Baker & McEnery, 2015; Partington et al., 2004; Partington, 2008). Many advocates of CACDA argue that this mythological synergy is more advantageous because it relies on the strengths of each approach to compensate for the weaknesses of the other and enhances discourse research in a variety of methodological aspects. First, CACDA relies on large-scale corpus with naturally occurring language, decreasing the influence of the researcher's bias (Baker, 2006). One of the key advantages of CADS is that CL can increase the objectivity of CDA (Taylor & Marche, 2018). Second, investigations using CACDA display data in authentic contexts, improving understanding of the discourse investigated and more comprehensively representing the discursive patterns. Third, CACDA relies on triangulation using the qualitative tools of CDA and the quantitative tools of CL. It thus yields more reliable results than traditional CDA methods because the corpus investigated is designed according to specific criteria ensuring the validity of generalizations (McEnery & Wilson, 2001). Accordingly, Marchi & Taylor (2009) explain that the methodological attributes of both CDA and CL are combined within the newly developed framework of CACDA. In Marchi & Taylor's (2009) words, CL is a quantitative approach grounded in a data-driven framework based on large samples with statistical significance, and descriptive analysis with great objectivity, leading to generalizable and reliable results. In contrast, CDA is a qualitative approach grounded in a theory-driven framework relying on individual, selected samples with social significance, and explanatory investigation that is often subjective.

The CAD field has grown rapidly, stimulating a beneficial combination of CL and discourse analysis. Following Baker et al. (2008), a growing body of CACDA research has emerged. Most of this research has targeted political conflicts, immigrants, and the marginalization of some minorities. The main data source in this growing body of research is the press and social media, because the media is the tool political and social authorities use to impose their laws and acts within social communities (Van Dijk, 1996). Kandil (2009) used keyword analysis, collocation, and concordance to investigate media representations of the conflict between Israel and Palestine. Salama (2011) relied on collocation in a CDA analysis of the ideological representations of Wahabi and anti-Wahabi campaigns in a corpus of two books after the 9/11 attacks. Su & Xiao (2015) investigated discursive representation of the Chinese dream among both officials and citizens based on a corpus of tokens compiled from the Chinese press. Similarly, Haider (2016) utilized the tools of keyword extraction, collocation and concordance in his investigation of representations of the Qaddafi regime in a corpus of 27 million words in the press before, during, and after the Libyan uprising periods.

The present study adopts the CACDA to investigate the discursive patterns of representations of immigration in the Brexit debate by both the Remain and Leave campaigns—an area that is not yet investigated. This paper is aimed at improving understanding of the representations of immigration issues in both positive- and negative- sentimentally classified sub-corpora. The present investigation also explores in what ways these representations are related to the socio-economic-political context through a qualitative concordance investigation grounded in CDA.

Research Questions

The present study aimed at addressing the following questions:

RQ (1): What are the most significant topics communicated in the (a) Brexit corpus as a whole, (b) Brexit—agreement corpus, and (c) Brexit—disagreement corpus?

RQ (2): What does concordance analysis of the word "*immigrants*" in the Brexit corpus reveal?

RQ (3): What does the word *immigration* collocate with? How do these collocations relate to the socio-economic-political context?

Methodological Framework

Methodology

The methodological framework applied in this study is based on the premises underlying the use of CL in discourse-related studies. According to Partington (2003), CL is used in CAD in a number of ways. Most simply, CL provides the analysts with instances of the phenomenon under investigation. At the other extreme, CL assists the analysts in reinforcing, refuting, or revising their initial assumptions. Baker et al. (2008) explain that CL helps the discourse analyst to "quantify discoursal phenomena already recognized in CDA" (p. 285). This study's methodology follows the methodological steps used by scholars conducting CAD in this field (Baker et al., 2008; Hardt-Mautner, 2009; Partington, 2003).

The first stage was "setting the scene" of the analysis (Baker et al., 2008, p. 284; Hardt-Mautner, 2009). A preliminary investigation of the corpus to formulate research questions involved extensive readings about the Brexit issue in the British press. Background investigation and formulation of the research questions were performed before the CAD analysis. The second stage was choosing and compiling the appropriate corpus to fulfill the study aim. The two representative sub-corpora, Opinion on Brexit—agreement and Opinion on Brexit—disagreement are important to yield significant results. The third stage was to select appropriate CL tools. The Brexit corpus selected was available at the Sketch Engine and could be analyzed via all the Sketch Engine CL tools. The fourth stage was detecting the emerging lexical patterns through frequency and keyword extractions and generating collocations grouped by semantic categorizations. This stage was crucial in identifying the most common themes and topics communicated in the corpus. The fifth stage was a qualitative investigation of these common themes and topics through concordance analysis. The final stage consisted of revising the research questions and drawing implications. (see Baker et al., 2008, p. 295, for an outline of these stages).

Data

The data investigated in this study came from the Brexit Corpus compiled as part of the EU-funded research project, a joint effort by the University of Trento, Websays.com, and Aix-Marseille University (see http://www.sense-eu.info/). The Brexit corpus consists of 108,452,923 words and 125,637,141 tokens compiled from 285,360 links and 506,808 documents mostly from blogs, tweets, and daily news related to Brexit topics. The corpus was compiled, classified, and annotated to enable searching by a specific sentiment (negative, neutral, or positive), topic (e.g., Leave and Leave AGAINST EU, Remain and Remain for EU, Immigration, Brussels), or opinion (agreement or disagreement). Appendix A displays the sub-corpora statistics for the Brexit corpus.

Results

Keyword Analysis

Keyness is illustrated by Baker et al. (2008) as "the statistically significant higher frequency of particular words or clusters in the corpus under analysis in comparison with another corpus" (p. 278). These words with significantly higher frequency guide the investigation by indicating the "aboutness" of the main corpus and the two sub-corpora (Scott, 1999). Considering the research questions of this study, the words with higher frequency indicate the most common themes represented in the corpus. The keyword analysis not only indicates the focus or aboutness of the corpus but also suggests the focus of further investigations based on with what the most common themes in the corpus collocate. The first 100 keywords were generated from each corpus using the English Web 2013 (enTenTen13) as the reference corpus. These keywords generated from these three corpora (Brexit Corpus, Opinion on Brexit–agreement and Opinion on Brexit—disagreement) are shown in Appendix B,C and D.

The keyword analysis shows two important things. First, the three corpora have great similarity in terms of their keywords, or the most significant themes/topics they discuss. Second, the Brexit—agreement corpus and the Brexit—disagreement corpus have great similarity in the lexical frequency of their keywords. The most frequent word in the three corpuses is *Brexit*, which occurred 352,529 times in the whole corpus 142,879 times in the Brexit—agreement corpus and 131,254 times in Brexit—disagreement corpus. The same was true with the other words, such as *Corbyn* (referring to Jeremy Corbyn, a leader of the Labour Party) *Remain*, *VoteLeave*, *EUref*, *VoteRemain*, *worryingly*, *marginally*, *job-destroying*, *unionists*, and *rightwing*.

Following Baker (2010), the generated keywords were grouped into semantic categories based on their semantic meaning. Preliminary readings of the historical and socio-political backgrounds helped mapping the generated keyword lists in the semantic categorization, as shown in table 1.

Table 1. *Semantic mapping of the Brexit Corpus keywords.*

Semantic Category	Examples
Voting regulations	Brexit, Corbyn, Remain, VoteLeve Sadiq, Davidson, Cameron, VoteRemain, Referendum, denounce
Immigration	Job-destroying, enfranchisement, havens, Mediterranean, transatlantic
Financial issues	Job-destroying, loopholes
Uncertainty	Underpowered, marginally, unwillingly, evaders, skepticism
International relations	EU, discredited, lambasting, opportunism

Table 1 lists the most common topics communicated in the Brexit corpus and the two sub-corpora. It is obvious and inevitable that the most frequent words in the three corpora are related to voting regulations in the national EU in/out Referendum. All the other significant topics

revealed are related to the influence of Brexit on the UK and the rest of EU. Among these significant topics is that of immigration. Relating the frequency analysis results to socio-economic background, the Leave campaign emphasizes that Brexit will enable the UK to decrease the flow of immigration and have more control over the movement of immigrants from other European countries. A crucial argument by the Leave campaign is that the more immigrants move to Britain, the more concerns they raise about job competitiveness, salaries, and overall quality of life. The counterargument is that immigrants increase consumption of services, leading to more job opportunities. In addition, a number of the immigrants are young and well educated and complement UK citizens' skills and professions.

The concordance analysis of the Opinion on Brexit—agreement sub-corpus and the Opinion on Brexit—disagreement sub-corpus yields significant results indicating opposing views on immigration. This analysis involves searching for the word *immigrants* in both sub-corpora and identifying the sentiment classification (positive or negative) of its usage. The word *immigrants* occurs 8,358 times (66.50 times per 1 million) in the Opinion on Brexit—agreement corpus. Its concordance analysis is illustrated in Figure 1.


Query immigrants, agreement 8,358 (66.52 per million)

Page 1 of 418 Go Next | Last

discussion...	risked descending into a plebiscite on whether immigrants are a good or a bad thing. To see what is at stake,
www.telegr...	Boris Johnson called for an amnesty on illegal immigrants in an attempt to "neutralise" extremists who
www.telegr...	" as he called for an amnesty on illegal immigrants who have been in this country for more than 12
www.telegr...	that it is "fair" and "humane" to grant illegal immigrants an amnesty. He said: "I am in favour of an amnesty
www.telegr...	. He said: "I am in favour of an amnesty of illegal immigrants who have been here for more than 12 years, unable
www.telegr...	with immigration and who are opposed to immigrants ." Mr Farage claimed that the establishment has
discussion...	risked descending into a plebiscite on whether immigrants are a good or a bad thing. To see what is at stake,
thebournem...	is folly. Firstly, it is well documented that immigrants pay £2.54bn more in income tax and national
thebournem...	they receive in benefits. Secondly, most of the immigrants we receive from the EU are from Western Europe,
thebournem...	world of trade, we will need to keep accepting immigrants to this country. Switzerland struggled and
thebournem...	' by singing Rule Britannia while kicking out immigrants , it was instead our morally questionable
discussion...	were vilified in the last century; when Jewish immigrants were targeted a century ago; and when Asian and
discussion...	risked descending into a plebiscite on whether immigrants are a good or a bad thing. To see what is at stake,
discussion...	were vilified in the last century; when Jewish immigrants were targeted a century ago; and when Asian and
discussion...	risked descending into a plebiscite on whether immigrants are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether immigrants are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether immigrants are a good or a bad thing. To see what is at stake,
www.telesu...	the Leave campaign it is this: "Blame immigrants for everything"-crime, lack of housing,
www.telesu...	the last decades. But this is nothing to do with immigrants and everything to do with the policies imposed
discussion...	risked descending into a plebiscite on whether immigrants are a good or a bad thing. To see what is at stake,

Figure 1. Concordance analysis of the word *immigrants* in the Opinion on Brexit—agreement corpus.

The concordance analysis clearly reveals that the Leave campaign has concerns about the impacts on quality of life that are resulted from the increasing number of immigrants. Examples are:

1- The EU are taking all our jobs, telling us what we can and can't do, making bananas straight, sending over immigrants, taking all our money, and worst of all, sending over immigrants. I'm sick of the health and safety brigade. (http://www.thelondoneconomic.com/satire/study-reveals-mainly-stupid-people-will-vote-brexit/20/06)
2- and white again, and children will play with spinning tops and carts made from fruit boxes and pram wheels instead of Nintardos. There won't be traffic jams anymore, and we'll be safer from those dangerous immigrants. Also, everyone will have a job again, like they used to in 1950. We need to put the Great back into England'(http://www.thelondoneconomic.com/satire/study-reveals-mainly-stupid-people-will-vote-brexit/20/06)

The word *immigrants* occur 7,274 (57.90 times per 1 million) in the Opinion on Brexit—disagreement corpus. Its concordance analysis is illustrated in Figure 2.

Page 1 of 364 Go Next | Last

discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	and genocide, seem to think only of money and **immigrants** , rather than the most important thing of all:
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,
discussion...	risked descending into a plebiscite on whether **immigrants** are a good or a bad thing. To see what is at stake,

Figure 2. Concordance analysis of the word *immigrants* in the Opinion on Brexit—disagreement corpus.

The concordance analysis clearly reveals that the Remain campaign is concerned about skilled immigrants' contributions to economic development that complement the UK nationals. The Remain campaign also emphasizes the cultural impact the diversity of communities can have. Examples are:

3- Immigration is 100% an economic positive for the UK. EU **immigrants** make up 5% of the country. Loads of Britins have left for Europe as well so the net population change is only a couple million people. They are in work and contributing tax as well as providing vital skills the country needs to continue growing and providing public services, particularly the NHS in which immigrants play a huge role. (http://www.eastdulwichforum.co.uk/forum/read.php?20,1658869,1690115,page=20#msg-1690104)

4- Most EU immigrants are in London, where most people are pro-Remain because they are best placed to see the advantages (economic and cultural) of EU migration. The supposedly lower skilled **immigrants** from Eastern Europe make up less than > half of total EU **immigrants** and are too small a population to have had any widespread impact on most working-class people. (http://www.eastdulwichforum.co.uk/forum/read.php?20,1658869,1690115,page=20#msg-1690104).

Others from the Remain campaign are concerned with the difficulties and suffering poor immigrants might face. The word *Mediterranean* is found to be among the most-frequent words in the corpus. The following extract using the word *Mediterranean* is a call to draw public attention to the hundreds of immigrants who have died trying to cross the Mediterranean.

5- who worked as a translator in his country and then came to the UK illegally, adds: "There's something strange that is going on at the moment. Hundreds of people have died trying to cross the **Mediterranean**. Has it had headline coverage day after day? No, it has had minimal coverage." (https://discussion.theguardian.com/comment-permalink/76731740

Collocation Analysis

Collocation analysis reveals the salient themes associated with the topic investigated (Baker, 2006; Baker et al., 2013). Collocation has always been associated with discourse prosody because it detects attitudes based on the association between words' meanings (Baker, 2006). In collocational analysis, an item is classified as positive, negative, or neutral prosody when it is frequently co-occurs with positive, negative, or neutral collocates.

A collocation analysis of the word *immigration* was conducted within the span (-5 to +5) of the whole Brexit corpus. The aim was to reveal the discursive patterns associated with the topics of immigration and Brexit. The analysis yields the top 24 collocates of these categories: modifiers of *immigration*, nouns and verbs modified by *immigration*, and verbs with *immigration* as their object. Figure 3 illustrates this collocation analysis. Each collocate generated has an indication of its lexical frequency score in the corpus calculated by the logDice formula. It is also accompanied

by an example of how the word occurs with its collocates. The concordance analysis of each collocate can also be generated in a compact form.

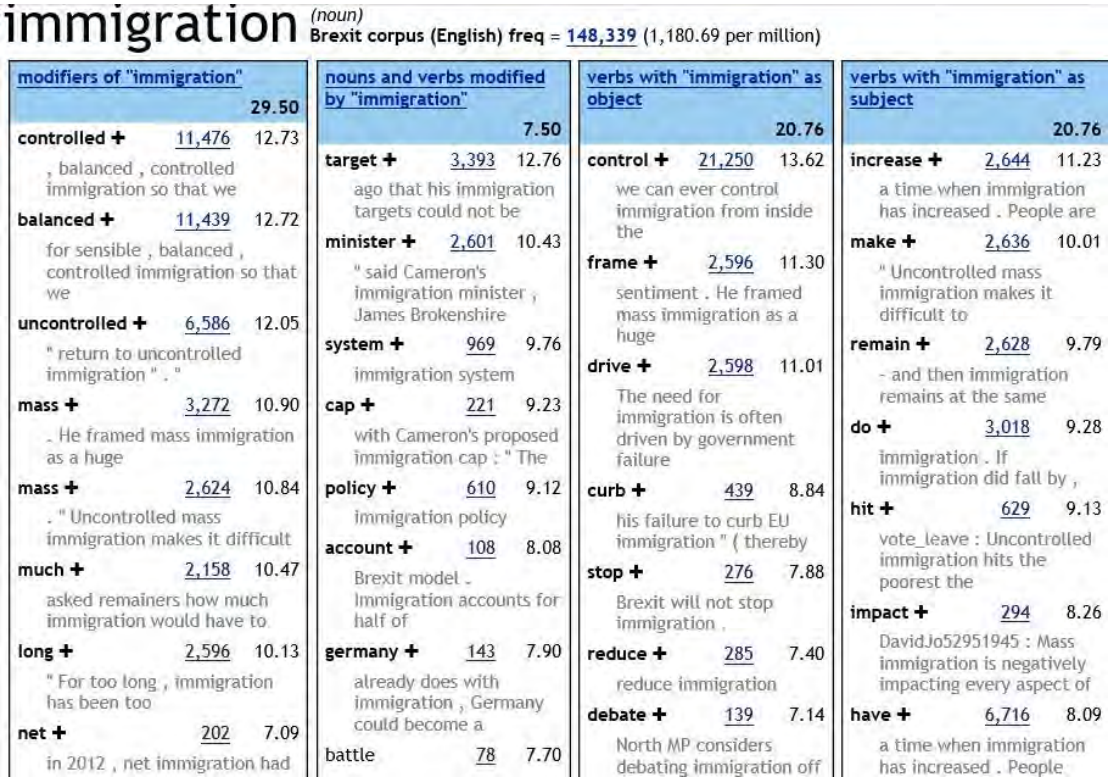

Figure 3. Screenshot of the collocation analysis of the word *immigration* in the Brexit corpus.

The word *immigration* occurs 148,339 times (118,069 times per 1 million) in the Brexit corpus. The top 24 collocates are displayed in Table2.

Table 2. *Top 24 collocate modifiers of the word immigration in the Brexit corpus.*

Rank	Modifier	Freq.	logDice	Rank	Modifier	Freq.	logDice
1	controlled	11,476	12.73	13	Muslim	92	6.05
2	balanced	11,439	12.72	14	control	73	5.75
3	uncontrolled	6,586	12.05	15	more	184	5.47
4	mass	3,272	10.90	16	sovereignty	62	5.34
5	much	2,158	10.47	17	Grady	49	5.19
6	long	2,596	10.13	18	permanent	45	5.04
7	net	202	7.09	19	Non-EU	40	4.87
8	controlling	182	7.08	20	shore	34	4.66
9	EU	707	6.66	21	security	34	4.56
10	illegal	117	6.40	22	low	61	4.49
11	unlimited	112	6.27	23	scale	30	4.48
12	high	171	6.11	24	controlling	28	4.38

Table 2 shows that the modifiers *controlled*, *balanced*, *uncontrolled*, *illegal*, and *unlimited* strongly collocate with *immigration*. The concordance analysis of these collocates revealed negative attitudes toward the immigration flow and its perception as uncontrollable affecting education and health services. Examples are below.

> 5- Some have tried to demonise me or others to say we've upped the rhetoric. Compared to the Scottish referendum we have done no such thing. All we've done is ask for sensible, **balanced**, controlled immigration so that we can have the right number of people to come to our country and benefit our society and we know we can't do that in the European Union.
>
> (https://discussion.theguardian.com/comment-permalink/76731740)
>
> 6-Extra people automatically lead to extra costs for the NHS, schools, housing and infrastructure (simply more cars on the roads in towns and cities) The problem this **uncontrolled** immigration creates is because we cannot really control our borders to EU immigrants, we cannot plan long term for our infrastructure, housing, and schooling. In addition, how much more crowded do we want to be.
> (http://www.bluebond.co.uk/2016/06/5332)

Table 3. Top 24 collocated nouns and verbs modified by immigration in the Brexit corpus.

Rank	Modifier	Freq.	logDice	Rank	Modifier	Freq.	logDice
1	target	3,393	12.76	13	question	137	6.63
2	minister	2,601	10.43	14	statistic	33	6.57
3	system	969	9.76	15	obsession	35	6.37
4	cap	221	9.23	16	number	72	6.35
5	policy	610	9.12	17	work	42	6.32
6	account	108	8.08	18	front	28	6.30
7	Germany	143	7.90	19	authority	35	6.28
8	battle	78	7.70	20	concern	37	6.24
9	control	207	7.26	21	level	91	6.21
10	rule	85	7.12	22	tonight	84	6.19
11	poster	99	7.01	23	cost	29	6.09
12	sovereignty	48	6.81	24	law	183	6.05

Table 3 also shows that the word *control* frequently co-occurs with the word *immigration* as a modified noun. Additionally, the verb *accounts* is found to collocate strongly with *immigration* referring to the positive impacts of young, educated, skilled immigrants on the UK. An example is given below.

6- Stricter immigration laws could be put in place, depending on the Brexit model. Immigration **accounts** for half of the UK's growth since 2005 and it has around 2.2 million EU skilled workers. (http://economictimes.indiatimes.com/news/international/world-news/brexit-referendum-is-warning-signal-for-eu-donald-tusk/articleshow/52835328.cms)

Table 4. *Top 24 collocated verbs with the word immigration as an object in the Brexit corpus.*

Rank	Modifier	Freq.	logDice	Rank	Modifier	Freq.	logDice
1	control	21,250	13.62	13	swap	68	6.16
2	frame	2,596	11.30	14	concern	73	6.10
3	drive	2,598	11.01	15	enable	62	6.00
4	curb	439	8.84	16	mention	56	5.81
5	stop	276	7.88	17	manage	50	5.67
6	reduce	285	7.40	18	regard	49	5.62
7	debate	139	7.14	19	increase	53	5.59
8	bring	154	7.09	20	call	102	5.52
9	cut	161	7.00	21	want	76	5.51
10	limit	121	6.86	22	need	73	5.48
11	fuel	97	6.34	23	accept	36	4.92
12	find	261	6.31	24	discuss	30	4.92

The verb *control* is also found to collocate frequently with *immigration* as an object. Other verbs that collocate strongly with *immigration* and express negative attitudes toward uncontrolled immigration to UK are *stop*, *reduce*, *cut*, *limit*, and *manage* in both sentiment and positive instances, that is, in the debate of both the Leave and Remain campaigns. Examples in contexts are given below.

9- UK Independence Party] argued that a Brexit would enable the UK to significantly **reduce** immigration, preventing both EU citizens from taking British jobs and non-EU citizens from sneaking in to commit terror attacks. (https://www.theguardian.com/tv-and-radio/2016/jun/20/john-oliver-brexit-britain-crazy-to-leave-european-union)

10-We could cut immigration into the UK by at least 100k people a year. without leaving the EU. By **cutting** non-EU immigration. Simples…if you don't speak English to a required standard, you cannot come in. (http://www.car4play.com/forum/post/index.htm?f=5&t=22676#503494)

Table 5. *Top 24 collocate verbs with the word immigration as a subject in the Brexit corpus.*

Rank	Modifier	Freq.	LogDice	Rank	Modifier	Freq.	LogDice
1	increase	2,644	11.23	13	appear	85	6.19
2	make	2,636	10.01	14	affect	52	5.71
3	remain	2,628	9.79	15	cause	50	5.64
4	do	3,018	9.28	16	enrich	46	5.61
5	hit	629	9.13	17	put	54	5.44
6	impact	294	8.26	18	level	37	5.23
7	have	6, 716	8.09	19	overwhelm	32	5.08
8	be	10, 460	7.66	20	pose	32	5.04
9	become	313	7.55	21	enhance	28	4.89
10	prove	129	6.87	22	strike	27	4.82
11	fall	116	6.53	23	happen	29	4.77
12	go	131	6.32	24	define	25	4.63

Some of the top collocate verbs with the word *immigration* as an object—*increase*, *affect*, and *overwhelm*—reveal the same attitude toward immigration. However, other collocates, such as *enrich*, emphasize the positive impacts of immigration and immigrants' contributions to the UK's economic development. Below is an extract as an example.

> 11-Privately, the remain campaign believes that the death of Cox, a strong supporter of the EU who believed immigration had **enriched** British society, will make it harder for their opponents to campaign as stridently as before on immigration issues. (https://www.theguardian.com/uk-news/2016/jun/20/leaveeu-donor-arron-banks-defends-polling-jo-cox-killing)

Conclusion

This study makes a response to a noticeable lack of uncovering the discursive patterns associated with the issue of immigration in the Brexit context. Following Baker et al. (2008), the CL and CDA approaches have complemented each other throughout the study by triangulating them in all aspects of the investigation. Within this framework, CL tools have been utilized to establish a pattern map of the corpus studied based on statistically significant results for keywords, collocations, and concordances. Meanwhile, CDA not only directed the analyst's attention to significant patterns examined with CL tools but also related the investigation to the socio-economic-political contexts. The investigation has been carried out carried out in accordance with the main assumption underlying CDA that discourse is a social practice (Fairclough, 1995a). The quantitative investigation follows the three phases proposed by Fairclough (1995a): describing, interpreting, and explaining. The analysis confirmed the observations made by Baker et al. (2008) about the fuzzy boundaries between the CL quantitative approach and the CDA qualitative

approach. Overall, the triangulated analysis utilizing CL and CDA tools has uncovered opposing views in EU in/out referendum corpus on the issue of immigration.

The keyword analysis reveals that immigration is one of the most-debated topics in the Brexit corpus. Words such as *job-destroying*, *enfranchisement*, *havens*, *Mediterranean*, and *transatlantic* are found among the top keywords. However, qualitative investigation is needed to uncover the discursive patterns associated with these words and the ideologies underlying their occurrences. Close observation of the concordance lines in which these words occur indicates that immigration is related in many ways to the other significant topics communicated by the main corpus and the sub-corpora, particularly finance, uncertainty, and international relations. Immigration, in some instances, is found to be related to the public fear that the UK is losing control over floods of immigrants, who negatively affect wages, education, and health services. In other instances, remarkable accounts point out that immigrants contribute positively to a growing and more productive economy. Immigrants not only take jobs but also create new jobs by establishing their own business and spending their money within the UK community. Empirical results show that EU immigrants, on one hand, take jobs and affect wages and education and health services but, on the other hand, give UK businesses access to professional, skilled, young employees who offer high-value added to businesses and public finance. A frequent remark made, even in the Opinion on Brexit—agreement corpus, is that the EU leave decision should be defined separately from the immigration issue.

The present study confirmed the relationship between language and media described by pioneering CDA scholars, such as Herman and Chomsky (1988) and Fairclough (1995). Fairclough (1995) emphasizes the importance of the relationship between the political and socio-economic backgrounds of a society and the language communicated in its media. This investigation of the Brexit corpus supports this argument that the media in any society reflects the ideologies of the political and social elites. This study also contributes to the work of research on language, media, and society demonstrating how the linguistic patterns in the media of a society are affected by their particular socio-economic-political contexts. CACDA facilitates both access to and investigation of the language of media. The Brexit corpus provides a large and authentic corpus of media language that, when analyzed with CL tools, quantifies the ideologies and views indicated in diverse media channels. Future studies might benefit from CACDA to undercover the relationships among language, media, and ideologies in under-researched political and social topics such as the Brexit and financial issues, uncertainty and UK international relations with EU.

About the Author:

Dr. Ebtisam Aluthman is an Assistant Professor of Applied Linguistics at the College of Languages at Princess Nourah bint AbdulRuhman University in Riyadh, Saudi Arabia. She holds a PhD in Linguistics from the University of Manchester in the United Kingdom. Her research interests focus on corpus linguistics, discourse analysis, language and media and TESOL. https://orcid.org/0000-0001-9643-9218.

References

Baker, P. (2006). *Using corpora in discourse analysis*. London, UK: Continuum.

Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh, UK: Edinburgh University Press.

Baker, P. (2011). Social involvement in corpus studies. In V. Viana, S. Zyngier, & G. Barnbrook (Eds.), *Perspectives on corpus linguistics* (pp. 17–28). Amsterdam, Netherlands: John Benjamins.

Baker, P., & McEnery, T. (2005). A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of Language and Politics, 4*(2), 197–97. doi:10.1075/jlp.4.2.04bak

Baker, P., Gabrielatos, C., & Khosravinik, M. (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.

Buckledee, S. (2018). *The language of Brexit: How Britain talked its way out of the European Union*. London, UK: Bloomsbury Academic.

Fairclough, N. & Wodak, R. (1997). Critical discourse analysis. In T. van Dijk (Ed.) *Discourse as social interaction* (pp. 258–284). London, UK: Sage.

Fairclough, N. (1995a). *Critical discourse analysis: The critical study of language*. London, UK: Longman.

Fairclough, N. (1995b). *Media discourse*. London, UK: E. Arnold.

Fairclough, N. (2001). *Language and power* (vol. 2). Harlow, UK: Longman.

Fairclough, N. (2003a). *Analyzing discourse: Textual analysis for social research*. London, UK: Routledge.

Fairclough, N. (2003b). 'Political correctness': The politics of culture and language. *Discourse & Society*, *14*(1), 17–28. doi:10.1177/0957926503014001927

Fontaine, L. (2017). The early semantics of the neologism BREXIT: A lexicogrammatical approach. *Functional Linguistics, 4*(6), 2–15.

Haider, A. (2016*). A corpus-assisted critical discourse analysis of the Arab uprisings: Evidence from the Libyan* case (Unpublished doctoral dissertation). University of Canterbury.UK.Hardt-Mautner, G. (2009). Checks and balances: How corpus linguistics can contribute to CDA. In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis* (pp. 122–143). London, UK: Sage.

Hobolt, S. (2016). The Brexit vote: a divided nation, a divided continent. *Journal of European Public Policy*. 23 (9): 1259–1277. doi:10.1080/13501763.2016.1225785. ISSN 1350-1763.

Herman, E. S., & Chomsky, N. (1988). Manufacturing consent: A propaganda model. In E. S. Herman & N. Chomsky (Eds.), *Manufacturing consent.* New York, NY: Pantheon.

Kandil, M. A. (2009). *The Israeli-Palestinian conflict in American, Arab, and British media: Corpus-based critical discourse analysis* (Unpublished doctoral dissertation). Georgia State University. USA.

Koller, V., & Mautner, G. (2004). Computer applications in critical discourse analysis. In C. Coffin, A. Hewings, & K. O'Halloran (Eds.), *Applying English grammar: Corpus and functional approaches* (pp. 216–228). London, UK: Arnold.

Kress, G. R. (1994). Text and grammar as explanation. In U. A. R. K. Meinhof (Ed.), *Text, discourse and context: Representations of poverty in Britain* (pp. 24–46). London, UK: Longman.

Marchi, A., & Taylor, C. (2009). If on a winter's night two researchers : A challenge to assumptions of soundness of interpretation. *CADAAD Journal [Critical Approaches to Discourse Analysis across Disciplines], 3*(1), 1–20.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics*. Edinburgh, UK: Edinburgh University Press.

McKay, S. B. (2009). Media and language: Overview. In J. L. Mey (Ed.), *Concise encyclopedia of pragmatics*. Oxford, UK: Elsevier Science.

Partington, A. (2004). Corpora and discourse, a most congruous beast. In A. Partington, J. Morley, & L. Haarman (Eds.), *Corpora and discourse* (pp. 11–20). Bern, Switzerland: Peter Lang.

Partington, A. (2006). Metaphors, motifs and similes across discourse types: Corpus- assisted discourse studies (CADS) at work. In A. Stefanowitsch & S. Gries (Eds.), *Corpus-based approaches to metaphor and metonymy* (pp. 267–304). Berlin, Germany: Mouton de Gruyter.

Partington, A. (2008). The armchair and the machine: Corpus-assisted discourse research. In C. Taylor Torsello, K. Ackerley, & E. Castello (Eds.), *Corpora for university language teachers* (pp. 95–118). Bern, Switzerland: Peter Lang.

Peers, S. (2016). The Brexit: The Legal Framework for Withdrawal from the EU or Renegotiation of EU Membership. Oxford: Hart Publishing.

Salama, A. H. Y. (2011). Ideological collocation and the recontexualization of Wahhabi-Saudi Islam post-9/11: A synergy of corpus linguistics and critical discourse analysis. *Discourse & Society, 22*(3), 315–342. doi:10.1177/095792651039544.

Scott, M. (1999). *Wordsmith tools help manual, version 3.0*. Oxford, UK: Oxford University Press.

Sinclair, J. (2004). *Trust the text*. London, UK: Routledge.

Stubbs, M. (1997). Whorf's children: Critical comments on critical discourse analysis (CDA). In A. Ryan & A. Wray (Eds.), *Evolving models of language: British studies in applied linguistics* (vol. 12, pp. 100–116). Clevedon, UK: Multilingual Matters.

Su, H., & Xiao, H. (2015). The discursive representation of Chinese dream: A corpus-based discourse analysis. *International Journal of Linguistics and Communication, 3*(1), 38–50.

Taylor, C., & Marche, A. (Eds). (2018). *Corpus approaches to discourse: A critical review*. London, UK: Routledge.

Van Dijk, T. A. (1996). Power and the news media. In D. Paletz & C. Vinson (Eds.), *Political communication and action* (pp. 9–36). Cresskill, NJ: Hampton Press.

Van Dijk, T. A. (1997). Discourse as interaction in society. In T. A. Van Dijk (Ed.), *Discourse as social interaction* (vol. 2, pp. 1–37). Newbury Park, CA: Sage.

Van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. London, UK: Sage.

Van Dijk, T. A. (2000). *Ideology and discourse: A multidisciplinary introduction.* Barcelona, Spain: Pompeu Fabra University.

Van Dijk, T. A. (2001). Critical discourse analysis. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 352–371). Oxford, UK: Blackwell.

Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied Linguistics, 21*(1), 3–25. doi:10.1093/applin/21.1.3.

Widdowson, H. G. (2001). Scoring points by critical analysis: A reaction to Beaugrande. *Applied Linguistics, 22*(2), 266–272.

Widdowson, H. G. (2004). *Text, context, pretext: Critical issues in discourse analysis*. Oxford, UK: Blackwell.

Wodak, R. (2001). What CDA is about—a summary of its history, important concepts and its developments. In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse analysis*. London, UK: Sage.

Appendices:

Appendix A: Sub-corpora Statistics of the Brexit Corpus.

Sub-corpora statistics			
Sub-corpus	Tokens	Words	%
Opinion on Brexit—agreement	54,691,630	~ 47,211,096	43.53
Opinion on Brexit—disagreement	47,377,577	~ 40,897,434	37.70
Sentiment—negative	78,968,685	~ 68,167,618	62.85
Sentiment—neutral	5,843,782	~ 5,044,489	4.65
Sentiment—positive	40,822,958	~ 35,239,333	32.49
Topic—Boris Johnson	1,083,477	~ 935,282	0.86
Topic—Brits abroad	23,975	~ 20,695	0.01
Topic—Brussels	650,633	~ 561,641	0.51
Topic—Immigration	3,056,045	~ 2,638,049	2.43
Topic—Leave and Leave AGAINST EU	7,763,358	~ 6,701,512	6.17
Topic—Remain and Remain FOR EU	27,257,911	~ 23,529,667	21.69
Topic—Scotland	449,636	~ 388,136	0.35
Topic—Terrorism	155,457	~ 134,194	0

Appendix B: Keywords in the Brexit Corpus.

Word	Frequency	Score	Word	Frequency	Score
Brexit	352,529	2798.3	campaigners	27,813	110.9
Corbyn	234,404	1771.9	rightwing	18,006	110.3
YouGov	106,712	708.0	Thursday's	35,508	110.0
Remain	178,959	700.2	unionists	20,587	109.2
VoteLeave	73,700	587.6	Gaulle	19,021	107.6
RT	266,924	579.9	pro-Remain	13,325	107.1
EUref	62,361	497.3	Britain's	79,337	105.8
Corbyn's	57,814	458.5	BBC's	27,502	105.4
BBCDebate	46,805	373.5	EURef	12,861	103.3
TTIP	47,626	362.6	pro-Brexit	12,820	103.0
Monnet	44,442	331.7	evaders	14,084	101.9
Sadiq	46,055	306.8	denounced	32,682	99.6
Vote	131,558	290.6	Tusk	14,292	97.2
Gove	47,319	285.1	Farage's	12,621	96.6
referendum	178,607	276.4	Stronger	21,699	93.5
Davidson	126,754	248.2	pro-Leave	11,559	93.0
Boris	94,163	237.2	job-destroying	11,571	92.1
Referendum	38,526	230.3	Britain	276,822	92.1
VoteRemain	27,419	219.2	GMB's	11,439	91.6
neutrals	40,858	217.1	Pidd	11,441	91.5
post-Brexit	26,454	211.5	Chipeta	11,439	91.4
Farage	34,318	206.7	Faragism	11,337	91.2
EU	641,453	205.8	Arkwright's	11,456	91.2
Ukip	25,792	187.4	seventh-century	11,451	89.3
Leadsom	22,419	178.3	instancing	11,439	89.2
StrongerIn	21,315	170.6	Roache	11,446	89.2
Gisela	23,210	167.7	avoiders	11,461	89.0

ISSN: 2229-9327

tikka	22,902	165.4	mayoral	20,533	87.4
GMB	22,930	159.9	Arkwright	11,596	86.7
vote_leave	19,565	156.7	Nigel	41,243	86.4
Worryingly	20,447	151.0	Grady	22,281	86.2
zingers	20,429	150.9	Mair	12,100	86.1
havens	34,630	148.7	burqas	11,491	86.0
Brexiters	18,507	148.3	Brussels	60,097	86.0
pre-prepared	20,433	143.0	demonise	11,484	85.7
robustly	23,098	140.3	Hoey	11,600	85.1
Leave	247,043	138.5	Attlee	11,238	84.8
Commentators	20,493	138.0	enfranchisement	11,446	84.4
Guardian's	23,010	137.8	Cox's	13,159	83.4
brexit	16,691	133.9	lambasting	11,440	82.7
retweets	22,988	133.1	Cameron	101,750	82.0
masala	22,905	131.4	discredited	21,112	81.4
underpowered	20,433	128.6	canvasses	11,439	80.4
opportunism	20,490	124.6	soundbites	11,501	79.8
tonight's	41,838	124.3	flamboyant	20,438	79.7
marginally	42,404	122.5	Tory	38,451	79.0
Labour	158,507	119.9	unification	21,487	78.9
Khan	129,671	117.9	riposte	11,467	78.9
bbcdebate	14,667	117.7	Britons	19,538	78.3
Sturgeon	21,581	117.1	polling	43,498	78.1

Appendix C: Keyword in Brexit-Agreement Corpus.

Word	Frequency	Score	Word	Frequency	Score
Brexit	142,879	2605.4	GMB's	6,990	109.8
Corbyn	122,895	2133.9	Pidd	5,457	109.7
Remain	80,879	726.9	Chipeta	5,448	109.6
Corbyn's	30,345	552.7	Arkwright's	5,415	109.3
YouGov	35,550	542.0	robustly	5,417	107.9
VoteLeave	26,875	492.4	seventh-century	5,415	107.0
RT	96,365	481.0	instancing	5,427	106.9
TTIP	24,808	433.6	Roache	5,422	106.9
EUref	22,419	410.9	rightwing	7,029	106.7
Monnet	22,936	393.1	avoiders	14,758	106.5
BBCDebate	19,811	363.2	pre-prepared	14,323	105.2
Gove	21,010	290.8	Mair	5,427	104.5
referendum	76,401	271.6	Arkwright	5,422	103.5
Vote	49,468	251.0	burqas	5,415	103.0
Sadiq	15,309	234.5	demonise	5,428	102.7
Referendum	16,304	223.9	Hoey	5,459	101.6
Farage	16,115	222.9	Commentators	5,430	101.4
Boris	37,238	215.5	enfranchisement	5,432	101.2
EU	282,448	208.2	Cox's	5,478	100.7
Ukip	12,409	207.1	lambasting	30,074	99.2
tikka	11,960	198.3	Nigel	7,360	96.5

GMB	11,987	191.9	canvasses	5,415	96.3
Davidson	41,160	185.2	Britain's	43,538	96.2
havens	18,123	178.6	soundbites	5,525	95.4
Guardian's	12,037	165.4	underpowered	5,415	94.6
post-Brexit	8,956	164.7	riposte	12,608	94.6
VoteRemain	8,904	163.8	minute's	5,415	93.6
neutrals	13,054	159.5	unification	18,100	93.3
retweets	12,003	159.5	EURef	5,447	92.4
masala	11,960	157.5	tonight's	7,029	91.7
StrongerIn	8,133	149.7	opportunism	10,169	91.6
Brexiters	8,028	147.8	marginally	5,415	91.6
vote_leave	7,605	140.1	scepticism	5,858	91.4
Leadsom	7,494	137.2	Britain	5,415	90.9
Leave	105,969	136.4	denounced	6,865	90.3
Gisela	7,838	130.3	Grandad	7,044	90.0
Thursday's	17,894	127.3	Transatlantic	9,048	88.5
Labour	70,923	123.2	Khan	9,375	88.2
pro-Remain	6,588	121.5	campaigners	5,431	87.5
pro-Brexit	6,586	121.4	Tory	106,467	87.4
evaders	7,170	119.0	Arabella	5,543	87.4
brexit	6,410	118.2	propping	5,447	87.3
Gaulle	8,862	115.2	Britons	5,415	86.8
Farage's	6,540	114.8	Sturgeon	4,297	86.8
Tusk	7,288	113.7	mediterranean	10,845	86.6
bbcdebate	6,160	113.6	loopholes	5,498	85.5
pro-Leave	6,044	111.5	atlantic	4,126	83.9
Worryingly	6,539	111.2	Gateshead	16,166	83.7
zingers	6,527	111.0	Rowena	5,415	83.6
job-destroying	6,041	110.3	Deutschmark	5,415	83.1

Appendix D: Keyword in Brexit-Disagreement Corpus

Word	Frequency	Score	Word	Frequency	Score
Brexit	131,254	2762.9	underpowered	5,980	116.7
Corbyn	109,603	2196.9	brexit	5,980	116.2
Remain	68,729	713.1	pro-Leave	5,980	116.0
YouGov	36,089	635.0	GMB's	5,985	114.7
Corbyn's	27,168	571.1	Pidd	7,717	114.7
RT	93,872	540.8	Chipeta	5,980	114.5
VoteLeave	24,656	521.4	Arkwright's	5,980	114.3
EUref	21,189	448.2	job-destroying	5,980	114.2
TTIP	22,147	446.9	opportunism	7,583	113.5
Monnet	21,220	419.7	marginally	5,980	113.1
Vote	50,037	293.1	tonight's	6,530	112.9
referendum	69,466	285.0	seventh-century	6,407	112.0
Sadiq	15,141	267.6	Roache	6,039	111.8
Gove	15,729	251.4	instancing	6,002	111.7
BBCDebate	11,402	241.7	avoiders	5,999	111.5

Boris	33,865	226.2	Arkwright	6,042	108.0
Ukip	11,716	225.6	burqas	6,544	107.5
Davidson	42,631	221.4	demonise	5,987	107.3
Farage	13,850	221.2	Hoey	6,935	106.3
VoteRemain	10,230	216.9	Britain's	5,981	106.3
tikka	10,854	207.7	Sturgeon	20,060	106.0
EU	240,540	204.7	enfranchisement	5,980	105.6
GMB	10,848	200.4	Khan	31,399	105.0
neutrals	13,980	197.0	Mair	5,995	104.0
post-Brexit	9,018	191.3	lambasting	6,531	103.6
havens	16,315	185.6	denounced	5,999	101.9
Referendum	11,441	181.5	canvasses	6,008	100.7
Guardian's	10,848	172.1	Nigel	11,071	100.5
retweets	10,885	166.9	soundbites	4,997	100.0
masala	10,854	164.9	unionists	13,412	98.9
Brexiters	7,502	159.3	unification	6,541	98.9
Leadsom	7,279	153.7	riposte	13,783	98.5
Gisela	7,455	143.0	Cox's	7,745	98.3
StrongerIn	6,722	142.9	minute's	118,978	97.3
Leave	93,457	138.9	propping	12,890	96.4
Worryingly	6,994	137.0	scepticism	6,001	96.0
zingers	6,990	137.0	campaigners	6,193	95.7
robustly	8,448	136.1	BBC's	42,223	95.3
vote_leave	6,198	131.8	Grandad	9,536	94.0
evaders	6,856	131.3	Britain	18,527	93.9
Thursday's	15,843	130.1	Transatlantic	6,028	91.4
pre-prepared	6,990	129.8	Arabella	7,174	91.1
Labour	64,403	129.1	mediterranean	9,438	90.5
pro-Remain	6,070	129.1	Deutschmark	6,950	89.3
rightwing	7,731	125.4	loopholes	5,980	89.2
Commentators	7,004	125.1	Gateshead	12,006	88.3
Gaulle	8,212	123.1	EURef	5,980	88.1
pro-Brexit	5,739	122.1	Tory	6,014	88.0
Farage's	5,887	119.2	atlantic	5,980	87.6
Tusk	6,588	118.6	Rowena	4,612	87.4