

A corpus-assisted approach to discursive news values analysis

Arash Javadinejad

Universidad Católica de Valencia – Universitat de València / Spain

Abstract – The main aim of this paper is the elaboration of an analytical tool for comparative studies. For this purpose, I used a combination of Discursive News Values Analysis (DNVA) and Corpus Linguistics (CL) to analyse a corpus of British Broadsheets' news coverage of the Brexit campaign. The four major British broadsheets which were analysed were *The Guardian*, *The Independent*, *The Times*, and *The Daily Telegraph*. A specific procedure was designed following previous studies on the topic and considering the challenges and opportunities that such a mixed-method approach (DNVA and CL) can face in exploring journalistic discursive practices and mapping the cultural and ideological discourses around certain topics. Some initial results of the case study are presented to show how the suggested procedure works in practice. From the present study's findings, the procedure seems to work in a reliable manner, although some challenges should be considered and addressed in future studies.

Keywords – news values; DNVA; corpus linguistics; Brexit; discourse analysis; CADS

1. INTRODUCTION¹

Discursive News Values Analysis (DNVA) is a framework recently designed and proposed to analyse how news values are used discursively to construct newsworthiness (Bednarek and Caple 2017). This model has already been widely adopted in journalism and discourse analysis studies (cf. Huan 2016; Molek-Kozakowska 2017, 2018; Fruttaldo and Venuti 2017; Lorenzo-Dus and Smith 2018; Fuster-Márquez and Gregori-Signes 2019; Makki 2019, 2020; Maruenda-Bataller 2021, among others). What still needs more exploration and research, however, is providing and populating it with further Corpus Linguistic (CL) tools, especially to aid its application to large bodies of data nowadays vastly available (Potts *et al.* 2015; Maruenda-Bataller 2021). The main objective of this

¹ I express my deepest gratitude to Patricia Bou Franch and Sergio Maruenda Bataller for their advice, guidance and mentorship during the research process. My special thanks also go to Monika Bednarek, Pascual Cantos Gómez, and Miguel Fuster-Márquez for their feedback and advice. Finally, I would like to express my appreciation to the two anonymous reviewers whose constructive comments improved the quality of the paper considerably.



paper is to explore such possibilities and offer some complementary CL and statistical tools for the model by applying it to a corpus of British Broadsheets' news coverage of the Brexit campaign.

The paper is organised as follows. Section 2 offers information on news values, Section 3 discusses the discursive construction of reality, while Section 4 deals with the challenges of combining DNVA with CL. After these preliminaries, Section 5 provides information on the corpus methodology used. Sections 6 and 7 are the core of the study and offer some considerations on the statistical and textual analysis of news values, and a case study on discursive constructions through news values, respectively. Finally, Section 8 provides some concluding remarks.

2. NEWS VALUES

News values are the criteria that determine the likelihood of an event being reported as news (Westerståhl and Johansson 1994; Palmer 2000). In essence, they determine what is news(worthy) (Bednarek and Caple 2014). News values contain a wide range of journalism assumptions that “prioritise the unexpected, the unusual, the conflictual, the discrete, the dramatic or the extraordinary, over consensual, the harmonious, the predictable” (Bell 1997: 10).

Despite this seemingly straightforward definition of news values, the study of newsworthiness has been a controversial and a much-researched topic in media studies. Initially, Galtung and Ruge (1965) conceptualised them as ‘common-psychology selection’ criteria that work heuristically in the mind of news practitioners. Later, Harcup and O’Neill (2001) completed the initial list proposed by Galtung and Ruge (1965) with other factors, such as ‘entertainment’ and ‘positivity’, while expanding ‘eliteness’ to organisations, institutions, and the paper’s agenda. Schultz (2007) proposed implicit news values related to the *doxa* of journalism as another dimension to be considered. However, this initial wave of studying news values seems to be marked by a dominant interest in journalistic practice rather than news text and linguistics. However, for the first time, Bell (1991) differentiates between news values embedded in the news events and news values related to producing a news story, that is, a text. Van Dijk (1988) and Fowler (1991) also introduce the socio-cognitive, intersubjective, and discursive elements of news values. As

pointed out in Bednarek and Caple (2017), such a multi-dimensional analysis causes a certain degree of conflation between different aspects of news values.²

3. THE DISCURSIVE CONSTRUCTION OF NEWSWORTHINESS

Bednarek and Caple (2017: 51) put forth a discursive framework that adopts a middle ground between constructionism and realism. ‘Social constructionism’ is the theory originally developed by Berger and Luckmann (1967), refuting the objectivity of social phenomena, attributing them to the shared assumptions, mental representation and habituated actions of social actors in the society. DNVA accepts the main tenet of the constructionist perspective on how reality is given meaning by the media. It is acknowledged that events are inherently endowed with newsworthiness to a certain extent. However, the media also play a crucial role in constructing the events as such. On the one hand, the media can emphasise or de-emphasise certain news values in texts (Bednarek and Caple 2014: 139). On the other hand, the potential news value of events depends on a given sociocultural system that assigns them value. Following a range of publications reviewing the existing literature on news values (Caple and Bednarek 2013; Caple and Bednarek 2016), Bednarek and Caple (2014, 2017) formulated their approach to news values as discursively constructed. They shift the focus from the news event *per se* (a cognitive decision, eventually) to the news text as a complex of texts plus image or visuals. They depart from previous approaches because they consider that a discursive approach is tangible and analysable, and it can also account for how newsmakers employ news values to construct newsworthiness. In this line, they propose to distinguish news values from news writing objectives and news selection factors. News values, for their part, are defined in relation to the concept of newsworthiness and as constructed through discourse in each community (Bednarek and Caple 2017: 42). In this sense, news values are communicated through discourse, which means that they are constructed and reconstructed through discourse in the processes of pre-news, during-news, and post-news production (Bednarek and Caple 2017: 43).

Following such a framework and focusing on the discursive aspect of news values, Bednarek and Caple (2017: 57–64) propose a comprehensive list of news values. Table 1 shows the news values constructed in discourse. Whenever possible, I have replaced

² See Caple and Bednarek (2013) for a full revision of the field prior to DNVA.

examples in the original with examples from my own corpus of the news discourse of Brexit campaign coverage.

News Value	Linguistic devices and examples
<i>Consonance</i> ([stereo]typical)	References to stereotypical attributes or preconceptions; assessments of expectedness/typicality (<i>typical, famed for</i>); similarity with past (<i>yet another, once again</i>); explicit references to general knowledge/traditions, and so on (<i>well-known</i>).
<i>Eliteness</i> (of high status or fame)	Various status markers, including role labels (<i>the Queen, Ministers, Economists</i>); status- indicating adjectives (<i>EU commission top analyst</i>); recognised names (<i>David Cameron, Boris Johnson</i>); descriptions of achievement/fame (<i>were selling millions of records a year</i>); use by news actors/sources of specialised/technical terminology, high- status accent or sociolect (esp. in broadcast news).
<i>Impact</i> (having significant effects or consequences)	Assessments of significance (<i>momentous, historic, crucial</i>); representation of actual or non-actual significant/relevant consequences, including abstract, material, or mental effects (<i>Brexit could mean for the economy, the economic Impact of Brexit, the effect on immigration, the outlook after leaving the EU</i>).
<i>Negativity/positivity</i> (negative/positive)	References to negative/positive emotion and attitude (<i>Brexit jitters, fears of Brexit, a safer UK</i>) negative/positive evaluative language (<i>shock, suffer, improve the economy</i>); negative/positive lexis (terrorism, economic damage, favour growth, brighter future); descriptions of negative (<i>uncontrolled immigration</i>) or positive behaviour (<i>has broken his promise, unveiled a cabinet with an equal number of men and women</i>).
<i>Personalisation</i> (having a personal/human face)	References to ‘ordinary’ people, their emotions, experiences (<i>domestic risk that our economy faces, people with disabilities and other ordinary people here and across Europe</i>); use by news actors/sources of ‘everyday’ spoken language, accent, sociolect (esp. in broadcast news).
<i>Proximity</i> (geographically or culturally near)	Explicit references to place or nationality near the target community (<i>British people</i>); references to the nation/community via deictics, generic place references, adjectives (<i>here, the nation’s capital, home-grown</i>); inclusive first-person plural pronouns (<i>our nation’s leaders</i>); use by news actors/sources of (geographical) accent/dialect (esp. in broadcast news); cultural references (<i>haka, prom</i>).

Table 1: The DNVA framework (adapted from Bednarek and Caple 2017: 79–80)

4. CHALLENGES IN COMBINING DNVA WITH CL

Bednarek and Caple (2017) point out some potential avenues for the application of DNVA and the possible ways in which the framework can be adjusted, modified, and enriched. Combining DNVA with CL tools is one of the most pertinent and potentially problematic areas among these avenues. Recently, several studies have combined DNVA with CL techniques to cover large-scale corpora. Bednarek and Caple (2017), for instance, highlight the possibility and prospects of combining DNVA as a qualitative method with CL as a quantitative method. However, one common problem with applying DNVA to a large corpus is that there are major overlaps between different categories, and many words and linguistic resources might construct different news values based on the context in

which they are used. This is mainly due to the fact that news values have an evaluative meaning, and they are not merely constructed by isolated words (Bednarek 2016: 229). This is probably the main reason why news values show considerable overlap with each other in practice. The sources of these overlaps are the evaluative aspect of news values and their culturally grounded nature. In operational terms and on the level of analysis and coding, these overlaps make it challenging to apply models like DNVA to large corpora.

Current CL tools have advanced significantly in recent years, especially with more tagging software available but many of these tools are still under development and have constraints (cf. Walsh *et al.* 2008, among others). Currently, the most commonly used software for semantic annotation includes the *USAS tagger*,³ *Wmatrix*⁴ and similar automatic taggers. However, using these tools for evaluative language, in which context and interpretation play a considerable role, proved problematic and in dire need of complementary tools such as the concordance analysis (López-Rodríguez 2022). This is especially extendable to DNVA since the previous research trying to use CL within this framework underlines such difficulties. Most importantly, Maruenda-Bataller (2021) highlights the potential overlap between the pointers of different news values due to the evaluative meaning aspect of news values, especially for particular audiences. Therefore, he specifically calls for a more nuanced analysis of particular prosodies among the linguistic pointers (Maruenda-Bataller 2021: 160). In previous studies, Potts *et al.* (2015) also drew attention to such challenges concerning CL techniques and software packages, especially since there is no closed list of news value devices due to their evaluative and highly culturally grounded nature. Therefore, it is not yet possible to use common semantic taggers to tag news values straightforwardly.

It seems wise to point out that the impediments mentioned above should not discourage researchers from using quantitative analysis in DNVA. As shown in the existing literature and underlined by Bednarek and Caple (2017), DNVA can be used to explore how news values are used to construct particular topics. Certain news values might be emphasised or used more than others or, by contrast, be rare or absent in the discourse, and all this may have serious ideological implications (Bednarek and Caple 2017: 239). Combining DNVA and CL as qualitative and quantitative methods to approach the analysis of a large corpus has proved fruitful in practice. Still, the whole

³ <https://ucrel.lancs.ac.uk/usas/>

⁴ <https://ucrel.lancs.ac.uk/wmatrix/>

endeavour seems to be in its very early stages and in need of further studies addressing this aspect, especially in elaborating more statistically sound tools for comparison and contrast in comparative, cross-cultural, and cross-linguistic studies. Potts *et al.*'s (2015) analysis of the news reporting on Hurricane Katrina in the United States, Fuster-Márquez and Gregori-Signes's (2019) research on the discursive representation of tourism in press news stories, and Maruenda-Bataller's (2021) study on news reporting on violence against women are three prominent examples of such endeavours. These studies show that DNVA specifically proves to be useful in identifying and coding recurrent linguistic pointers that articulate certain discursive constructions. Consequently, it can shed light on (dis)similarities in how news values are used to convey newsworthiness cross-linguistically and cross-culturally. However, they also make it clear that caution should be considered when DNVA is applied to a large corpus, including the manual coding of the corpus and possible crossovers between news values, as well as the subjective nature of coding news values by the researcher.

In this paper, I pursue the line of CL plus DNVA analysis, specifically with the aim of elaborating a reliable statistical tool for comparative studies. Following previous studies, I designed a procedure taking advantage of some of the CL tools that proved their usefulness in previous studies (including frequency, collocation and concordance analysis) and added some other tools and analytical steps to address some of the existing challenges. To do so, I apply DNVA to a different news environment (a referendum campaign), socio-political topic (Brexit), and journalism type (a campaign coverage) to map the ideological discourses present in that specific setting. I compiled a corpus (Section 5) from different ideological orientations and political stances to apply DNVA in a cross-ideological and comparative context and observe the potentials of this model in this regard, following the suggestion of Bednarek and Caple (2017) and subsequent research discussed above.

5. CORPUS NATURE AND DESIGN

A corpus of four major British broadsheets —*The Guardian*, *The Independent*, *The Times*, and *The Daily Telegraph*— was collected using *Nexis UK News* databases,⁵ as illustrated in Table 2. The search word used for data retrieval was *Brexit*. The results were

⁵ <https://www.lexisnexis.com/uk/legal/news?>

downsampled by limiting search timespan [22 February to 23 June 2016], news type [articles], and managing duplicities (i.e., articles repeated in digital and paper editions). The same procedure was used for each daily, resulting in four different sub-corpora.

Newspaper	Political Stance	Brexit Stance	Number of articles	Corpus tokens
<i>The Guardian</i>	Left	Remain	3,584	4,549,153
<i>The Independent</i>	Left	Remain	2,272	1,709,259
<i>The Times</i>	Right	Remain	1,696	1,071,314
<i>The Daily Telegraph</i>	Right	Leave	1,233	814,048
Total			8785	7,329,726

Table 2: Corpus description

For maximum representativeness and accuracy, I selected the newspapers with the most readers and circulation according to the *Readership Average Issue Reach Index* (Figure 1).⁶ I decided to exclude tabloids and regional press because their journalistic style differs from national broadsheets.

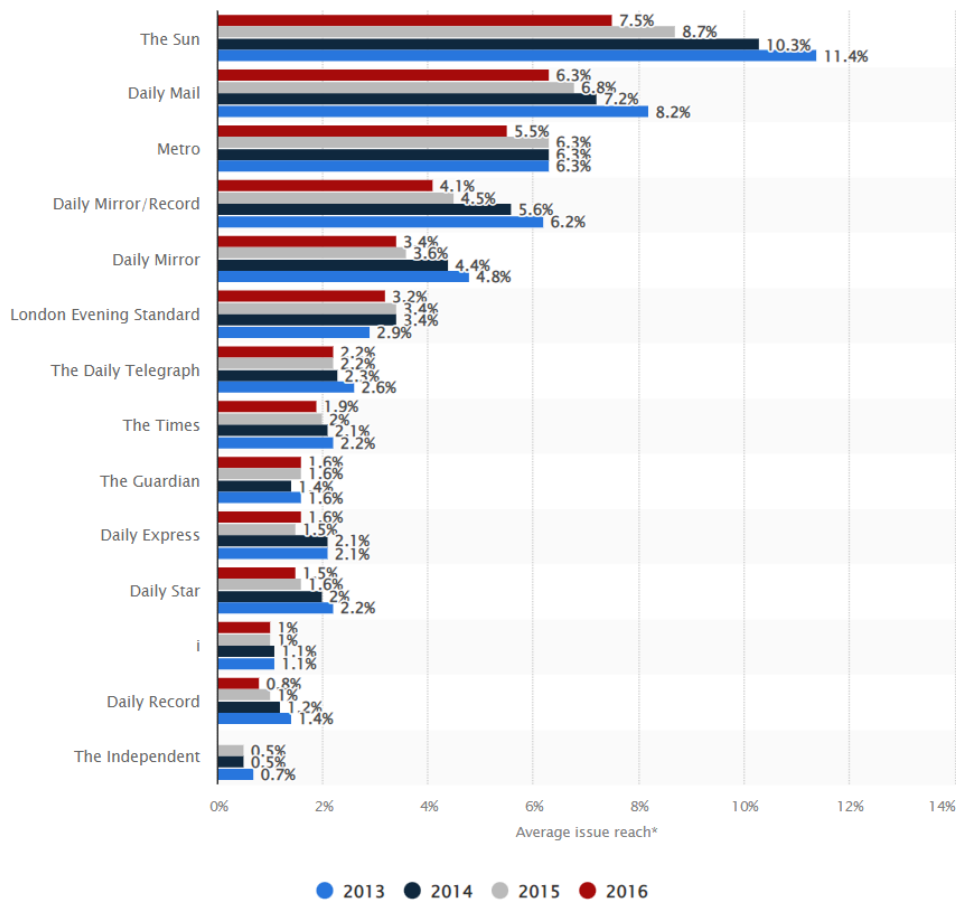


Figure 1: Newspapers rank in the UK in 2013 and 2016

⁶ <https://www.statista.com/statistics/290086/newspapers-ranked-by-penetration-in-the-united-kingdom/>

In choosing the broadsheets, I looked for a balance between two important factors: traditional political affiliation and Brexit stance and selected two prominent left-wing newspapers, *The Guardian* and *The Independent*, and two right-wing broadsheets, *The Times* and *The Daily Telegraph*. As for the Brexit stance, the first three newspapers officially backed a remain vote, while *The Daily Telegraph* was the only one supporting a leave vote. There was no left-wing broadsheet officially backing a leave vote.

The coding language *R* was used to clean and prepare the corpus, i.e., to manage the textual data with more precision and minimise ‘noise’ in the corpus that could affect the results. *R* is mostly known for its statistical capacities and has gained popularity within CL in recent years thanks to the introduction of several useful packages explicitly designed for this purpose (Gries 2009). Therefore, after compiling the corpus using *Nexis UK*, the corpus was cleaned using the *R Software Package* (R Core Team 2013) with the help of *tm library* (Feinerer and Hornik 2018). Once the corpus was cleaned, it was saved in a plain text format (TXT) so that it could also be imported to other CL software.

6. STATISTICAL AND TEXTUAL ANALYSIS OF NEWS VALUES: SOME CONSIDERATIONS

6.1. Frequency analysis and cut-off points

In the frequency analysis of a corpus, one of the most critical decisions is to establish the cut-off point to extract the main search terms for the next stages of the analysis. This involves determining the threshold that separates the words in a frequency list that should be further scrutinised and those that would not be considered in the analysis. There is, however, no consensus in the literature on how to decide on a cut-off point. For instance, Baker *et al.* (2008) and Bednarek and Caple (2014, 2017) set this cut-off point at the 100 most frequent words. The decision seems to be based on common sense and experience rather than any statistical criterion. There is also no clear consensus among those scholars who make use of a statistical yardstick. For instance, Biber *et al.* (1999) and Scott and Tribble (2006) suggest a cut-off point of 20 per 1,000,000 words, but others, such as O’Keeffe *et al.* (2007), indicate a completely different number of 20 for a five-million-word corpus. In this paper, however, I put forward a more precise statistical criterion to fix a cut-off point. The reason for this is two-fold. First, when faced with a sizeable corpus, the sheer volume of the corpus requires a more reliable method of analysis.

Second, in many cases, the size of the sub-corpora to be compared is considerably different, and using the same number for all subsets could affect the comparison.

To address this concern, I used a cluster analysis technique to differentiate the most frequent words in the corpus. Cluster analysis refers to any type of multivariate analysis used for the categorisation and classification of a vast number of items, in which advanced statistical metrics are used to put different items in hierarchical trees, also labelled a ‘dendrogram’, according to the desired variable or variables (Saraçlı *et al.* 2013). Cluster analysis is already a well-known and widely used tool in CL (Qian 2017). Still, so far, its use has been limited to semantic clustering, grammatical research and variation analysis (Moisl 2015). It is a sophisticated statistical tool but, thanks to openly available coding packages such as *R*, it can be widely used in different areas of applied linguistics. Since it is already available as an open-source, downloadable package in *R*, it is easy for linguists to use and it does not entail considerable training time or additional coding knowledge. In this study, to identify clusters of recurrent words, hierarchical clustering based on *parametrised finite Gaussian mixture models*⁷ was performed using *R*’s *mclust4* package (v5.4.5), and histograms were drawn with *ggplot2* package (v3.2.1). Figure 2 below shows an example of the code for running the cluster analysis and extracting the search terms in *The Daily Telegraph*.

```
#Making cluster
library(mclust)

fit <- Mclust(table(Telegraph_plot$top))

fit$data[max(fit$uncertainty)] #select max uncertainty
plot(fit) # plot results

#Index to plot by colours
Telegraph_plot$cluster <-
fit$classification[match(Telegraph_plot$top,names(fit$classification))]

#Making histogram
ggplot(Guardian_plot,aes(x=factor(top,levels=names(sort(table(top),decreasing =
TRUE)))
))) +
  geom_bar(aes(fill = factor(colores)), alpha = 0.7, width = 1, col = 'black')
+ theme_classic()+ theme(axis.text.x = element_text(angle = 90)) +
  scale_fill_manual(values = c('darkred', 'darkgreen', 'blue', 'black')) +
  guides(fill = 'none') + ylab('Frequency\n') + xlab('') +
  scale_y_continuous(expand = c(0,0.1), breaks =seq(0,60000,by=10000))
```

Figure 2: *R* code for running the cluster analysis and extracting the search terms

⁷ A Gaussian mixture model is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters.

Figure 3 depicts the result of the cluster analysis in a histogram using the *ggplot* tool of the *R* package. Each colour in the figure represents a cluster of frequent words, from the most frequent to the least frequent ones. The full results of the cluster analysis and words that were chosen for further analysis are shown in Section 7.3.

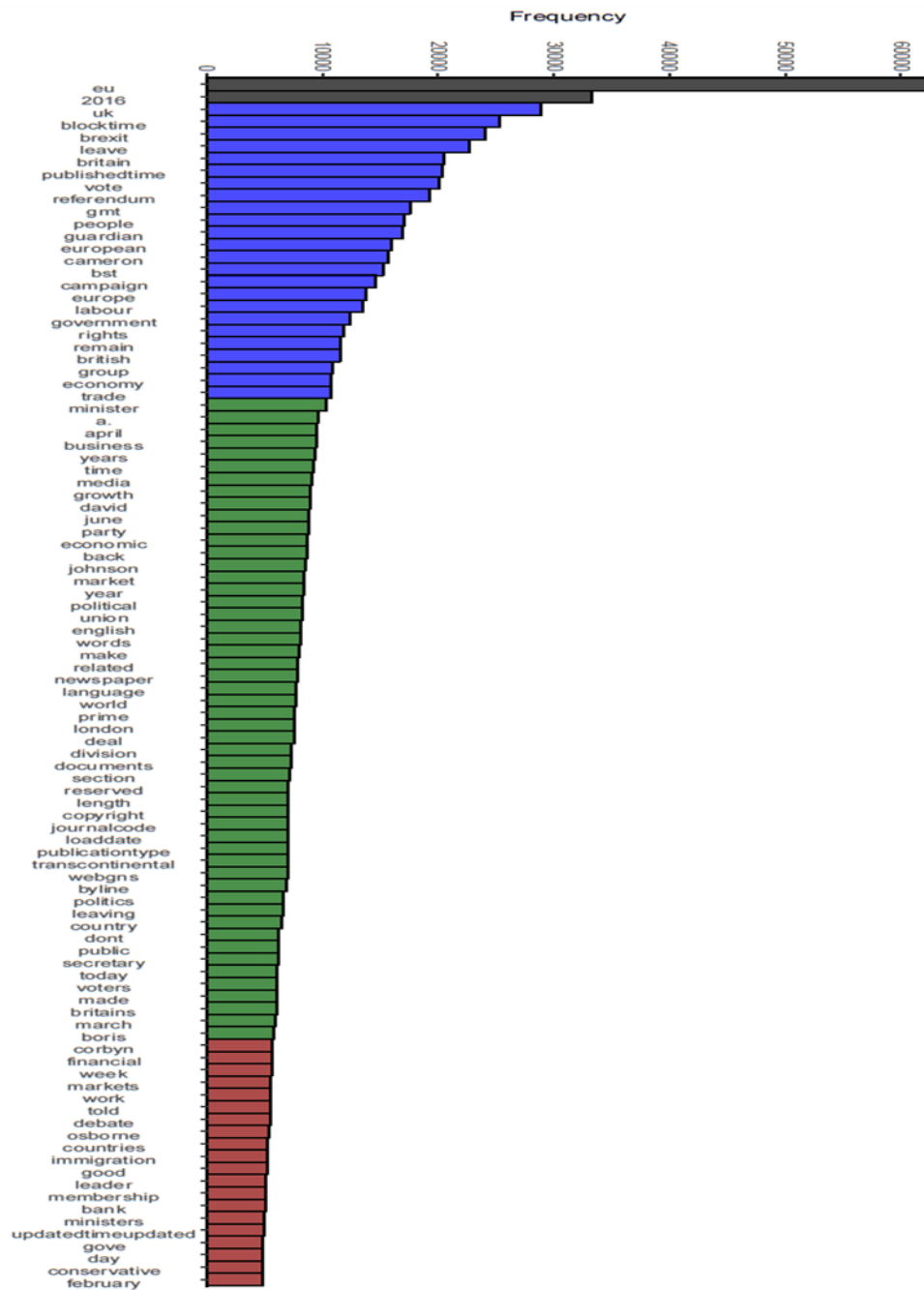


Figure 3: Hierarchical clustering to extract the most frequent words of the corpus

6.2. Coding collocations: Word bundles and what associations can tell about discourse

Firth (1967) noted that meaning is usually constructed by association and not by individual units. With this idea in mind, the importance of collocations was established (Halliday 1966; Sinclair 1991), and soon, a widespread interest in multi-word units emerged (Greaves and Warren 2010: 212–213). The test of collocability traditionally referred to the fact that two words would occur together in a text with statistically sufficient frequency so that the co-occurrence could not be ascribed to mere chance. With modern CL tools, however, we can go further than direct collocation. This is possible in two ways. First, instead of looking at the traditional meaning of a collocation that is adapted for educational purposes (words that tend to be used together although there is no dictating grammatical rule), collocations can be considered further by association in discourse structure. In this sense, collocations do not simply refer to the co-occurrence of two words together, but they are instead the tendency of several words to occur in a certain proximity (with on average three to five words to left and right). This shows how, in constructing a discourse, word associations work to bring specific discourse patterns and convey a message by semantic association.

Collocation analysis, in this sense, is used in Bednarek and Caple (2017) and, more extensively, by Potts *et al.* (2015). The discursive construction of a specific topic by means of news values, as Potts *et al.* (2015) and Maruenda-Bataller (2021) also highlight, is highly context-based. This means that the same word could construct different news values in different contexts. For this reason, almost all CL tools and software packages include the option of concordance lines, giving the researcher the possibility of carrying out a qualitative analysis by considering the context in which the collocation occurs. Therefore, collocations/pointers should be coded into different news values by qualitatively examining their concordance lines in each case. This adds a considerable amount of work in the analytical stage of the research and requires adding sizeable qualitative, manual work since the concordance lines should be checked one by one by the coder. However, due to the decisive role of context and co-text in conveying news values, qualitative analysis is an absolute necessity at this stage and cannot be replaced by any machined-based procedures yet.

All things considered, and drawing upon the previous literature, the procedure which is used in the present study is shown in Table 3.

Stages	Steps of procedure	Rationale	Tools
First stage	Frequency analysis.	First step in CL, used by Bednarek and Caple (2017) to select most frequent words for analysis.	Frequency analysis package; <i>R</i> .
	Most common words selection. Codification of these words into semantic fields.	No existing consensus on the cut-off point. So that words belonging to the same semantic field can be analysed together.	Cluster analysis and <i>ggplot packages</i> ; <i>R</i> Qualitative analysis.
Second stage	Collocation analysis.	To find the most frequent pointers to news values; following Potts <i>et al.</i> (2015) and Maruenda-Bataller (2021).	Collocation analysis; <i>AntConc</i> (Anthony 2014).
	Concordance analysis and coding pointers.	There is no pre-existing tagging for pointers to different news values. Since a single pointer can constitute different news values in different contexts, all pointers were checked and coded by the researcher.	Concordance analysis; <i>AntConc</i> . Qualitative analysis by the researcher.
	Calculating news values distribution and statistical testing for differences.	In order to compare different sub-corpora in terms of news values distribution in each semantic field.	Statistical testing; Chi-square using <i>R</i> .
Third stage	Analysing news values distributions.	To analyse how patterns of news value use were different between the subcorpora.	Tables and histograms.
	In-depth analysis of selected excerpts.	For in-depth analysis of how news values were used in different contexts.	Qualitative analysis.

Table 3: Summary of the procedure and rationale

7. CASE STUDY: DISCURSIVE CONSTRUCTION OF BREXIT THROUGH NEWS VALUES

7.1. Cluster analysis determining semantic fields

Table 4 below shows the result of the cluster analysis and the possible semantic fields to be considered in the study. It can be observed that the most frequent words across the

newspapers are very similar. The most frequent words belonging to the same semantic fields were grouped together to be considered as search terms. The final established semantic fields of the corpus could be the following. Firstly, the search term *Brexit* must be considered separately, due to its evident importance and its appearance as one of the most frequent words in all newspapers. The second semantic field is *economy*. In this case, a range of different search terms related to the same semantic field are found in the data: *economy, economic, economy, economic, trade, business, financial, growth, and market*. The third semantic field is *immigration*. The salience of *immigration* in Brexit debates is evident and confirmed by the results of the cluster analysis. The fourth semantic field is the representation and discursive construction of *EU* vs. *UK*. The duality between the two entities is of utmost importance in this discourse, as is confirmed by the occurrence of a range of words related to these two areas. The search terms of this semantic field are *EU, Europe, European, UK, Britain, British*. Finally, the fifth semantic field contains general references to people with constituting search terms, such as *people* and *public*.

The similarity observed in the most frequent words and topics across the four data sets is noteworthy both on the quantitative and qualitative levels. On the quantitative level, it shows a degree of intra-reliability of the proposed tool. However, this should be tested in future research and *vis-a-vis* other corpora. On the qualitative level, it is an indication that the whole mediatic discourse of Brexit was probably driven by very similar forces and discursive practices. This falls beyond the scope of this paper, but each of the semantic fields discovered at this stage of analysis is worthy of an independent future discursive analysis within the broader societal and discursive practices around the Brexit referendum and its news discourse.

The Guardian		The Independent		The Daily Telegraph		The Times	
Word	NF	Word	NF	Word	NF	Word	NF
EU	6,2331 (6.24%)	EU	18,006 (6.49%)	EU	7,494 (5.76%)	EU	8,301 (4.95%)
UK	28,934 (2.9%)	Brexit	9,273 (3.34%)	Brexit	3,415 (2.63%)	Brexit	4,243 (2.53%)
Brexit	24,128 (2.42%)	UK	9,226 (3.32%)	Britain	2,857 (2.2%)	Britain	3,720 (2.22%)
Britain	20,531 (2.06%)	Britain	5,025 (1.81%)	UK	2,747 (2.11%)	UK	2,964 (1.77%)
People	17,035 (1.71%)	People	4,882 (1.76%)	European	2,081 (1.6%)	European	2,437 (1.45%)
European	15,936 (1.6%)	British	3,761 (1.35%)	People	1,757 (1.35%)	English	2,383 (1.42%)
Cameron	15,638 (1.57%)	Cameron	3,749 (1.35%)	Cameron	1,672 (1.29%)	People	2,351 (1.4%)
Europe	13,783 (1.38%)	Europe	3,274 (1.18%)	English	1,606 (1.23%)	Europe	2,138 (1.28%)
British	11,503 (1.15%)	English	3,066 (1.1%)	Europe	1,509 (1.16%)	British	1,946 (1.16%)
Economy	10,774 (1.08%)	Johnson	2,904 (1.05%)	British	1,438 (1.11%)	Business	1,843 (1.1%)
Trade	10,675 (1.07%)	Boris	2,488 (0.9%)	Business	1,346 (1.03%)	Cameron	1,673 (1%)
Business	9,421 (0.94%)	David	2,412 (0.87%)	Market	1,283 (0.99%)	Trade	1,658 (0.99%)
Growth	8,883 (0.89%)	Economic	2,095 (0.75%)	Economic	1,243 (0.96%)	Market	1,547 (0.92%)
David	8,875 (0.89%)	Trade	1,970 (0.71%)	Trade	1,097 (0.84%)	Economy	1,191 (0.71%)
Economic	8,627 (0.86%)	Business	1,965 (0.71%)	Economy	1,092 (0.84%)	David	1,174 (0.7%)
Johnson	8,456 (0.85%)	Public	1,881 (0.68%)	Johnson	1,014 (0.78%)	Economic	1,161 (0.69%)
Market	8,389 (0.84%)	Economy	1,856 (0.67%)	David	999 (0.77%)	Johnson	1,153 (0.69%)
English	8,104 (0.81%)	Immigra- tion	1,500 (0.54%)	Financial	806 (0.62%)	Boris	925 (0.55%)
Public	6,179 (0.62%)	Market	1,466 (0.53%)	Boris	742 (0.57%)	Growth	914 (0.55%)
Boris	5,742 (0.57%)	Osborne	1,249 (0.45%)	Growth	685 (0.53%)	Financial	794 (0.47%)
Financial	5,685 (0.57%)	Financial	1,127 (0.41%)	Public	672 (0.52%)	Public	785 (0.47%)
Markets	5,493 (0.55%)			Immigra- tion	601 (0.46%)	Immigra- tion	728 (0.43%)
Oborne	5,337 (0.53%)			Osborne	588 (0.45%)	Osborne	706 (0.42%)
Immigration	5,179 (0.52%)			Gove	562 (0.43%)		
Gove	4,876 (0.49%)						

Table 4: Categorising the most frequent words of the sub-corpora into related semantic fields

7.2. Distribution of news values for Brexit

The overall distribution of news values is shown in Tables 5 and 6 below.

	<i>Guardian_c</i>	<i>Independent_c</i>	<i>Times_c</i>	<i>Telegraph_c</i>	P-value
<i>Negativity</i>	1,761	865	560	140	<0,001
<i>Eliteness</i>	1,478	483	140	71	<0,001
<i>Impact</i>	1,102	4,832	763	33	<0,001
<i>Positivity</i>	0	0	10	46	<0,001
<i>Timeliness</i>	368	0	203	61	<0,001
Total Collocations	4,709	6,180	1,676	351	

Table 5: Absolute frequency and statistical test of news values in Brexit

	<i>The Guardian</i>	<i>The Independent</i>	<i>The Times</i>	<i>The daily Telegraph</i>
<i>Negativity</i>	37%	14,00%	33%	40%
<i>Eliteness</i>	31%	7,82%	8%	20%
<i>Impact</i>	23%	78,19%	46%	9%
<i>Positivity</i>	0%	0,00%	1%	13%
<i>Timeliness</i>	8%	0,00%	12%	17%
Total	100%	100%	100%	100%

Table 6: Normalised frequencies of the news values for Brexit

The *R* code used to calculate the P-value for the statistical significance of the table is shown in Table 7, whereas the normalised results are visually represented in Figure 4, for ease of comparison.

```
#brexit
prop.test(c( 1761, 865, 560, 140), c( 4709, 6180, 1676, 351))
prop.test(c( 1478, 483, 140, 71), c( 4709, 6180, 1676, 351))
prop.test(c( 1102, 4832, 763, 33), c( 4709, 6180, 1676, 351))
prop.test(c( 0, 0, 10, 46), c( 4709, 6180, 1676, 351))
prop.test(c( 368, 0, 203, 61), c( 4709, 6180, 1676, 351))
```

Table 7: *R* code used in the study

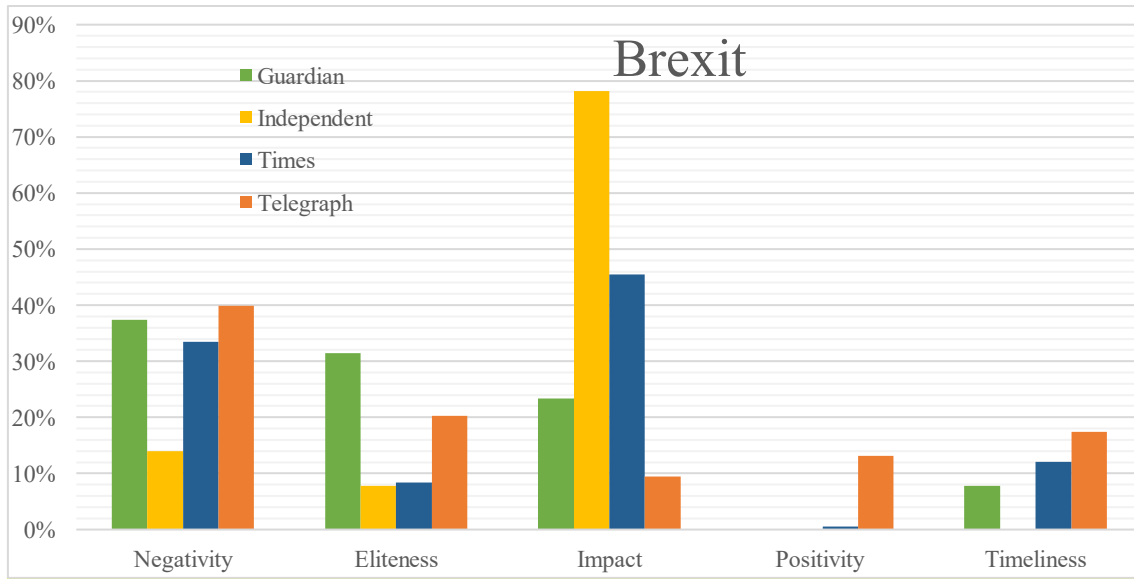


Figure 4: Normalised frequencies of news values for the semantic field of Brexit

As observed in Figure 4, in the discursive construction of Brexit, ‘impact’ is the most frequent news value in *The Independent* and *The Times*, and ‘negativity’ is most frequent in *The Guardian* and *The Daily Telegraph*. However, the pro-leave *Telegraph* used ‘impact’ far less frequently in comparison with the other newspapers. On the other hand, *The Independent* makes a considerably higher use of ‘impact’ when compared to the other newspapers. This is followed by ‘eliteness’, with a lightly higher usage in the left-wing and the pro-remain *Guardian*, on the one hand, and right-wing and the pro-leave *Telegraph* on the other, and similar normalised frequencies for the other two newspapers. ‘Timeliness’ shows a significant difference of use in the right-leaning and left-leaning press, with a higher frequency by the right, especially by the leave backing *Telegraph*. As for ‘positivity’, it is almost non-existent in the pro-remain newspapers, with a small exception of the right-wing *Times*, but significantly present in the pro-leave *Telegraph*.

The previous discussion makes it clear that the CL tools applied here can successfully map the ideological patterns in the discursive construction of Brexit in the analysed corpus of news discourse. The results specifically reveal important differences and similarities across political affiliation (left-right), and stance towards Brexit (leave-remain), revealing the type of discourse each outlet constructed around Brexit in their campaign coverage.

On another level, the analytical tools used here can also assist in delving into the specific discursive strategies for constructing newsworthiness around Brexit. In what follows, I present those results for the most significant news values observed in the

previous stage of analysis ('negativity/positivity', 'impact', 'eliteness'), and a more in-depth analysis of some excerpts from the corpus.

7.3. Negativity/positivity

'Negativity' is constructed through four major discursive strategies. The linguistic pointers related to each strategy are shown in Table 8.

Strategies	<i>The Guardian</i>	<i>The Independent</i>	<i>The Times</i>	<i>The Daily Telegraph</i>
Fear and danger	<i>Fears</i>			
	<i>Fear</i>			
	<i>Dangers threaten</i>		<i>Dangers</i>	<i>Dangers</i>
	<i>Threat</i>	<i>Fears</i>	<i>Jitters</i>	<i>Fears</i>
	<i>Hit</i>		<i>Fears</i>	<i>Threat</i>
	<i>Worries</i>		<i>Threat</i>	<i>Fear</i>
	<i>Concerns</i>		<i>Hit</i>	
Uncertainty and risk	<i>Trigger</i>		<i>Puts (at risk)</i>	
	<i>Risks</i>		<i>Risks</i>	
	<i>Risk</i>	<i>Risk</i>	<i>Trigger</i>	<i>Risks</i>
	<i>Uncertainty</i>		<i>Concerns</i>	
			<i>Uncertainty</i>	
			<i>Risk</i>	
Negative outcomes	<i>Negatively</i>		<i>Implications</i>	
	<i>Harm</i>		<i>Consequences</i>	
	<i>Consequences</i>	<i>Consequences</i>	<i>Blow</i>	<i>Consequences</i>
	<i>Hurt</i>	<i>Damage</i>	<i>Damage</i>	
	<i>Cost</i>	<i>Cost</i>	<i>Cost</i>	
	<i>Damage</i>			
	<i>Costs</i>			
Accusations and admonitions	<i>Blame</i>		<i>Warns</i>	
	<i>Warns</i>	<i>Warn</i>	<i>Warn</i>	<i>Warnings</i>
	<i>Warn</i>	<i>Warns</i>	<i>Warnings</i>	<i>Warned</i>
	<i>Warnings</i>	<i>Warning</i>	<i>Warning</i>	<i>Warning</i>
	<i>Warning</i>	<i>Lead</i>	<i>Warned</i>	<i>Concerns</i>
	<i>Accused</i>	<i>Warned</i>	<i>Accused</i>	
	<i>Warned</i>			

Table 8: Linguistic pointers of 'negativity' for Brexit across strategies and newspapers

The discursive strategies used to construct 'negativity' fall into four major categories. The first category is 'fear and danger', in which we observe a range of words/pointers that associate Brexit with fearful scenarios, including words such as *fear* and *danger*, as well as *threat*, *threatening*, and *hit*. Another salient discursive strategy of 'negativity' is 'uncertainty and risk'. In this strategy, a range of pointers are used with Brexit, including *risk*, *uncertainty*, and other words (see Table 8) which induce the sense of imminent risk, worries, and concern with a high emotional charge over the future. The third discursive

strategy is related to the ‘negative outcomes’ of voting for or against Brexit. This includes pointers related to *consequences/implications*, or *damage*, *cost*, and *harm*. The fourth discursive strategy is ‘negative prediction or admonitions’ about different scenarios related to Brexit, mainly indicated by pointers related to different people or specific reports (see Table 8).

Inducing the sense of ‘fear and danger’ seems to be a pervasive strategy in this semantic field. In many news stories, the growing ‘fears of Brexit’ and the negative consequences it would bring about were mentioned to construct the ‘negativity’ around Brexit, as shown in excerpt (1).

Excerpt 1: *The Guardian*, Business section, June 7, 2016, Tuesday
Sterling’s value has become increasingly volatile as **fears of a Brexit** have increased among investors.

This news story concerns the growing volatility in Sterling’s value, in which a highly fearful scenario was constructed about the value of the pound. In the text of the news story, the pointer of ‘fears’ conveys a sense of ‘negativity’. Although ‘negativity’ is the most salient news value used in excerpt (1), it is not the only news value adopted. There also are some pointers related to ‘superlativeness’ as in the case of *hits*, *peaks*, or *height*, which, at the same time, highlight ‘negativity’. In addition, the phrases *increasingly volatile* and *have increased* can be attributed to pointers of the news value of ‘superlativeness’. It seems then that, in (1), ‘negativity’ and ‘superlativeness’ are combined to enhance the negative outcome of Brexit constructed in this discourse. Hence, we might say that news values are used in a synergistic manner. This is in line with other studies, which also showed that news values may co-occur in different contexts (Fruttaldo and Venuti 2017; Fuster-Márquez and Gregori-Signes 2019; Makki 2019, 2020; He and Caple 2020).

The strategy of associating fears with Brexit (and consequently with other adverse outcomes) is frequently found in the data. Excerpt (2), taken from *The Guardian*, illustrates the use of same strategy when discussing a different topic.

Excerpt 2: *The Guardian*, Business section, May 17, 2016, Tuesday
[Stamp duty rush boosts March house prices, says ONS] High-end London homes have seen prices fall since April, according to some reports, as the higher stamp duty rates and **fears of a Brexit** deter wealthy buyers. The International Monetary Fund is one of many economic forecasters to warn that UK house values will plummet should Britons vote to leave the EU in the June referendum.

The news story, which revolves around a potential fall in house prices in London, is heavily constructed in negative terms. The use of collocations between *Brexit* and *fears* warns, in worrying terms, about a possible plunge in housing prices. Once again, in addition to ‘negativity’, the news values of ‘impact’ (*deter*), ‘superlativeness’ (*plummet*), and ‘proximity’ (*London*) are used. *Plummet* conveys a very steep and sudden fall, not an ordinary decrease in prices, and *UK*, *London*, and *Sterling* construe ‘proximity’.

The examples above highlight the fact that constructing newsworthiness is social and ideological, as has been shown in the literature (Bednarek and Caple 2017; Fruttaldo and Venuti, 2017; Makki 2019, 2020; Maruenda-Bataller 2021). Both a drop in the value of Sterling and in London housing prices can actually be described as positive for some social groups. Future home buyers would indeed benefit from such lower prices, and a drop in Sterling price would be desirable for the buyers of imported food products. However, in both cases, Brexit was constructed as a very negative phenomenon with significant adverse consequences. Considering that *The Guardian* is a high-end quality paper, this makes sense, as the majority of its readers are probably upper-middle-class liberals; the news stories might be constructed to be ‘negative’, ‘impactful’, and with relatively enhanced outcomes (‘superlativeness’) for their readership, at least as perceived by the paper’s editorial. This especially concurs with discursive approaches that underline the role of audience, such as Bell’s (1991) audience-design model.

By contrast, in the discourse of *The Daily Telegraph*, the negative impact was downplayed quite strategically, as shown in (3).

Excerpt 3: *The Daily Telegraph* (London), June 11, 2016, Saturday, Edition 1; National Edition

[Voters **fear Brexit** will spoil our holidays]

The debate about Britain’s place in Europe in the run-up to the referendum on June 23 has seen a variety of doomsday scenarios voiced by both the Remain and Leave camps. But not even the most cynical politician has yet suggested that, in the wake of Britain voting to break ties with the EU, UK holidaymakers would be banned from visiting the beaches of Greece, the bars of Amsterdam, or the restaurants of Paris.

In excerpt 3, the news story that contains the pointer of *fear* is dedicated entirely to the possible negative outcome of Brexit on the price of holidays. The headline uses irony to dismantle the idea of Brexit affecting foreign holidays, as observed in the body of the news story. Although the lexis conveys ‘negativity’ (*fear*) and ‘impact’ (*will*), ‘negativity’ is considerably downplayed in the discourse when compared to excerpts (1) and (2).

‘Superlativeness’ is absent in the news story. The topic itself has some role in downplaying the ‘impact’: other news stories are about day-to-day and primary needs (housing) or serious economic matters (stock market and investment), but in this case, the topic is a rather luxury item (holidays abroad). The ideological and social aspects of the news story are also considerable.

The analysis of (3) reflects back on two critical aspects of the news value. On the one hand, the use of irony in (3) could entirely cast doubt on coding ‘fear’ as conveying ‘negativity’, therefore, confirming Potts *et al.* (2015) and Maruenda-Bataller’s (2021) remarks on the difficulties of quantifying news value usage relying exclusively on CL tools. However, in large corpora and in mass media, the sheer fact of repeated association of certain lexical items can be meaningful and effective in influencing the audience. The effect of repetition on public opinion in mass media has been studied extensively (cf. Lecheler *et al.* 2015; Liu *et al.* 2019). However, textual subtleties such as those observed in this piece should always be considered, highlighting the importance of an in-depth, qualitative analysis in the DNVA model. On the other hand, the above-mentioned point shows how news values usage is intertwined with journalistic social practices, as related to their potential audiences and interest groups (cf. Huan 2016; Fuster-Márquez and Gregori-Signes 2019). Similar patterns may be observed in all other strategies in constructing ‘negativity’ in association with Brexit.

In terms of ‘positivity’ around Brexit, there are few statistically significant collocations. The only ones considered as potential pointers are *improve* and *favour*. However, a noteworthy finding is the amount of constructed ‘positivity’ in the discourse found in *The Daily Telegraph*, compared to the almost absence of this news value in the discourse of the pro-remain newspapers. It is clear from the news values distribution that the pro-remain newspapers shied away from constructing any positivity in their coverage of Brexit. Curiously, *The Daily Telegraph* is the newspaper that uses the highest number of news values of both ‘positivity’ and ‘negativity’. The pointer *favour*, for example, is used in *The Daily Telegraph* discourse to construct the ‘positivity’ of some particular groups endorsing Brexit, as illustrated in (4).

Excerpt 4: *The Daily Telegraph* (London), May 25, 2016, Wednesday, Edition 1, National Edition

[Women rightly see the EU as a threat to family]

Could this explain why a poll by Netmums shows that women are more likely to see the EU as a threat to family life, and mothers **are inclined to favour Brexit**?

The story in (4) clearly represents mothers and families as supporting Brexit and, at the same time, constructs a very ‘negative’ and anti-family picture of the EU. Referring to *Netmums* also constructs ‘eliteness’. This is particularly interesting since *Netmums* is a well-known parenting advice institution and, therefore, backs up the intended narrative of the news story. In (4), different news values are also combined to enhance the message. In addition to the pattern observed in previous examples, where news values were used synergistically, news values are used antagonistically in excerpt (4) to enhance the degree of ‘positivity’ about Brexit intended by the newspaper. That is to say, a negative picture of the EU is constructed adjacent to a positive tone on Brexit, which might enhance the resonance and saliency of both narratives by juxtaposition.

Additionally, excerpt (4) also shows how ‘negativity’ and ‘positivity’ are constructed in relation to the ideological values of the newspaper. Since *The Daily Telegraph* has a more conservative readership, it might be exalting (traditional) family life and motherhood as very positive values with potential ideological orientations. This shows how news values usage is highly charged with ideological and social implications both in terms of representation (how a particular view is represented in discourse) and legitimisation of certain ways of life over others (traditional conservative family life over other ways that are not mentioned in this text).

7.4. Impact

Three discursive strategies are adopted to convey ‘impact’. The first one is the straightforward mentioning of the ‘effects and impacts’ of Brexit. The second is ‘prediction’, which includes statements with high certainty about what will happen after Brexit. The third strategy includes a more speculative aspect, that is to say, conjecturing about the possible future effects of Brexit, mostly using modals. Some parts of this news value usage actually parallel with ‘negativity’ through uncertainty, but not strongly enough to be categorised as constructing risk and uncertainty in most cases. In other words, the main difference between prediction and ‘speculation’ is the degree of certainty expressed in the reporting on the consequences or repercussions of Brexit, as shown in

Table 9. Once again, a noteworthy finding is how the coverage of the three pro-remain newspapers is similar in this aspect, irrespective of whether they are left or right wing. *The Daily Telegraph*, however, consistently downplays the impacts of Brexit by avoiding the use of such news value in its discourse.

Strategies	<i>The Guardian</i>	<i>The Independent</i>	<i>The Times</i>	<i>The Daily Telegraph</i>
Effects and impacts	<i>Effects</i> <i>Effect</i> <i>Impact</i> <i>Affect</i>	<i>Affect</i> <i>Cause</i> <i>Impact</i> <i>Make</i>	<i>Factor</i> <i>Impact</i> <i>Effect</i>	<i>Impact</i>
Prediction	<i>Predicts</i> <i>Prospect</i> <i>Mean</i> <i>Bring</i>	<i>Will</i> <i>Following</i> <i>after</i>	<i>Happens</i> <i>Mean</i> <i>Put</i>	
Speculation (modals)	<i>Could</i> <i>Might</i>	<i>Could</i> <i>Would</i> <i>If</i> <i>How</i>	<i>Would</i> <i>Could</i> <i>Might</i>	

Table 9: Linguistic pointers of ‘impact’ for Brexit across strategies and newspapers

As can be seen in the discursive strategies and pointers, and as observed previous examples, ‘impact’ is highly intertwined with ‘negativity’ in the corpus. In excerpt (5), an example from the pro-leave *Telegraph*, it may be noticed that the outlet, apart from quantitatively downplaying ‘impact’, sometimes attempts to mitigate the negative impacts of Brexit.

Excerpt 5: *The Daily Telegraph* (London), June 16, 2016, Thursday, Edition 1; National Edition
Most independent economic studies suggest Brexit **would have a short-term impact** on economic growth.

The use of *short-term* actually mitigates the degree of ‘negativity’ that is constructed in this instance, especially shifting the modality from deontic to intrinsic (Schulze and Hohaus 2020) and using hypothetical modal *would* instead of *will*. In addition, ‘eliteness’ is also used in a two-fold manner in the example, which presents a highly ideological use of this news value. Both of the pointers, *Osborne* and *Most independent economic studies*, indicate ‘eliteness’ in this piece but in different manners. The outlet indeed distances its discourse from Osborne’s position (who was a staunch Remainer during the campaign) by citing and associating itself with other sources of authority (*economic studies*) that refute his position, also because it is in the fact-checking section which implies that Osborne’s views are not really based on facts.

7.5 Eliteness

The construction of ‘eliteness’ includes three main discursive strategies, as shown in Table 10.

Strategies	<i>The Guardian</i>	<i>The Independent</i>	<i>The Times</i>	<i>The Daily Telegraph</i>
Proper names	<i>Carney</i> <i>George</i> <i>Osborne</i> <i>Johnson</i>		<i>Hammond</i> <i>Boris</i>	
Authority roles (social deixis)	<i>Queen</i> <i>Ministers</i> <i>Economists</i>		<i>Ministers</i>	<i>Economists</i>
Support and endorsement	<i>Backs</i> <i>Backing</i> <i>Supporting</i> <i>Thinks</i> <i>Backed</i> <i>Leading</i>	<i>Backs</i> <i>Backing</i> <i>Support</i> <i>Says</i> <i>Say</i>	<i>Backs</i> <i>Supporting</i> <i>Leading</i>	<i>Backing</i> <i>Back</i>

Table 10: Linguistic pointers for ‘eliteness’ across strategies and newspapers

Some of the previous excerpts illustrated the way ‘eliteness’ is used synergistically with other news values to construct newsworthiness. Here, however, I focus on another layer of discursive practices in relation to the news value of ‘eliteness’: on how these newspapers construct a for/against position towards Brexit through attributed discourse and, thus, using ‘eliteness’. ‘Eliteness’ is generally used in two ways. In some cases, it is employed to construct support and endorsement of the intended positions in discourse. In other instances, it is adopted to distance from certain opinions by quoting an external source. Overall, as seen before in the distribution of news values, *The Guardian* especially emphasises ‘eliteness’ to construct Brexit. However, the case of the *Leave-backing Telegraph* seems to be more interesting. *The Daily Telegraph*, in general, uses ‘eliteness’ quite frequently in quantitative terms but it specifically tends to stay away from leaders and prominent figures, and the only times they construct ‘eliteness’ is either by quoting economists in general as a source of authority or showing the support and endorsements through the *back* and *backing*. This could show how the pro-leave side coverage tends to be consistent with the properties of populist discourse, in this case, by staying away from the elite as much as possible. This is illustrated in excerpt (6), taken from *The Daily Telegraph*:

Excerpt 6: *The Daily Telegraph* (London), February 23, 2016, Tuesday, Edition 2; National Edition

[One in three Tory MPs confirm they will be **backing Brexit**]

Downing Street had thought fewer than 80 Conservative MPs **would back a Brexit**, but many appear to have been emboldened by the decisions of Boris Johnson and Michael Gove, the Justice Secretary, to vote to leave.

The news story about the vote intention plays around with the notion of endorsement by a number of MPs (MPs would back a Brexit), and even names (*Boris Johnson* and *Michael Gove*). Therefore, the newsworthiness of Brexit, in this case, is constructed around ‘eliteness’ in a specific way. On a deeper level, such an endorsement actually gives voice to an allegedly neglected group of MPs that now, following two prominent MPs, *have been emboldened* and dare to speak out. The ways ‘eliteness’ is used to associate with an ideological stance or otherwise distance from it in other outlets has been shown in various previous examples throughout this section.

8. CONCLUDING REMARKS

In this paper, by following one of the research lines suggested by Bednarek and Caple (2017), I have aimed at applying DNVA to a different news environment and exploring potential developments in the model. The application of a combined method in this study has some implications worth mentioning. A combined cluster-frequency analysis helped in the identification of the principal semantic fields covered in the sizeable corpus under scrutiny. Based on the results achieved in the present study, the procedure seems to be working in a reliable manner. The cluster analysis tool provided the research with a statistically valid apparatus that helped extract the main topics covered by the British press during the referendum, organised into clear-cut semantic fields and their constituent search terms. These tools provided a solid structure for analysing a sizeable corpus of about eight million words. That was not viable only by manual analysis. On the other hand, it helped in identifying the appropriate search terms, which can be considered as linguistic pointers for exploring the ways in which news values were used in the discourse to create certain representations and construct particular discourses around the topics being covered. Furthermore, the adopted statistical codes provided the analysis with an additional layer of information that, combined with tools of qualitative analysis in DNVA, made it possible to detect noteworthy discursive practices across the four data sets, with considerable ideological and political implications. All this showed that DNVA,

especially when combined with well-designed CL tools, can be a powerful analytical tool for detecting discourses in the coverage of crucial socio-political events in the press, such as the Brexit referendum. It was already shown that the patterns by which news values are used are a beneficial tool in mapping the cultural and ideological discourses around certain topics (cf. Fruttaldo and Venuti 2017; Venuti and Fruttaldo 2019; Maruenda-Bataller 2021). However, the proposed tools facilitated the investigation of ideological discourses by offering a statistically reliable picture of the variations in patterns of use of news values both quantitatively and qualitatively.

Nevertheless, there also are challenges of applying DNVA to a large corpus. The most salient of them is the possibility of overlap between the categories of news values and the subtleties and indirect ways in which news values are used in many cases (Potts *et al.* 2015; Maruenda-Bataller 2021). It must be admitted that the offered statistical tools do not address such challenges thoroughly. In this regard, some points should be taken into consideration.

First, the importance of concordance analysis in categorising news values should be highlighted. Potts *et al.* (2015) show that DNVA could be further developed by taking advantage of a complementary framework in which coding collocations based on the context is possible (or more straightforward). However, both Potts *et al.* (2015) and subsequent research (Maruenda-Bataller 2021) also admit the challenges of such a combination. Practically, as observed in many examples in our data, a single word or expression can convey different news values based on the context in which it is used. This goes further than the models and references already proposed. For example, Potts *et al.*'s (2015) suggestions are mainly based on the supplementary categorisation of linguistic resources, according to additional information and tags such as part of speech. In addition, Maruenda-Bataller (2021) proposes populating the DNVA model with further linguistic devices. However, in many cases, even such additional clues do not work or are not sufficient, and the only way to decide how to code a specific word is to check the concordance lines directly. Concordance analysis showed great potential in previous research on the discourse of news values (cf. Fuster-Márquez and Gregori-Signes 2019). In the case of the present analysis, consulting concordance lines before coding news values helped to avoid many possible misinterpretations and miscategorisation of collocations. This is a manual and time-consuming process, but one that resolves many problems in the coding phase, which is part of the qualitative component of the procedure.

Second, the issue of news values co-occurrence and intensity should be considered: one aspect that should be brought into attention is that the construction of newsworthiness does not take place by adopting single, independent, or isolated news values. On the contrary, it is a large-scale and contextual discursive practice. Delving into the news value distribution in large corpora could show us how each news value is used discursively and how news values can be used synergistically to create a specific bigger picture. This specific discursive practice should be considered in news values analysis, especially because of the effect they have on what could be called the intensity of news value usage. In this paper, I tried to quantify the frequency of the appearance of different news values in the discourse, which would let us compare different sub-corpora for cross-ideological analysis. However, the point that we should be cautious about is that the discursive practices in constructing newsworthiness are much more complex. Using news values is a multi-layered discursive practice. Therefore, the intensity of news values is as important as their frequency; offering quantitative ways of measuring this aspect in the discourse of news values is a task that is yet to be done (if possible at all).

Finally, it should be noted that the findings of this study once again underline the highly cultural, ideological, and interpretive nature of newsworthiness construction in discourse. As illustrated in excerpts (1)–(6), the ideological agenda and inclination of the newspaper and the interpretive processes of its readership can determine the ways in which newsworthiness is constructed in the discourse. This is highly related to Bell's (1984) audience design and the basic premise of how sociolinguistic variation in news text can be explained as a strategy to accommodate different target audiences. This point is not nuanced and was addressed in previous research (cf. Bednarek and Caple 2014; Fuster-Márquez and Gregori-Signes 2019; Makki 2019, 2020; and Maruenda-Bataller 2021). Nevertheless, in the case of this study and in a topic with profound socio-political and ideological challenges, it once again proved to be vital in analysing the discourse of news values.

REFERENCES

- Anthony, Lawrence. 2014. *AntConc* (Version 3.4.4). Tokyo: Waseda University.
<https://www.laurenceanthony.net/software/antconc/>
- Baker, Paul, Costas Gabrielatos, Majid Khosravini, Michał Krzyżanowski, Tony McEnery and Ruth Wodak. 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.
- Bednarek, Monika. 2016. Investigating evaluation and news values in news items that are shared through social media. *Corpora* 11/2: 227–257.
- Bednarek, Monika and Helen Caple. 2014. Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society* 25/2: 135–158.
- Bednarek, Monika and Helen Caple. 2017. *The Discourse of News Values: How News Organisations Create Newsworthiness*. New York: Oxford University Press.
- Bell, Allan. 1984. Language style as audience design. *Language in Society* 13/2: 145–204.
- Bell, Allan. 1991. *The Language of News Media*. Oxford: Blackwell.
- Bell, Philip. 1997. New values, race and ‘The Hanson Debate’ in Australian media. *Asia Pacific Media Educator* 1/2: 4: 38–47.
- Berger, Peter and Thomas Luckmann. 1967. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Michigan: Anchor books.
- Biber, Douglas, Geoffrey Leech and Stig Johansson. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Caple, Helen and Monika Bednarek. 2013. *Delving into the Discourse: Approaches to News Values in Journalism Studies*. Oxford: Reuters Institute for the Study of Journalism.
- Caple, Helen and Monika Bednarek. 2016. Rethinking news values: What a discursive approach can tell us about the construction of news discourse and news photography. *Journalism* 17/4: 435–455.
- Feinerer, Ingo and Kurt Hornik. 2018. *Ttm: Text Mining Package*. R package version 0.7–6. <https://cran.r-project.org/web/packages/tm/index.html>.
- Firth, Raymond. 1967. Ritual and drama in Malay spirit mediumship. *Comparative Studies in Society and History* 9/2: 190–207.
- Fowler, Roger. 1991. *Language in the News: Discourse and Ideology in the Press*. London: Routledge.
- Fruttaldo, Antonio and Marco Venuti. 2017. A cross-cultural discursive approach to news values in the press in the US, the UK and Italy: The case of the supreme court ruling on same-sex marriage. *ESP Across Cultures* 14: 81–97.
- Fuster-Márquez, Miguel and Carmen Gregori-Signes. 2019. La construcción discursiva del turismo en la prensa española (verano de 2017). *Discurso y Sociedad* 13/2: 195–224.
- Galtung, Johan and Mari H. Ruge. 1965. The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research* 2/1: 64–90.
- Greaves, Chris and Martin Warren. 2010. What can a corpus tell us about multi-word units? In Anne O’Keeffe and Michael McCarthy eds. *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 204–220.
- Gries, Stefan Th. 2009. What is corpus linguistics? *Language and Linguistics Compass* 3/5: 1225–1241.

- Halliday, Michael A. K. 1966. Some notes on 'deep' grammar. *Journal of Linguistics* 2/1: 57–67.
- Harcup, Tony and Deirdre O'Neill. 2001. What is news? Galtung and Ruge revisited. *Journalism Studies* 2/2: 261–280.
- He, Juan and Helen Caple. 2020. Why the fruit picker smiles in an anti-corruption story: Analysing evaluative clash and news value construction in online news discourse. *Discourse, Context & Media* 35: 100387. <https://doi.org/10.1016/j.dcm.2020.100387>
- Huan, Changpeng. 2016. Leaders or readers, whom to please? News values in the transition of the Chinese press. *Discourse, Context & Media* 13: 114–121.
- Lecheler, Sophie, Linda Bos and Rens Vliegenthart. 2015. The mediating role of emotions: News framing effects on opinions about immigration. *Journalism & Mass Communication Quarterly* 92/4: 812–838.
- Liu, Jiawei, ByungGu Lee, Douglas McLeod and Hyesun Choung. 2019. Effects of frame repetition through cues in the online environment. *Mass Communication and Society* 22/4: 447–465.
- López-Rodríguez, Clara Ines. 2022. Emotion at the end of life: Semantic annotation and key domains in a pilot study audiovisual corpus. *Lingua* 277: 103401. <https://doi.org/10.1016/j.lingua.2022.103401>
- Lorenzo-Dus, Nuria and Philippa Smith. 2018. The visual construction of political crises. In Marianna Patrona ed. *Crisis and the Media: Narratives of Crisis across Cultural Settings and Media Genres*. Amsterdam: John Benjamins, 76–151.
- Makki, Mohammad. 2019. Discursive news values analysis of Iranian crime news reports: Perspectives from the culture. *Discourse & Communication* 13/4: 437–460.
- Makki, Mohammad. 2020. The role of culture in the construction of news values: A discourse analysis of Iranian hard news reports. *Journal of Multicultural Discourses* 15/3: 308–324.
- Maruenda-Bataller, Sergio. 2021. The role of news values in the discursive construction of the female victim in media outlets: A comparative study. In Miguel Fuster-Márquez, José Santaemilia, Carmen Gregori-Signes and Paula Rodríguez-Abrunheiras eds. 2021. *Exploring Discourse and Ideology through Corpora*. Bern: Peter Lang, 141–165.
- Moisl, Hermann. 2015. *Cluster Analysis for Corpus Linguistics*. Berlin: Walter de Gruyter.
- Molek-Kozakowska, Katarzyna. 2017. Communicating environmental science beyond academia: Stylistic patterns of newsworthiness in popular science journalism. *Discourse & Communication* 11/1: 69–88.
- Molek-Kozakowska, Katarzyna. 2018. Popularity-driven science journalism and climate change: A critical discourse analysis of the unsaid. *Discourse, Context & Media* 21: 73–81.
- O'Keeffe, Anne, Michael McCarthy and Ronald Carter. 2007. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.
- Palmer, Jerry. 2000. *Spinning into Control: News Values and Source Strategies*. London: Leicester University Press.
- Potts, Amanda, Monika Bednarek and Helen Caple. 2015. How can computer-based methods help researchers to investigate news values in large datasets? A corpus linguistic study of the construction of newsworthiness in the reporting on hurricane Katrina. *Discourse & Communication* 9/2: 149–172.

- Qian, Lu. 2017. Cluster analysis for corpus linguistics. *Journal of Quantitative Linguistics* 24: 245–248.
- Saraçlı, Sinan, Nurhan Doğan and İsmet Doğan. 2013. Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications* 1: 1–8.
- Schultz, Ida. 2007. The journalistic gut feeling: Journalistic *doxa*, news habitus and orthodox news values. *Journalism Practice* 1/2: 190–207.
- Schulze, Rainer and Pascal Hohaus eds. 2020. *Re-Assessing Modalising Expressions*. Amsterdam: John Benjamins.
- Scott, Mike and Christopher Tribble. 2006. *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.
- Sinclair, John M. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/>
- Van Dijk, Teun A. 1988. *News as Discourse*. Hillsdale: Erlbaum.
- Venuti, Marco and Antonio Fruttaldo. 2019. Contrasting news values in newspaper articles and social media: A discursive approach to the US Ruling on same-sex marriage. In Barbara Lewandowska-Tomaszczyk ed. *Contacts and Contrasts in Cultures and Languages*. Cham: Springer, 147–161.
- Walsh, Steve, Anne O’Keeffe and Michael McCarthy 2008. Post-colonialism, multiculturalism, structuralism, feminism, post-modernism and so on and so forth. In Annelie Ädel and Randi Reppen eds. *Corpora and Discourse: The Challenges of Different Settings*. Amsterdam: John Benjamins, 9–31.
- Westerståhl, Jörgen and Folke Johansson. 1994. Foreign news: News values and ideologies. *European Journal of Communication* 9: 71–89.

Corresponding author

Arash Javadinejad
 Universidad Católica de Valencia
 Faculty of Educational Sciences
 C/ Sagrado Corazón, 5.
 Godella 46110
 València
 Spain
 E-mail: arash.Javadinejad@ucv.es

received: December 2022
 accepted: April 2023