ARTÍCULOS

Círculo de Lingüística Aplicada a la Comunicación
ISSN: 1576-4737

EDICIONES COMPLUTENSE

https://dx.doi.org/10.5209/clac.72672

The Europe of Brexit: a corpus-assisted discourse study of identities in the press

Carmen Pena-Díaz[1]; Mª del Mar Sánchez Ramos[2]

Recibido: 25 de noviembre de 2020 / Aceptado: 20 de julio de 2021

Abstract. Drawing on what is known as a corpus-assisted discourse study (CADS) approach (Baker et al., 2008), this article will research the construction of different identities by means of the language used in two newspaper articles on Brexit from the Spanish *El País* and the British *The Guardian*, to examine how these identities are constructed through media discourse at the time following the Brexit referendum (2016-2018). Media discourse surrounding Brexit is examined under the consideration of media power. A comparable corpus made up of original newspaper articles about Brexit was used to carry out the analysis, identifying statistically significant keywords compared with a reference corpus with the aim of providing an example of how the British and Spanish press construct identity.
Keywords: corpus-assisted discourse analysis, media discourse, identities, Brexit, keyness.

[es] La Europa del Brexit: un estudio del Discurso Asistido por Corpus sobre las identidades en la prensa

Resumen. Partiendo de la perspectiva de los estudios del discurso asistido por corpus (Baker et al., 2008), este artículo investigará la construcción de diferentes identidades mediante el lenguaje utilizado sobre el Brexit en el diario español *El País* y el británico *The Guardian* para así examinar cómo se construyen dichas identidades a través del discurso mediático en el momento posterior al referéndum del Brexit (2016-2018). El discurso mediático en torno al Brexit se examina bajo el prisma del poder mediático. Para llevar a cabo el análisis se ha utilizado un corpus comparable formado por artículos periodísticos originales sobre el Brexit, identificando palabras clave estadísticamente significativas en comparación con un corpus de referencia con el objetivo de proporcionar un ejemplo de cómo la prensa británica y española construyen la identidad alrededor de un mismo tema.
Palabras clave: análisis del discurso asistido por corpus, discurso mediático, identidad, Brexit, *keyness*.

Cómo citar: Pena-Díaz, Carmen; Sánchez Ramos, Mª del Mar (2021). The Europe of Brexit: a corpus-assisted discourse study of identities in the press. *Círculo de Lingüística Aplicada a la Comunicación* 88, 225-234, https://dx.doi.org/10.5209/clac.72672

1. Introduction

The UK voted to exit the European Union in June 2016 by a very close outcome. The nation was dramatically divided into those who had voted to leave the EU and those who wanted to remain, giving way to quite distinctive points of view. Since then, there have been many different political stages and great controversy has arisen from all sides, thus many discourses from different political perspectives which, in turn, are reflected in the media coverage (Lutzky & Kehoe, 2019).

The term "Brexit" is the combination of "Britain" and "exit" and represents Britain leaving the EU, therefore constituting a national unit in relation to the European Union as a different unit: "us" versus "others". The latter, however, did not only reflect the EU but also inside national differences, thus "others" referring to the UK vs. the EU but also the "pro-EU" vs "pro Brexit". Back in 2016, many politicians were not fully aware that the Brexit vote would end up reflecting multiple aspects of modern Britain which were interacting with each other and that society was not yet conscious of. After the extensive analysis produced since then, in-depth studies have shown that the Brexit vote was about pluralism in culture, nationalism, and citizenship. On the one hand, those who attempt to build a 'multicultural' British national identity and who embrace being part of a wider socio-political area such as the EU have encountered resistance from those who see Britain under threat, as their understanding of the construction of a British

[1] Universidad de Alcalá. Correo electrónico: carmen.pena@uah.es; https://orcid.org/0000-0002-2329-756X
[2] Universidad de Alcalá. Correo electrónico: mar.sanchezr@uah.es; https://orcid.org/0000-0001-5684-1552

identity is very different to their view of a continental European one and they consider themselves as a separate entity in their geographical area. For the latter, the empowerment of the EU in Britain means a loss of their British identity. Pluralism was thus central in the vote, as it reflected conflicts over the effects of multiculturalism, between different views of national identity, and also between the different economic and political assessments regarding foreign citizenships. In this paper the discursive construction of Brexit with reference to the EU will be analyzed with the purpose of studying the role of media in the construction of identity.

The media (e.g. the written or digital press) are often analysed through the prism of discourse analysis (DA), and more specifically through approaches known as critical discourse analysis (CDA). These are qualitative approaches which clearly provide information about the symbolic load that the press constructs, while the press also reproduces ideologies among the dominant groups. This is the main reason why these qualitative approaches make increasing use of quantitative methodologies, such as corpus linguistics (CL), which analyses large volumes of texts with tools that provide statistical information about the frequency of certain words or lexical and grammatical patterns. Drawing on a corpus assisted discourse approach (CADS) approach (Baker et al., 2008), and based on a comparable corpus of English and Spanish newspapers articles, our work will research how language is used by the press to construct different identities. The comparable corpora draw on different articles on Brexit from two daily liberal, left-of-centre tendencies and pro-European newspapers, the Spanish *El País* and the British *The Guardian*. The aim of this paper is to examine how identities were constructed through media discourse at the time following Brexit (2016-2018), when leaders were deciding how to plan and organize life outside the EU through CADS and how this is represented in two newspapers with similar ideologies but different perspectives. The initial hypothesis was that the main focus which constitutes the "other" (i.e. exiting the EU) will be preserved in both newspapers and that the UK will be portrayed as a part of the EU whereas the "others" will refer to Brexiteers. However, the starting point for each media (British identity for *The Guardian* and European identity for *El País*) reflect different perspectives and, in fact, a national identity with a common "us" in the UK (versus the "others" referring to the EU) in the British press sheds light into how even the British pro-European left leaning press provided elements for the Brexit cause.

2. Theoretical framework

2.1. Political discourse and the media

Political discourse is the language used by institutions to communicate with other institutions or with society in which language is used as the means to transfer political action. This political action is a struggle for power from different ideologies and social spectres to try to resolve conflicts of interest and persuade others and the basis of this persuasion lies in speech (Partington & Taylor, 2018). Different political actors have different views and thus language varies accordingly to their perspective. Communication is, therefore, never neutral, as it is manipulated in order to persuade and interfere in favour of a certain point of view or ideology.

News media focuses on delivering information to the public and reflects and reproduces power structures within a society (Fairclough, 1995), acting as a barometer of public opinion. Although this information is apparently objective and non-biased, it is the carrier of ideological discourses which is used as a power tool. As Baker & McEnery stated (2005, p. 200) "newspapers, which are widely read on a daily basis, can help to shape/reflect public opinion". Media thus produces an ideological vision of the world both in its original form but also when it is translated or edited in another language.

Since the studies carried out by the Frankfurt School in modern society, the role of the press is to create a so-called 'public opinion' by transmitting or amplifying the opinion generated by public debate, but media has also often become a space of 'representative publicness' or 'representative publicity' (Habermas 1989, p. 46), a 'public' sphere that generates social and cultural values, serving thus as a transmitter and generator of public debate. This double function leads us to propose another level of analysis for, as a generator of public debate, the different types of discourse used in the media will imply a significantly different perspective from being a mere transmitter of public debate to acting as the generator of that same public debate in other instances. If we take into account that what we are analyzing in this article are national News Agencies, that is, sources of management and distribution of news that, beyond the transmission of those news, also produce an implicit and inevitable edition and selection of the published news, we could say that every national News Agency is, by definition, an institution in charge of generating and representing national public opinion.

This approach takes us to the old question of whether any intermediary in the transmission of a message (as happens, for example, with translation) is only a means or if, in fact, alters that message, as Marshall McLuhan explained back in 1962 with his theory on the 'global village' and the world becoming interconnected. With this modern concept of globalisation, back in mid 1900s, individual nations seemed to be about to be taken over by a larger and more complex entity and the media was thought to become an instrument to interconnect different public opinions. However, as we have seen lately, individual national identities have taken on a stronger position and nations have not let go of their sovereignty even when they belong to greater geographic and political entities, such as the European Union, for example. Media has also localized to offer these different perspectives in a global context, but it is not clear that the perspective is global referring to an international perspective or global referring to the localization in perspective to other states, geographical areas, etc.

2.2. National identity

National identity can be defined in different ways, amongst which is the sense of a nation as a cohesive whole, represented by different traditions, culture and language. It is also a mechanism for giving people a sense of individual and collective worth, providing a strong sense of belonging, beliefs, values, assumptions and expectations shared by group members. Fox & Miller-Idriss (2008) describe the concept of 'nation' not simply as the product of macro-structural forces, but, simultaneously, as the practical accomplishment of ordinary people engaging in routine activities (2008, p. 537).

As Mole (2007) explains, "identity seeks to convey who we are or are perceived to be and the way we, as individual or groups, locate ourselves and others in the social world", thus categorizing the individual or group in relation to other individuals or groups and therefore focusing on the differences between them. According to Hall & Du Gay (1996, pp. 4-5) identities are constructed through "the relation to the Other, the relation to what it is not".

The construction of a national identity implies building a coherent, linear and collective narrative that responds to the basic elements of every *bildungsroman* and that makes it translatable, comparable to other narratives under dispute for the same sphere of the construction of identity. As White (1980, p. 1) points out 'we may not be able to fully comprehend specific thought patterns of another culture, but we have relatively less difficulty understanding a story coming from another culture'. As Bhabha (1994, pp. 139-140) states in his famous *The location of culture*, in every national narrative construction there is a metaphorical process in which the sense of home and belonging to the imaginary community of the people or the nation is transferred. News agencies (the national ones) play an important metaphorization and transfer function that materializes in the very fact of the transmission of informative pieces and, much more, in the translation of all the news that may feel installed in the very center of the identity debate. In the cases studied, it may not be so much about translations from one language to another, as about translations of what would seem not to require translation: the narrative discourse itself. We might think that we find two discourses, with two narratives that run in very different senses (the British and the European) and that, confronted within the translation space, show their differences. Two different narratives, two different metaphorization processes that point, however, to the same scheme of common representativeness, to the same debate that disputes the presence and dominance of public space derived from that same debate.

The construction of any narrative that claims to have some validity in the construction of identity needs to be projected in the public space. What's more, it has to mold 'public opinion' so that the 'we' is not empty of content. That is one of the main practical functions that the management of any National News Agency will have, much busier in the creation and representation of a coherent discourse that materializes the national project, than in the mere objective dissemination of events occurred in the planet. Much more in the case of issues that directly affect the heart of that project, such as Brexit. This will always start from the artificial articulation of a 'we' facing a 'them', two pronouns that in their definition will mark the national boundaries both geographically and psychologically and stress 'intra-national sameness' (Wodak et al., 2009, p. 37).

Both narration and representation lead to the creation of a space of supposed consensus in which the exclusion or belonging of every individual to a model, to a heritage that, in the words of Guillory (1993, p. 40) is delimited and is nothing other than cultural capital. News Agencies are important institutional instruments that contribute to the creation and maintenance of such cultural capital, as much or more than schools, academies, publishers or other areas commonly associated with the public cultural space.

2.3. From corpus to discourse: a corpus-assisted discourse approach

Qualitative approaches such as CDA (Fairclough, 1989, 1995; Van Dijk, 1998, 2001; Wodak, 2001) have quickly spread as methods of analysis in recent years. The origin of CDA can be traced back to Fairclough himself and his work in the research group *Language, power and ideology* (Lancaster University, United Kingdom), while his work *Language and power* (1989) is the cornerstone on which the study of language from a critical perspective is built (Blommaert & Bulcaen, 2000, p. 454). CDA serves as a text analysis tool in order 'to investigate critical social inequality as it is expressed, signaled, constituted, legitimized and so by language use (or in discourse)' (Wodjak, 2001, p. 2). In other words, CDA offers a systematic approach to understand the different relations between language, power and identity: 'the significance of language in relations of power, as well as in resistance to abusive power and struggles for power, in one issue. Another is language aspects of the emergence, enactment and change of social and personal identities, including national identities and gender identities' (Fairclough, 2009, p. 520). Baker et al. (2013) have also acknowledged the benefits of using CL in CDA as CL allows to process large quantitites of data and identify salient discursive patterns. These authors advocate that CL is based on a descriptive study characterized by its objectivity and based on data, which leads to reliable and generalizable results, whereas CDA is an eminently qualitative approach grounded in a theory-driven framework and relying on samples selected subjectively. This is why the combination of CL and CDA can provide a line of investigation into a variety of social issues, such as ideology, political conflicts, immigrants and refugees or ethnic minorities, which is known as corpus-assisted discourse studies (CADS).

CADS originates from the work published by Partington (2004), in which the author draws on the research that was being undertaken in the context of DA and which employed CL techniques. As Jaworska states (2016, p. 157): 'insights derived from corpus research have increased our understanding of language use by providing empirical evidence for the existence of regularities and patterns that are not immediately visible to the naked eye'. As suggested by Liu & Zhang (2018, p. 2), CL will allow to process large quantities of data 'by focusing on frequency and saliency, and thus enhance the credibility of analytic results'. Corpus analysis tools usually provide quantitative and statistical information about word frequency, wordlists, keywords, concordances and collocations. First, frequency and wordlists provide information about the words used most in the corpus. Thus, with respect to the most frequent words, certain patterns in the regularity of their appearance in the text may be observed. The wordlists are 'lists of corpus words in alphabetical or frequency order, upon which further data (such as keywords) may be generated' (Calzada Pérez, 2017, p. 8). Next, the keywords are the result of statistical comparison (e.g. log-likelihood or chi-square tests) between the terms in the wordlist and a reference corpus (Scott, 2010; Scott & Trible, 2006), and they are usually words that denote the specificity of a corpus. Keyword analysis has been the focus of different studies on British press (Jeffreys & Walker, 2017) and the main methodological approach to content analysis of Spanish political speeches (Duque, 2014). It is also important to highlight that keyword analysis can present some limitations, as suggested by the recent study by Egbert & Biber (2019). Thirdly, there are the concordances, through which words or groups of words selected in context can be studied. The word searched appears in the middle and is known as a key word in context (KWIC). Finally, there is the function for analysis of collocations, which have been defined by Stubss (2001, p. 35) as elements that 'are also widely shared within a speech community', and also are associated to 'nodes around which ideological battles are fought'.

Therefore, CADS-based methodology uses quantitative methods drawn from CL as referred to earlier, which tie together with the qualitative techniques employed in CDA. In this way, together with the data provided by the corpus analysis tools, such as wordlists or word frequency, a more qualitative reading of the texts can be monitored, with the aim of explaining the reasons behind these statistics. This explanation is frequently examined taking close account of the social, political and historical context (Partington, 2014). The use of CADS has yielded results in the field of political discourse (Aluthman, 2018), the construction of nationalism and identity (Jaworska, 2016; Vessey, 2013, 2016), and discourse analysis in the press (Branun & Charteris-Black, 2015; Lutzky & Kehoe, 2019; Sánchez Ramos, in press/2022; Pena Díaz, in press/2022), in addition to the depiction of immigrants or refugees (Baker et. al. 2013; Cap, 2019; Taylor 2014).

Although our study might be labelled as following the 'critical' tradition due to the ideological component of media when trying to identify identities, we follow Partington et al. (2013) and more recent studies as Glaas (2016). Our main purpose is not focused on social justice or the representation of marginalized groups, but on the definition of DA as the discipline that 'studies how language is used to (attempt to) influence the beliefs and behavior of other people (Partington et al., 2013: 5). In fact, we advocate for CADS as 'the use of corpus linguistics methodology in discourse analysis studies offers considerable benefits in three ways; in the reduction of the impact of researcher bias, through quantification of the 'incremental effect of discourse' and in the identification of counterexamples showing 'resistant' or changing discourse' (Glaas, 2016, p. 11).

3. Methodology

Our research is based on the analysis of a comparable corpus (English-Spanish), which is made up of original newspaper articles —opinion and editorial articles— about Brexit published in *El País* (Spain) and *The Guardian* (United Kingdom) from 2016-2018, two daily liberals, left-of-centre tendencies newspapers. As reference corpora to carry out our analysis, we used the Bank of English (BoE) and CREA (Corpus de Referencia del Español Actual), which allowed us to identify keywords, that is, statistically significant words that are unusually frequent in our corpus if compared with the reference corpus. The comparable corpus consists of 6,141,731 words. The subcorpus *The Guardian* (EN) had 3,384,262 while 2,757,469 belonged to *El País* (ES). The corpus was compiled semi-automatically using Nexis database. Although the corpus was not compiled manually, the results obtained were carefully checked and revised in order to avoid inappropriate data such as additional metadata or irrelevant information.

There is a wide variety of corpus and concordance software (i.e., Wordsmith Tools, LancBox, Sketch Engine). We have chosen Laurence Anthony's AntConc as the software to study our comparable corpus. It is one of the most popular and user-friendly computer tools employed in CL. It allows to generate keywords, wordlists, clusters and concordances. This software also provides the statistical functions needed for our analysis, such as the frequency of words. AntConc calculates the 'keyness' by comparing the relative frequencies of words within two corpora (one of them the reference corpus) and carrying out statistical tests (log-likelihood) in order to determine whether words have occurred more or less often than expected by chance. Based on this frequency, we will be able to focus our qualitative analysis on the most repeated (frequent) words of the corpus.

A CADS analysis was carried out, which allowed us to triangulate quantitative and qualitative tools and generalize the results (Marchi & Taylor, 2009). AntConc helped us analyze our corpus and identify statistically significant keywords (quantitative analysis). On the other hand, our qualitative analysis is mainly based on the interpretation and description of the most prominent linguistic features, and it is grounded on the discourse methods employed in CDA.

4. Results

Two keywords lists result from the comparison of the terms of each of our corpus (*The Guardian* and *El País*) with the two references corpora used for our analysis, such as BoE and CREA respectively. The way in which words are grouped provides representations of socio-culturally significant concepts and help identify discursive patters. Examining and comparing keywords will 'reveal aspects of identity constructions' (Jaworska, 2016). Also, keywords 'may be shown to be indicative of the writer's position and identity, as well as the discourse community with its values and beliefs about the subject matter' (Bondi, 2001, p. 7).

After comparing each of our corpus with the two reference corpora, our keyword analysis shows that our comparable corpus have great similarity on the most significant topics covered. Following Baker (2010), the generated keywords were grouped into semantic categories based on their semantic meaning. In doing so, we obtained quite similar semantic categories ('Brexit, 'negotiation', 'referendum', 'consequences'), with the only exception of the word *crisis* related to Brexit, which is close to the problem of 'Gibraltar' in *El País*.

Table 1 shows the keywords identified in our comparable corpus. As can be observed, both subcorpora offer a similar number of words related to the referendum of Brexit. Among them, 'Brexit' is the mostly used keyword in both corpora. One interesting finding for our research is the keyword 'EU' (European Union) in both corpora, as we try to find out how identities are discursively constructed in both corpora in terms of a national or supranational identity. Another relevant result from the keyword list is the use of 'our' in *The Guardian*. Other words are related to the referendum, such as 'government', 'deal', 'leaving', 'negotiations'; and geographical regions, such as 'Britain', 'UK', or 'Gibraltar' for *El País*.

Table 1. Top 25 keywords in the Spanish and English corpora.

Rank	Freq.	Keyness		Rank	Freq.	Keyness	
1	43245	418.61	Brexit	1	48707	425.61	Brexit
2	39255	320.52	unido	2	31851	265.42	EU
3	31308	332.43	EU	3	26856	222.02	UK
4	29254	317.86	reino	4	25915	205.18	time
5	25877	290.36	Unión	5	17567	141.79	government
6	19732	260.87	negociaciones	6	16424	124.21	deal
7	16693	240.86	Gibraltar	7	14941	119.25	European
8	15504	170.32	peñón	8	11650	91.13	union
9	13742	160.81	laboristas	9	11525	90.07	leaving
10	11464	140.44	acuerdo	10	10280	89.08	May
11	7830	135.97	permanencia	11	9651	83.62	our
12	4769	120.35	país	12	10675	82.84	labour
13	6423	104.76	referéndum	13	11391	81.9	block
14	5827	80.97	Bruselas	14	10010	77.19	Britain
15	3825	78.94	británico	15	10326	73.05	vote
16	5790	64.67	colonia	16	9790	68.61	trade
17	3860	60.81	salida	17	7573	65.58	inmigration
18	2407	58.34	Parlamento	18	8637	65.57	people
19	2399	50.5	Cameron	19	8627	65.49	leave
20	2398	48.56	europeo	20	8279	62.55	minister
21	2398	47.21	Corbyn	21	8257	62.36	party
22	2385	43.53	Escocia	22	6707	58.07	exit
23	2384	40.12	divorcio	23	7301	54.32	referendum
24	2384	38.55	eurozona	24	5693	49.28	rights
25	2384	35.66	partido	25	6658	48.93	British

The next step was thus to examine words that were identified as being quantitatively important in the texts and interpreting them with social and background events which could have influenced the patterns found in the data. As our study tries to elicit information about the construction of identities during Brexit, the following sections give a close analysis of the keyword 'EU' in both corpora to examine in more detail if there are any relevant differences and if there is any type of discourse around it. Also, and as it also occupies a considerable high position in *The Guardian*, we decided to give a close analysis to the term 'our'.

As we mentioned above, our initial hypothesis was that the main focus which constitutes the 'other' (i.e. exiting the EU) will be preserved in both newspapers. However, the starting point for each media (British identity for *The Guardian* and European identity for *El País*) reflects different perspectives. *The Guardian* might have been thought to encourage a pro-European stand regarding Brexit, but in fact, from a transnational focus, it was emphasizing precisely the 'us' versus 'we' dichotomy, but from a national perspective, i.e. whereas we hypothesized that *The Guardian* would refer to the 'others' in the same way as *El País*, as the 'exit from the EU', it held a national perspective and thus referred to the EU as the others, creating a significant strangement from the supranation.

Although both corpora keep a common political perspective, both are left-of-centre and pro-European newspapers, a close look at the term 'EU' offers quite relevant results. It seems that the two newspapers differ in their preferences for naming the impact of Brexit in different social stages. A close examination of the corpus enables us to discover lexical patters and offer several clues about the use and meanings of the 'EU' term.

Regarding *The Guardian*, 'EU' is constructed through the idea of challenge and uncertain change. Throughout the corpus, concepts are positive towards the EU, however, the lexical items surrounding 'EU' are very frequently represented in a negative sense using words such as 'depart', 'leave', 'outside', 'exit'. Indeed, it is also referred as an economic challenge:

a) Investments (EFSI) triggered around (EURO) 76bn of investment in Europe. If the UK left the EU, access to this fund for financing purposes would be difficult if not impossible
b) Have long wanted British fish. They wanted to link UK access to the single market to EU access to British fishing waters. But they did not get that.

And although there are positive nominal constructions ('pro EU', 'all EU', 'remaining EU'), the keyword 'EU' is mainly characterized by being modified by words such as 'thorny', 'anti', 'leaving', 'another', 'their', 'difficult', 'incoherent', 'old', nullify', or 'quitting'.

c) EU officials wish to nullify Britain's power to retain membership of the single market without the EU's consent
d) […] government to expect two more thorny negotiations with the EU.
e) The real costs to British business run into many millions. Time has always been a variable concept in the corridors of Brussels power. The old EU practice of "stopping the clock" in negotiations is legendary.

The use of a considerable number of negative modifiers in *The Guardian* transform the potential pro-European attitude towards a more distant and inevitable negative one towards the EU (i.e, the use of 'the old EU'). A close collocation analysis reveals nouns and verbs related to the EU such as 'legislation', 'membership', 'regulation' or 'leader'. However, it is worth noting constructions such as, 'EU economic instability' or 'EU uncontrolled immigration', as an example of collocates that construct the EU negatively (Table 2). Regarding verbal collocations, EU is used as an object and pictured as an entity disempowering UK. Some examples are 'imposed', 'hold', 'leave, 'exit', 'blame', 'pay', 'quit' or 'take'. On the other hand, EU is also used as subject with the following verbs: 'want', 'have', 'impose', 'take'.

Table 2. Top 10 nouns and verbs modifiers of the term 'EE' in *The Guardian (L5-R5 occurrences)*.

Rank	Freq.	MI Score	Collocates
1	851	8.96539	have
2	728	8.79323	membership
3	708	7.90087	exit
4	652	7.68340	leaving
5	613	7.60339	quit
6	534	7.38101	economic
7	477	7.44462	inmigration
8	345	7.04419	vote
9	301	6.95209	legislation
10	267	5.33146	pay

By contrast, *El País* prefers to use more neutral nominal constructions, such as 'integración europea' [European integration], 'economía europea' [European economy], 'familia europea' [European family], 'comunidad europea' [European community], 'identidad europea' [European identity], 'institución europea' [European institution] or 'coherencia europea' [European coherence]. In all of these linguistic constructions the Spanish press is reluctant to label the EU in a negative way—in fact 'EU' is related to the meaning of family or community—, thus contributing to a pro-European identity and stressing UK as a nation inside that same community.

a) […] ni siquiera el Brexit es capaz de aminorar la querencia de Reino Unido por torpedear la integración europea
b) el Gobierno quiere evitar que este pequeño territorio de 30.000 habitantes [Gibraltar], con 55.000 empresas registradas, amplíe su ventaja competitiva cuando abandone la familia europea.

On the contrary, as Table 3 shows, a collocational analysis of the Spanish corpus shows nouns and verbs modified by 'EU' with a common feature of 'integration' into the community, which indicates a positive and constructive attitude towards the EU ('alcanzar' [to reach], 'inversion' [investment], 'regulación' [regulation], 'campaña' [campaign], 'permanecer' [to stay], 'ayudar' [help], 'votar' [to vote], 'reformar' [to reform], 'dar' [to give], 'necesitar' [to need]). Indeed, EU is portrayed as being more beneficial to the development of UK within this institution.

Table 3. Top 10 nouns and verbs modifiers of the term 'Eu' in *El País (L5-R5 occurrences)*.

Rank	Freq.	MI Score	Collocates
1	1851	9.90087	comisión
2	1245	9.72827	cumbre
3	1008	9.60339	alcanzar
4	1013	8.38101	reformar
5	952	7.44462	inversión
6	887	7.68340	votar
7	778	7.55209	permanecer
8	656	7.23324	necesitar
9	576	7.04419	alcanzar
10	487	1.93146	ayudar

As mentioned above, we consider also relevant to provide a close examination for the key term 'our' as it outstands as one of the main keywords in *The Guardian* (9,651 occurrences). Table 4 provides a cluster analysis, where 'our' collocates with words that denote identity, such as 'people', 'national', 'families', 'nation', 'citizens', 'government' or 'Prime minister'.

Table 4. 2-3n grams for 'our' in the English subcorpus.

Rank	Freq.		Rank	Freq.	
1	481	our country	11	65	Our democracy
2	328	our economy	12	64	Our trade
3	306	our future trading	13	63	Our people
4	192	our own people	14	60	Our interests
5	173	our Brexit	15	57	Our money
6	87	our politics	16	54	Our nation
7	89	Our national life	17	52	Our government
8	88	Our borders	18	49	Our membership
9	84	Our nation	19	48	Our children
10	77	Our laws	20	44	Our constitution

This finding is particularly interesting as the use of 'our' tries to add a human and personal dimension to the discourse and gives a human personification of the UK, a result also found in parliamentary debates when using 'we' (Wenzl, 2019), bringing it closer to readers, but once again shifting it further from the EU. The use of 'our' in such an intense way portrays the construction of a British identity as being opposed to a European one and moves towards transforming a pro-European identity and elevating the discussion of a nation as opposed to the supranational body.

5. Discussion and conclusions

Social constructed discourse distributes meaning through media channels and connects media discourse with public discourse (Johnson-Cartee 2005; Fowler 1991). In the case of this study, the media discourse surrounding Brexit is examined under the consideration of media power (Freedman 2014). The diverse nature of media discourse about Brexit becomes visible through the different ways of setting Brexit in relation to either the world, the EU or the UK.

As we have seen, texts exist in their social context and sociocultural background as creators of public opinion for specific audiences who identify with them. The relationships between the different participants involves differences in power which the present analysis confirms linguistic patterns in the media are determined by their target audience and texts adapt to their specific target society. Thus, the British media on Brexit have used a domesticating perspective, constructing Brexit as an eminently British issue which contributes to reinforce the remoteness of the EU, at the same time as it creates a loss of empowerment of the EU and increases national sovereignty.

The British and Spanish press describe different ideologies concerning sovereignty regarding the UE as has been proved by the CADS analysis, which has identified the dominant discourses in both corpora. There are limitations in our corpus, but it reveals the different constructions of British identity in the English discourse.

Brexit has meant not only a European crisis because of its economic repercussions - which are the ones that have mainly occupied the information space – but a redefinition and identity crisis. Weren't we European? Didn't we share a common legacy that is now in doubt? Hasn't the old definition of the 19th Century nation state not been surpassed? Have we not transferred sovereignty? Were we building a new identity discourse that would exceed the old borders? It seems that the immediate answer to all the previous questions has been a resounding no. National flags continue to weigh in the definition of the various communities that make up what we know as Europe. The recovery of old frontiers and values entails the recentralization of the debate, the return to the terms of theoretical discussion that shaped modernity and that seemed definitely surpassed.

Britain's withdrawal from the EU has taken us back to the traditional principles upon which the hegemonic centers of power in the West were built: the old national-colonial discourse. In a way, it is the last and most recent attempt to return to the center (did it ever exist?), to avoid the influence of margins in the definition of "we", of national identity. Brexit (and its spirit embodied in different ways in different contexts) is the maximum expression of fear of the dissolution of the center in a global context that presents itself as liquid, open to the contradiction of geography and time. We could establish the following corollary: in the face of the threat of plurality, the recovery of the narrative of unity, the pilgrimage to the hegemonic center. Something like the reverse trip to which Homi Bhabha posed when he spoke of the new reality constituted by the peoples and people who, banished, had populated the western capitals. What he called 'dissemi-nation' (1994, p. 139) playing with the idea of building a new national identity from individuals scattered by the various national diasporas.

The media (i.e the press) have a direct influence on the social representation of different sectors and the way that society today perceives reality. This research has sought to provide an example of how the British and Spanish press construct identity. Therefore, our work makes a contribution to other research into the extremely important role of media such as the digital press in the construction and representation of ideologies and identities. In turn, the research conducted serves as a starting point for future research that will help to consolidate the results obtained, for example, through data triangulation with other qualitative techniques, such as ethnographic methods. We are aware, however, of the limitations of our work, as it could be further expanded with a larger corpus including more newspapers from different ideologies which would throw deeper insights into media discourse and the distinct perspectives.

Acknowledgements

This research was carried out within the framework of the project Original, translated and interpreted representations of the refugee cris(e)s: methodological triangulation within corpus-based discourse studies (PID2019-108866RB-I00 / AEI / 10.13039/501100011033) funded by the Spanish Ministry of Science and Innovation.

References

Aluthman, E. (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. https://doi.org/10.24093/awej/vol9no4.2

Bhabha, H. (1994). *The location of culture*. London: Routledge.

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), 197-226. https://doi.org/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. https://doi.org/10.1177/0957926508088962

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

Baker, P., Costas, G., & McEnery, T. (2013). Sketching Muslims: a corpus driven analysis of representations around the word 'Muslim' in the British press 1998-2009. *Applied linguistics 34*(3), 255-278. https://doi.org/10.1093/applin/ams048

Blommaert, J., & Bulcaen, C. (2000). Critical discourse analysis. *Annual review of anthropology 29*(1), 447-466. https://doi.org/10.1146/annurev.anthro.29.1.447

Bondi, M. (2001). Perspectives on keywords and keyness. In M. Bondi & M. Scott (eds.), *Keyness in texts* (pp. 1-18). Amsterdam: John Benjamins Publishing. https://doi.org/10.1075/scl.41.01bon

Branun, J., and Charteris-Black, J. (2015) The Edward Snowden affair: a corpus study of the British press. *Discourse and communication 9*(2), 199-220. https://doi.org/10.1177/1750481314568544

Calzada Pérez, M. (2017). Corpus-based methods for comparative translation and interpreting studies: mapping differences and similarities with traditional and innovative tools. *Translation and Interpreting Studies 12*(2), 231-252. https://doi.org/10.1075/tis.12.2.03cal

Cap, P. (2019). 'Britain is full to bursting point': immigration themes in the Brexit discourse on the UK independence party. In . Koller, S. Kopf & M. Miglbauer (eds.), *Discourses of Brexit* (pp. 69-85-120). London: Routledge. https://doi.org/10.4324/9781351041867-5

Duque, E. (2014). Análisis de contenido mediante análisis de palabras clave. *Mediaciones sociales. Revista de ciencias sociales y de la comunicación*, *13*, 39-73.

Egbert, J., & Biber, D. (2019). Incorporating text dispersion into keyword analyses. *Corpora*, *14*(1), 77-104.

Fairclough, N. (1989). *Language and power*. Harlow: Pearson Education.

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

Fairclough, N. (2009). Language, reality and power. In J. Culpeper, F. Katamba, P. Kerswill, R. Wodak & T. McEnery. (eds.). *English Language: Description, Variation and Context* (pp. 512-522). London: Palgrave Macmillan.

Fox, J., & Miller-Idriss, C. (2008). The `here and now' of everyday nationhood. *Ethnicities 8*(4), 573-576. https://doi.org/10.1177/1468796808008004103

Fowler, R. (1991). Language in the news: discourse and ideology in the press. London: Routledge.

Freedman, D (2014). The contradictions of media power. London: Bloomsbury.

Glaas, S. (2016). *A corpus-assisted study of British newspaper discourse on the European Union and a potential membership referendum*. PhD Thesis, University of Birminghan, UK.

Guillory, J. (1993). *Cultural capital, the problem of literary canon formation*. Chicago and London: The University of Chicago Press. https://doi.org/10.7208/chicago/9780226310015.001.0001

Habermas, J. (1989). *The structural transformation of the public sphere*: *an inquiry into a category of bourgeois society*. Cambridge, Massachusetts: MIT Press.

Hall, S., & Du Gay, P. (1996) *Questions of cultural identity*. London: Sage Publications.

Jaworska, S. (2016). Using a corpus-assisted discourse studies (CADS) approach to investigate constructions of identities in media reporting surrounding mega sport events: the case of the London Olympics 2012. In I. Lamond & L. Platt (eds.), *Critical events studies: approaches to research* (pp. 149-174). London: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-52386-0_8

Jeffreys, L., & Walker, B. (2017). *Keywords in the press: the new labour years*. London: Bloomsbury.

Johnson-Cartee, K. S. (2005). *News narratives and news framing.* Lanham, MD and Oxford: Rowman and Littlefield.

Liu, M., & Zhang, Y. (2018). Discursive constructions of scientific (un) certainty about the health risks of China's air pollution: a corpus-assisted discourse study. *Language and communication 60*, 1-10. https://doi.org/10.1016/j.langcom.2018.01.006

Lutzky, U., & Kehoe, A. (2019). 'Friends don't let friends go Brexiting without a mandate': changing discourses of Brexit in *The Guardian*. In V. Koller, S. Kopf & M. Miglbauer (eds.), *Discourses of Brexit* (pp. 104-120). London: Routledge. https://doi.org/10.4324/9781351041867-7

Marchi, A., & Taylor, Ch. (2009). If on a winter´s night two researchers: a challenge to assumptions of soundness of interpretation. *Critical approaches to discourse analysis across disciplines (CADAAD) Journal 3*(1), 1-20.

Mole, R. (2007). *Discursive constructions of identity in European politics*. Basingstoke: Palgrave MacMillan. https://doi.org/10.1057/9780230591301

Partington, A. (2004). Corpora and discourse, a most congruous beast. In A. Partington, J., Morley & L. Haarman (eds.) *Corpora and discourse* (pp. 9-18). Frankfurt: Peter Lang.

Partington, A., Duguid, A., & Taylor, Ch. (2013). *Patterns and meanings in discourse. Theory and practice in corpus-assisted discourse studies (CADS).* Amsterdam: John Benjamins Publishing. https://doi.org/10.1075/scl.55

Partington, A. (2014). Mind the gaps. The role of corpus linguistics in researching absences. *International journal of corpus linguistics 19*(1):,118-146. https://doi.org/10.1075/ijcl.19.1.05par

Partington, A., & Taylor, Ch. (2018). *The language of persuasion in politics. An introduction*. London: Routledge. https://doi.org/10.4324/9781315177342

Pena- Díaz, C- In press/2022. Transeditando el Brexit. *Discurso y Sociedad.*

Sánchez Ramos, M. M. In press/2022. Constructing identities in crisis situations: a study of the volunteer in the Spanish and English press. In. I. E. Chiluwa (ed.), Discourse, media and conflict. Cambridge: Cambridge University Press.

Scott, M. (2010). Problems in investigating keyness, or clearing the undergrowth and marking out trails. In M. Bondi & Scott (eds.), *Keyness in texts* (pp. 43-58). Amsterdam: John Benjamins Publishing. https://doi.org/10.1075/scl.41.04sco

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. Amstendam: John Benjamins Publishing.

Stubbs, M. (2001). *Words and phrases. Studies in lexical semantics.* London: Blackwell.

Taylor, Ch. (2014). Investigating the representation of migrants in the UK and Italian press: A cross-linguistic corpus-assisted discourse analysis. *International journal of corpus linguistics 19*(3), 368-400. https://doi.org/10.1075/ijcl.19.3.03tay

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

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

Vessey, R. (2013). Challenges in cross-linguistic corpus assisted discourse studies. *Corpora 8*(1), 1-26. https://doi.org/10.3366/cor.2013.0032

Vessey, R. (2016). *Language and Canadian media: representations, ideologies, policies.* London: Palgrave Macmillan.

Wenzl, N. (2019). 'This is about the kind of Britain we are': national identities as constructed in parliamentary debates about EU membership. In V. Koller, S. Kopf & M. Miglbauer (eds), *Discourses of Brexit* (pp. 32-47). London: Routledge. https://doi.org/10.4324/9781351041867-3

White, H. (1980). The value of narrativity in the representation of reality. *Critical Inquiry 7(1),* 5-27. https://doi.org/10.1086/448086

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* (pp. 1-13). London: Sage Publications. https://doi.org/10.4135/9780857028020.n1

Wodak, R., De Cilia, R., Reisigl, M., & Liebhart, K. (2009). *The discoursive contstruction of national identity*. Endinburgh: Edinburgh University Press.