

A Keywords Corpus Analysis of ‘Taking the Knee’ and its Representation by the UK Press

Martin SPIVEY¹

University of Birmingham

Abstract

In recent years, the violent treatment of black citizens by police officers in the United States has received global attention. This has led to various forms of protest, one of which is the ‘taking the knee’ gesture performed by professional athletes prior to sports events. This paper presents the findings of a small-scale, exploratory keywords corpus analysis of 35 UK newspaper articles on ‘taking the knee’ in European football between June 2020 and November 2021. The aim is to discover how the issue has been covered by the UK press and there is some evidence to show that, while it creates some controversy, the gesture has been presented in a mostly positive (or at least neutral) light. However, more forensic and detailed analysis is required in order to gain a greater understanding of the issue at hand.

Keywords: corpora, keywords, race, critical discourse analysis, media representation

Introduction

In corpus linguistics, a keyword ‘is a word (or word cluster) which is found to occur with unusual frequency in a given text or set of texts’ (Scott, 2010a, p.149). Creating a keywords list is beneficial for studying the type of language that is used in a collection of texts or on a particular topic and has been applied in critical discourse analysis to research the media representation of various social and political issues (Al-Hejin, 2015; Baker et al., 2013; Branum & Charteris-Black, 2015). This study presents a keywords analysis of UK press reports on the ‘taking the knee’ gesture in European football from June 2020 to November 2021 using a small, specialized 25,915-word corpus of UK newspaper articles. This paper presents the results of the keywords analysis which seeks to answer the following research question:

What are the most positive keywords found in articles on the ‘taking the knee’ gesture in European football matches and what do these suggest about how the gesture is represented by the UK press?

¹ The author can be reached at: mart.spiv@gmail.com

I will address this question after presenting some of the keywords corpora studies that have investigated social and political issues and how they have been represented in the media. This is followed by a brief overview of the Black Lives Matter (BLM) movement along with an explanation of its influence on the act of Europe-based football players ‘taking the knee’.

Keywords corpora studies on media representation of social and political issues

In the past couple of decades, a growing number of studies have used corpus techniques in the critical analysis of discourse (Baker et al., 2008). Corpora use allows hundreds of texts to be investigated at once, rather than just one particular text or a small number over a limited period of time (O’Halloran, 2010). In the same paper, O’Halloran (2010) studied how the UK tabloid newspaper *The Sun* reported on the expansion of the European Union in 2004, using a 26,350-word corpus of 77 texts. There was a particular focus on how Eastern European migrants were discussed. The researcher compared the corpus with the BNC-baby (a 4-million-word sub-corpus of the British National Corpus) to extract a list of keywords. Through examining the keywords, he found that several referred to key British political figures at the time, as well as immigration, however he could not always find a direct link to Eastern Europeans. Baker et al. (2013) used keywords to compare tabloids and broadsheets as part of a broader study into how Muslims and Islam were represented in the British national press from 1998 to 2009. A list of the top 100 keywords for each corpus was compiled and the words were separated into categories in order to highlight any major differences between the two, such as language use or writing style. Subtirelu & Baker (2017) compared the online articles of US media outlets CNN and FOX and how they framed the federal budget negotiations between the Democrat and Republican parties in 2012. A keywords analysis found the top 10 keywords in each corpus when compared against the other, ranking the words based on their keyness scores. This was followed by a qualitative concordance analysis to interpret the findings and make generalizations. The political ideology of the UK press and the effect this has on its reporting was examined by Branum & Charteris-Black (2015), who conducted a comparative keywords analysis on how the Edward Norton story was reported by *The Sun*, *The Guardian*, and *Daily Mail*. They found that each newspaper presented the same events in strikingly dissimilar ways depending on their ideology, with *The Sun* and *Daily Mail* displaying nationalist ideals and evidence of framing which apparently attempted to discredit Norton.

The following section will now provide some background on the Black Lives Matter (BLM) movement and its influence on the ‘taking the knee’ gesture at football matches.

Background to Black Lives Matter & ‘taking the knee’

The origins of the BLM movement can be traced back to February 26th 2012 when 17-year-old African-American Trayvon Martin was shot dead by George Zimmerman (a Caucasian male) in Florida, USA. Zimmerman was charged with second-degree murder, and in a high-

profile court case was found not guilty on July 13th 2013 (Chase, 2018). This perceived injustice led to the creation of BLM by three black female activists seeking social justice for black people (Carney, 2016). The movement gained national recognition the following year when widespread protests took place after Michael Brown and Eric Garner were both killed by police officers (Skoy, 2021).

Moving forward to August 26th 2016, American football quarterback Colin Kaepernick decided to kneel during the American national anthem prior to a pre-season game for the San Francisco 49ers, having previously sat through the anthem at two other recent games (Graber et al., 2020). Kaepernick claimed that he was using his status to protest for the oppressed and to fight for change so all people were properly represented in the country (Trimbur, 2019).

On May 25th 2020, George Floyd, a 46-year-old African-American man, was killed by a white police officer in Minneapolis, USA. This triggered several demonstrations across the US under the BLM banner and led to global protests against racial discrimination (Barrie, 2020). On June 17th in the English Premier League (EPL), Sheffield United played Aston Villa after a Covid-enforced break and both sets of players and officials took the knee for 10 seconds prior to kick-off in support of the BLM movement at a ground empty of spectators (McNulty, 2020). The gesture was performed prior to several matches throughout the English leagues on their return and continued into the 2020-2021 season. In June 2021, the Euro 2020 tournament took place in several European cities after a year's delay due to the COVID-19 pandemic. In England's opening game against Croatia at Wembley Stadium in London on June 13th, a number of supporters booed the 'taking the knee' protest after doing so in their pre-tournament warm-up games. A YouGov survey showed that 54% of 547 English football supporters claimed to be in favour of the protest while 39% were opposed. Responses from Scotland and Wales supporters were of a similar ratio (Welton, 2021).

This suggests that there is some controversy over the 'taking the knee' gesture across British society and therefore I feel it is worthwhile to conduct a keywords analysis to study how the UK press has generally covered this issue.

Keywords analysis of 'taking the knee' in the UK press

The purpose of this study is to use keywords analysis to gain insight into how the 'taking the knee' gesture in European football matches has been portrayed by the UK press. The analysis highlights the benefits of conducting such research but also exposes the limits of this approach. The following sections show how the study was formed and present the results and analysis.

Corpus construction

A specialized corpus was created of 35 UK newspaper articles (25,915 words) dating from June 2020 to November 2021. This was done by searching for words linked to the topic through the LexisNexis database (LexisNexis, 2021) at four different time periods to ensure a fairly even spread of coverage, as shown in the table below.

PERIOD	WORDS SEARCHED	KEY EVENTS
A June/July 2020	taking the knee, football, Premier League	Start of ‘taking the knee’ at UK football grounds
B January/February 2021	same as A	
C June/July 2021	A plus EURO 2020	EURO 2020 tournament
D October/November 2021	same as A	

Table 1: Search details for the *Taking the Knee* corpus

Eight national newspapers and one major regional newspaper were chosen and articles were selected based on the level of relevance to the topic where ever possible. A minimum threshold of 100 words per article was set and where a relevant text could not be found, a Scottish/Sunday edition or online version was used. The table below shows the selected newspapers, their political leanings and the number of articles and word counts in the corpus:

	<i>LEFT-LEANING</i>	<i>RIGHT-LEANING</i>
TABLOID	<ul style="list-style-type: none"> ▪ Daily Mirror (4 articles:3653 words) 	<ul style="list-style-type: none"> ▪ The Sun (4:3816) ▪ Daily Express/Sunday Express (3:1001) ▪ Daily Mail/Mail on Sunday /MailOnline (4:3814)
BROADSHEET	<ul style="list-style-type: none"> ▪ The Guardian (4:3532) ▪ The Independent (4:1826) 	<ul style="list-style-type: none"> ▪ The Daily Telegraph (4:2478) ▪ The Times (4:2173) ▪ Yorkshire Post (4:3783)

Table 2: Newspapers in the corpus (adapted from Baker et al., 2013)

Keywords extraction

To create a set of keywords for the specialized corpus, a reference corpus needs to be made for comparative purposes. It is important to select a suitable reference corpus for this procedure (Scott, 2010b). Typically, keywords studies will use a general corpus such as the British National Corpus, however Geluso & Hirsch argue that ‘...a reference corpus of the same sub-register will highlight words that are particular to a specific target corpus’ (2019, p.231). I used Sketch Engine (Sketch Engine, 2021) to create the specialized *Taking the Knee* corpus and for the reference corpus I made a sub-corpus from the English Broadsheet Newspapers 1993-2013 corpus called *UK Newspapers 2013* (7 UK national newspapers totalling 136.1 million words).

What are the most positive keywords found in articles on the ‘taking the knee’ gesture in European football matches and what do these suggest about how the gesture is represented by the UK press?

This question will be answered through an analysis of the top 100 keywords found in the *Taking the Knee* corpus when compared to the sub-corpus outlined in the previous section. Rather than finding statistically significant keywords using, for instance, the log-likelihood ratio in AntConc, Sketch Engine uses a ‘simple maths’ method to create a keyness score (Kilgarriff, 2009). This is done by adding a number ($N=1/10/100$ etc) to the word frequencies so a normalized frequency ratio can be computed for all words in the corpora and thus deal with the null hypothesis issue (Jaworska & Kinloch, 2018). I have decided to select the default option of $N=1$ for this keywords analysis as we are looking for relatively rare words that feature in the focus corpus compared to the reference corpus. Due to space limitations, only the top 20 keywords are presented in the table below.

To answer this question, an appropriate next step is to categorize the words in the list so we can see more clearly which types of words feature in the top 100 keywords. Broad categories were chosen based on their semantic meanings and the words placed in order of keyness with their ranking in parentheses.

The focus corpus is compared to a 2013 corpus of UK newspapers so there is little surprise that some of the keywords are related to the COVID-19 pandemic which runs through the entire time period of the corpus. There are also the names of several prominent sport and political figures that form part of two large categories. Some of these keywords include both the first and second names of the same person, such as *Wilfried Zaha* (the Crystal Palace football player) and *Sanjay Bhandari* (chairman of football’s anti-discrimination group, Kick It Out). During the process of categorization, on a few occasions the concordance data needed to be studied to ensure that the words were placed in the most appropriate word group. For instance, the word *diversity* appears 16 times in 6 different texts. A look at the concordances shows that while *taking the knee* appears close to the KWIC (Key Word in Context) in five lines, in most

cases it is referring to diversity within football administration and in three lines it forms part of a job title so I felt it more appropriate to include in the ‘Race and discrimination’ section.

RANK	WORD	KEYNESS SCORE	FREQUENCY	FREQUENCY/ MILLION	DOCUMENT FREQUENCY
1	lives	772.9	52	1762.59	24
2	blm	708.4	21	711.82	9
3	kneel	465.6	46	1559.22	10
4	rashford	441.6	13	440.65	8
5	kaepernick	429	15	508.44	7
6	bame	415.7	13	440.65	3
7	efl	369.2	11	372.86	9
8	toney	298.6	9	305.06	6
9	boyer	266.6	10	338.96	3
10	boo	219.4	82	2779.47	12
11	anti-racism	214.3	14	474.54	9
12	covid-19	170.5	5	169.48	3
13	bullingham	170.5	5	169.48	1
14	fsu	169.4	5	169.48	1
15	bhandari	166.3	5	169.48	3
16	coronavirus	147.3	6	203.38	6
17	kock	145.6	8	271.17	1
18	racism	145	63	2135.45	24
19	nate	144.4	7	237.27	3
20	gesture	143.1	88	2982.85	24

Table 3: Top 20 keywords in the *Taking the Knee* corpus

CATEGORY	KEYWORDS
Taking the knee	kneel (3), efl (7), boo (10), gesture (20), knee (24), apolitical (26), player-led (48), brentford (58), degrading (59), solidarity (62), platitude (64), connotation (65), quarterback (66), ticket-holder (67), marxism (68), pre-match (71), anthem (73), nfl (75), restart (81), kick-off (83), disapproval (84), taking-the-knee (88), then-49ers (90), boo-er (98)
Race & discrimination	lives (1), blm (2), bame (6), anti-racism (11), racism (18), defund (36), anti-discrimination (52), cottager (54), minneapolis (55), anti-racist (57), racial (60), basketball (61), diversity (70), african-american (72), equality (74), racist (74), virtue-signalling (96)
Athletes, coaches etc	rashford (4), kaepernick (5), toney (8), boyer (9), kock (17), nate (19), southgate (25), mings (28), mahrez (32), ogbene (33), saka (34), sancho (37), zaha (39), wilfried (40), raheem (41), kudela (42), bukayo (47), havertz (49), jaden (51), assombalanga (56), ondrej (63), foden (76), bridcutt (79), britt (82), nedelcearu (91), nketiah (92)
Political/administrative figures & organisations	efl (7), bullingham (13), fsu (14), bhandari (15), dowden (23), edleen (31), priti (35), molango (43), clarke-smith (50), sanjay (53), burnett (78), elliot (80), rupinder (87), maheta (99)
Other figures	teigan (21), floyd (22), chauvin (44), marvina (46)
COVID-19 pandemic	covid-19 (12), coronavirus (16), lockdown (27), covid (30), pandemic (38), stadia (69), pre-covid (89), socially-distanced (94)
Miscellaneous (errors etc)	sunsport (29), findyours (45), cottager (54), maine (85), aepernick (86), starsant (93), stadiumpicture (95), etcetera-etcetera (97)

Table 4: Semantic categories of the top 100 keywords

We now need to take a closer look at some of the keywords to establish how ‘taking the knee’ is represented in UK newspapers in general. Determining the context in which the keyword is used is imperative for answering the question satisfactorily. The lemma *boo* has a raw frequency of 82 in the corpus and appears in 12 texts (34% text coverage). It has the third highest relative frequency in the keyword list which suggests it is a major feature in the discussion around ‘taking the knee’. Booming is typically used to show strong disagreement with somebody or something but how is this framed in the different corpus texts? All 82 concordance lines were downloaded, the context in which *boo* is used was examined and each instance was coded and placed into one of three categories:

- In support of booing ‘taking the knee’ (20/82 – **24.3%**)
- Neutral stance (31/82 – **37.8%**)
- Against booing ‘taking the knee’ (31/82 – **37.8%**)

These figures imply that most press articles show support or are neutral in their reporting of the issue, however caution is advised here. 35 instances of *boo* appear in only two articles and there are a variety of views given within one article. In a MailOnline article, *boo* appears 28 times – 6 in support of booing, 10 neutral, and 12 against, many of which appear in quotes from various figures. Without reading the entire article, it is hard to judge the writer’s overall framing of their argument, and this is a limitation of keywords analysis. What can be stated with some degree of confidence is that the issue of booing the ‘taking the knee’ protest has received fairly mixed coverage by the UK press, however it has largely reported objective facts or strong opinions *against* the act of booing.

I wish to now discuss the word *gesture*, which has a raw frequency of 88 in 24 different texts (68.5% text coverage) and is 20th in terms of keyness score. This word has a neutral meaning on its own and one would expect it to have a high frequency in articles discussing ‘taking the knee’ but in what context is it used? Do the concordance data highlight a particularly supportive or unsupportive stance towards the gesture? Each instance referring to the ‘taking the knee’ gesture was examined and categorized as follows:

- Mostly positive stance towards ‘taking the knee’ (20/86 – **23.2%**)
- Neutral/unclear stance (25/86 – **29%**)
- Mostly negative stance towards ‘taking the knee’ (41/86 – **47.6%**)

Almost half the instances of *gesture* in the corpus occur in a negative stance towards the gesture. Further enquiry shows that some opinions expressed in the articles are not necessarily a fundamental disagreement with the gesture itself, as the concordance lines show in Table 5.

A small minority of articles have expressed a fairly negative, cynical view of the gesture, appearing to purposefully undermine the footballers’ intentions by mentioning alleged links the BLM movement has to Marxism. This task has highlighted the importance of checking concordances for nuance and connotation when conducting a keywords analysis. In answer to the question, one could make an argument that the top 100 keywords in the *Taking the Knee* corpus show that, while as a whole the UK press has presented a fairly mixed range of viewpoints on the issue, the majority of the coverage has been mostly neutral or outwardly positive.

12. #doc 9 footballers among them Les Ferdinand, Wilfried Zaha and Britt Assombalonga have started to rail against the <u>gesture</u> , which they believe has become devalued and is covering up a lack of real change in anti-racism policies.
17. doc#10 The debate comes a week after Championship club Brentford claimed the <u>gesture</u> had lost its impact.
22. doc#13 The Crystal Palace talisman has reiterated his belief that the <u>gesture</u> is “degrading” and that black players should “stand tall”.
88. doc#34 QPR players also elected not to take the knee, with director of football Les Ferdinand arguing that the <u>gesture</u> had become little more than a gimmick in the fight against racial inequality and discrimination in his opinion.

Table 5: Concordance lines of *gesture* in the corpus

Conclusion

This paper has examined how a keywords corpus analysis can be used to study media discourse on a controversial issue. This has been done by investigating the UK press coverage of the ‘taking the knee’ protest prior to football matches, through an analysis of a 25,915-word corpus of 35 articles. An examination of the top 100 keywords found several different perspectives but a majority had a seemingly positive (or at least neutral) view of the gesture as a means to protest racial discrimination. This form of analysis should be viewed as a ‘gateway’ into more forensic study of a collection of texts. As Subtirelu & Baker (2017, p.114) argue, the keyword analysis results ‘...offer the analyst an entry point into the data rather than a readily interpretable set of findings’. It helps reduce researcher bias as the corpus software runs the calculations and determines which words are statistically significant. A skilled researcher can then select the most pertinent words, examine their collocations, and take a closer look at the texts to judge nuance and context.

In conclusion, keywords analysis has been an effective tool for an exploratory study into UK press representation of the ‘taking the knee’ protest prior to European football matches. However, more detailed analytical study of the corpus is required in order to gain a greater understanding of how UK newspapers have covered this issue.

References

- Al-Hejin, B. (2015). Covering Muslim women: Semantic macrostructures in BBC News. *Discourse and Communication*, 9(1), 19–46.
<https://doi.org/10.1177/1750481314555262>

- Baker, P., Gabrielatos, C., Khosravini, M., Krzyzanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and Society*, 19(3), 273–306. <https://doi.org/10.1177/0957926508088962>
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge University Press.
- Barrie, C. (2020). Searching Racism after George Floyd. *Socius*, 6, 1-3. <https://doi.org/10.1177/2378023120971507>
- Branum, J., & 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>
- Carney, N. (2016). All Lives Matter, but so Does Race: Black Lives Matter and the Evolving Role of Social Media. *Humanity & Society*, 40(2), 180–199. <https://doi.org/10.1177/0160597616643868>
- Chase, G. (2018). The Early History of the Black Lives Matter Movement, and the Implications Thereof. *Nevada Law Journal*, 18(3), 1091-112.
- Geluso, J., & Hirsch, R. (2019). The reference corpus matters. *Register Studies*, 1(2), 209–242. <https://doi.org/10.1075/rs.18001.gel>
- Graber, S. M., Figueroa, E. J., & Vasudevan, K. (2020). Oh, Say, Can You Kneel: A Critical Discourse Analysis of Newspaper Coverage of Colin Kaepernick’s Racial Protest. *Howard Journal of Communications*, 31(5), 464–480. <https://doi.org/10.1080/10646175.2019.1670295>
- Jaworska, S. & Kinloch, K. (2018). Using multiple data sets. In C. Taylor & A. Marchi (Eds.). *Corpus Approaches to Discourse: A Critical Review*. (1st ed., pp. 110-129). Routledge.
- Kilgarriff, A. (2009). *Simple maths for keywords*. In *Proc. Corpus Linguistics*.
- LexisNexis (2021). <https://advance.lexis.com/BisNexisNewsSearch/>
- McNulty, P. (2020, June 17). *Aston Villa 0 Sheffield United 0*. BBC. <https://www.bbc.com/sport/football/51595065>
- O’Halloran, K. (2010). How to use corpus linguistics in the study of media discourse. In A, O’Keeffe & M, McCarthy (Eds.). *The Routledge Handbook of Corpus Linguistics*. (1st ed., pp. 563-577). Routledge.
- Scott, M. (2010a). What can corpus software do? In A, O’Keeffe & M, McCarthy (Eds.). *The Routledge Handbook of Corpus Linguistics*. (1st ed., pp. 136-151). Routledge.
- Scott, M. (2010b). Problems in investigating keyness, or clearing the undergrowth and marking out trails.... In M, Bondi & M, Scott (Eds.). *Keyness in Texts* (1st ed., pp. 43-57). John Benjamins.
- Sketch Engine (2021). <https://app.sketchengine.eu/#dashboard>
- Skoy, E. (2021). Black Lives Matter protests, fatal police interactions, and crime. *Contemporary Economic Policy*, 39(2), 280–291. <https://doi.org/10.1111/coep.12508>

- Subtirelu, N. C., & Baker, P. (2017). Corpus-based approaches. In J. Flowerdew & J. E. Richardson (Eds.), *The Routledge Handbook of Critical Discourse Studies* (1st ed., pp. 106–119). Routledge.
- Trimbur, L. (2019). Taking a Knee, Making a Stand: Social Justice, Trump America, and the Politics of Sport. *Quest*, 71(2), 252–265.
<https://doi.org/10.1080/00336297.2018.1551806>
- Welton, B. (2021, June 25). *England vs Croatia: Gareth Southgate's side booed for taking a knee but majority of fans applaud gesture*. Sky Sports.
<https://www.skysports.com/football/news/12016/12331391/england-vs-croatia-gareth-southgates-side-booed-for-taking-a-knee-but-majority-of-fans-applaud-gesture>