

Investigating media representations of the coronavirus in the UK, USA and Germany: what can a comparative corpus-based discourse analysis contribute to our understanding of the Covid-19 pandemic?

Book or Report Section

Accepted Version

Jaworska, S. ORCID: <https://orcid.org/0000-0001-7465-2245> (2021) Investigating media representations of the coronavirus in the UK, USA and Germany: what can a comparative corpus-based discourse analysis contribute to our understanding of the Covid-19 pandemic? In: Jones, R. H. ORCID: <https://orcid.org/0000-0002-9426-727X> (ed.) *Viral Discourse. Elements in applied linguistics*. Cambridge University Press, Cambridge, pp. 26-37. ISBN 9781108986465 Available at <https://centaur.reading.ac.uk/97219/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Publisher: Cambridge University Press

All outputs in CentAUR are protected by Intellectual Property Rights law,

including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

4. Investigating media representations of the coronavirus in the UK, USA and Germany: What can a comparative corpus-based discourse analysis contribute to our understanding of the Covid-19 pandemic?

Sylvia Jaworska

Why a comparative, corpus-based discourse analysis?

Through naming, describing and evaluating, media play a key role in creating representations of illness for public ‘use’. Since many aspects of a disease, for example a virus and its spread, are not visible to the naked eye and difficult to understand by non-experts, media reporting is for most people the prime source of health information; through the choice of language and images the media make the invisible visible influencing imaginaries, opinions and, in turn, responses to a health crisis (e.g. Dorfman and Wallack, 2007; Thibodeau and Boroditsky, 2011).

When at the beginning of January 2020 many people from across the world were returning from the Christmas break, while others were looking forward to celebrating the Chinese New Year, reports about a new and deadly virus started to circulate. Officially named SARS-CoV-2 and widely referred to as the coronavirus, the virus has since spread to most countries across the world leading to unprecedented health, social, and economic consequences.

This small study explores some of the ways in which the coronavirus has been discursively constructed in the popular public media in three distinctive cultural and linguistic contexts: the UK, US, and Germany. Such comparisons are important for several reasons. With some exceptions (e.g. Antanasova and Koteyko, 2017), most discourse-analytical research on health and illness has focused mainly on one national context and one language. While such perspectives can offer rich insights into the representations of a discursive phenomenon, these representations are always ‘bespoke’ and restricted to that context and language, thus limiting our understating of how the phenomenon is ‘seen’ elsewhere (Partington et al., 2013; Leuschner and Jaworska, 2018). The coronavirus knows no borders and while the biological properties of the pathogen are everywhere the same, the ways in which the virus is talked about can be influenced by distinctive societal, cultural and linguistic factors. If we want to better understand the social reality of the Covid-19 pandemic, it is essential to compare how the virus has been represented in different contexts. Widely disseminated media outlets such as national newspapers can offer some important insights into such representations. Moreover, comparisons across different contexts are important not only because they can limit some of the generalizations that are sometimes made (based on research on representations in English only) about the ways in which we talk about health and illness, but also for epistemological reasons. Exploring how the coronavirus has been represented in public media discourse across different national contexts could uncover different ways of reasoning in relation to the pandemic and how they are reflected and reinforced through the language choices people make, potentially leading to a better understanding of the pandemic and stimulating knowledge exchange.

This study examines media representations of the coronavirus by adopting the approach of a corpus-assisted discourse analysis (Baker, 2006; Partington et al., 2013). Corpus methods can

be useful here for at least two reasons. First, a corpus approach gives us the opportunity to study collocations, that is, recurrent lexical choices occurring in the vicinity of the term under study. The concept of collocation goes back to the idea that the meanings of words are not just inherent in the prime word form but arise from the typical combinations of the word with other words in their context of use (Firth, 1957). The combination ‘killer virus’ can evoke different perceptions or associations than, for example, ‘mortal virus’. Moreover, although in principle text producers have a larger pool of lexical items at their disposal, the choices that they make are likely to reflect deeper societal beliefs as well as ideological positions that in turn can reinforce particular kinds of representations and trigger particular outcomes (Stubbs, 2001). When the Trump administration has consistently referred to the coronavirus as the ‘China virus’, this choice serves, among other things, to blame the people of China for the spread of the disease, and led to a rise in prejudiced attitudes and stigmatisation of Asians (The Editorial Board, 2020).

Secondly, a corpus-based approach relies on a quantitative investigation of large collections of texts based on automated frequency counts. This allows for lexical choices to be explored consistently across corpora. Findings obtained in this way can point to the existence of recurrent discursive patterns and signal salient discourses around the studied phenomenon that an analysis of a few texts may not reveal (Baker, 2006); replicating the same procedures, we can then compare our findings systematically across corpora, shedding light on how the phenomenon in question is discursively constructed in different sources and languages (Partington et al., 2013; Jaworska and Leuschner, 2018).

Data and methods

For the purpose of this study, three corpora of press coverage of the pandemic in the UK, US and Germany were compiled using the database Nexis and two search terms ‘coronavirus’ and ‘Covid-19’. Both terms were also used to extract data from German sources, since they have exact equivalents and are used so in German. 7 January was selected as the starting point for the data collection because on that day the virus was officially confirmed to be a novel type of coronavirus. The end point was 31 March when the virus began to spread across the world leading to unprecedented lockdowns. Therefore, this study covers the initial outbreak of the Covid-19 pandemic.

Articles were collected from major national newspapers from the three countries that are available on Nexis. For the UK, the broadsheets *The Guardian*, *The Daily Telegraph* and *The Times* as well as ‘regular’ and middle-range tabloids such as *The Daily Mail*, *The Mirror* and *The Sun* were included. For the US, *The New York Times*, *The New York Post*, *USA Today* and *Los Angeles Times* were included. The German corpus comprised articles from *Der Spiegel* (the most popular news magazine published weekly), *Die Tageszeitung (TAZ)*, *Die Welt* and *Bunte* (a popular magazine similar in content and design to a tabloid). Table 4.1 shows the sizes of the three corpora. The differences reflect the availability of newspapers on Nexis and their status as daily or weekly newspapers or news magazines. For example, the two sources included in the DE-Corpus *Der Spiegel* and *Bunte* are weekly magazines and hence, fewer articles are expected than in daily newspapers, which explains partially the smaller size of the German corpus.

Corpus	No. of Articles	Corpus size
UK-Corpus	4,095	4,356,873

USA-Corpus	3,190	1,058,809
DE-Corpus	1,424	950,491

Table 4.1: The sizes of the corpora

To identify the dominant discursive construction of the coronavirus across the three national media contexts, collocations of the term ‘coronavirus’ were explored. The term was selected because at the time of the data collection, it was the most frequently used term in public discourse associated with the virus and the pandemic as evidenced in a study conducted by the team of the Oxford English Dictionary (OED, 2020); it has been more widely used than the official scientific term SARS-CoV-2, and the disease caused by the virus was not dubbed Covid-19 until early February 2020. To retrieve the collocations of ‘coronavirus’, the software programme Sketch Engine was used. In contrast to other similar programmes, Sketch Engine includes the tool Word Sketch, which specifies grammatical patterns of collocations allowing for the identification of, for example, the most salient verbal collocates depending on whether the search term is in the subject or object position (though errors occur in the classification and it is recommended to check the concordance lines to identify items that have been categorised inaccurately). Another benefit of using Sketch Engine is that it identifies collocates based on the Log Dice (LD) score, which is a measure of a collocation’s salience and it does not depend on the total size of the corpus. This allows the researcher to have a consistent comparison measure across corpora of unequal sizes. Only collocations with a minimum frequency of 5 and LD score of 7 or above were considered, parameters also used in similar previous research.

Results

Table 4.1 shows the most salient collocations of ‘coronavirus’ in the UK press corpus. Only modifiers and verbs are included, since they emerged as the ‘combat’, ‘kill’, ‘strike’ and ‘hit’ in the vicinity of coronavirus suggesting a similar approach as in the UK press.

Modifiers of ‘coronavirus’			Verbs with ‘coronavirus’ as subject			Verbs with ‘coronavirus’ as object		
Modifier	Raw Freq.	LD Score	Verb	Raw Freq.	LD Score	Verb	Raw Freq.	LD Score
novel (noun)	115	11.51	cause	138	10.07	contract	184	11.04
new	435	10.92	affect	132	10.03	catch	83	10.05
novel (adjective)	49	10.28	spread	118	9.79	fight	66	9.77
deadly	44	9.91	hit	99	9.5	contain	80	9.73
suspected	23	9.01	pose	61	9.12	tackle	60	9.64
Covid-19	49	8.87	have	598	8.85	combat	51	9.5
terrible	10	8.08	reach	54	8.84	spread	57	9.48
fast-spreading	8	7.81	be	976	8.74	have	237	9.11
			continue	50	8.5	declare	33	8.63
			infect	38	8.45	treat	25	8.38
			die	37	8.01	get	47	8.32
			appear	29	7.88	regard	21	8.31
			impact	23	7.77	detect	18	8.08
			kill	22	7.64	confirm	42	7.91
			rise	28	7.61	defeat	15	7.9
			force	20	7.54	beat	16	7.88
			mean	21	7.47	call	21	7.86
			become	22	7.35	surround	13	7.66
			strike	17	7.34	transmit	12	7.55
			do	28	7.15	discuss	13	7.48
						warn	15	7.48
						battle	11	7.46
						discover	11	7.44

Table 4.1: Collocations of ‘coronavirus’ in the UK-Corpus: modifiers and verbs

largest grammatical categories of all collocations identified. Striking is the use of many words from the domain of warfare or violence. Out of the 51 collocates shown in Table 4.1, 12 are specific references to military conflict (‘battle’, ‘combat’, ‘surround’, ‘defeat’, ‘strike’) or violence/physical fighting (‘beat’, ‘fight’, ‘force’, ‘hit’, ‘tackle’, ‘strike’). These are shown in bold and seem to be particularly prominent (also in terms of raw frequencies) in the category of verbal collocates that occur with the search term as an object. This suggests that in this

British press discourse measures of halting the pandemic were often framed as similar to those needed in a war or a fight.

A similar tendency, although with fewer instances of war references (8 in total), can be observed in the US press coverage (Table. 4.2). Here too, we find words such as ‘fight’, ‘combat’, ‘kill’, ‘strike’ and ‘hit’ in the vicinity of coronavirus suggesting a similar approach as in the UK press.

Modifiers of ‘coronavirus’			Verbs with ‘coronavirus’ as subject			Verbs with ‘coronavirus’ as object		
Modifier	Raw Freq.	LD Score	Verb	Raw Freq.	LD Score	Verb	Raw Freq.	LD Score
novel (noun)	226	12.54	cause	171	11.23	contract	82	10.89
new	458	11.75	spread	106	10.49	spread	63	10.34
novel (adjective)	59	10.79	continue	52	9.44	fight	40	9.92
deadly	42	10.06	affect	45	9.42	contain	38	9.7
current	8	7.51	hit	37	9.1	catch	30	9.6
			infect	29	8.9	combat	23	9.28
			pose	28	8.86	have	133	8.96
			have	271	8.84	prevent	22	8.85
			be	455	8.57	treat	19	8.77
			appear	21	8.24	get	40	8.69
			become	21	8.07	battle	11	8.31
			break	15	7.95	declare	13	8.24
			grow	16	7.94	transmit	9	8.02
			strike	13	7.78	beat	8	7.81
			kill	13	7.72	address	9	7.79
			disrupt	10	7.42	handle	8	7.67
			threaten	10	7.4	believe	8	7.64
			make	15	7.38	confront	7	7.63
			emerge	10	7.35	call	11	7.58
			begin	13	7.28	know	10	7.53
			turn	10	7.28	stop	8	7.49
			force	9	7.23	dismiss	6	7.49
			go	15	7.19	regard	6	7.42
						discuss	6	7.34
						cure	5	7.23
						regard	6	7.42
						think	8	7.41
						limit	7	7.22
						surround	5	7.15

Table 4.2: Collocations of ‘coronavirus’ in the US-Corpus: modifiers and verbs

Looking at the German data (see Table 4.3), we find different items as salient collocations. Most striking is the absence of references to the militaristic language prominent in the two other corpora. Among the strongest verbal and modifying collocates of ‘Coronavirus’, only two verbs come from the domain of warfare ‘bekämpfen’ (to combat) and ‘besiegen’ (to defeat). But other words are much more salient and these include items that are, in turn, absent from the top collocates in the UK and US corpora. These include verbs such as ‘testen’ (to test) and ‘schützen’ (protect) as well as ‘informieren’ (inform), ‘untersuchen’ (investigate) and ‘entwickeln’ (develop) (the item ‘test’ is a noun collocate of ‘coronavirus’ in the UK-Corpus but is less salient than other collocates).

Modifiers of ‘coronavirus’			Verbs with ‘coronavirus’ as subject			Verbs with ‘coronavirus’ as an accusative object			Verbs with ‘coronavirus’ as a dative object		
Modifier	Raw Freq.	LD Score	Verb	Raw Freq.	LD Score	Verb	Raw Freq.	LD Score	Modifier	Raw Freq.	LD Score
neuartig	182	12.94	ausbreiten	13	10.53	testen	35	11.37	infizieren	42	12.05
neu	114	10.98	verbreiten	8	9.78	eindämmen	7	9.25	anstecken	19	11.15
positiv	16	9.65	wüten	6	9.62	entwickeln	6	8.87	schützen	11	10.4
ausbreitend	9	9.26	sorgen	6	9.31	bekämpfen	4	8.65	sterben	6	9.3
grassierend	5	8.38	treffen	7	9.15	treffen	6	8.55	verhindern	5	8.79
aktuell	8	8.13	legen	6	9.15	stehen	6	8.48	erkranken	6	8.76
gefährlich	5	7.8	grassieren	5	9.04	besiegen	5	8.38	umgehen	5	8.66
rund	7	7.68	spielen	6	8.83	vergleichen	6	8.34			
kurz	5	7.47	bedeuten	5	8.82	untersuchen	5	8.27			
			greifen	5	8.69	informieren	6	8.08			
						nutzen	5	7.83			

Table 4.3: Collocations of ‘Coronavirus’ in the German press

At the time of writing (July 2020), the pandemic situation in these three countries was different. In Germany, the spread of the virus seemed largely under control, while in the UK the infection rate was going down slowly and, in the USA, it was still rising. This has to do, in part, with the

different national responses to the pandemic. Germany was relatively quick in introducing a lockdown and widespread testing, and hence it should not be surprising to see ‘testen’ (to test) as one of the top verbal collocates. The prominence of ‘untersuchen’ (investigate) and ‘entwickeln’ (develop) points further to an approach grounded in science, while ‘schützen’ (protect) places an emphasis on a care and protection. The UK reacted more slowly to the pandemic and had a limited testing capacity. Its response was characterized by an invocation of the ‘Blitz spirit’ of the WWII. Examples of the war language around the pandemic from the UK-Corpus include headlines such as:

1. Why our hero health staff are no1 defence against covid-19 (*The Sun*)
2. NHS and private hospitals join forces to fight coronavirus (*The Guardian*)
3. We are at war against the virus (*Daily Mail*)

The fewer instances of war and violence metaphors in the US-Corpus may also be a reflection of the pandemic timeline and the rather inactive stance on the part of its government during the initial outbreak. The conspicuous absence of war and violence metaphors in the DE-Corpus may have historical reasons; war rhetoric and specifically words coined during the Nazi period are not common as source domains in the German press discourse and are used predominantly in original historical references (Schröter, 2018)¹.

The use of war metaphors in the context of a pandemic is problematic. They can sometimes be constructive in that they can mobilize public health efforts. Yet, when it comes to patients, particularly those who suffer from deadly conditions, they can be distressing, and even unethical, especially if the patients or their doctors are not ‘winning the battle’ (e.g. Sonntag, 1989; Semino et al., 2018). Also, the kind of qualities that a military mobilization requires such

as strong character and physical strength are, in and of themselves, not going to weaken the virulence of the pathogen which spreads quickly through close human contact. The virus does not distinguish between ‘friends’ and ‘enemies’; it can affect us all, whatever our physical strength and moral stance. As we have seen, testing and prevention measures as well as clear public communication are more successful.

Conclusion

The study illustrates the usefulness of studying collocations to uncover dominant ways of representing the coronavirus; it demonstrates how they can shed light on discourses that are salient but not necessarily obvious to the researcher. Metaphors were not a specific focus of this study; yet they have emerged as an important linguistic device in the discourses around ‘coronavirus’ in the UK and US corpora. This study has also shown the benefits of a comparative approach clearly indicating that what can be said about the pandemic and how it can be said depends on the linguistic and cultural factors including different local histories. The relative absence of war metaphors in the German public discourse around ‘coronavirus’ is a single but compelling case in point. Since the national media are for most people the first source of health information, studying how the media report on the novel virus is critical if we want to understand public responses to the pandemic. The way in which the media choose words to frame the virus may itself lead the population to respond in more or less effective ways. Yet, this study is small in scope and does not utilise the full potential of corpus-based methods to study discourse (see Baker, 2006; Partington et al., 2013). It has considered collocations only and as de-contextualised items presented in lists; further research is needed to explore how they were used in the press discourse investigated here taking into consideration a wider range of lexico-grammatical properties of the studied languages including, e.g. the genitive objects in German and noun collocations in both English and German.

References

- Antanasova, D. and Koteyko, N. (2017). Obesity frames and counter-frames in British and German online newspapers. *Health*, 21(6): 650–669.
- Baker, P. (2006). *Using corpora in discourse analysis*. London: Continuum.
- Dorfman, L. and Wallack, L. (2007). Moving nutrition upstream: the case for reframing obesity. *Journal of Nutrition Education and Behavior*, 39(2): 45–50.
- The Editorial Board. (2020, March 23). Opinion: Call it ‘coronavirus’. *The New York Times*. Retrieved 24 March 2020 from <https://www.nytimes.com/2020/03/23/opinion/china-coronavirus-racism.html>
- Firth, J. R. (1957). Modes of meaning. In *Papers in Linguistics* (pp. 1934–951). London, New York: Oxford University Press.
- Jaworska, S. and Leuschner, T. (2018). Crossing languages - crossing discourses: a corpus-assisted discourse study of Kulturkampf in German, Polish and English. *Pragmatics and Society*, 9(1): 119–149.
- Leuschner, T. and Jaworska, S. (2018). Anglo-German discourse crossings and contrasts: Introduction to the special issue. *Pragmatics & Society*, 9 (1):
- Oxford English Dictionary (OED). (2020). Corpus analysis of the language of Covid-19, <https://public.oed.com/blog/corpus-analysis-of-the-language-of-covid-19/#note1> [last access 15 November 2020]
- Partington, A., Duguid, A. and Taylor, C. (2013). *Patterns and meanings in discourse. Theory and practice in corpus-assisted discourse studies (cads)*. Benjamins, Amsterdam.

Schröter, M. (2018). How words behave in other languages: the use of German Nazi vocabulary in English. *Pragmatics and Society*, 9(1): 93–118.

Semino, E., Demjen, Z., Demmen, J. E. (2018). An integrated approach to metaphor and framing in cognition, discourse and practice, with an application to metaphors for cancer. *Applied Linguistics*, 39(5): 625–645.

Sontag, S. (1989). *AIDS and its metaphors*. Farrar, Straus and Giroux, New York.

Stubbs, M. (2001). *Words and phrases: Corpus studies in lexical semantics*. Blackwell, Oxford.

Thibodeau, P. and Boroditsky, L. (2011). Metaphors we think with: the role of metaphor in reasoning. *PLoS ONE* 6(2): e16782.

Notes

¹ There seems to be a high degree of sensitivity in Germany when it comes to the use of language around the pandemic; in a recent debate streamed on the acclaimed Coronavirus-Update Podcast (produced by NDR – Norddeutscher Rundfunk ‘Northern German Broadcast’), which includes prominent scientists and social scientists, the use of the word ‘Bewältigung’ (overcoming) in the context of dealing with the Corona crisis was criticised as inappropriate because it is strongly associated with the ‘Vergangenheitsbewältigung’ (overcoming the past), where the past means the Nazi past (see <https://www.ndr.de/nachrichten/info/podcast4684.html>)