Anti-vaccine discourse under the COVID-19 context: A corpus-assisted discourse analysis

Xinmei Sun

1. Introduction

Anti-vaccine communities are built around a deep belief that vaccines are extremely harmful (Kata, 2012). They are characterised by mistrust of the government(Kata, 2012). Anti-vaccine movements were present since the introduction of vaccination in the 19th century (Maci,2019). These movements heightened in the 1990s after the publication of a study that demonstrated a correlation between the measles, mumps and rubella (MMR) vaccine and autism (Wakefield et al., 1998). Despite that the medical community has confirmed the absence of such a correlation in further studies, the study still incited misinformation and mistrust towards vaccines among lay people (Baker, 2008). The latest iteration of anti-vaccine movements is the resistance to COVID-19 vaccines. Smith and Reiss (2020) contend that such movements have developed rapidly and have the potential to do enormous harm. The current study aims to contribute to the understanding of anti-vaccine movements as a social phenomenon through a close examination of the patterns of language use in a specific online community on the website of the anti-vaccine broadcast "The Highwire". Adopting a framework of corpus-assisted discourse analysis, the study starts with an exploratory approach by identifying the salient themes and recurrent language practices in the anti-vaccine community, it then focuses its lens on community members' language practices that construe COVID-19 and vaccines.

2. Background

In the past, the repository of medical information rests in the medical community, whereas at present, through digital media, medical information is easily accessible by lay people (Hussain et al., 2018). Digital media plays a crucial role in anti-vaccine movements, as it enables the efficient dissemination of medical misinformation on a large scale (Numerato et al., 2019). Under such a context, online platforms such as Facebook, Twitter and discussion forums have become the main source in studies on anti-vaccine discourse. This section starts by laying out the recurring themes in anti-vaccine discourse, it then focuses on the linguistic patterns that characterise anti-vaccine discourse

2.1 Recurring themes in anti-vaccine discourse

There are three recurring themes in anti-vaccine discourse that are relevant to the current study. The main theme is that vaccines are extremely dangerous. They are believed to cause autism, idiopathic diseases and even death (Zimmerman et al., 2005; Tangherlini et al., 2016). Such adverse effects are not discussed in terms of probability, but as an indisputable result of vaccination (Senier, 2008). A second theme is that the reason the truth about vaccines is not spoken about is connected to the economical profits of the pharmaceutical companies, frequently referred to as the 'big Pharma' (Wolfe et al., 2002). The government and scientists are also after economic profits and are also involved in concealing the truth (Wolfe et al.,2002). A third theme is that the obligation to vaccinate is seen as violating civil liberties (Zimmerman et al., 2005). Tracking of individual's vaccinations is seen as an intrusion of citizens' lives (Wolfe et al., 2002).

Smith and Reiss (2020) examined the latest iteration of anti-vaccine movements that concerns COVID-19 vaccines through a content analysis of the anti-vaccine broadcast "The Highwire". They uncovered that the three themes are reiterated in discussions of both COVID-19 measures and COVID-19 vaccines. For example, COVID-19 measures such as quarantining and wearing masks and COVID-19 vaccines are both considered as violating civil liberties. The current study builds on Smith and Reiss's (2020) content analysis of the anti-vaccine broadcast and examines the comment sections under the broadcast, aiming to uncover community members' subjective accounts and reiterations of anti-vaccine opinions as well as their recurrent language practices.

2.2 Linguistic patterns in anti-vaccine discourse
Studies that focus on linguistic patterns that characterise anti-vaccine discourse have found two contrasting practices: appealing to rationality and appealing to emotions.

The practice of appealing to rationality is highlighted in studies that find characteristics compatible with that of scientific discourse, such as: presenting statistical data and official documents (Simona & Maria-Steluța, 2018; Maci, 2019); using a technical register, involving technical vocabularies and grammatical features such as nominalisation (Richardson, 2003). Both serve to present the information as scientific-looking as possible to the readers. This is also reflected in studies that find frequent references to the one's expert status or the author of the quoted sources' expert status,

such as a scientist or a doctor, to demonstrate the reliability of the information (Richardson, 2003; Myers, 2004). In addition, Myers (2004) and Orpin (2014) also identified differences to scientific discourse in that anti-vaccine discourse tends to express claims in greater certainty, usually dropping hedges in the original texts.

The practice of appealing to emotions is highlighted in studies that find a consistent co- occurrence of narratives of children injured by vaccines and images of children on anti-vaccine websites, Facebook pages and tweets (e.g., Wolfe et al., 2002; Simona & Maria-Steluța, 2018; Maci, 2019). Employing a synergy between corpus linguistics and multimodal critical discourse analysis in examining anti-vaccine tweets, Maci (2019) uncovered uses of accusations such as 'murder' alongside compilations of images of children died of vaccination, usually gazing at the camera. In addition, Maci (2019) also uncovered uses of possessive pronoun in 'murder our children', which serves to establish a relationality process with the readers who are also parents. The use of possessive pronouns is also discussed as a discursive strategy that separates 'us' and 'them' in Numerato et al. (2019), in which 'our children' is being 'poisoned and sterilised' by 'them' – the government and the elites, inciting strong emotions towards the opposing side.

3. Methodology

3.1 Data collection and processing

"The Highwire" (https://thehighwire.com) is an anti-vaccine broadcast based in the US with more than 40 million views. It posts one episode (approx. 2h) and several short videos (approx. 4-6 mins) on its website each week. Under each video, there is a section that enables viewers to post comments, though they could not comment on other viewers' comments. Since the contents produced by "The Highwire" are censored on platforms such as Youtube, Facebook and Twitter, the comment sections on the website have become an important space for the viewers to express their opinions and formed a discourse community. The data consists of the comments under videos posted from July 31st to December 11th 2020, because the website only displays the videos and their comment sections posted in the last 20 weeks. The study supposes that viewers perceived the comment sections of the website as a public space and that it is their intention to express their opinion. A total of 3,124 comments were collected. Nine of the comments were removed as they were duplicated comments posted by the same user under the same comment section. Spelling errors and spelling variations

were corrected and standardised manually. Abbreviations were changed to full words, e.g., 'u' to 'you', 'yrs' to 'years'. The 3,115 comments were saved in separate text-only files and named according to the username of the viewer, the video title and the post date. The 3115 comments (112,514 running words) form a specialised corpus for the study of patterns of language use in this anti- vaccine community.

3.2 Analytical framework and procedures

The current study adopts corpus-assisted discourse analysis as its analytical framework, employing computational tools and statistical measures to the study of language patterns in a large body of texts (Baker, 2006). It starts with an exploratory approach and generates a list of keywords of the specialised corpus to identify the salient themes and recurrent language practices in the anti-vaccine community. It then targets certain keywords and examines their collocation profiles and concordance lines in detail to identify community members' language practices that construe COVID-19 and vaccines. The detailed analytical procedures and statistical choices are outlined below.

The analysis was conducted in LancsBox 5.1.2 (Brezina et al., 2015). It starts with generating the keywords for the specialised corpus. Keywords are defined as words that are considered as statistically salient in a text compared to a reference corpus (Baker, 2004). They point researchers to important concepts in a text and lay the foundation for further analysis of the discourses embedded in the text (Baker, 2004). The Brown corpus (1,023,634 words) (Francis and Kucera, 1979) is chosen as the reference corpus based on the fact that it is compatible with the specialised corpus in the sense that both consist of texts in American English. Log Ratio, a combination of a measure of statistical significance and effect size (Hardie, 2014), was chosen as the measure for keywords. To ensure the keywords are reflective of the recurrent patterns in the comments, keywords that occur in less than 2 percent of the comments were excluded. In addition, to enable further thorough qualitative analysis, only the top 30 keywords were considered.

The top 20 keywords were grouped into thematic categories and those that are considered as most helpful in identifying community members' language practices that construe COVID-19 and vaccines are subjected under collocation analysis. Collocations are defined as repeated co- occurrences of

words (Baker, 2006). As two words co-occur repeatedly in discourse, they form a close association

that constructs discourse in a certain way (Baker 2006). MI3 was used as a statistical measure for

collocates as it takes both exclusivity and frequency into consideration (Brezina, 2018). The analysis

employs a span of 5 words to the left and right side of the word. To keep the amount of collocates

manageable, the analysis employed a frequency threshold of 5 times and only the top 10 collocates

were considered.

3.3 Advantages and limitations

The main advantage of the analytical framework is that it enables the identification of recurring

linguistic patterns that are diffused in a large body of texts. The quantification of the linguistic patterns

also offers a relatively objective way of uncovering dominant language practices in the texts (Baker,

2006). However, it is important to acknowledge that corpus methods cannot completely remove

researchers' subjectivity. Researchers still need to select statistical measures and thresholds as well

as complete further interpretations of the linguistic patterns (Baker, 2006).

4. Analysis

4.1 Overview

The top 30 keywords and their statistical information are displayed in Appendix. After examinations

of the concordance lines, the keywords are categorised into 6 thematic categories (see Table 1).

This provides us with an overview of what is being talked about and what people are doing in the

comment sections. The keywords indicate discussions of both COVID-19 and measures under the

pandemic as well as vaccines, with frequent reference to medical professionals, political figures and

business figures. Commenters frequently express their appreciation towards the content of the

broadcast, with frequent reference to 'del' – the host of the broadcast and 'highwire team' – the

broadcast team. Commenters also frequently express their plans in watching the videos from the

broadcast and sharing them. The keyword 'I'm' in particular, suggest a dominant personal voice in

the comments. The following analysis examines the collocation profile of the keywords in the first

three categories.

Table 1 Keywords in thematic categories ranked in order of log ratio value, with frequency in brackets

COVID-19	'covid' (209), 'masks' (148), 'mask' (142), 'virus' (222), 'test' (98), 'stop' (91)
Vaccines	'vaccines' (209), 'vaccine' (397), 'flu' (145), 'truth' (167), 'health' (124)
People and institutions	'fauci' (100), 'trump' (125), 'del' (504), 'media' (100), 'gates' (90), 'doctors' (78), 'highwire' (105) , 'team' (80), 'everyone' (81), 'dr.' (141)
Appreciation and politeness	'thank' (363), 'thanks' (118), 'love' (168), 'please' (216)
Broadcast-related	'video' (109), 'videos' (210), 'watch' (114), 'share' (71)
Grammatical	'I'm' (196)

4.2 Framing COVID-19: constructing expertise

Discussions surrounding COVID-19 and measures under the pandemic are both framed under mistrust and a topic of civil liberties, with frequent attempts in constructing one's expertise. One of the most prominent grammatical collocates of 'covid' is the pronoun 'they' (27), used as a generic reference to the medical community, the government, people who are pro-vaccine, or a combination of these. Extracts containing 'they' are presented below. *The Highwire* commenters repeatedly present 'they' as the agent of actions that are considered problematic and/or harmful and seek to discredit what 'they' think and say. For example, in 'have they ever isolated covid 19' and 'the virus they detect is not even covid', the commenters directly question the competence of the mainstream medical community, conveying a sense of mistrust as well as placing themselves as experts. In 'they would just let him die', the commenter construe the mainstream medical community as those that are out to harm people. This repeated framing of a complex group of social actors as 'they' forms an opposition between the anti-vaccine community and the rest.

The lexical collocates of 'covid' (see Table 2) demonstrated a discussion of both the COVID-19 measure and COVID-19 vaccines. The collocates 'death', 'rate' and 'deaths' occur in the compounds 'covid death rate' (4) and 'covid deaths' (6). They are used in contexts that accuse the officials of exaggerating the seriousness of the pandemic through lying about the number of deaths. Seven of 'covid death rate' and 'covid deaths' co-occur with statistics reported by the officials that are then discredited. By drawing from scientific discourse and then discrediting such information, the commenters again place themselves as experts.

Table 2 Top 10 lexical collocates of 'covid'

L	'mandatory' (7), 'symptoms' (5)
R	'vaccine' (21), 'deaths' (9), 'death' (9), 'tests' (8), 'vax' (6), 'isolated' (5), 'rate' (5), 'test' (7)

The compound 'mandatory covid vaccine/vax/tests/test' (7) occurs in contexts that consider them as a violation of civil liberties. This topic in vaccines continues in the contexts surrounding the keywords 'masks' and 'mask'. Extracts containing 'wear masks' are presented

below. Phrases such as 'force to' and 'have to' are frequently combined with 'wear a mask' and 'wear masks'. In these cases, commenters places themselves or other people such as 'drivers' and 'kids' in a powerless position that has no control over their own choices. The comments can be categorised as complaints or accusations, in which commenters or other people are construed as victims and the government as the abuser.


```
         Our country is forcing employees to wear   masks   in work places where it is very
              Our Bus Drivers are forced to wear   masks   even though they have plastic shields separating
           drivers continue to be forced to wear   masks.  I reported this accident and expressed my
                  that is not forcing guests to wear   masks   and as a result getting bad reviews
                 in England kids are having to wear   masks   that 's ABUSE good show
         Healthy persons who are having to wear   masks   are becoming Not Well. I do n't
    ignorant by following the REQUEST to wear   masks!  I also told Robin Vos office this
```

4.3 Framing 'vaccine' and 'vaccines': scientific grounds and personal tone

Discussions surrounding vaccines are constructed under a mixture of scientific grounds and personal voice. The keyword 'vaccine' in singular form (209) is either used to refer to a type of vaccine in particular or used as a premodifier in a noun phrase (e.g. 'vaccine manufacturers' (9), 'vaccine ingredients' (6), 'vaccine injury' (5)). The Keyword 'vaccines' in plural form (397) is used to

refer to vaccines in a general sense. The top 10 collocates of'vaccine' and 'vaccines' are displayed

in Table 3.

Table 3 Top 10 lexical collocates of 'vaccine' and 'vaccines

		'vaccine' (397)		'vaccines' (209)	
	Shared collocates	Different collocates	Shared collocates	Different ollocates
L	'flu' (24), 'covid' (21)	'get' (19), 'got' (13), 'received' (7), 'reaction' (6)	'flu' (9),'covid' (7)	'bad' (5), 'need' (6), 'dangerous' (5)
R	'safe' (10)	'manufacturers' (10), 'ingredients' (9), 'injury' (5)	'safe' (5)	'do' (11), 'just' (8), 'children' (5), 'years'

The shared collocate 'safe' (10) and its concordance lines demonstrate a negative evaluation that there are no safe vaccines. The phrases such as 'there is no such thing as' and 'so called' are juxtaposed with 'safe vaccine(s)'. High percentages such as 95% were drawn upon in arguments that because vaccines cannot be 100% safe, they are problematic. This demonstrates that commenters are familiar with the probabilistic information and the scientific modelling of vaccine risks; however, although commenters seem to construct arguments on scientific grounds, they take up and propagate worst-case thinking that is based on fear

Collocates unique to the singular form 'vaccine' point to expressions of personal stances and references to vaccine trials. The collocate 'get' (19) is used most frequently in comments that express a personal stance, such as 'I'll never', 'I'm not willing to', or 'I have no plan' to get the COVID-19 vaccine, expressing a firm resistance. In contrast, the collocate 'got' (13), 'received' (7) and 'reaction' (6) are largely associated with references to vaccine trials. These include highly specialised vocabulary in empirical experiments such as 'threshold', 'control' and 'placebo' and references to statistical information. In these cases, commenters are questioning the validity of vaccines through questioning the experiment design in vaccine trials, with the source of the information coming from medical professionals outside the mainstream. This demonstrates that although commenters seem to construct their arguments on scientific grounds, their sources of information are cherry-picked from those that agree with their view.

threshold for those that got their experimental	vaccine	and the cycle threshold set to 45
no challenge test? How many got the	vaccine	and how many got the control? Do
only 10% of the 43538 got the	vaccine	then in terms of catching covid19 they
10% of the group actually got the	vaccine	candidate and hence 90% got the placebo.
basically, about 4,400 people… Worldwide… Got this	vaccine	candidate. That means 39,600 for the placebo.
Sonnie Bates, who refused to take the	vaccine	and got a dishonorable discharge from the
of people who got it. But in	vaccine	trials, it means something different.
died. Their daughter got meningitis from that	vaccine.	Keep repeating until

placebo… But wait. 0.205% that received the	vaccine	still contracted it and 0.209% in the
with Covid(and recovered) or received the	vaccine	this will NEVER end. How are they
all been normal .Have I received a	vaccine	implanted with something that did not take

Collocates unique to the plural form 'vaccines' point to a concern for harmful effects and arguments against vaccines in general. This is highlighted in the collocates 'bad' (5) and 'dangerous' (5), which are used in evaluations of vaccines. The collocate 'years' was used in arguments against vaccines drawing from either personal experiences or a broader perspective of

human history. In the first case, the commenters argue that they spent 'x years' of their life without any vaccine and are perfectly fine. In the second case, the commenters argue that humans have spent 'thousands of years' without any vaccine and are perfectly fine. The collocate 'children' (5), in light of these evaluations and arguments, are used in contexts that express heightened emotions of anger, characterised by accusations as well as uses of exclamation markers and capital letters. In 4 of the 5 occurrences, 'children' is premodified by the collective possessive pronoun 'our'. The use of 'our' activates all parents in this anti-vaccine community, creating a sense of belonging together.

4.4 Framing 'truth' and social actors: forming associations

The keyword 'truth' (167) refers generally to a set of alternative facts such as the harmful effects and the economical profit behind vaccines. It is frequently presented in parallel with other concepts using the connective 'and' (27). Extracts containing examples of this are presented below. It could be seen that 'truth' frequently co-occurs with valuable qualities such as 'leadership', 'courage' and 'integrity' and important concepts such as 'health' and 'free speech'. These frequent co-occurrences of concepts create entrenched associations between the set of alternative facts and positive concepts, construing them as something valuable and important. 'Truth' is also parallel to 'common sense'. This establishes the set of alternative facts as something that should be known by the majority of people, construing them as mainstream within the anti-vaccine community.

```
                    thank God we have a Health and  Truth  soldier like you fighting for us. Very
proving that REAL leadership, courage, integrity, and  truth  are very much alive! Thank you for
             He is part of your preparedness and  truth  solution.
                  a rebut to the common sense and  truth  of it all. They are so hell
      in our nation, information, free speech and  truth.  God bless you all. Your Jewish fan
```

A similar strategy is used in grouping social actors together in the comments. The keyword 'fauci' (100), referring to Dr. Anthony Fauci, the main adviser to the US president on health issues, is frequently represented along with a number of other social actors and institutions through the connective 'and' (26). Extracts containing examples of this are presented below. It could be seen that 'fauci' often occur with 'big pharma' and 'bill gates', reiterating the idea embedded in anti-vaccine discourse that scientists are the same as 'big pharma' in that are

also corrupted by economic profits. Phrases such as 'Fauci and his hitler youth' and 'Fauci and the children in power' in particular, construe Fauci as the lead in the conspiracy, the government as mindless and unintelligent followers. These co-occurrences and the associations they form serve to discredit medical professionals and the government.

swamped brains of people the likes of	Fauci	and any other puppet of the cha-ching
or his ilk. Sociopaths, like Gates and	Fauci,	and big business, big Pharma or even
Tony	Fauci	and Bill Gates should be tried for
Dr. Judy Mikovitz. She worked with both	Fauci	and Birx and claims that her lab
n't force a vaccination. That is why....Dr.	Fauci,	and CDC, Bill Gates and the WHO
a pathway to an AIDS cure but	Fauci	and friends confiscated all her work and
does not bode well for people like	Fauci	and Gates.
about COVID-19 and clearly he has failed.	Fauci	and Gates have proven to be more
men intervene... How stupid are we really??!!	Fauci	and Gates have the permission to go
the vaccine and therefore a target for	Fauci	and Gates. It became the subject of
money. Trump has nothing to gain financially.	Fauci	and his cohorts stand to make millions
Atlas is a genius and will crush	Fauci	and his equally corrupt sidekick, Birks. Del
please address the fact that if they(	Fauci	and his hitler youth) can n't test
companies who stand to make billions gives	Fauci	and his ilk more legitimacy than their
Why are n't	Fauci	and his wife in jail for violating
the mark of the beast. Controlled by	Fauci	and no man has the latest technology
	Fauci	and PA Gov. Wolf need to be
the process. At the very least fire	Fauci	and put one of these front line
	Fauci	and the CCP planned this fake pandemic.
but considering the mentality of people like	Fauci	and the children in power that we

The keyword 'dr.' (142) is used as an important marker of expertise for social actors. It is frequently used in front of the names of the cherry-picked scientists who hold a view that is compatible with the anti-vaccine view. It can thus be used to emphasise their expertise and increase the reliability of the presented view. In referring to experts they disagree with, e.g. Dr. Andrew Fauci, 'dr.' is often dropped. In the 100 occurrences of 'fauci', only 10 has 'dr.' in the front.

5. Discussion

Content-wise, similar to Smith and Reiss's (2020) discovery in terms of the broadcast, the comments under the broadcast, regardless of whether its topic is on COVID-19 measures or vaccines, are also framed under a mistrust towards the government and medical professionals as well as a topic of civil liberties.

Language-wise, similar to the language practices – appealing to rationality and appealing to emotions – identified in past studies, the current study uncovered a mix of scientific voice and

personal voice. As uncovered in past studies such as Richardson (2003), Simona & Maria- Steluța (2018) and Maci (2019), commenters made use of statistical data and a technical register, serving to discredit information from the government and medical professionals and place commenters themselves as the experts. Commenters also frequently refer to one's expert status to demonstrate the reliability of the information (Richardson, 2003; Myers, 2004). They do it selectively, adding the title 'dr.' only in front of the scientists they agree with. The current study also uncovered differences from the scientific discourse. For example, commenters build their arguments on scientific grounds in terms of vaccine risks, yet appeal to worst-case thinking that is based on fear. In addition, comments take on a strong personal voice, characterised by frequent use of 'I'm', often referring to personal experiences as evidence in the arguments. Comments including 'our children' and the generic reference 'they' for governments, medical professionals and those who are pro-vaccine, in juxtaposition to the accusations, also actively create a separation between 'us' and 'them' and incite strong emotions (Numerato et al., 2019).

A further discovery is that the use of parallel structure – connecting an array of concepts or social actors using 'and' – can be used in forming associations. For example, the juxtaposition of alternative facts with concepts such as 'leadership', 'courage' and 'integrity' frames it as valuable and important, discrediting and marginalising the mainstream information.

6. Conclusion

The current study has examined the salient themes and recurrent language practices in a particular anti-vaccine community. An advantage of choosing comments on "The Highwire" website is that it is uncensored compared to comments on major social media platforms, which enable us to gain a full picture of anti-vaccine discourse. A limitation is that as the data consists of the discourse of only one anti-vaccine community, the results could not be generalised to all anti-vaccine communities. Further study could examine the comments in combination with the broadcast content, focusing on how anti-vaccine views from the broadcast are take up and reiterated in the comments.

Reference

Baker, J. P. (2008) 'Mercury, Vaccines, and Autism One Controversy, Three Histories', *Public Health Then and Now*, 98(2), 244–253.

Baker, P. (2004). 'Querying Keywords', *Journal of English Linguistics*, 32(4), 346-

359. Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.

Brezina, V. (2018). Statistical choices in corpus-based discourse analysis. In C. Taylor & A. Marchi. *Corpus Approaches To Discourse* (pp. 259–280). Routledge.

Brezina, V., McEnery, T. & Wattam, S. (2015). 'Collocations in context: a new perspective on collocation networks.', *International Journal of Corpus Linguistics*, 20(2), 139-173

Francis, W.N. & H. Kucera. (1979). *Brown Corpus Manual*. Providence: Brown University.

Hardie, A. (2014). Log ratio – An informal introduction. *ESRC Centre for Corpus Approaches to Social Science*. Available at: http://cass.lancs.ac.uk/log-ratio-an-informal-introduction/ (Accessed: 13th January, 2021).

Hussain, A., Syed A., Ahmed, M. & Sheharyar, H. (2018). 'The Anti-vaccination Movement: A Regression in Modern Medicine', *Cureus*, 10(7), e2919.

Kata, A. (2012). 'Anti-vaccine activists, Web 2.0, and the postmodern paradigm – An overview of tactics and tropes used online by the anti-vaccination movement', *Vaccine*, 30(25), 3778–3789.

Maci, S. (2019). 'Discourse Strategies of Fake News in the Anti-vax Campaign', *Languages Cultures Mediation*, 6(1), 15–43.

Myers, G. (2004) *Matters of Opinion: Talking about public issues*. Cambridge: Cambridge University Press.

Numerato, D., Vochocová, L., Štětka, V., & Macková, A. (2019). 'The vaccination debate in the "post-truth" era: social media as sites of multi-layered reflexivity', *Sociology of Health & Illness*, 41(S1), 82–97.

Richardson, K. (2003). 'Health risks on the internet: establishing credibility on line', *Health, Risk and Society*, 5(2), 171-184.

Orpin, D. (2014). *Discussing Science in the Public Sphere: A Corpus-assisted Study of Web-based Interaction Concerning the Measles Mumps and Rubella (MMR) Triple Vaccine* (Unpublished doctoral dissertation). University of Birmingham, Birmingham.

Senier, L. (2008). '"It's Your Most Precious Thing": Worst-Case Thinking, Trust, and Parental Decision Making about Vaccinations', *Sociological Inquiry*, 78(2), 207–229.

Smith, T.C. & Reiss, D.R. (2020). 'Digging the rabbit hole, COVID-19 edition: anti-vaccine themes and the discourse around COVID-19', *Microbes and Infection*, 22(10), 608–610.

Simona, V., & Maria-Steluța, S. (2018). 'Vaccines: saving lives or depopulating the world? A discourse analysis', *Journal of Comparative Research in Anthropology and Sociology*, 9(1), 67–87.

Tangherlini, T. R., Roychowdhury, V., Glenn, B., Crespi, C. M., Bandari, R., Wadia, A., Falahi, M., Ebrahimzadeh, E. and Bastani, R. (2016). '"Mommy Blogs" and the Vaccination Exemption Narrative: Results From A Machine-Learning Approach for Story Aggregation on Parenting Social Media Sites', *JMIR Public Health and Surveillance*, 2(2), e166.

Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., Berelowitz, M., Dhillon, A. P., Thomson, M. A., Harvey, P., Valentine, A., Davies, S. E. and Walker-Smith, J. A. (1998). 'Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children', *Lancet*, 351(9103), 637-641.

Wolfe, R. M., Sharp, L. K. and Lipsky, M. S. (2002). 'Content and Design Attributes of Antivaccination Web Sites', *Journal of the American Medical Association*, 287(24), 3245–3248.

Zimmerman, R. Z., Wolfe, R. M., Fox, D. E., Fox, J. R., Nowalk, M. P., Troy, J. A. & Sharp, L. K. (2005). 'Vaccine Criticism on the World Wide Web', *Journal of Medical Internet Research*, 7(2), e17.

Appendix: Top 30 keywords in order of the log ratio value			
Type	Frequency (raw frequency)	Dispersion (range in percent)	Log Ratio
covid	209.000000	5.393258	11.8927522212532
vaccines	209.000000	5.040128	11.8927522212532
vaccine	397.000000	8.860353	11.8185359753416
videos	109.000000	3.113965	10.95357738046
highwire	105.000000	3.178170	10.8996386201527
fauci	100.000000	2.664526	10.8292493154504
trump	125.000000	2.760835	10.1513251161779
video	210.000000	5.842697	9.89978629352832
masks	148.000000	3.531300	8.81003168150759
del	504.000000	15.280899	8.70337566813048
flu	145.000000	3.210273	7.36543138988625
virus	222.000000	4.686998	7.27949420828461
mask	142.000000	3.402889	7.16534652421775
thank	363.000000	10.690209	6.51942511373803
media	100.000000	2.953451	6.12893455072065
gates	90.000000	2.311396	5.77048366322216
please	216.000000	6.099518	4.98621529622651
thanks	118.000000	3.595506	4.85871452561981
doctors	78.000000	2.118780	4.51672770615333
watch	114.000000	3.531300	3.67856431315564
truth	167.000000	4.526485	3.5919475277723
health	124.000000	3.338684	3.42547472452015
team	80.000000	2.471910	3.09806050251766
everyone	81.000000	2.407705	2.91066433539891
test	98.000000	2.407705	2.90541544941982
I'm	189.000000	4.847512	2.86512266372192
stop	91.000000	2.600321	2.78642751976678
dr.	141.000000	3.434992	2.77818031397777
share	71.000000	2.118780	2.7205615141103
love	168.000000	4.686998	2.70747641971771