


Please cite the Published Version

Heritage, Frazer  (2022) Public and private discourses of lesbians: exploring the discourses surrounding lesbians in the Spoken BNC2014 and the British press in 2017. *Journal of Corpora and Discourse Studies*, 5. pp. 61-89.

DOI: <https://doi.org/10.18573/jcads.18>

Publisher: Cardiff University Press

Version: Published Version

Downloaded from: <https://e-space.mmu.ac.uk/632852/>

Usage rights:  [Creative Commons: Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Additional Information: This is an open access article published in *Journal of Corpora and Discourse Studies*, by Cardiff University Press.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

PUBLIC AND PRIVATE DISCOURSES OF LESBIANS: EXPLORING THE DISCOURSES SURROUNDING LESBIANS IN THE SPOKEN BNC2014 AND THE BRITISH PRESS IN 2017

FRAZER HERITAGE
BIRMINGHAM CITY UNIVERSITY

CITATION

Heritage, F. (2022). Public and private discourses of lesbians: Exploring the discourses surrounding lesbians in the Spoken BNC2014 and the British press in 2017. *Journal of Corpora and Discourse Studies*, 5: 61-89.

This paper explores the language used to discuss lesbians in both informal spoken British English and British written news media. The research draws on two data sources: the Spoken BNC2014 and a specialised corpus of 1.2 million words from newspapers written between January 2017 and December 2017 inclusive. Discourses (re) produced in both informal conversations and the mass media are ways of sustaining and maintaining ideological stances towards groups marginalised due to their sexuality (see Baker, 2014; Gupta, 2016). In this paper, I argue that the discourses surrounding lesbians found within both corpora are problematic and archaic. I discuss two prominent discourses in spoken British English (that lesbians are typically seen as masculine, which in turn is judged negatively, and that lesbianism is a choice), two prominent discourses in written British English (that lesbians are a threat to children and that lesbianism is a marked identity which is delegitimised), and one discourse which overlaps across both corpora (that all gender, sexual, and romantic minorities are a homogenised group). The discourses in both corpora appear to be damaging to how lesbians are viewed in general society and tend to mask underlying homophobic ideologies.

KEYWORDS

gender and sexuality; lesbians; queer linguistics; representation; media discourse; British English

CONTACT

Frazer Heritage, Department of Criminology and Sociology, School of Social Sciences, Birmingham City University, 4 Cardigan Street, Birmingham, B4 7BD, UK.
frazer.heritage@bcu.ac.uk

DOI

10.18573/jcads.18

ORCID

0000-0002-2788-3208

ISSUE DOI

10.18573/jcads.v5

LICENSE

© The authors. Available under the terms of the CC-BY 4.0 license

Public and private discourses of lesbians: Exploring the discourses surrounding lesbians in the Spoken BNC2014 and the British press in 2017

Frazer Heritage
Birmingham City University

1. Introduction

This paper explores the discourses surrounding the search term *lesbian**¹ within the Spoken BNC2014 and a specialised corpus of British newspaper articles compiled around the search term. Throughout this paper, the acronym ‘GSRM’² is used instead of acronyms like ‘LGBTQ+’ (*Lesbian, Gay, Bisexual, Transgender, Queer, and more*), as ‘GSRM’ appears to be more encompassing of all identities³ (the exact origin of the term is unknown, and not yet pervasive in academic literature, though see for example Skinner, 2016; Phelps, 2017). Although this paper is particularly concerned with the representation of lesbians (as a sexual and romantic minority), it hopes to provide a starting point for change in the language used about GSRM individuals in a more general sense.

This study draws on data from two corpora: the Spoken BNC2014 (Love et al., 2017; Love, 2018; 2020) and a specialised corpus of 1.2 million words compiled of British national newspapers published between January 2017 and December 2017 inclusive (hereafter, Lnews17). Throughout this paper, the search term *lesbian** is used, which is shown in italics to highlight the difference between the search term and discussion of lesbians as a collective noun. The proceeding asterisk is also included to signify the inclusion of variations of the term *lesbian* in the search, such as *lesbians* and *lesbianism*.

The term ‘lesbian’ originates from the Greek *Lesbos*, which was the home of Sappho, who was known for her erotic and romantic verse about women. The etymology of the word, thus, has a history of discussing women with reference to their sexual practices.

-
- 1 Similar to Baker (2005) I recognise the issues of the term *lesbian* to refer to gay women because it is also a noun. However, similar to Baker’s work, the author was unable to find an adjectival alternative. Therefore, the term lesbian is used based on a series of compromises and personally subjective decisions which reflect existing inconsistencies in the ways that terms to refer to different sexual groups are currently used.
 - 2 I use both terms “sexual” and “romantic” in order to account for lesbians who may be asexual or lesbians who may be aromantic
 - 3 Some have argued for the use of GSRI (Gender, Sexual, and Romantic Individuals), but while I respect that GSRI does indeed highlighting how personhood is always at the very centre of any debate around gender and sexuality, I would argue that by marking the term as an ‘individual’, it removes why it is important to look at lesbians: mainly because they are a minority group who face systemic power imbalances. A GSRI could, in theory, also relate to a heterosexual man who is not a member of any minority categories.

The nature of the term signifies two things: first, that a lesbian is a woman; second, that these women are sexually and/or romantically attracted to other women. In present-day English, nouns like *lesbian* are often used interchangeably with adjectival modifiers and a noun denoting gender identity such as *gay woman*.

Previous research has shown that the word *gay* has been used as a derogatory adjective in constructs like *that's so gay* (for example see Armstrong, 1997; Lalor and Rendle-Short, 2007; Woolley, 2013). By contrast, terms like *lesbian* do not appear to have the same specific pejorative usages. As Armstrong (1997: 329) argues “young people have expropriated the preferred neutral sexual identity marker for **males with a homoerotic orientation** and, through connotative extension, given it a negative value” (my emphasis in bold). This same process of expropriation does not seem to have occurred with lesbians, however. Though, it is imperative to highlight the wealth of literature which has raised attention to the language which is used to discriminate against GSRM individuals (see, for example, Gabriel, 2017). As well as other GSRM identities, gay men and lesbians are often targeted with derogatory language, and insinuation of a non-heterosexual identity can be used as an insult (see Brown and Alderson, 2010).

Discussions about lesbian individuals, therefore, are not just discussions about sexuality, but they are also discussions about gender identity and the point at which these identities intersect. This paper draws on feminist and queer theory to explore how lesbians are both spoken about in private contexts by and among assumed non-lesbians, and represented in mass media. The following section highlights a dearth of corpus research into the representation of lesbians, a gap which this paper aims to contribute to filling.

The paper is structured as follows: first, the paper is situated in the broader scholarly fields of language, gender and sexuality, and corpus-based studies of language, gender and sexuality. This is followed by an outline of the corpora and corpus software used in this investigation. This leads to a discussion of the findings: I start by discussing the findings from the Spoken BNC2014 before examining the findings from Lnews17. Once the findings from each corpus have been presented separately, I present a discourse which overlaps across the corpora for a more comprehensive image of ideologies towards lesbians across both informal and public settings. Finally, the paper concludes by drawing attention to the implications of the findings and discusses some possible avenues of future research.

2. Situating the research

While there are plenty of corpus-based studies which analyse the representation of gender in the mass media, corpus analyses which examine the representation of sexuality within this context are comparatively rarer (though, see Bachmann, 2011; Baker, 2005; 2012; Motschenbacher, 2018; Bailey, 2019; Wilkinson, 2019; see also Zottola, 2018 for an

overview of corpus research into transgender identities). In this section, the current research paper is situated within the broader field of discourses used by and about GSRM individuals. I discuss how previous research has used an array of techniques to not only analyse the language used by and about lesbians but also about GSRM identities in a more general sense. The focus of this section is then narrowed to look particularly at corpus studies of sexuality and demonstrate their fruitful nature while highlighting the dearth of literature which looks specifically at corpus approaches to the discourses around lesbians in both public and private contexts.

There has been a considerable amount of research into lesbian discourse from a linguistic perspective (for example see Koller, 2011; 2013; Morrish and Sauntson, 2011; Jones, 2012; Bailey, 2019). However, the nature of “lesbian discourse” is multifaceted and needs to be compartmentalised in order to be more fully appreciated. Research into “lesbian discourse” encompasses work on multiple topics, including but not limited to: how lesbian groups construct identity (for example, Jones, 2012); how lesbians challenge heteronormative assumptions by making their sexual identity apparent (for example, Land and Kitzinger, 2005), how lesbian characters are written about, especially within erotic literature (Baker, 2005; Morrish and Sauntson, 2011); and how lesbians construct their identity in texts aimed at other lesbians (see Koller, 2011, 2013; Bailey, 2019). A commonality among all this research into “lesbian discourse” is a focus on lesbians as the creators of the analysed texts.⁴ What appears to be rarer, with regards to analyses of lesbian discourse, is how lesbians are represented in informal conversations among assumed non-lesbians and in mass media, where the creators of texts about lesbians may not identify as lesbians.

One of the few studies to explore the language used by assumed non-lesbians via an analysis of a large general corpus comes from Motschenbacher (2018), who explored the representation of sexual identity labels within the Corpus of Contemporary American English (COCA) (Davies, 2010). Motschenbacher points out how the main collocates for the term *lesbian* relate to their identity, relationships, gender, partners, and practices. In particular, he notes that collocates of family and political affiliations play large roles in the construction of a lesbian identity. Motschenbacher looked specifically at COCA alone, which raises questions about how lesbians are represented not just within corpora of British English, but also in informal contexts – where the language may not have undergone rigorous editing. That is to say, the texts in COCA are largely written or written to be spoken, and thus may have been edited to give a more (dis) favourable representation of sexuality.

There is a surprisingly small number of studies which explore the representation of GSRM identities in the British written mass media (though see for example Wilkinson, 2019 for a discussion on the diachronic change in the representation of bisexuality in the

4 In analyses of written material, there is often an assumption that the author(s) are lesbians.

British newspaper *The Times*). Therefore, this section now turns to investigations which have examined the representation of lesbians in other mass media. In particular, there are three closely related studies to the present research. These are Koller's (2011; 2013) investigations of magazines specifically written by and aimed at lesbians and Bailey's (2019) research into how lesbians construct homonormative ideals on websites. In Koller's research, specific texts were taken and analysed in depth from a socio-cognitive perspective, as opposed to using a corpus approach. In her 2011 study, Koller focuses on a group identity construction within a single text. In her 2013 study, she specifically focuses on two texts. Koller argues that the writers of the texts use language to create a collective group identity. Bailey (2019) used both multimodal and corpus-assisted critical discourse analysis in order to explore how lesbian, bisexual, and other queer female writers constructed in-group identities. The research revealed that the language used by these writers typically constructed normativity around young, cisgender, lesbian women.

There is also a growing body of corpus research on the representation of lesbians in pornographic short stories and erotic narrative (see Baker, 2005; Marko, 2008; Morrish and Sauntson, 2011; Wilson, 2012; see Motschenbacher, 2018 for an overview). Typically, the kind of discourses which appeared within the pornography would position lesbian women in heteronormative roles – with one 'butch' lesbian and one 'femme' lesbian, in which the 'butch' lesbian would typically perform gender in a way associated with masculinity, and the 'femme' would perform typically perform gender in a way associated with femininity (see Butler, 1990 for a discussion on gender performativity). However, the function of the narratives is vastly different from the functions of the British press. As characters in pornographic stories, lesbians are expected to be portrayed in explicitly erotic ways. However, this raises questions about how they are seen in a general sense and in contexts where they are not inherently sexualised. Therefore, this paper aims to look at an aspect of the representation of lesbians which has similarities to previous studies, but which has not yet been covered.

I would also argue that the texts analysed by Koller (2011; 2013), Bailey (2019), and Morrish and Sauntson (2011) are different to the kind of media investigated by others such as Baker (2008) and Gupta (2016; 2018). Both Baker and Gupta explore the representation of gender and sexuality in texts which are aimed at the general public. Although the discourses typically portrayed GSRM individuals in negative ways, both Baker and Gupta note the importance of editors of the mass media, which influenced these representations. The participants in the context of a magazine aimed at lesbians, which are likely to be lesbian writers and lesbian consumers, would probably encourage positive representations of lesbians than in contexts where lesbians are not the main discourse participants. With regard to the current investigation, the representation of lesbians is likely to be different in lesbian magazines than in British national newspapers.

Taken together, the research reviewed in this section indicates that there are a number of gaps in corpus research about lesbians. While the gaps are too extensive to list here, two are of particular relevance to the current study. The first gap is that corpus techniques have not yet been employed when investigating the discourses surrounding the search term *lesbian** in British English reference corpora built around informal spoken conversations. Secondly, that the search term *lesbian** has not been explored in corpora compiled of British national newspapers.

These gaps in the literature lead to the following research questions:

- 1) What kind of discourse(s) surrounds the search term *lesbian** in a reference corpus of informal spoken British English?
- 2) What kind of discourse(s) surrounds the search term *lesbian** in a corpus compiled of newspaper articles from the British press?
- 3) What, if any, overlaps are there between the discourse(s) in the BNC2014 and the newspaper corpus?

3. Methodology

The Spoken BNC2014 (Love et al., 2017; Love, 2018; 2020) was used as it was the most up to date reference corpus of informal spoken British English at the time of writing. While it would have been useful to compare the data to the Written BNC2014 (Hawtin, 2019), at the time the research was conducted, it was not publicly available. The Spoken BNC2014 contains 1251 files, all of which were transcribed from private informal conversations by British native speakers. Thus, they represent the private discourses in which lesbians are discussed. These files contain over 1000 hours' worth of audio transcription, which created a corpus of 11,422,617 words. The corpus is hosted online at CQPWeb (Hardie, 2012).

Lnews17 was generated using 'Nexis'⁵. The search term input in the software was "lesbian*", and this was searched for within all UK national newspapers. Nexis automatically removed duplicate texts and newswires. Articles were excluded if they contained any of the following terms anywhere in the text: "Edition 1; Ireland OR Edition 1; Scotland OR Eire Edition OR Ulster Edition OR Edition 2; OR Edition 1; Northern Ireland OR 3 Star Edition". This further removed any duplicate files which were not automatically identified. These terms were used as they have proved successful in other projects (for example, Baker, 2018). Only articles published in 2017 were collected as this provided the most up to date data available at the time of writing. Lnews17 consists of 1.2 million words. There were between 1038 and 1413 files provided for each month. Cumulatively,

5 <https://www.nexis.com> is a professional website which allows for the collection of newspaper articles from both online and print sources.

this totalled 24,601 articles within the corpus. As these articles come from publicly available texts, Lnews17 provides an insight into public discourses about lesbians in 2017.

Wordsmith 7 (Scott, 2016) was used to process the Lnews17 corpus. One of the advantages of Wordsmith 7 is that it added an extra feature to remove any duplicate files which Nexis was unable to detect. The software has multiple functions, but in this project, three are used in particular: generating concordance lines, keyword lists, and collocate lists, all of which were used in the analysis.

Overall, the Spoken BNC2014 provides an insight into how lesbians are discussed within informal conversations, while Lnews17 demonstrates the public discourses about lesbians. The triangulation of these two different data sources provides a comprehensive view of the discourses around lesbians in Britain in both public and private contexts (similar to the work of Anderson, 2016; 2019 on the representation of gender and ageing in public and private contexts).

4. Findings in the Spoken BNC2014

A search for the term *lesbian** in the Spoken BNC2014 only generated 77 results across 38 different texts (6.74 instances per million words). There are various explanations for the frequency at which *lesbian** occurs: firstly, that discussions of lesbians are not salient in casual conversation. Alternatively, it may be that the topic of lesbians was not central to the conversations recorded within the Spoken BNC2014. Given that the Spoken BNC2014 is meant to be representative of informal spoken British English in general, it appears as though the reason for the low occurrence is the former. This is not to say that discussion of gender and sexuality are not spoken about in a more general sense. For example, a search for the term *gay** generated 450 results across 141 different texts (39.40 instances per million words). This higher frequency of *gay** could possibly suggest that speakers are not specifically discussing lesbians, but are discussing GSRM identities in different ways. However, it could also be that many of the ways lesbians are discussed are relatively problematic, and speakers may want to avoid these problematic terms under research conditions. As predicted, some of the cases of *gay** were also used to refer to lesbians in addition to queer men. Although there were occurrences of *gay* being used as an adjective to describe some women, given the limited space, this paper retains a focus on the search term *lesbian**.

The low frequency of this word meant that a collocational analysis would not be viable. Therefore, this section specifically utilises concordance line analysis and discusses the kind of discourse prosodies surrounding *lesbian**.

4.1. *Lesbians, women, and masculinity*

One common discourse occurring in the spoken data was that lesbians are often judged for performing masculinity (12 occurrences in 77 concordance lines). When lesbians are discussed with reference to their performances of masculinity, they are judged negatively (10 occurrences in 77 concordance lines). Some examples of this are presented in bold in the following examples:

- (1) not a **butch lesbian** tattoo on your arm (.) imagine on your wedding day (S8X7 858)
- (2) but **she and she came across quite lesbian** [...] do you s- so you think he's thinking of [name retracted] **cos she's got short hair?** (S8X7 814)

In these examples, both women are judged due to their visual performances of masculinity. In the first example, a performance of masculinity appears to 'ruin' a wedding, and this statement appears to rest on a heterosexist and gender-normative assumption of the roles in which different participants are to take (a bride who typically performs a stylised type of femininity, and a groom who performs a masculine role). While in Example (2), short hair appears to be associated with lesbians. In the second example, the fact that someone is 'coming across quite lesbian' and is marked as doing so because of the shortness of her hair could reinforce heteronormative and problematic ideals about visual performances of gender and sexuality. The use of *lesbian* in these examples bears a resemblance to other slurs based on sexuality, such as *that's so gay*, which appears to indicate that labels for different marginalised sexualities can be used in a derogatory way.

No concordance lines discussed lesbians who perform femininity, which suggests that lesbians who perform masculinity are marked as a specific type of woman and are more 'prototypical' to the speakers in the Spoken BNC2014. Therefore, lesbians who perform masculinity appeared to be marked as disrupting normativities associated with women, and this is usually discussed in negative ways. This finding is interesting given Jones' (2012) previous ethnographic research on a lesbian community of practice. Jones' research showed that lesbians constructed those who perform masculinity as 'authentic' and those who perform femininity as 'inauthentic'. Thus, there appears to be a difference in how certain lesbian communities of practice judge lesbians who perform masculinity and how these performances are viewed by speakers in the Spoken BNC2014, who are presumably not lesbians. Thus, this brings to the fore issues about whether visibility and authenticity in lesbian identities are valued or not, and what social actors are placing value (or not) on this visibility and authenticity.

Although it cannot be said with certainty that the women discussing lesbians are heterosexual, it is possible to suggest that their lack of indexing their identity as lesbians, in addition to their use of negative prosodies around the term, would suggest that they are not lesbians. Furthermore, lesbian speakers in these conversations who are not 'out' are

placed in a double-bind: if they come out as a lesbian or steer the conversation away from discussion of lesbians, the language used about lesbians may only change in that specific context and may not change in conversations between those who do not identify as a lesbian. Indeed, this notion bears some resemblance to work on coming out narratives (such as those by Chirrey, 2003), as well as the work which suggests that lesbians often face negative experiences when they come out (as outlined by Morrish and Sauntson, 2007: 55)

These kinds of discourse appear to be an attempt at policing women's gender performances (see Butler, 1990). By marking lesbians who perform masculinity as negative, women appear to be encouraged to perform femininity. By positioning women who perform masculinity as negative, participants in the conversation may wish to distance themselves from these performances or face similar judgment. Scholars such as Koller (2011; 2013) and Jones (2012) have noted that there often appears to be a binary distinction within lesbian discourses: that lesbians are marked as either 'butch' or 'femme'. In the Spoken BNC2014, this appears to be slightly different in so far that only 'butch' lesbians are recognised, and lesbians who perform different types of femininity are not foregrounded. Thus, lesbians appear to be faced with a quandary: either perform masculinity and make their sexuality clear but be judged negatively, or perform femininity and have their lesbian identity backgrounded but be treated more favourably than butch lesbians.

One finding which does arise from the data is what constitutes a 'masculine' performance from lesbians. In all the concordance lines which discuss lesbians in relation to masculinity, the conversation appears to centre around a woman's negotiation of physical movement or physical attributes (see also, Koller, 2013). For example:

- (3) S0330: you should've erm (.) challenged
 S0331: yeah
 S0330: them to an arm wrestle afterwards
 S0331: well obviously we'd've won
 S0331: any of us (.) V erm any of them
 S0328: I don't know those lesbians were pretty tough
 S0331: the lesbians were a bit tough (.) yeah (...) to wrestle (SUWR 162)

In this example, an exchange between two males (S0330; S0331) and one female (S0328), the physical strength of the lesbians, and their 'tough' nature is used as a basis for humour. However, this raises questions about why these lesbian's identities are seen as the cause of their physical strength. There are many physically strong heterosexual women, but the speakers appear to be subtly implying a causal link between the sexual identity of these women and their physical strength. The judgment placed on the lesbians within this extract is a rather complex issue. Speaker S0328 appears to be using the lesbians as a method of attacking speaker S0331's physical strength, and thus his masculinity.

This attack can be seen as humorous, but the humour masks a problematic ideology: that, assumingly heterosexual, men should be physically stronger than women. Lesbians who perform masculinity present a challenge to this covert ideology: they are women, but their masculine performances prevent men from performing their ideal masculinity. In essence, if a lesbian performs femininity, they are less likely to be perceived as a threat to a man's masculinity.

Taken together, it appears as though the lesbians who perform masculinity are not only marked as authentically lesbian but negatively evaluated, and this can become the basis for attacking a heterosexual men's masculinity.

4.2. *Lesbianism is a choice*

A further common but no less problematic discourse arising from the concordance lines in the Spoken BNC2014 was that lesbianism is some form of flippant choice (15 out of 77 occurrences).⁶ For example:

- (4) S0439: same erm the woman who played Miranda in *Sex and the City*
 S0441: yeah she she's got a kid hasn't she?
 S0439: she was married for fourteen years or something like that and then
 S0441: **turned**
 S0439: **decided to become a lesbian** (.) never know --ANONnameF one day one day it could be you (S6A7 533)
- (5) S0198: not this week is it that changeable?
 S0197: yeah it is changeable
 S0192: yeah she is
 S0198: yeah (.) I look forward to meeting her
 S0192: >> Jewish one week **lesbian the next** (SBTC 1849)

In Example (4), both speaker S0441 and speaker S0439 use verbs which suggest that Cynthia Nixon⁷ flippantly changed her sexuality as easily as she could change her hair colour. This kind of superficial change appears similar to the change implied in Example (5) in which the two speakers are talking about speaker S0192's sister's shallow nature.

Within 13 of the 15 concordance lines which suggest that lesbianism is a temporary state or a choice, the verb indicates a level of choice. These verbs are presented in Table 1.

6 The author would like to draw attention to the fact that while this ideology manifested as a common prosody, they in no way align themselves with the idea that sexuality is a choice.

7 An American actress who is known for her role as Miranda in *Sex and the City*.

Root Verb	Frequency	Concordance lines
Become	3	<ul style="list-style-type: none"> I just might become a lesbian let's not become a lesbian nun decided to become a lesbian
Decide	1	<ul style="list-style-type: none"> she decided she was a lesbian apparently
Get	1	<ul style="list-style-type: none"> the night before Rome you'd get a bit lesbian
Have	2	<ul style="list-style-type: none"> we need to have a lesbian experience god er and she had a lesbian phase
Is	4	<ul style="list-style-type: none"> she wasn't originally a lesbian but now she is thank god you were never a lesbian then his mum split up with his dad and is now a lesbian how there has to be someone for her not to be a lesbian
Pretend	2	<ul style="list-style-type: none"> yeah pretends she pretended to be a lesbian tell when there was a guy in there pretending to be a lesbian
Total	13	

Table 1: Verbs co-occurring with Lesbian*

There appear to be three nuanced discourses which feed into a broader discourse about choice here. The first discourse acknowledges the choice lesbians undertake to 'come out'. For example, "she **wasn't** originally a lesbian but now **she is**" suggests that the person came out after having previously not discovered or disclosed her identity. Although possibly worded in a way which suggests inauthenticity, there is appears to be an awareness for shifting nature of identities. The lexical choice of "originally" suggests that there has been some form of change in the way she identifies. This also bears some resemblance to the work conducted on coming-out narratives, and how the process of 'coming out' is conceptualised as an irreversible change (for example, see Chirrey, 2003). The second discourse appears to position lesbians as choosing their sexuality in a flippant way. For example, "she **decided** she was a lesbian apparently". Finally, the third discourse is even more nuanced and suggests that there is some level of fluidity with sexuality. For example, "we need **to have** a lesbian experience". In each of these verbs, the semantic agency of the women is particularly important: existential verbs such as "is" suggests a lower degree of flippantness in comparison to more active mental process verbs such as "decided" (see Darics and Koller, 2019 for a discussion of semantic agency).

The first of these discourses occurs less frequently in comparison to the second and third discourses, but seems to acknowledge the historical struggles of GSRM individuals,

which could be argued to be a positive representation of GSRM identities in informal conversations. That is to say, that some people may be acknowledging the socio-political struggles of lesbians, and might in turn ultimately vote and become active in a way to put an end to these struggles. However, there are issues with this kind of verb in terms of how it constructs sexuality in relation to temporality (discussed in more detail in the analysis of Example (6)).

By contrast, the second of these discourses is problematic for multiple reasons. The idea that a person can simply “**decide**” to be queer or experience same-sex attraction sustains problematic homophobic ideologies. It delegitimises a lesbian identity as it makes the identity appear to be a flippant statement. However, on a nuanced level, it also adds to the erasure of bisexual and pansexual identities. By not labelling a person with an appropriate label acknowledging sexual fluidity, such as bisexual, it may add to the idea that people are either heterosexual or homosexual⁸ with no other alternatives (Baker, 2008: 149-150; see also Motschenbacher, 2018 for discussion of how sexual identity labels are used in binary opposition to each other). Even if a person does discover that they experience same-sex attraction, this does not necessarily mean they exclusively experience this. The speakers in the BNC2014 appear to assume that women who experience same-sex attraction are lesbians, when they may be bisexual or pansexual.

There did not seem to be much resistance to this discourse, except for one speaker. The conversation is as follows:

- (6) S0285: and it was one of the funniest things to see ever and I remember his dad came the day after erm (.) the day after cabin stripping day cos he had (.) his mum split up with his dad and **is now** a lesbian and got married to
 S0287: >> well she probably was in the first place but
 S0285: well I I don't know
 S0287: >> then realised and then
 S0282: --UNCLEARWORD
 S0285: >> anyway you hear [...] (SE4X 464)

In this example, speaker S0287 seems to attempt to correct speaker S0285's language. The language is not explicitly homophobic but sustains homophobic ideologies through the use of the verb phrase “is now”, which suggests a binary distinction: that one can only be either heterosexual or homosexual at any given point. However, speaker S0285 seems to sustain this idea, despite S0287's attempt to change the way they talk about lesbians. The conversation ultimately reverts to the original topic without this homophobic ideology being resolved.

8 Note, although terms like “homosexual” have been used to medicalise GSRM individuals, here it is used to contrast with the notion of heterosexuality.

The final discourse appears to suggest that sexuality is fluid, such as “the night before Rome you'd **get a bit lesbian**”. This kind of language appears to carry significant implications for queer linguistic research. This use of language seems to blur the categories of sexuality, suggesting that while there is some degree of flexibility and choice, it is not always clear-cut. This kind of language also appears to suggest that different situations can induce different performances of sexuality. Although this has been argued by scholars for a long time (for example, Butler, 1990), this research demonstrates how such concepts are integrated into private discourses. Therefore, this linguistic construction could suggest that members of the general British public have adopted (some of) the notions from queer theory and have begun to view sexuality in a more fluid way, such as the notion that sexuality, like gender, is something which a person performs, rather than being a stable, fixed identity.

Overall, the discourses surrounding the term *lesbian** within the Spoken BNC2014 are relatively negative, with only two concordance lines referencing lesbians in an explicitly positive way:

- (7) [talking about American politicians trying to stop same-sex marriage] muslims support (.) marriage equality (.) as do both of our Muslim elected (.) officials in the United States Congress (.) one even serves as vice chair of the LGBT equality (.) caucus (.) there are many faithful gay and lesbian Muslims in the US and we love and support all of them (SXR9 643)
- (8) We 're doing this film with like these like Japanese lesbian couple that are like they like they 're like voguers and they Madonna recruited them and cos they were over here they were over here doing that and then we like (SXCB 968)

There were also only two examples which explicitly noted the homophobia experienced by lesbians:

- (9) S0530: what are you reading ?
 S0529: don't know something about a police officer punching a woman in the face
 S0530: >> oh (.) I hate police
 S0529: mm (.) me too (.) a lesbian couple saw their dream vacation turn into a nightmare when they were harassed by an off-duty police officer
 S0530: oh my god (S7NV 722)
- (10) S0487: I yeah me and --ANONnameF were like arch enemies and I remember once she like
 S0488: >> was --ANONnameF like?
 S0487: and she thought I was like a lesbian she was like really homophobic and weird erm (STK7 47)

Ultimately, it appears as though in most of the private conversations present in the BNC2014, lesbians are referred to in generally negative ways. It is important to reiterate

that the number of occurrences of *lesbian** within the Spoken BNC2014 was relatively low (a finding that echoes that of Love and Baker, 2015). Thus, to get a better understanding of how lesbians are represented in British English, larger and/or more specialised corpora must be consulted.

5. Findings in Lnews17

The word *lesbian* had over 300 statistically significant collocates within a 5L 5R window ($MI \geq 6$) (see Durrant and Doherty, 2010 for justification of the MI significance threshold). This list was still too extensive for manual processing and to report on given the limitations of space. Therefore, this paper implemented a limit of a minimum of 10 occurrences with an $MI \geq 6$. This yielded a total list of 105 collocates, of which only the top 50 collocates are reported on. These collocates, as organised by MI score, are listed below in Table 2:

Number	Frequency	Word	Number	Frequency	Word
1	27	transgendered	26	13	solomon
2	13	gluckstein	27	26	affair
3	26	hideous	28	123	weidel
4	123	bi	29	793	daenerys
5	793	bisexual	30	11	tina
6	11	fingersmith	31	62	watershed
7	62	fling	32	10	steamy
8	10	simeon	33	64	defamation
9	64	mardi	34	24	applicants
10	24	romps	35	64	vampire
11	64	gras	36	21	luxa
12	21	tryst	37	19	storylines
13	19	tyrion	38	19	knox
14	19	brookside	39	34	painter
15	34	romp	40	55	atomic
16	55	intersex	41	12	rated
17	12	interfaith	42	79	marx
18	79	tea	43	17	eastenders
19	17	transsexual	44	34	ass
20	34	hints	45	22	mick
21	22	strips	46	149	transgender
22	149	lover	47	28	dramas
23	28	jaw	48	17	suspected
24	17	inmate	49	23	calman
25	23	butch	50	85	kiss

Table 2: Top 50 collocates of *lesbian* in *Lnews17*

Broadly speaking, the collocates within Table Two can be categorised into three different groups: GSRM identities, pop culture references, and references to sexual practices. In the first group of collocates, GSRM identities, *lesbian** occurs with words about transgender people, though there are also problematic and politically incorrect collocates, such as *transgendered* (which implies that being transgender is a process that has a completion, rather than an ongoing identity and performance) and *transsexual* (which historically has been used as a synonym for transgender. This term predicates an authentic trans identity as one which relies on surgery and hormones, rather than as an identity). There is some conflation of lesbians and other GSRM identities: issues which may affect transgender people are positioned as also affecting lesbians. Indeed, a common lexical bundle which occurred within the concordance lines of some of these collocates was “lesbian, gay, bisexual, and transgender”. This grouping of lesbians and other GSRM identities will be discussed in more detail in a subsequent section.

In the second category of the collocates, pop culture references, there is a discussion of how lesbians are portrayed in television shows. However, this is usually in regards to sexual practices, such as: “20 book bits they cut from the show, from Daenerys's lesbian sex to Tyrion's hideous crimes [...]” (*The Telegraph*, 2017-08-28). This suggests that when lesbians appear in popular TV shows (or indeed have their identity erased in some), the media reports it in an attempt to sensationalise that TV show.

Finally, and related to the representation of lesbian characters, the third category of collocates discusses lesbians in relation to sexual intercourse. This semantic category within the list of collocates could suggest that lesbians are sexualised and that there is language which focuses on their sexual practices: something which does not necessarily happen to heterosexual characters. This kind of discourse is also reminiscent of the discourses emerging in lesbian erotica (as covered by Baker, 2008; Morrish and Saunston, 2011), in so far as lesbians' sexual practices are eroticised or marked as an important characteristic.

These findings warranted further research, and thus a keyword list was generated, in order to see the discourses *lesbian** is embedded within. As Baker (2004: 347) argues: “Keywords will therefore not reveal discourses, but will direct the researcher to important concepts in a text (in relation to other texts) that may help to highlight the existence of types of (embedded) discourse or ideology”. In other words, while the statistical tests for collocation have revealed some ways that lesbians were represented, I wanted to see if these results resemble the findings from the keyword analysis.

The currently available version of the written BNC ([1994] 2007) was used as a reference corpus to generate a keyword list, which is displayed in Table 3.

Number	Frequency	Log-likelihood	Word
1	25976	4164.888	gay
2	22248	2810.608	rights
3	18153	2059.337	sex
4	14607	1989.733	marriage
5	10277	1849.634	trump
6	7106	1278.923	lgbt
7	5395	956.7	transgender
8	5147	871.502	lesbian
9	6178	840.465	australia
10	5234	759.067	equality
11	6909	749.618	sexual
12	28177	738.962	people
13	4044	727.831	twitter
14	9396	652.303	party
15	3844	570.635	sexuality

Table 3: Top 15 keywords in the *Lnews17* corpus

The keywords broadly suggest that lesbians are typically represented in two ways: either they are seen as part of a homogenous mass of GSRM individuals who are affected by legal and social policies (evidenced in *rights*, *marriage*, *[T]rump*, *[A]ustralia*, *equality*, *[T]witter*, *party*), or they are discussed within articles that talk about sexual practices (*sex*, *sexual*, *sexuality*). Both these discourses appear to echo the previous literature, in which Moschenbacher (2018) found that lesbians were positioned with regard to political legislation and Baker (2008) found that lesbians were frequently sexualised. While it is good to see lesbians discussed in relation to the social rights that they are afforded, the fact lesbians are so frequently discussed with regard to sexual practices may be interpreted as problematic. There may still be some sexualisation of lesbians, and possibly women in a more general sense, within the media. Indeed, the concordance lines for these keywords reveal that usually, these terms are used to discuss the actions of sexual practices. This echoes Gill's (2009) work on the representation of lesbians in advertising, in which she argues that the sexualisation of lesbians is still prevalent. In other words, there is still sexualisation of women in the media and this dataset suggests that lesbians are embedded

within a broader discourse which sexualises both their identity as a woman and their identity as a lesbian.

Although the term lesbian is a sexual identity label and is linked to sexual attraction, the emphasis on sex, as opposed to say relationships, is problematic. The focus on lesbian's sexual behaviour, as opposed to their romantic behaviour, appears to contribute to a discourse which sexualises lesbians.

5.1. *Lesbians are perverse and a sexual threat (especially to children)*

One of the common discourses within the concordance lines was the idea that lesbians are sexually deviant, and this sexual deviance is a threat, especially within contexts where they have positions of power (12 occurrences in 100 concordance lines⁹). For example:

- (11) Teacher escapes jail for second time after lesbian romps with 16-year-old pupil
- (12) [name retracted], 25, is suspended after 'having lesbian sex with a 17-year-old pupil
- (13) DOCTOR'S LESBIAN SIDEKICK IN 'PERVING' STORM¹⁰

These extracts seem to associate lesbianism and sexual deviance. There is a negative judgement towards these women for abusing positions of power, but there is also unnecessary marking of the type of sex: lesbian sex.

In the above cases, a lesbian's identity is brought to the fore while mentioning her profession. In the first two examples, there is an implication that the lesbian has abused her power over children. In order to examine whether or not this kind of language was stereotypical of teachers who abuse their power, the findings were compared against data in the current written BNC ([1994] 2007). A search for the terms *teacher* and *sex* within a 10L 10R window¹¹ of each other yielded a total of 17 concordance lines, 12 of which discussed teachers having sex with pupils. In particular, a teacher's sexual identity was not used as an adjective when discussing their sexual activity with pupils. Additionally, typically it only reported on when male teachers had sex with female students. For example:

- (14) A 36-year-old married teacher, who admitted having unlawful sex with a 15-year-old pupil on a sailing trip was found guilty of raping the girl after a three-day trial (CS1 1109)

9 Throughout the analysis of how keywords are used in context, a random sample of 100 concordance lines was used. This is because it was a similar number to those available in the spoken BN2014 and was a number large enough to give a representative sample while still being small enough for the author to analyse.

10 Given the need for anonymity, I have deliberately not provided citations to papers which provide names of real-life people the concordance lines taken from Lnews17.

11 This window size was chosen as in the first example, *teacher* and *sex* within 9 words of each other.

- (15) A TEACHER'S EVIL LUST He took sex snaps of kids (CS1 1432)
- (16) The teacher who allowed drugs and underage sex in his home. (K1E 1323)

The lack of occurrences of an adjective for sexuality modifying the noun *teacher*, especially when the co-texts alludes to the sex being between someone who identifies as male and someone who identifies as female, suggests that the teachers being described are heterosexual, and therefore their sexuality is not marked as it is seen as the “norm”. This appears to resonate with the idea of compulsory heterosexuality, within not just an educational setting but also in general society (for an overview of heteronormativity, see Gray, 2015; for an overview of compulsory heterosexuality see Baker, 2008: 107-109). These findings could also echo the, now slightly dated, work of Tierney (1997: 96), who explains that 75% of gay and lesbian academics are not ‘out’. Tierney argued that one reason for this lack of visibility was the negative experiences faced by the out teachers. More recently, a poll by the teachers’ union NASUWT revealed that approximately 30% of school teachers are not out in their workplace (see Duffy, 2018). Given that the lesbian teachers featured in the Lnews17 were all mentioned for abusing positions of power and were viewed as threats to children, some lesbian teachers might feel under more scrutiny; that if they do something wrong, they will not just be negatively judged because they are a teacher, but also because they are a lesbian.

Regardless of if the teachers in this sample are heterosexual or not, the fact remains that within Lnews17, lesbian teachers and other professionals who are lesbians, have their sexual identity marked, and this sustains a homophobic ideology: that lesbians should be ‘othered’ for their sexual behaviour. Linking back to the keywords, this reiterates the written media’s obsession with lesbian’s sexual practices. It appears as though marking the kind of sex these people have serves to delegitimise people with similar sexual identities within the same profession.

5.2. *Lesbianism is marked compared to heterosexuality and thus delegitimised*

In a similar vein to *lesbian* being used as a premodifier for sexual practices, and this language being linked to abuse of power as a way to negatively evaluate lesbian identities, *lesbian* is used as a premodifier for a number of other words. These words relate to a large range of topics, not just sexual acts (30 occurrences in 100 concordance lines related to non-sexual acts; 39 occurrences in 100 concordance lines including occurrences where *lesbian* is a premodifier for sexual behaviours). For example:

- (17) [name retracted] and [name retracted], a lesbian couple from West Virginia
- (18) [name retracted] was called a "sperm donor" by the lesbian mother of his baby

- (19) Yes, well done, she did have a lesbian affair with [name retracted]

A commonality in all the examples is that *lesbian** is used as a premodifier when the noun which it premodifies does not necessarily relate to the person's sexual identity. That is to say that when this co-occurrence appears, the sexuality of the person has little to no relevance to the content of the articles (27 occurrences in 100 concordance lines related to non-sexual acts). For example, "[name retracted] and [name retracted], a lesbian couple from West Virginia" could equally be constructed as "*Jane and Rachel, a couple from West Virginia*"¹². Similarly, "Yes, well done, she did have a lesbian affair with [name retracted]" could be constructed as "*Yes, well done, she did have an affair with Sally*". The fact that a lesbian has her identity needlessly highlighted could suggest that a lesbian's identity is somehow marked as an 'other' in a heteronormative society, and therefore less legitimate compared to a heterosexual identity. It is also possible that the writers of the newspapers could be presenting a heteronormative view that couples are heterosexual, mothers are heterosexual women, and that affairs happen between (heterosexual) men and women. Nevertheless, in these examples, *lesbian* is being used as an adjective to delegitimise other aspects of lesbian's identities, such as their motherhood status or their relationship status.

It is also important to note that there were only three cases where *lesbian** was used as a premodifier within a context where a lesbian identity was important to the context of the story:

- (20) becoming the first openly lesbian Cabinet minister (*The Daily Mail*, 2017-04-20)
- (21) her acts signalled a powerful new voice in lesbian fiction (*The Guardian*, 2017-04-08)
- (22) ruled that the right of gay and lesbian couples to marry is protected" (*The Daily Mail*, 2017-04-20)¹³

In these examples, changes in policy and society are the focus of the articles, and therefore lesbianism is brought to the fore. In comparison to earlier examples, the identity of the women is important here, as these are landmark cases for GSRM history, which shows that the way lesbians are treated is different compared to the past. Taken with the occurrences of *lesbian* as a premodifier and indeed the data from the collocational analysis, this echoes Love and Baker's (2015) observation that, despite socio-political changes, the language used to discuss "homosexuality" is still relatively negative. Not only is the language used to discuss this topic still negative, the language specifically used to describe lesbians is too.

12 These names are fictitious.

13 The two examples from the Daily Mail in these concordance lines come from different articles.

6. An overlapping discourse: Lesbians are similar to other GSRM individuals and together form a homogenous category

So far, this paper has outlined the discourses in the Spoken BNC2014 and Lnews17 separately. In this section, I draw attention to a discourse which is shared across the corpora that I have not yet fully explored in the separate analyses. In other words, while I have highlighted the discourses which are unique in each corpus, this section discusses a shared discourse.

In Section 5, I touched upon how lesbians were conflated with other members of the GSRM ‘community’¹⁴. In this section, I expand on this idea and highlight occurrences of this discourse in both the Spoken BNC2014 and Lnews17.

In the 77 concordance lines for *lesbian* in the Spoken BNC2014, there are 15 instances of co-occurrence with the word *gay*. For example:

- (23) S0246: oh I don't mind Bates Motel to be honest
 S0245: >> I thought it started off okay and then it went really bad (.) and it kind of ruined Psycho
 S0244: oh gay and lesbian
 S0246: well it can't really ruin Psycho cos Psycho 's just good" (S2W4 1303)
- (24) S0202: no wait right so I went to Niagara falls stayed there for a day and a half got a taxi to erm from our hotel to the bus station in Niagara to go to Toronto and the guy in the taxi as soon as I got in like oh where you going? like where you going when you get to the bus station like oh we're going to Toronto blah blah blah and like oh okay are you going for the queer festival ? I were like excuse me? and like ah are you going to the queers they're having a parade
 S0203: do they use that all the time up there?
 S0202: wait for it so he was saying oh are you going to --UNCLEARWORD parade? are you going to see the queers? all this all that and I were like excuse me but what are you talking about? like that's really offensive and he was like ah I 'm talking about like lesbians gays and stuff like that they 're having a a two million strong parade in Toronto like it's Canada Day and erm the the gay pride thing like oh okay (STZ3 372)

In Example (23), speaker S0244 appears to be referring to a broader genre of films and television shows – “gay and lesbian” films and television shows. There appears to be a layer of judgment in this: that the media unnecessarily makes characters GSRMs. While scholars like Trivette (2018) have argued that GSRMs would like more GSRM representation on television, it appears as though these, assumed heterosexual, people are arguing for the opposite. There is a wealth of literature on the performances of gender and iden-

14 I have used the term ‘community’ as GSRM individuals are often viewed as part of an ‘imagined’ community (see Anderson, 1983). In other words, not all lesbians are members of a physical community, but they are part of a socially constructed community, imagined by the people who perceive themselves as part of that group (Nagel, 2003; Eckert and McConnell-Ginet, 2007).

tity in gay and lesbian films (for example, see Morrish and Saunston, 2007; Baker, 2008). These representations of gay men and lesbian women tend to be more diverse than found in films made for audiences seeking heteronormativity (see Baker, 2008). Therefore, there may be a disconnect in how, presumably heterosexual speakers, identify with the characters and storylines, and this could explain why speaker S0246 was able to connect with (largely) heteronormative characters in the film *Pyscho*, but not in the explicitly gay and lesbian television show *Bates Motel*.

In Example (24), “queer” people appear to be composed of “lesbians”, “gay”, and “stuff like that”. The language used appears to background bisexual, transgender, and queer identities. While the speakers acknowledge that *lesbians* and *gays* are queer identities, there appears to be a general amalgamation of GSRM identities. One potential reason for this could be that the acronyms LGBTQIAA or LGBTQ+, which are currently commonplace compared to the acronym GSRM, are longwinded and do not currently encapsulate all possible identities. However, it could also be the conflation of many queer identities within “stuff like that”, though it is interesting to see how “gay and lesbian” are highlighted as the identities that are named.

This discourse is also present in a sample of 100 random concordance lines from Lnews17 (22 occurrences within 100 concordance lines). For example:

- (25) unisex toilets will stop bullying and stop lesbian, gay, bisexual and transgender pupils from having to worry about using the ‘wrong’ bathroom (*Express Online*, 2017-04-20)
- (26) Claimed it would prevent gay, lesbian and transgender pupils from being bullies (*Mailonline*, 2017-04-20)

There are a number of cases where the homogenisation of GSRM individuals appears to be in relation to protection, rights, and safeguarding, which can create some form of political agency for GSRM individuals. While the debates around socio-political issues such as access to gender-neutral toilets is an important issue, it is often more significant and more pertinent for transgender and genderqueer people who experience body dysmorphia (see Beemyn, 2005). In other words, while gender-neutral toilets are a step in the right direction for equality for transgender students, the media still emphasises the implementation of gender-neutral bathrooms are primarily helping *lesbian*, *gay*, and *bi-sexual* students, who do not face the same issues as transgender individuals. This is not to say that lesbian women do not experience toilet-harassment, but while the move to gender-neutral bathrooms could prevent a GSRM individual from experiencing toilet-harassment, for transgender people, this kind of movement allows them to destabilise and disentangle the sex/gender conflation that is prevalent in society and allows them the freedom to use bathrooms which match their own identities. Nevertheless, it is important to note the well-meaning intentions behind this kind of language and how easy it is to

criticise people for using LGBT as a conflated concept, I would argue that simple changes such as placing *Transgender* first in the list would show which group the policy is more likely to affect, and why the policy may have been implemented in the first place, while still including the terms together to allow for political mobility and to demonstrate solidarity among GSRMs.

7. Discussion and conclusions

The various discourses surrounding the term *lesbian** in the Spoken BNC2014 and Lnews17 generally appear to be negative. The findings suggest that lesbians still face marginalisation and are linguistically ‘othered’ in both informal spoken data and in publicly available written news texts. Considering the Spoken BNC2014 is formed of informal conversations, this could suggest that lesbians still face prejudice in both public and private contexts. This echoes Love and Baker’s (2015) argument that even though legal practices have been implemented to protect GSRM individuals and criminalise against discrimination, there is still linguistic evidence to suggest a bias against this group. While Love and Baker explored the language used only in a public context, this research has shown that homophobic discourses, at least towards lesbians, are also present in private contexts too.

The discourses in the BNC2014 provide a window into embedded ideologies towards the group in general private contexts. The language used within Lnews17 reveals some of the underlying ideologies of the authors of the texts, the editorial team of the newspapers, and people quoted in the texts. The discourses surrounding *lesbian** and the way lesbians are represented in the media may either be the product of or the creator of language which sustains ideologies in general. That is to say, media texts are shaped by the private views which journalists are exposed to, and private views are similarly shaped by media texts. The editorial team would write and edit the papers in order to present not only their ideologies but would also tailor the ideologies to align with their own target audiences. Thus, these discourses may be circular, which is worrying, given the problematic nature of some of the representations.

The discourses I have presented in this paper are by no means an exhaustive list of the discourses surrounding the term *lesbian** in either the Spoken BNC2014 nor Lnews17. However, the discourses presented represent some of the more dominant discourses within each corpus. The frequency of positive discourses compared to negative discourses was staggering, with both corpora showing more negative evaluation towards lesbians. The corpora were able to demonstrate how lesbians were discussed in both public and private contexts, which suggests that in both spheres, they face prejudice views.

8. Directions for future research

It is hoped this research will act as a springboard for other avenues of future research. Firstly, during the final stages of the publication process, the BNC2014 was made publicly available, therefore it would be interesting to triangulate the data within that corpus with the findings in these corpora. Secondly, this research could spark additional future research into the discourses surrounding other search terms such as *bisexual**, *transgender**, and *queer** in the British mass media in comparison to spoken corpora (though, see the work of Wilkinson, 2019 for a discussion of the representation of bisexuality in *The Times*, and Zottola, 2018 for a discussion of the representation of transgender identities in the British press). The research presented in this article opens up avenues of research which could explore whether or not the same or similar discourses about gender non-conformity, sexuality as a choice or sexual predation are found in public discourses about bisexual women, femme gay men, and/or non-binary individuals.

This research opens new avenues for further research, specifically within the sub-field of research into lesbian discourse, such as comparing how lesbians are represented on social media (such as Twitter) in comparison to the written media (though, see the work of Bailey, 2019 for a discussion of lesbian identity construction on specific internet fora), and how lesbian authors for mass media outlets construct lesbians in comparison to heterosexual authors.

Ultimately, the homophobic ideologies uncovered in this research call for more research into how these ideologies are sustained and maintained in other contexts, in order to start dismantling them as a step towards equality.

Competing interests

The authors have no competing interests to declare.

References

- Anderson, B. (1983). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Anderson, C. (2016). *Public voices, private voices: an investigation of the discourses of age and gender and their impact on the self-identity of ageing women*. Unpublished doctoral dissertation, University of Birmingham, Birmingham.

- Armstrong, J. (1997). Homophobic slang as coercive discourse among college students. In A. Livia., & K. Hall (Eds.), *Queerly Phrased: Language, Gender, and Sexuality* (pp. 326–334). New York & Oxford: Oxford University Press.
- Bachmann, I. (2011). Civil Partnership – Gay marriage in all but name: A corpus-driven analysis of discourses of same-sex relationships in the UK Parliament. *Corpora*, 6(1), 77–105. doi:10.3366/cor.2011.0005
- Bailey, A. (2019). “Girl-on-girl culture”: Constructing normative identities in a corpus of sex advice for queer women. *Language and Sexuality*, 8(2), 195–220. doi:10.1075/jls.18013.bai
- Baker, P. (2004). Querying Keywords. *Journal of English Linguistics*, 32(4), 346–359. doi:10.1177/0075424204269894
- Baker, P. (2005). *Public Discourses of Gay Men*. London: Routledge. doi:10.4324/9780203643532
- Baker, P. (2008). *Sexed Texts: Language, Gender and Sexuality*. London: Equinox.
- Baker, P. (2012). From gay language to normative discourse: a diachronic corpus analysis of Lavender Linguistics conference abstracts 1994–2012. *Journal of Language and Sexuality*, 2(2), 179–205. doi:10.1075/jls.2.2.01bak
- Baker, P. (2014). *Using corpora to analyze gender*. London: Bloomsbury.
- Baker, P. (2016). Gendered Discourses. In P. Baker, & J. Egbert. (Eds.) *Triangulating Methodological Approaches in Corpus Linguistic Research* (pp. 138–151). London: Routledge. doi:10.4324/9781315724812
- Baker, P. (2018). *Which techniques of down-sampling best complement a corpus-assisted discourse analysis? A case study on press representations of obesity*. Paper presented at the Corpora and Discourse Conference 2018, Lancaster.
- Baker, P., Gabrielatos, C., & 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. doi:10.1093/applin/ams048
- Beemyn, B. (2005). Making Campuses More Inclusive of Transgender Students. *Journal of Gay & Lesbian Issues in Education*, 3(1), 77–87. doi:10.1300/j367v03n01_08
- Brown, T., & Alderson, K. (2010). Sexual identity and heterosexual male students' usage of homosexual insults: an exploratory study. *The Canadian Journal of Human Sexuality*, 19(1-2), 27–42.

- Butler, J. (1990). *Gender Trouble*. New York: Routledge.
- Chirrey, D. (2003). 'I hereby come out': What sort of speech act is coming out?. *Journal of sociolinguistics*, 7(1), 24–37. doi:10.1111/1467-9481.00209
- Darics, E., & Koller, V. (2019). Social actors “to go”: An analytical toolkit to explore agency in business discourse and communication. *Business and Professional Communication Quarterly*, 82(2): 214–238.
- Davies, M. (2010). The corpus of contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing*, 25(4), 447–464. doi:10.1093/llc/fqq018
- Duffy, N. (2018). *A third of LGBT teachers stay in the closet*. Retrieved from: <https://www.pinknews.co.uk/2018/02/10/a-third-of-lgbt-teachers-stay-in-the-closet>
- Durrant, P. & Doherty, A. (2010). Are high frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory*, 6(2), 125–155. doi:10.1515/cllt.2010.006
- Eckert, P., & McConnell-Ginet, S. (2007). Putting communities of practice in their place. *Gender and Language*, 1(1), 27–37. doi:10.1558/genl.2007.1.1.27
- Gabriel, U. (2017). Is Saying “No” to a Request a Subtle Form of Discrimination Against Lesbians and Gay Men? A Fresh Look at Old Findings. In E. van Leeuwen., & H. Zagefka (Eds.), *Intergroup Helping* (pp. 45–63). Switzerland: Springer. doi:10.1007/978-3-319-53026-0_3
- Gill, R. (2009). Beyond the 'Sexualization of Culture Thesis: An Intersectional Analysis of 'Sixpacks', 'Midriffs' and 'Hot Lesbians' in Advertising. *Sexualities*, 12(2), 137–160. doi:10.1177/1363460708100916
- Gray, J. (2015). Language and Sexuality: heteronormativity and queer identities. In S. Preece (Ed.), *Routledge Handbook of Language and Identity* (pp. 225–240). London: Routledge.
- Gupta, K. (2016). *Representation of the British suffrage movement*. London: Bloomsbury.
- Gupta, K. (2019). Response and responsibility: Mainstream media and Lucy Meadows in a post-Leveson context. *Sexualities*, 22(1-2), 31–47. doi:10.1177/1363460717740259

- Hardie, A. (2012) CQPweb - combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17(3), 380–409. doi:10.1075/ijcl.17.3.04har
- Hawtin, A. (2019). *The Written British National Corpus 2014: design, compilation and analysis*. Unpublished doctoral dissertation, Lancaster University, Lancaster
- Jones, L. (2012). *Dyke/girl: Language and identities in a lesbian group*. Basingstoke: Palgrave Macmillan. doi:10.1057/9781137271341
- Koller, V. (2011). Analysing lesbian identity in discourse: combining discourse-historical and socio-cognitive approaches. In C. Hart (Ed.), *Critical Discourse Studies in Discourse and Cognition* (pp. 97–141). Amsterdam: Benjamins. doi:10.1075/dapsac.43.07kol
- Koller, V. (2013). Constructing (non-)normative identities in written lesbian discourse: a diachronic study. *Discourse and Society*, 24(5), 551–568. doi:10.1177/0957926513486166
- Lalor, T., & Rendle-Short, J. (2007). 'That's So Gay': A Contemporary Use of Gay in Australian English. *Australian Journal of Linguistics*, 27(2), 147–173. doi:10.1080/07268600701522764
- Land, V., & Kitzinger, C. (2005). Speaking as a lesbian: Correcting the heterosexual presumption. *Research on Language and Social Interaction*, 38(4), 371–416. doi:10.1207/s15327973rlsi3804_1
- Love, R. (2018). *The Spoken British National Corpus 2014: design, compilation and analysis*. Unpublished doctoral dissertation, Lancaster University, Lancaster
- Love, R. (2020). *Overcoming Challenges in Corpus Construction: The Spoken British National Corpus 2014*. New York: Routledge. doi:10.4324/9780429429811
- Love, R., & Baker, P. (2015). The hate that dare not speak its name? Contemporary Discourses of Hate and Radicalism across Space and Genres. *Journal of Language Aggression and Conflict*, 3(1), 57–86. doi:10.1075/jlac.3.1.03lov
- Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017). The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics*, 22(3), 319–344. doi:10.1075/ijcl.22.3.02lov
- Marko, G. (2008). *Penetrating Language. A Critical Discourse Analysis of Pornography*. Tübingen: Gunter Narr.

- Morrish, L., & Sauntson, H. (2007). *New perspectives on language and sexual identity*. London: Palgrave MacMillan. doi:10.1057/9780230599406
- Morrish, L., & Sauntson, H. (2011). Gender, desire and identity in a corpus of lesbian erotica. In B. Scherer., & M. Ball (Eds.) *Queering Paradigms II: Interrogating Agendas* (pp. 63–81). Oxford: Peter Lang Press.
- Motschenbacher, H. (2018). Corpus linguistics in language and sexuality studies: Taking stock and future directions. *Journal of Language and Sexuality*, 7(2), 145–174. doi:10.1075/jls.17019.mot
- Nagel, J. (2003). *Race, ethnicity, and sexuality: Intimate intersections, forbidden frontiers*. Oxford: Oxford University Press.
- Phelps, E. (2017). Transgender identity disclosures via Youtube narratives. In N. Gordon., & L. Finley (Eds.), *Reflections on Gender from a Communication Point-of-View: GenderSpectives* (pp. 61–74). Cambridge: Cambridge Scholars Publishing.
- Scott, M. (2016). *WordSmith Tools version 7*. Stroud: Lexical Analysis Software.
- Skinner, S. (2016). 'Queer Subjectivity, Transmedia, and Embodiment in the Carmilla Fandom'. Unpublished Bachelor's dissertation, Georgia, Georgia State University.
- Stubbs, M. (1994). Grammar, Text, and Ideology: Computer-assisted Methods in the Linguistics of Representation. *Applied Linguistics*, 15(2), 201–223. doi:10.1093/applin/15.2.201
- Stubbs, M. (1996). *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*. Oxford: Blackwell.
- The British National Corpus*. ([1994] 2007). *The British National Corpus, version 3 (BNC XML Edition)*. Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. Retrieved from: <http://www.natcorp.ox.ac.uk/>
- Tierney, W. (1997). *Academic outlaws: Queer Theory and Cultural Studies in the Academy*. London: Sage. doi:10.4135/9781483327877
- Trivette, E. (2018). *Our Relationships with LGBTQ+ Characters' talking about how LGBTQ+ people respond to LGBTQ+ representation on television*. Paper presented at the 2018 Birmingham English Language Postgraduate Conference, Birmingham.
- Wilkinson, M. (2019). 'Bisexual oysters': A diachronic corpus-based critical discourse analysis of bisexual representation in The Times between 1957 and 2017. *Discourse & Communication*, 13(2), 249–267. doi:10.1177/1750481318817624

- Wilson, A. (2012). Using corpora in depth psychology. A trigram-based analysis of a corpus of fetish fantasies. *Corpora*, 7(1), 69–90. doi:10.3366/cor.2012.0018
- Woolley, S. (2013). *'Identity and Difference: Negotiating Gender and Sexuality in High School Contexts'*. Unpublished doctoral dissertation, California, University of California Berkeley.
- Zottola, A. (2018). Transgender identity labels in the British press (2013–2015). *Language and Sexuality*, 7(2), 237–262. doi:10.1075/jls.17017.zot