A Diachronic Corpus-Based Analysis of LGBTQ-Related Research Articles

Sarunya Tarat[a,*], Nawarat Siritararatn[b], Woravut Jaroongkhongdach[c]

[a] sarunya.tar@ku.th, English for International Communication (EIC) program, Kasetsart University, Thailand
[b] nawarat2000@hotmail.com, Department of Foreign Languages, Faculty of Humanities, Kasetsart University, Thailand
[c] woravutj@gmail.com, Department of Foreign Languages, Faculty of Humanities, Kasetsart University, Thailand
[*] Corresponding author, sarunya.tar@ku.th

ABSTRACT

This study uses a diachronic corpus-based analysis to explore the topics presented in LGBTQ-related research articles published between 2001 and 2020, and to examine diachronic changes in these topics over time. The data are from 400 LGBTQ-related research articles which are divided into four time periods (2001-2005, 2006-2010, 2011-2015, 2016-2020) with 100 articles per period. We analyzed the data by looking at keywords, comparing keywords across time, and investigating the concordance lines. The findings from the keywords analysis revealed thirteen LGBTQ-related topics: sexuality, gender, relationship, hate, research, theory, culture, socio-politics, standard, science, place, sport, and nationality. The findings from the comparison across the four time periods show the diachronic changes in topics by looking at keywords identified as always key, previously key, and recently key. The concordance lines show in-depth investigation of the topics to discuss about LGBTQ. The in-depth investigation of the concordance lines indicates that LGBTQ people have difficulty in coming out due to traditional values. This analysis of LGBTQ-related research articles provides new insights into the way LGBTQ people are discussed which can be useful for further studies regarding application of a methodological framework to investigate LGBTQ-related topics in other types of data sources.

Keywords: corpus analysis, diachronic analysis, LGBTQ, research articles

Introduction

In the past twenty years, there has been a noticeably increasing number of research studies exploring topics related to the LGBTQ community which contribute to our better understanding of LGBTQ people. Some of these studies have attempted to explore LGBTQ-related topics using spoken data. For instance, Love and Baker (2015) studied changes in the debates on LGBT equality in the British parliament. Similarly, Santonocito (2020) investigated the speeches of prime ministers and uncovered a problematic condition for LGBT rights based on the broader field of civil rights.

Others have explored LGBTQ topics from the written data. Looking at Gay Times/Gay News magazine, for example, Baker (2014) pinpointed negative representations of homosexuality in the Daily Mail as being scandalous or shameful, leading to homophobic practices. More recently, Jones and Collins (2020) found that homosexual men were portrayed as sexually irresponsible people and HIV/AIDS carriers in British newspapers.

Regarding websites, Wang and Ma (2020) found that there was little attention paid to promoting LGBTQ+ rights in English-language version websites. Baider (2018) has also suggested that the construction of LGBT identities on websites was based on hegemonic heteronormativity and the influence of religion. In online newspaper archives, Zottola (2018) revealed that the use of the terms *transgender*, *transsexual* and *trans* were related to sex change, dysphoria, and gender reassignment.

By looking at different data sources, these previous studies provide us with a better understanding of LGBTQ-related topics. As the LGBTQ community has become more visible in societies, there have been more LGBTQ research articles published, especially in the past twenty years. These research articles can be a source for studying LGBTQ-related topics, but they seem to be underexplored. This study, then, aims to explore LGBTQ-related topics in research articles between 2001 and 2020 and to examine diachronic changes in the topics over time.

Literature Review

Research into LGBTQ-Related Topics

To understand LGBTQ-related topics, we need to understand heteronormativity, which is defined as a set of beliefs about women and men's roles (Kollmayer et al., 2018) and the actions that are acceptable and appropriate for them (Cislaghi & Heise, 2020). Based on this set of beliefs, being heterosexual is therefore perceived as acceptable and preferable (Baker, 2014).

The term LGBT has been used to refer to those who "share the common status of 'other' because of their members' departures from heterosexuality and gender norms" (Institute of Medicine of the National Academies [IOM], 2011, p. 13). The term *queer* was later included as part of the LGBT group "to pay homage to a queer history and perspective that is equally deserving of respect, dignity, and recognition" (Mekler, 2018, p. 156).

A breaking point for the LGBTQ rights movement, according to Nelson (2015), was the Stonewall Riots, which took place in 1969 in New York. Since then, the LGBTQ community has gained wider attention in both media and academic contexts. In media contexts, we can see that LGBTQ people have been increasingly visible on many different media platforms such as TV series, online media, and streaming media.

In academic contexts, there are research articles exploring LGBTQ-related topics in various fields such as family studies, education policy, and law. For example, in family studies, Kuvalanka (2013) investigated integrating LGBTQ issues into family courses, pointing out that their inclusion could help reduce LGBTQ-associated stigma. Another example of studies on LGBTQ and family studies involves LGBTQ parents. Lev (2010) studied the impact of heteronormativity on LGBTQ parenting and suggested that LGBTQ parents face pressure when raising children due to LGBTQ parents' experiences while growing up and societal condemnation.

Similarly, Goldberg and Allen (2018) discussed teaching about LGBTQ identities and family issues to make connections between academic content and the reality of family life.

LGBTQ-related topics have also been discussed in education. For instance, discussing LGBTQ-inclusive curricula and policies, De Witte et al. (2019) agreed that these can ensure that "young people have a good experience during their time at school; it is essential that their identity is nurtured, and that diversity is celebrated" (p. 490). Thus, inclusive policies can provide students with understandings of gender diversity in ways that construct students' views of LGBTQ groups as a social phenomenon.

There are studies that have looked into LGBTQ in law (Eichler, 2010; Lewis & Kern, 2018). For example, Eichler (2010) pointed out that law enforcement in some countries has an impact on LGBTQ people in terms of political struggles for rights and inclusion, so these people may be unfairly treated. This is because negative perceptions of same-sex relationships have led to an unfriendly environment or climate in which LGBTQ people may face discrimination by law enforcement and providers of public services. By taking a macro-level look, Lewis and Kern (2018) found that the laws associated with LGBTQ rights are complex and uncertain.

From these studies, we can see that LGBTQ-related topics have been studied, investigated, and discussed across different fields. To understand these topics, we now turn to the definition of the term *topic*. Topic is defined as "semantic or discourse prosodies" (Gabrielatos & Baker, 2008, p. 11). Thus, in this study, a topic is a subject of discussion of a specific discourse. There are different methods for exploring topics including text analysis. In this study, we aim to use text analysis across time periods which is called diachronic research.

Diachronic Research

Diachronic research is research that focuses on changes over time (Gabrielatos & Baker, 2008). Given that diachronic research can also trace the development of discourses and ideologies in documents at different periods of time, several researchers (e.g., Baker, 2013; Love & Baker, 2015; Wilkinson, 2022) have used corpus-based diachronic analysis to explore representations or ideologies.

Baker (2013) conducted a diachronic keyword analysis to study the development of LGBTQ-related issues in conference abstracts published in the Lavender Languages and Linguistics Conference between 1994 and 2012. Love and Baker (2015) used corpus-based methods to examine changes in British Parliamentary debates about LGBT equality to decrease social acceptability in using discriminatory language against LGBT people. The data used in this analysis were debates from the late 1990s/early 2000s and from the year 2013.

Similar to previous research by Love and Baker (2015), Wilkinson (2022) has recently done a diachronic corpus-based critical discourse analysis of queer representation in The Times between 1957–1967 and 1979–1990. These two historical corpora were compared to explore how The Times used language to represent queer people during the two time periods.

To build on these previous works on LGBTQ using diachronic research, the current study will address the following two research questions.

1. What are the LGBTQ-related topics in LGBTQ-related research articles between 2001 and 2020?
2. What are the diachronic changes in the LGBTQ-related topics in these research articles between 2001 and 2020?

Methodology

Data Collection

Based on the purposes of this study, we built our four corpora from LGBTQ-related research articles between 2001 and 2020 taken from SAGE Journals Online to investigate

LGBTQ-related topics in research articles using three steps. To begin with, we followed the guidelines from Zottola (2018) and Gabrielatos and Baker (2008) to search for research articles and limit search results. In other words, Zottola used search terms such as *transgender*, *transexual*, and *trans* to search newspaper articles related to transgender issues, while Gabrielatos and Baker collected online news about refugees by using the search terms with the Booleans OR and AND NOT in the search query to limit their search results. To this end, the words *lesbian*, *gay*, *bisexual*, *transgender*, and *queer* were used as search terms with the Boolean NOT to search for LGBTQ-related research articles. For example, *gay NOT lesbian NOT bisexual NOT transgender NOT queer* was the search query for gay-related research articles.

It was decided to use 400 LGBTQ-related research articles from between 2001 and 2020 from SAGE Journals Online because SAGE has a commitment to the social sciences (SAGE, n.d.). In terms of data comparability, Love and Baker (2015) asserted that researchers should consider the size of corpora for making quantitative comparisons between corpora. To account for this, we divided the 400 LGBTQ-related research articles equally into four separate files based on the four time periods (2001-2005, 2006-2010, 2011-2015, 2016-2020). Thus, there were 100 research articles contained in each file, and the research articles from the four separate files were used to build our four corpora.

Finally, the four sets of research articles were inserted into AntConc (Anthony, 2022), a corpus tool, which can be used to build and analyze corpora. Before building corpora with AntConc, all research articles were cleaned to exclude unwanted parts including titles, names and affiliations of authors, abstracts, headings, tables, figures, reference lists and acknowledgements.

Data Analysis

Data analysis involved four steps. Firstly, the LGBTQ-related research articles in the four separate files were used to build the corpora using AntConc. The number of word tokens in the four corpora of LGBTQ-related research articles between 2001 and 2020 is provided in Table 1.

Table 1

Number of Word Tokens in Four Corpora of LGBTQ-Related Research Articles between 2001 and 2020

LGBTQ-related research articles corpora	Word tokens
2001-2005	639,379
2006-2010	667,659
2011-2015	661,609
2016-2020	666,532

From Table 1, we can see that each of the corpora contained a similar number of word tokens so that they can be considered as being of equal size for the purpose of being compared.

Secondly, the keyword analysis tool in AntConc was utilized to elicit words that could be keywords for each of the corpora. We obtained four lists of keywords from the four corpora. However, there were several hundred words in each keyword list. According to Baker (2014), the top 50 keywords were used to investigate linguistic similarities and differences between two sets of data. In the same vein, Pojanapunya and Watson Todd (2018) proposed that the top 50 keywords are a sufficient cut-off point for genre analysis and critical applied linguistics research. Therefore, it was decided to select the top 50 keywords with the highest keyness value (log-likelihood) in each keyword list to focus on topics which seemed to be most salient between 2001 and 2020.

In this step, the search terms used in the data collection were excluded from the analysis, although these words were found in the top 50 keywords. In addition, we used only content words in the analysis because they carry meanings and can be categorized into semantic groups. The

selected keywords were transferred to an Excel spreadsheet. This allowed us to group keywords and make comparisons across corpora. Keywords that have similar meanings were grouped together using the Longman Dictionary of Contemporary English, as this dictionary is corpus-based (Gabrielatos & Baker, 2008) and gives corpus-based examples.

Thirdly, a set of criteria used by Baker (2013) was adapted to plot diachronic patterns. It included *always key* (occurring in all four time periods), *recently key* (occurring in the latter two or three periods), and *previously key* (occurring in the former two or three periods; occurring in the first period and becoming key again in the latter one or two periods). Keywords that were only key in one time period were excluded from the analysis because "the fact it only occurs once is perhaps less indicative that this is a word that is becoming more popular over time" (Baker, 2013, p. 185).

Finally, we followed Cohen (2017) to conduct an intra-rater reliability process by reading and evaluating each list more than once. Given that corpus analysis deals with large amounts of text data, it helps to check the validity of analysis in which research findings can be valid and generalizable (Biber & Reppen, 2015). Therefore, corpus analysis was considered to check the validity of the analysis.

Findings

In this section, we separated the findings into two parts based on the two research questions: the LGBTQ-related topics in LGBTQ-related research articles between 2001 and 2020 and the diachronic changes in the topics over time. We first presented LGBTQ-related topics in the research articles obtained from grouping keywords. This was followed by the diachronic changes of these topics over time. Table 2 shows the LGBTQ-related topics in research articles between 2001 and 2020.

Table 2

The LGBTQ-Related Topics in LGBTQ-Related Research Articles between 2001 and 2020

Topic	Longman Dictionary of Contemporary English	Keywords
Sexuality	the things people do, think, and feel that are related to their sexual desires	sexual, sex, sexuality, homosexuality, masculinity, heterosexuality, bisexuality, masculine, feminine
Gender	the fact of being male or female	heterosexual, men, gender, women, homosexual, male, female, straight, LGBT, gendered
Relationship	the way in which two people or two groups feel about each other and behave towards each other	relationship, couples, family, partner, mothers, community, individuals, kinship
Hate	to dislike something very much	HIV, AIDS, violence, homophobia, homophobic, ambivalence, discrimination, oppression, different, stigma, minority, negative, internalized

Research	the activity of finding information about something that you are interested in or need to know about	participants, respondents, research, study, meanings, interviews, interviewees
Theory	general principles and ideas about a subject	identity, orientation, discourses, theory, counselling, described, binary
Culture	the attitudes and beliefs about something that are shared by a particular group of people or in a particular organization	cultural, culture
Socio-politics	relating to society and the government, politics, and public affairs of a country	social, politics, political, feminist, experience, disclosure, status, workers, practices, belonging
Standard	a level that is considered acceptable, or a level that someone or something has achieved	norms, heteronormative, normative, role, heteronormativity
Science	a particular part of science for example biology, chemistry, or physics	desire, donor, biological, body, transition, psychological
Place	a space or area to do something	space, ways
Sport	a physical activity in which people compete against each other	athletes
Nationality	the state of being legally a citizen of a particular country	Cuban

Research Question 1: What are the LGBTQ-Related Topics in LGBTQ-Related Research Articles between 2001 and 2020?

To investigate LGBTQ-related topics, we grouped keywords from the four corpora into groups. These groups of keywords were then identified as topics. As shown in Table 2, there were 13 LGBTQ-related topics in LGBTQ-related research articles between 2001 and 2020: *sexuality, gender, relationship, hate, research, theory, culture, socio-politics, standard, science, place, sport,* and *nationality.* These topics were explained with the use of the Longman Dictionary of Contemporary English (Longman Dictionary of Contemporary English Online, n.d.).

Given that sexuality refers to sexual feelings, we grouped keywords signifying sexual feelings such as *homosexuality, masculinity,* and *heterosexuality* under the topic of *sexuality.* Meanwhile, the keywords *homosexual, male, female, straight,* and *LGBTQ* were linked to the topic of *gender* in which gender can be a person's identity based on social construction. Another topic is *relationship.* It was found that keywords like *couples, family, mothers, community,* and *individuals* showed the ways that people behave with each other. In the same way, the keyword *kinship* was key in the topic of *relationship* to show family relationship.

There were keywords referring to hate. For example, *HIV*, *violence*, and *homophobia* can be seen as forms of hate towards members of LGBTQ groups. They also included *discrimination*, *minority*, and *oppression* as keywords signifying that there would be increased focus on hate, referring to the topic of *hate*. In terms of research procedures, we put the keywords related to conducting research into the topic of *research*, i.e., *participants, respondents, research, study, meanings, interviews,* and *interviewees*. By grouping keywords relating to the topic of *theory*, it was possible to group *identity*, *orientation*, and *binary* together because these keywords refer to one's sense of self. The keyword *discourses* was also included in this topic as the keyword was about discussion on a particular subject.

For the topic of *culture*, it was clear that two keywords, *cultural* and *culture*, were directly about shared attitudes about something. Furthermore, we found that there were keywords relating to the topic of *socio-politics*, for instance, keywords like *social*, *politics*, *political*, *feminist*, and *status* showed a relationship between society and politics. Moving on, we categorized *heteronormative*, *norms*, *normative*, and *role* into the topic of *standard* as these keywords could be considered as standards or values. Regarding place, the two words *space* and *ways* were key under the topic of *place*.

It was surprising to see keywords about a particular area of science such as *donor*, *transition*, and *psychological*. Therefore, we put the keywords into the topic of *science*. For the topics of *sport* and *nationality*, there was only one keyword under each topic, i.e., athletes and Cuban, respectively.

For RQ 1, we presented the LGBTQ-related topics in LGBTQ-related research articles between 2001 and 2020. However, this study also aimed to examine changes in the topics over time. To do this, we tracked the keywords under each topic to plot the changes in the topics between 2001 and 2020 with the use of diachronic analysis. The findings from the diachronic analysis of keywords in each topic were presented under RQ 2.

Research Question 2: What are the Diachronic Changes in the LGBTQ-Related Topics in these Research Articles between 2001 and 2020?

We examined the changes of keywords under the following topics: *sexuality*, *gender*, *relation*, *hate*, *research*, *theory*, *culture*, *socio-politics*, *standard*, *science,* and *place*. However, the topics of *sport* and *nationality* were not included in the diachronic analysis of keywords due to only one keyword being found for each of the topics. Keywords that occurred once were not analyzed because we could not track their changes.

In addition, keywords were subjected to concordance analyses for interpretation. The tick used in the table refers to a word that was key in each time period. Table 3 shows the keywords under the topic of *sexuality* over time.

Table 3

Keywords under the Topic of Sexuality over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Sexual	✓	✓	✓	✓
Sex	✓	✓	✓	✓
Sexuality	✓	✓	✓	✓
Masculinity	✓	✓	✓	✓
Masculine	✓	✓		✓
Heterosexuality	✓	✓		✓
Bisexuality	✓		✓	✓
Feminine		✓	✓	✓
Homosexuality	✓	✓	✓	

As shown in Table 3, it is clear that *sexual, sex, sexuality, masculinity,* and *masculine* were always key. The examination of concordance lines revealed that LGBTQ people were involved in sexual

activity, as in Example 1. These people were also depicted as persons having the preferable characteristics of men, as in Examples 2-3.

> Example 1: bisexual women appear as particularly risky **sexual partners**, because they are assumed to have a history or even current practice of **sex with men**.
> Example 2: think I helped them realize they could be Black, strong, **masculine**, and gay.
> Example 3: ...constructions of **masculinity in the gay adult** film industry...

However, the keyword *feminine*, considering a key term since 2006, was recently key, signifying that LGBTQ identity has also been discussed in terms of having the qualities of women, as in Example 4.

> Example 4: The second most common description for **feminine gay men** was not being able to hide one's sexual orientation or not being able to pass as heterosexual.

Based on Table 3, it was found that keywords relating to the state of being sexually attracted to someone were previously key. Keywords including *heterosexuality*, *bisexuality*, and *homosexuality* were found to initially occur in 2001-2005. Nevertheless, usage of the three keywords changed at different points. The use of *bisexuality* and *heterosexuality* declined in 2006-2010 and 2011-2015, respectively, but became key again in 2016-2020. In contrast, the keyword *homosexuality* remained key until 2015.

Table 4

Keywords under the Topic of Gender over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Heterosexual	✓	✓	✓	✓
Men	✓	✓	✓	✓
Gender	✓	✓	✓	✓
Women	✓	✓	✓	✓
Homosexual	✓	✓	✓	✓
Male	✓	✓	✓	✓
Straight	✓	✓		✓
LGBT		✓	✓	✓
Female	✓	✓	✓	✓
Gendered		✓		

Table 4 shows the set of keywords under the topic of *gender*. The following keywords were seen as always key: *heterosexual*, *homosexual*, *men*, *women*, *male*, *female*, and *gender*. Interestingly, the two terms *heterosexual* and *straight*, which carry the same meaning, were used interchangeably to mention straight or heterosexual, and LGBTQ people together in a particular context, as seen in Examples 5-6. However, there was a decline in the use of *straight* during the period between 2011 and 2015.

> Example 5: ...body dissatisfaction appears to be normative for both **heterosexual and lesbian women**.
> Example 6: The experiences of **straight and gay stepfamilies** may be similar in many ways...

Another interesting point in the topic of *gender* was that the keyword *LGBTQ* has emerged as key since 2006 and can be considered recently key. In Example 7, the concordance lines for the keyword showed that the use of *LGBTQ* has identified LGBTQ as a group rather than an individual.

Example 7: In this time of greater social and public acceptance of **LGBTQ people** and identities, …

Table 5

Keywords under the Topic of Relationship over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Relationship	✓	✓	✓	✓
Couples	✓	✓	✓	✓
Partner	✓	✓	✓	✓
Mothers	✓	✓	✓	✓
Family	✓		✓	✓
Individuals		✓	✓	✓
Community	✓	✓	✓	✓
Kinship			✓	

As seen in Table 5, another set of keywords was about the topic of *relationship*. There were different forms of relationship under this topic. To begin with, the keywords *couples* and *partner* were identified as romantic relationships. The keywords *mothers* and *family* referred to family relationships. Looking in more detail, it was found that the keywords *couples*, *partner*, and *mothers* were always key while the keyword *family* that can be seen in 2001-2005, 2011-2015, and 2016-2020 was identified as previously key. The analysis of concordances of these keywords demonstrated that it might be unacceptable for LGBTQ people to have family, as the following examples attest.

Example 8: Although the **lesbian mothers** were often able to renegotiate their relationship with their parents, this did not however necessarily mean that the grandparent generation fully accepted the lesbian relationship as a proper relationship or as a basis for a family.
Example 9: These social decisions were sometimes made by the **transgender partner** and other times were mutually decided.
Example 10: …exclusion processes that contributed to **LGBT family** members feeling like outsiders within their own families.
Example 11: …a relationship wellness checkup may benefit **lesbian couples** in similar ways…

Moreover, LGBTQ people were considered as a group in terms of relationship. The keyword *community* was found to be always key over time, as shown in Example 12. It was argued that the keyword *individuals* was likely to refer to LGBTQ people having individuality, as seen in Example 13. From Table 5, the word *individuals* was considered as recently key as it has been key in the corpora since 2006.

Example 12: …there was a role for the **LGBT community** in supporting older members…
Example 13: **Lesbian and gay individuals** are embedded in a social system, and their distress regarding sexual orientation.

Table 6

Keywords under the Topic of Hate over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
HIV	✓	✓	✓	✓
AIDS	✓	✓		
Homophobia	✓	✓		
Discrimination				

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Minority		✓	✓	✓
Violence	✓		✓	✓
Homophobic	✓			
Stigma	✓			
Negative				✓
Internalized				✓
Ambivalence				✓
Oppression	✓			
Different		✓	✓	

Table 6 presents the keywords under the topic of *hate* with the following keywords under examination: *HIV*, *AIDS*, *homophobia*, *discrimination*, and *minority*. These keywords have changed differently over time. For example, the keyword *HIV* was identified as always key. The concordance analysis showed that the representation of *HIV* was different in each time period. In Example 14, *HIV* was related to treatment and prevention from 2001 to 2015, but there has been a shift away from prevention to stigmatization in the period 2016-2020, as seen in Example 15.

> Example 14: ...**HIV prevention** and caring for its customers health, while acknowledging different levels of commitment.
> Example 15: ...the **stigmatization of HIV-positive individuals**, while maintaining the public health discourses on individual responsibility, risk, and HIV.

In terms of being a disease, *HIV* is generally understood as *AIDS* and vice versa. *HIV* was always key, but the keyword *AIDS* was considered previously key, occurring in 2001-2005 and 2006-2010. In Example 16, *AIDS* is particularly associated with gay. It seems that there were more negative representations of *AIDS*, compared with *HIV* representations.

> Example 16: **AIDS** is still viewed as a gay disease and is associated with gay men who are seen as carriers of the plague.

The keyword *homophobia* was previously key. It was used between 2001 and 2010. The use of *homophobia* used with the keyword *internalized* in Example 17 revealed negative attitudes of being LGBTQ toward self-acceptance as an LGBTQ person. Also considered as previously key, the keyword *minority* was initially key in 2001-2005, but not key until 2011. From the concordance lines, it was seen that *minority* was used with another word to discuss a particular context, as seen in Example 18.

> Example 17: ...link between **internalized homophobia** and distress operated through low self-esteem and perceptions of a lack of social support.
> Example 18: ...gay people as **a minority group** in terms of both cultural difference and non-normativity.

However, the keyword *discrimination* was recently key since it has been found in corpora since 2006. In Example 19, *discrimination* was discussed in line with another issue to represent a form of hate.

> Example 19: ...gay/lesbian persona was seen as necessary to protect themselves from **judgement, discrimination, and conflict** with others.

Table 7

Keywords under the Topic of Research over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Participants	✓	✓	✓	✓
Research	✓	✓	✓	✓
Study	✓	✓	✓	✓
Respondents	✓			
Interviews			✓	
Interviewees			✓	
Meanings	✓			

Table 7 shows the keywords under the topic of *research*. There were only three keywords in the analysis, *participants*, *research*, and *study*, which were considered always key. It was found that these keywords directly referred to research procedure by looking at the concordance lines in Examples 20-22.

> Example 20: ...empirical **research on lesbian and gay families** can successfully dispel homophobic attitudes and eliminate legal obstacles faced by such families.
> Example 21: In these groups, **participants** were asked to watch three separate 20-minute collections of edited clips from the second season of the US Queer as Folk.
> Example 22: ...conducting **a qualitative study of gay men** in New York to investigate the idea that people belong to different groups within a community, ...

Table 8

Keywords under the Topic of Theory over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Identity	✓	✓	✓	✓
Orientation	✓	✓	✓	✓
Theory	✓	✓	✓	✓
Discourses	✓	✓	✓	✓
Counselling		✓		
Described			✓	
Binary			✓	

As shown in Table 8, the following keywords in the topic of *theory* were included in the analysis: *identity*, *orientation*, *theory*, and *discourses*. These four keywords were found to be always key in all periods. Additionally, these keywords were discussed with another area of study, such as mental health, sociology, and law, in Examples 23-26.

> Example 23: ...to have others recognize their **sexual identity** and to be visible in solidarity with other queer Instagrammers.
> Example 24: ...**sexual orientation** and mental health are essentially independent dimensions of psychological function...
> Example 25: ...we engage **queer theory** and sociology with more careful attention to the possibilities and limitations of these respective approaches.
> Example 26: gender sexuality – is a master concept that is central to all legal, academic, and popular **discourses of gender and sexuality**.

Table 9

Keywords under the Topic of Culture over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Culture	✓	✓		
Cultural	✓	✓		✓

Table 9 shows the two keywords *culture* and *cultural* grouped into the topic of *culture*. These terms were considered as previously key, that is, *culture* was key in the former two periods (2001-2005 and 2006-2010). Similarly, *cultural* was discussed in the periods 2001-2005 and 2006-2010, and again in 2016-2020. From concordances in Examples 27-28, the keywords *culture* and *cultural* were discussed with another topic.

> Example 27: Lesbian and gay historiography has been routinely mapped onto capitalist industrialization and urbanization: **desire, identity, community or culture, and political movement**.
>
> Example 28: ...a deep **cultural and political disconnect** with lesbian women who do not share and understand her experiences and history as an African woman who identifies as a lesbian.

Table 10

Keywords under the Topic of Socio-Politics over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Social	✓	✓	✓	✓
Experience	✓	✓	✓	✓
Feminist	✓	✓		✓
Politics	✓			
Political	✓			
Practices		✓		✓
Disclosure		✓		
Status		✓		
Workers		✓		
Belonging			✓	

Table 10 shows the keywords in the topic of *socio-politics* between 2001 and 2020. The four keywords, *social*, *experience*, *feminist*, and *practices,* were examined to see their changes. From Table 10, we can see that the keywords *social* and *experience* were key in all periods. Furthermore, both *social* and *experience* were used as a way of discussing political perspective in Examples 29-30.

> Example 29: ...today's **social and political struggles** of marginalised and excluded minorities are unable to connect to each other.
>
> Example 30: That fragility is also recognised in the context of the wider landscape and **experience of inequality** between gays and straights.

As found in 2001-2005, 2006-2010 and 2016-2020, the keyword *feminist* was then identified as previously key. As seen in Example 31, *feminist* was discussed in the context of political movements.

> Example 31: ...the priorities of **feminist politics** are rooted in redefining the concepts of power and privilege and thus reversing and rebuilding a patriarchal world.

Finally, the keyword *practices* was considered as recently key, as it became key in 2006-2010 and 2016-2020. In Example 32, the keyword *practices* was used to show political issues associated with LGBTQ people.

> Example 32: ...the **practices of queer politics** in India have provided some space for the LGBT groups towards a coalitional and open political networking with women's groups.

Table 11

Keywords under the Topic of Standard over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Norms		✓	✓	✓
Heteronormative		✓	✓	✓
Normative			✓	✓
Heteronormativity				✓
Role			✓	

The keywords in the topic of *standard* between 2001 and 2020 were presented in Table 11. The three keywords, *norms*, *heteronormative*, and *normative*, were recently key; the keywords *norms* and *heteronormative* have been used since 2006, and the keyword *normative* has been used since 2011. As shown in Examples 33-35, the use of these keywords indicated acceptable standards or values based on the concordance analyses.

> Example 33: ...prevailing **norms of heterosexuality** can be understood as operating on multiple social levels.
> Example 34: ...**heteronormative practices** have negative repercussions beyond gay and lesbian subjects.
> Example 35: ...queer people, in particular, navigate their everyday lives and make sense of their identities and practices in relation to **normative social centers**.

Table 12

Keywords under the Topic of Science over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Donor	✓		✓	
Body		✓	✓	
Desire	✓		✓	✓
Transition			✓	✓
Biological		✓		
Psychological			✓	

Another set of keywords relating to the topic of *science* is shown in Table 12. The following keywords were included in the analysis: *donor*, *body*, *desire*, and *transition*. By tracking their changes, the keywords *donor*, *body*, and *desire* were seen as previously key while the keyword *transition* was recently key. Although the three keywords were identified as previously key, they changed differently over time. The keyword *donor* was found in 2001-2005 and 2011-2015 while *body* was key in 2006-2010 and 2011-2015. The keyword *desire* was discussed in three periods: 2001-2005, 2011-2015, and 2016-2020. *Transition* has emerged as a keyword since 2011.

The concordance lines indicated that *donor* and *transition* could be understood as a way of discussing advances in medical practices in Examples 36-37. Meanwhile, the keyword *body* was related to science in terms of aesthetics in Example 38. This was different from the keyword *desire*.

Associated with the psychological process, the keyword *desire* indicated a field of science studies in Example 39.

> Example 36: ...celebration of interesting new types of family; especially father-free families constituted through **donor insemination**.
> Example 37: A transitioning partner can bring tremendous change in the form of **physical transition** and sexual complexities.
> Example 38: **Body image** may be one way for gay men to highlight differences between gay and straight men.
> Example 39: The liberation of **homosexual desire** to be conducive to new modes of relation (a group identity) opposed to an individualised notion of the sexual self.

Table 13

Keywords under the Topic of Place over Time

Keyword	2001-2005	2006-2010	2011-2015	2016-2020
Ways	✓	✓		
Space	✓			✓

Finally, Table 13 shows a set of keywords under the topic of *place*. There were only two keywords in this topic: *ways* and *space*. These two words as previously key keywords emerged in 2001-2005, but the two keywords changed at different periods of time. In other words, *ways* was discussed until 2010 while *space* became key again in the period of 2016-2020. In Examples 40-41, it is suggested that *ways* and *space* referred to a place for LGBTQ people.

> Example 40: The **ways** that gender play and infusions of sexual feeling lead to such ambiguities of both sexual and cultural personas have been central to the development of Tropicana's performance art, and of the Carmelita Tropicana persona in particular.
> Example 41: Any aspiring competitive city must have a themed gay **space** and the only type of gay space that an aspiring competitive city can have is a themed one.

Based on the findings for RQ 1, we can see that the LGBTQ-related topics in LGBTQ-related research articles between 2001 and 2020 were *sexuality*, *gender*, *relationship*, *hate*, *research*, *theory*, *culture*, *socio-politics*, *standard*, *science, place*, *sport*, and *nationality*. Regarding RQ 2, we examined the changes of the LGBTQ-related topics between 2001 and 2020 by looking at the keywords in each topic. The findings for RQ 2 revealed that there were changes in the LGBTQ-related topics over time by using the criteria in tracking keywords, i.e., always key, previously key, and recently key.

Discussion

Prior to discussing the findings and drawing implications, there are some limitations to the findings that should be noted. Firstly, the data sets are solely of results for LGBTQ-related research articles. Secondly, the analysis is strictly on LGBTQ-related research articles between 2001 and 2020, spanning 20 years. Thirdly, the analysis has focused on limited sets of data of LGBTQ-related research articles from SAGE Journals Online. Despite these limitations, this study has provided an insightful analysis of the LGBTQ-related topics in research articles between 2001 and 2020 to present a better understanding of LGBTQ needs and concerns.

The findings have revealed the LGBTQ-related topics in LGBTQ-related research articles between 2001 and 2020 (RQ1) and the changes of topics over time (RQ2). RQ 1 uncovers the 13 LGBTQ-related topics including *sexuality*, *gender*, *relationship*, *hate*, *research*, *theory*, *culture*, *socio-politics*, *standard*, *science, place*, *sport*, and *nationality*. Six of the thirteen topics primarily support previous studies demonstrating that LGBTQ is discussed in terms of sexuality (Baker, 2014), gender (Wilkinson, 2022; Zottola, 2018), hate (Jones & Collins, 2020; Eichler, 2010), theory (Goldberg &

Allen, 2018), socio-politics (Love & Baker, 2015; Santonocito, 2020), and standard (Baider, 2018; Lev, 2010). Such topics may indicate the prominence of certain topics and ways of discussing them. Additionally, our analyses reveal that LGBTQ people are also discussed under the topics of *relationship*, *research*, *culture*, *science*, *place*, *sport*, and *nationality*. This may emphasize more wide-ranging LGBTQ discussions, regarding LGBTQ-related topics. To examine the LGBTQ-related topics, we conducted a diachronic corpus analysis focusing on changes of the topics over time.

Regarding diachronic comparison, RQ 2 reveals changes of the LGBTQ-related topics over time. As mentioned above, some LGBTQ-related topics are consistent with previous studies, which could imply that these topics have become well-established LGBTQ topics. This seems to indicate that the majority of research articles generally tended to center on the topics related to inter-relatedness of binary oppositions (normality and abnormality) and homophobic practices. Thus, we look at other LGBTQ-related topics to provide additional and more extensive discussions about the LGBTQ community. Due to space constraints, the discussions focus on the topics of *science*, *relationship*, and *culture*. The reasons for discussing the three topics are related to the keywords in these three topics emerging as key in the same period and the linkages between the topics.

For the topic of *science*, the findings that emerged from the analysis identify the two diachronic patterns in this topic: previously key and recently key. The keyword *donor* was previously key while the keyword *transition* was recently key in the corpora. From a diachronic perspective, the topic has changed over time. That is, the two keywords particularly indicate a shift away from the focus of donor to transition. Rather, the keywords from concordance lines show the ways that LGBTQ people normalize their lives, for example, "...celebration of interesting new types of family; especially father-free families constituted through *donor* insemination". From this concordance line, it can be interpreted that LGBTQ topics have challenged how society defines normality. At the same time, it seems very difficult to do so due to LGBTQ-related laws. This is consistent with Santonocito (2020) who pointed out that LGBT may be in a problematic condition because of civil rights issues. A possible explanation is about legal complexity (Lewis & Kern, 2018) and law enforcement (Eichler, 2010) which may also force LGBTQ people into uncomfortable situations.

Another discussion is about the topic of *relationship*. The diachronic data indicates the changes of the topic based on its keywords. The use of keywords over time points to the three diachronic patterns: previously key, recently key, and always key. Some keywords under the topic of *relationship* have disappeared in a period of time, indicating that the topic has changed over time. For in-depth investigations, it is similar to the topic of *science* that LGBTQ people may be in uncomfortable situations. In terms of family relationship, it is perhaps difficult to come out as LGBTQ in a family as seen in this instance: "although the lesbian *mothers* were often able to renegotiate their relationship with their parents, this did not however necessarily mean that the grandparent generation fully accepted the lesbian relationship as a proper relationship or as a basis for a family". This may be an impact of value orientations in heterosexual relationships. This in turn is likely to reflect beliefs of people about which actions are preferable (Baker, 2014), acceptable, and appropriate for men and women (Cislaghi & Heise, 2020; Kollmayer et al., 2018).

However, the topic has remained constant over time in particular contexts, i.e., romantic relationships and in-group relationships. It is possible to frame romantic relationships and in-group relationships as dominant issues over time. Based on concordances, it seems that LGBTQ couples or partners can maintain their relationships which can be seen in the following concordance line: "These social decisions were sometimes made by the transgender *partner* and other times were mutually decided". Additionally, the concordance line of in-group relationships also shows that LGBTQ members support one another when they are in a group. For example, "...there was a role for the LGBT *community* in supporting older members...". This suggests that LGBTQ people have supportive relationships, which in turn may underpin respect among LGBTQ members (Mekler, 2018), leading to the solidarity of the LGBTQ community.

Finally, the topic of *culture* has changed over time. In this topic, keywords are previously key across the four corpora. From concordance lines of the keywords, LGBTQ culture can be considered as an example of solidarity, for instance: "Lesbian and gay historiography has been routinely mapped onto capitalist industrialization and urbanization: desire, identity, community or *culture*, and political movement". This suggests that LGBTQ culture can lead to social movements, i.e., the Stonewall Riot as a gay liberation movement (Nelson, 2015).

As discussed, we can see that coming out as LGBTQ people may not be easy. There are conditions leading to bias and discrimination. Such practices would make LGBTQ people different from other people. One way to bridge these gaps is inclusion. Particularly, it is suggested that LGBTQ inclusion helps in gaining a better understanding of LGBTQ people, particularly in academic contents (Kuvalanka, 2013). Considered as a source of information, a research article gives greater insight into the way LGBTQ people are discussed and represented. It may be useful to apply the method or framework to other textual data to explore LGBTQ topics in more contexts.

Conclusion

This study has focused on LGBTQ-related research articles between 2001 and 2020 with the aims to explore the LGBTQ-related topics in the articles and to examine the diachronic changes in the topics over time. The analysis has revealed groups of keywords, which in turn refer to the topics and their diachronic changes. The findings support the view that LGBTQ communities are under social pressure. On the other hand, the community has strong in-group relationships, leading to solidarity.

This exploratory study sheds new light on studying LGBTQ needs and concerns by looking into research articles. Further, the study increases our knowledge of LGBTQ-related topics to have a better understanding of LGBTQ people and issues. It should be also noted that further work may consider other types of written data such as scripts, blog posts, and comment threads as other possible ways of exploring LGBTQ-related topics.

Caveats and Recommendations for Further Research

Our findings have revealed a better understanding of LGBTQ-related topics. One of the limitations is that this study used only the intra-rater reliability check, leading to a potentially lower level of reliability, compared to inter-rater reliability checks. Thus, future research may take into account inter-rater reliability which may yield more convincing results. Another caveat is that the data used in this study was limited to research articles between 2001 and 2020 collected from SAGE Journal Online. These research articles were not representative of research articles from other online databases. To address this limitation, future research may use data from multiple sources, which will allow for a more comprehensive analysis leading to a clearer understanding of LGBTQ communities. Finally, the LGBTQ-related topics discussed in research articles were only related to the academic context. Future research may explore LGBTQ-related topics in media contexts and compare them to those in academic contexts. Particularly given the powerful influence of media over people's attitudes, this would provide a broader perspective on the issues and enhance our understanding of LGBTQ communities.

About the Authors

Sarunya Tarat: A PhD candidate from the English for International Communication (EIC) program, Kasetsart University, Bangkok, Thailand.

Nawarat Siritararatn: An assistant professor at the Department of Foreign Languages, Faculty of Humanities, Kasetsart University, Bangkok, Thailand.

Woravut Jaroongkhongdach: An assistant professor at the Department of Foreign Languages, Faculty of Humanities, Kasetsart University, Bangkok, Thailand.

References

Anthony, L. (2022). *AntConc* (Version 4.1.4) [Computer Software]. Waseda University. Available from https://www.laurenceanthony.net/software/antconc/

Baider, F. (2018). "Go to hell fucking faggots, may you die!" framing the LGBT subject in online comments. *Lodz Papers in Pragmatics*, *14*(1), 69–92. https://doi.org/10.1515/lpp-2018-0004

Baker, P. (2013). From gay language to normative discourse. *Journal of Language and Sexuality*, *2*(2), 179-205. https://doi.org/10.1075/jls.2.2.01bak

Baker, P. (2014). *Using corpora in discourse analysis*. Bloomsbury Academic

Biber, D., & Reppen, R. (2015). *The Cambridge handbook of English corpus linguistics*. Cambridge University Press.

Cislaghi, B., & Heise, L. (2020). Gender norms and social norms: Differences, similarities and why they matter in prevention science. *Sociology of Health & Illness*, *42*(2), 407-422. https://doi.org/10.1111/1467-9566.13008

Cohen, Y. (2017). Estimating the intra-rater reliability of essay raters. *Frontiers in Education, 2*, 1-11. https://doi.org/10.3389/feduc.2017.00049

De Witte, K., Iterbeke, K., & Holz, O. (2019). Teachers' and pupils' perspectives on homosexuality: A comparative analysis across European countries. *International Sociology*, *34*(4), 471-519. https://doi.org/10.1177/0268580919854295

Eichler, M. A. (2010). Joining the family: Experiences of being and becoming ally activists of LGBTQ people. *Journal of Transformative Education*, *8*(2), 89-102. https://doi.org/10.1177/1541344611406904

Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding. *Journal of English Linguistics*, *36*(1), 5-38. https://doi.org/10.1177/0075424207311247

Goldberg, A. E., & Allen, K. R. (2018). Teaching undergraduates about LGBTQ identities, families, and intersectionality. *Family Relations*, *67*(1), 176-191. https://doi.org/10.1111/fare.12224

Institute of Medicine of the National Academies. (2011). *The health of lesbian, gay, bisexual, and transgender people: Building a foundation for better understanding*. National Academies Press.

Jones, L., & Collins, L. (2020). PrEP in the press. *Journal of Language and Sexuality*, *9*(2), 202-225. https://doi.org/10.1075/jls.20002.jon

Kollmayer, M., Schober, B., & Spiel, C. (2018). Gender stereotypes in education: Development, consequences, and interventions. *European Journal of Developmental Psychology*, *15*(4), 361-377. https://doi.org/10.1080/17405629.2016.1193483

Kuvalanka, K. (2013). The "second generation": LGBTQ youth with LGBTQ parents. In A. Goldberg & K. Allen (Eds.), *LGBT-Parent Families* (pp. 163-175). https://doi.org/10.1007/978-1-4614-4556-2_11

Lev, A. I. (2010). How queer!-the development of gender identity and sexual orientation in LGBTQ-headed families. *Family Process*, *49*(3), 268-290. https://doi.org/10.1111/j.1545-5300.2010.01323.x

Lewis, M. M., & Kern, S. (2018). Using education law as a tool to empower social justice leaders to promote LGBTQ inclusion. *Educational Administration Quarterly*, *54*(5), 723-746. https://doi.org/10.1177/0013161x18769045

Longman Dictionary of Contemporary English. (n.d.) https://www.ldoceonline.com/

Love, R., & Baker, P. (2015). The hate that dare not speak its name? *Journal of Language Aggression and Conflict*, *3*(1), 57-86. https://doi.org/10.1075/jlac.3.1.03lov

Mekler, A. G. (2018). LGBTIQ (in)visibility: A human security approach to SOGIESC. In C. L. Mason (Ed.), *Routledge handbook of queer development studies* (pp. 155-168). Routledge.

Nelson, T. R. (2015, April 10). *A movement on the Verge: The spark of Stonewall* [Paper presentation]. MAD-RUSH Undergraduate Research Conference, Harrisonburg, VA, United States.

Pojanapunya, P., & Watson Todd, R. (2018). Log-likelihood and odds ratio: Keyness statistics for different purposes of keyword analysis. *Corpus Linguistics and Linguistic Theory*, *14*(1), 133-167. https://doi.org/10.1515/cllt-2015-0030

SAGE. (n.d.). *About us — SAGE publishing*. https://group.sagepub.com/about

Santonocito, C. S. (2020). LGBT* people in the speeches of Italian and British PMs: A corpus-assisted critical discourse analysis. *Critical Approaches to Discourse Analysis across Disciplines*, *11*(2), 187–212.

Wang, G., & Ma, X. (2020). Representations of LGBTQ+ issues in China in its official English-language media: A corpus-assisted critical discourse study. *Critical Discourse Studies*, *18*(2), 188–206. https://doi.org/10.1080/17405904.2020.1738251

Wilkinson, M. (2022). Radical contingency, radical historicity and the spread of 'homosexualism': A diachronic corpus-based critical discourse analysis of queer representation in the times between 1957-1967 and 1979–1990. *Discourse, Context & Media*, *48*, 1-9. https://doi.org/10.1016/j.dcm.2022.100623

Zottola, A. (2018). Transgender identity labels in the British press. *Journal of Language and Sexuality*, *7*(2), 237-262. https://doi.org/10.1075/jls.17017.zot