A Corpus-based Study of Sexist Language in the Hashtag #everydaysexism on Twitter

Wanwen Wang, Hong Kong Metropolitan University, Hong Kong

Abstract

This paper examines the Twitter hashtag #everydaysexism as the discourse of sexist language about the digital feminist movement, focusing on the fourth wave of feminism. Twitter users use #everydaysexism to "shouting back" and expose their experiences relating to sexism in daily life. The corpora consist of 1118 tweets in the hashtag #everydaysexism that include all the English tweets posted within 12 months (from April 1, 2020, until March 31, 2021). After conducting Melville et al.'s (2019) thematic model and drawing on Mills' (2008) sexist language framework, the workplace and customer service domains were found to have an overwhelming share, with 24.14% being overt and 75.86% indirect sexism. Even though overt sexism is not the primary sexist form in this domain, but it is still relatively apparent in its quantity, especially in first names, surnames, titles and generic pronouns, nouns. This study concludes by analyzing the different sexist language markers to reflect some issues regarding gender differences and to signal people to think about their behaviour and speech.

Keywords: Sexist Language, Hashtag Feminism, #Everydaysexism, Overt Sexism, Indirect Sexism

Introduction

Sexist language is the bias against women in language structures and in its use (Weatherall, 2016). The sexism in language has been debated within feminist circles since the 1960s (Mills, 2008). Over the past six decades, women have increased awareness of gender equality. In recent years, thousands of women have been using online media platforms to discuss, promote, and activate gender equality and social justice (Femfuture, 2021). Micro-blogging websites, especially Twitter, are increasingly becoming forums for public debate and articulation of feminist protest (Ricarda & Elke, 2016). Hashtags (i.e. a hash mark # followed by a thematic word or phrase) can make Twitter searchable and find a digital home for discussion (Drüeke & Zobl, 2006). This study aims to provide some insight into the sexist language features around the hashtag #everydaysexism to reflect some issues regarding gender differences and to signal people to think about their behaviour and speech. Firstly, I shall review the current literature and the background of this study's object. Next, I shall introduce a detailed description of the methodological overview. Finally, I will explain the results via data sampling and elaboration.

Literature Review

Hashtag Feminism and the Everyday Sexism Project

Over recent years, hashtag feminism has become a central part of the feminist media repertoire; digital spaces also define fourth-wave feminist practices and allow for global engagement (Clark-Parsons, 2021; Guillard, 2016; Turley & Fisher, 2018). In recent years a growing number of scholars (e.g. Antunovic, 2019; Brantner et al., 2020; Golbeck et al., 2017; Horeck, 2014) have turned their attention to feminism in social media. Many hashtags (e.g. #YesAllWomen, #SolidarityIsForWhiteWomen, #heforshe, #FemFuture, #TwitterFeminism, #Everydaysexism) provide a visible platform for feminists to "shout back" to expose the experiences relating to sexism while sharing and reacting to it and provoking responses (Turley & Fisher, 2018). #everydaysexism is a hot hashtag initiated by the *Everyday Sexism Project* to inspire women not to be passive recipients of sexism. Thus, this research project investigates everyday dominant discourses of sexism under this hashtag. While the database contains multiple stories and experiences posted by disparate Twitter users (known as collective storytelling), naturally, there is a mix of sexist practices related to various topics. Some researchers have attempted to categorise the themes under this hashtag, notably Melville, Eccles and Yasseri (2019), who derived two topic modelling. Here, I will build on their (2019) findings to further explore which domains and what forms sexism is mainly present in present-day society. Last but not least, although the hashtag #everydaysexism provides millions of tweets for readers, for the time being, it remains under research and deserves more attention, especially from a discourse-analytical perspective.

Overt and Indirect Sexism

Feminist linguist Sara Mills (2008) has a creative look at the sexist language; she creates a sexist language framework on overt and indirect criticism. In her (2008) book, Language and Sexism, she stresses "overt sexism is more difficult to articulate these days", more indirect sexism has developed "which manages to express sexism whilst at the same time denying responsibility for it" (p.12). However, Bates (2013), founder of the *Everyday Sexism Project*, has reported many overt sexisms, ranging from domestic violence to threats of torture, death and rape, that exists in the database. As Jenkins and Finneman's (2018) study also shows, workplace abuse and harassment are overt and always slip between existing laws and policies.

Even in the animated film Zootopia, there is a more pronounced overt sexism (Naovaratthanakorn, 2017) (see also Fiaunillah (2015) The Lord of the Rings). Another study also analyses the feminist hashtag #aufschrei on Twitter and finds that anti-feminists use it to post many misogynistic messages (also of the overtly sexist variety) (Drüeke & Zobl, 2006). This discrepancy can be attributed to Mills' argument based on the fact that overt sexism is no longer socially acceptable, but the media feminist movement and the anti-sexist movement were less influential in the third wave of feminism than the current fourth wave. Tellingly, but not surprisingly, representations of sexism are largely context-specific; for example, direct sexism in some formal settings is difficult to articulate. Therefore, this study aims to analyse whether overt sexism is more challenging to express in a particular domain (e.g. media, workplace, school).

It is worthy to note that Sarrasin, Gabriel, and Gygax (2012) found that "research has so far only focused on forms of sexism characterised by antipathetic/negative attitudes toward women" (Sarrasin et al., 2012, p.114); this type of sexism, namely, overt sexism. Therefore, this paper will highlight previous studies' inadequacies and broaden the focus by considering more invisible forms of sexism - indirect sexism. Unlike overt sexism, indirect sexism is not intentionally harmful. Hence, indirect sexism might go unnoticed if those indirect behaviours are not defined as gender discrimination might not get more attention than it deserves. The detailed methodology for addressing the following research questions will be provided in the next section to achieve those intentions.

1)   Under the hashtag #everydaysexism on Twitter, which domain or topic has more sexism existing? If so, is there more overt sexism or indirect sexism?
2)   How are the linguistic markers employed in these overt and indirect sexism?

Methodological Overview

A corpus-based methodology is adopted in this paper: the corpus data consists of 1118 tweets that include all the English tweets posted within 12 months (from April 1, 2020, until March 31, 2021) in the hashtag #Everydaysexism on Twitter. It, in total, involves 37,579 tokens. Additionally, for the objectives of this paper, the corpus has to satisfy four criteria: those tweets only focus on women's experience of sexism; the contents are semantically clear and in English; those tweets are not replies and links to avoid duplicate data collection. Finally, 835 tweets from 1118 have been selected as research data. This paper involves, first and foremost, a quantitative analysis to provide an overview of the research object, and then it covers a qualitative analysis. The topic modelling (Melville et al., 2019) is used to conduct a thematic analysis. The sexist language framework (Mills, 2008) is used to reveal the overt and indirect sexist linguistic features. The first step is to analyse which areas or topics appear mostly sexist under this feminist hashtag, and in what ways is sexism present, directly or indirectly? Finally, the analysis of how sexism is manifested in the linguistic markers. The following sections describe the findings and discussion of this study.

Results and Discussion

To address the first research question, investigate which domain or topic under this hashtag has the most sexism in quantity, i.e. inferring which domains sexism is mainly occurring today. Accordingly, after conducting a thematic analysis using Melville, Eccles, and Yasseri's (2019) model, it was found that topic S2 (work, office, company, customer) had the most significant number of sexism cases, as shown in Table 1.

Topic number	Topic	Number	%
S0/S3	public space/transport/street harassment	110/1118	9.84%
S1	online/comments	197/1118	17.70%
S2	work/office/company/customer	290/1118	25.94%
S4	school/teacher/uniform	41/1118	3.67%
S5	media	182/1118	16.37%
S6	domestic abuse/relationships/home	15/1118	1.34%
Total		835/1118	74.78%

Table 1: The Distribution of the Topics

Note: The original topic modelling had seven topics, while this study codes S0 (public space/street harassment) and S3 (transport/street harassment) together because they all belong to sexism in the public sphere.

From the results in the above table, it emerges that indirect sexism has overtaken overt sexism in number as the dominant sexism in topic S2 (work, office, company, customer). As mentioned previously, Mills (2008) stresses that "overt sexism is more difficult to articulate these days." (p.12). However, overt sexism at 24.14% is not a small percentage; therefore, it is still relatively evident in the work domain. Table 2 provides an overview of the features of sexist language based on Mills' (2008) framework. Regarding the second research question, this section will combine a quantitative and qualitative analysis to reveal the indirect and overt sexist linguistic markers employed in the workplace and customer service domains ("S2" work, office, company, customer). As is clearly shown in table 2, the total number of indirectly sexist language found was 220 out of 290 tweets. The 127 tweets (43.79%) are presupposition; 51 tweets (17.59%) are androcentric perspectives; 18 tweets (6.21%) are humour; 15 tweets (5.17%) are collocation; 9 tweets (3.10%) are scripts and metaphor.

Overt Sexism			
Subtypes	Rank	Occurrence	%
(f) First names, surnames and titles	1	32/290	11.03%
(c) Generic pronouns and nouns	2	11/290	3.79%
(d) Insult terms for women	3	10/290	3.45%
(e) Semantic derogation	4	8/290	2.76%
(a) Naming	5	5/290	1.72%
(g) Transitivity	6	2/290	0.69%
(b) Dictionaries	7	1/290	0.34%
(i) Jokes	8	1/290	0.34%
(h) Reported speech	9	0/290	0.00%
TOTAL		**70/290**	**24.14%**

Indirect Sexism			
Subtypes	Rank	Occurrence	%
(b) Presupposition	1	125/290	43.79%
(f) Androcentric perspectives	2	49/290	17.59%
(a) Humour	3	16/290	6.21%

(e) Collocation	4	15/290	5.17%
(d) Scripts and metaphor	5	9/290	3.10%
(c) Conflicting messages (ISC)	6	6/290	0.00%
TOTAL		**220/290**	**75.86%**

Table 2: Overview of Overt and Indirect Sexism (Mills, 2008)

Presupposition

The frequencies of masculine terms that appeared in the 127 tweets are related to sexism at the level of presupposition, such as "he" and "husband". The following datum is taken to reveal the masked sexism. These findings show that when men and women are in the same work setting, and men are more likely to be assumed to be directors, have higher status and greater power - can be attributed to the fact that here, "director" is always presupposed to be a male referent (consider example 1) (Eckert & McConnell, 2003). Moreover, those presuppositions also evoked the notion of stereotypes. In this sense, stereotypical assumptions assume that women's primary working role is subordinate to men.

Rank	Frequency	Words
1	49	he
2	41	husband
3	22	male
4	20	him
5	15	man
6	11	guy
7	11	his
8	7	boyfriend
9	5	men
10	4	boys
11	4	dad
12	2	boy

Table 3: Male Words Used in Women's Sexist

1)   Is this #everydaysexism? I set up a company with my husband as a co-director. I am principal shareholder and only fee earner. He's receiving letters from accountants and service providers. I'm not."

Androcentric Perspectives

To draw on an example, it constitutes a form of indirect sexism that focuses on the male perspective. The women who are allowed to come into the boards, but simply that they only have less than half the quotas. It implies that men are more inclined to be accepted as board members in this demographic imbalance of gender-skewed working fields. It carefully avoids allegation of sexism; as Mills (2008) stresses when the working setting's androcentrism is not foregrounded, this type of indirect sexism is hard to identify.

2)   Getting boards to have quotas for #women is great. But why are they setting them at 40% when we are 50% of the population? #everydaysexism.

Humour

The following datum can illustrate that some high-frequency words seem to be "positive" appraisals of women's appearance rather than their workability, such as "lovely" and "cute". This argument is also supported by Jenkins and Finneman (2018), who indicates the audiences care more about female journalists' physical appearance; the high frequent evaluation words are "attractive", "glamorous", "look trustworthy", "new hairstyle", or "clothes". These words serve the indirect semantic derogation of women. In example 3, the utterance "14-year-old daughter" is an operation of irony in the expression of sexism that exaggeratively lower women's age. As Mills (2008) indicates, the irony is a typical strategy for humorous remarks. Additionally, the phrase "make-up" is found in 4 out of the 16 tweets. All of them refer to the cosmetics applied to the face after checking the context of each tweet (consider example 4). The male character describes "paintbrushes" as the cosmetic applicator, which is the overt playfulness on women. Mills (2008) expresses that humour usually exaggerates certain features associated with a group or uses and gives play to stereotypical knowledge for comic effect. "make-up" is a stereotypical image for women who are only concerned with surface appearance (Mills, 2008), which creates a sense that women only can be good make-up girls, not good painters. In the same scene, it rarely uses similar linguistic markers towards male customers.

Frequency	Words/Word phrase
4	Make up
3	lovely
2	Wearing
2	Look
2	cute

Table 4: Valuing Women Based on Their Appearance

3)　　"You look like my 14 year old daughter" - unprofessional reply to a young entrepreneur seeking investment #r4today #everydaysexism.
4)　　Buying art supplies including new paint brushes..old guy behind counter: "now remember, these aren't for make up " Me: ....!!! #everydaysexism.

Collocation

As Mills (2008) defines, collocations are concerned with the company that words keep, and table 5 shows 12-word collocations of the word "women" in the collocation corpus. The word "woman" generally keeps collocating with other words like "single woman" and "unmarried woman" that deliberately emphasize women's emotional states with negative connotations in the working scene. Here is discriminatory because "single men" and "unmarried men" are rarely used in the same way. In addition, "looks", "beautiful", and "beauty" seems to be female-referent words, and they appear in high collocational patterns with the word women. While the same findings by Carroll and Kowitz (as cited in Mills, 2008), they also find "rich", "famous", and "brave" are assigned to male-referent words. While examining lexical collocates can provide a preliminary insight into sexist linguistic markers, it is impossible to understand comprehensively without putting them in an authentic context. Similarly, "childcare", "housework", and "cook" seem to be a uniquely female familial responsibility. As a result, for example (5), women are asked to prepare "tea and coffee" for colleagues in meetings, a scope of work that is always limited to women only, in opposites compared to men.

Words	Collocates
woman women, she her	looks, childcare, unmarried, beautiful, tea and coffee, single, baby, beauty, housework, cook, appearance, marriage

Table 5: The Lexical Collocates of Women

5)   #JackieWeaver asserts authority in a meeting, the men try to silence and belittle her. This is #everydaysexism. Every woman has an example. When I was a young woman working in a senior role I would routinely be asked to make the tea & coffee in meetings. No men ever were.

Scripts and Metaphor

In what follows, scripts are brought into the hospital about women. In this male doctor's utterance, he gives a scenario script to the female patient and positions women as "liars". The women's speech is widely regarded as untrustworthy in a male-centred society. Such a case also often appears in political elections. As Alingasa and Ofreneo's (2020) case study shows, one president states that the next ombudsman must not be a woman because they cannot be trusted and lack integrity. Thus, this narrative also can be interpreted as a maligning storyline against women (Alingasa & Ofreneo, 2020). As Mills (2008) expresses, those speeches are difficult to characterize as overt sexism.

6)   Every woman expiriances #everydaysexism I saw it when I worked in male dominated environment. I felt it when I went to the drs saying I had stomach pains and the dr said 'you're over exaggerating because you're a woman'. I hear it when men say 'smile love, it might never happen'.

Conclusion

This paper is a corpus-based study of sexist language in the hashtag #Everydaysexism on Twitter. An English corpus is collected and classified into six domains/topics and sexist language types. After two quantitative studies, the main conclusion is that, surprisingly, the workplace and customer service domain have more sexism cases, which mainly manifests indirectly. Subsequently, it analyzes the linguistic markers of sexism. Another notable result is that indirect sexism highly occurs at the level of presupposition and androcentric - an overwhelming proportion in this paper, and they become the most common forms of indirect sexism. Nevertheless, overt sexism is still relatively apparent in its quantity, proving that overt sexism is also easily articulated, especially in first names, surnames, titles and generic pronouns, nouns.

These findings in this study expose some of the micro-aggressions of sexism in the workplace and remind individuals to consider appropriate language and behaviour to prevent the possibility of sexism. As Mills (2008) points out, indirect sexism is more difficult to challenge than overt sexism because it is impossible to identify words and phrases as inherently or explicitly sexist (Mills, 2008). However, as this study demonstrates, indirect sexist discourses

are still widespread in the workplace, so more work is to be done in this area of research. The limitation of this paper is that the interpretation of discourse/pragmatically meaning is individual. In a word, there are differences in the discursive interpretation of sexism between individuals. As a brief example, describing a working woman as a "young lady" is perceived by some as an over-interpretation of sexism; the speaker may not be sexist, but by others as indirect sexism - again, this explains why indirect sexism is difficult to challenge.

Acknowledgements

References

Alingasa, A. P. T., & Ofreneo, M. A. P. (2020). "Fearless, powerful, Filipino": identity positioning in the hashtag activism of# BabaeAko. *Feminist Media Studies*, 1-17.

Antunovic, D. (2019). "We wouldn't say it to their faces": online harassment, women sports journalists, and feminism. *Feminist Media Studies*, 19(3), 428-442

Bates, L. (2013, April, 16). *The Everyday Sexism Project: A year of shouting back*. Retrieved from https://www.theguardian.com/lifeandstyle/the-womens-blog-with-jane-martinson/2013/ap r/16/everyday-sexism-project-shouting-back

Brantner, C., Lobinger, K., & Stehling, M. (2020). Memes against sexism? A multi-method analysis of the feminist protest hashtag #distractinglysexy and its resonance in the mainstream news media. *Convergence (London, England)*, 26(3), 674-696. https://doi.org/10.1177/1354856519827804

Clark-Parsons, R. (2021). "I see you, I believe you, I stand with you":# MeToo and the performance of networked feminist visibility. *Feminist Media Studies*, 21(3), 362-380.

Drüeke, R., & Zobl, R. (2016) Online feminist protest against sexism: the German-language hashtag #aufschrei. *Feminist Media Studies*, 16(1), 35-54. https://doi.org/10.1080/14680777.2015.1093071

Eckert, P., & McConnell-Ginet, S. (2013). *Language and gender*. Cambridge University Press.

Femfuture (2021). Available online at: http://www.femfuture.com/why-now/ (Accessed Mar 18, 2021).

Fiaunillah, T. K. (2015). Sexist language in "The Lord of The Rings" film trilogy. *Language Horizon*, 3(1), 143-149. https://ejournal.unesa.ac.id/index.php/language-horizon/article/view/11866/11094

Golbeck, J., Ash, S., & Cabrera, N. (2017). Hashtags as online communities with social support: A study of anti-sexism-in-science hashtag movements. *First Monday*.

Guillard, J. (2016). Is feminism trending? Pedagogical approaches to countering (Sl)activism. *Gender and Education*, 28(5), 609-626. https://doi.org/10.1080/09540253.2015.1123227

Horeck, T. (2014). AskThicke: "Blurred Lines," Rape Culture, and the Feminist Hashtag Takeover. *Feminist Media Studies*, 14(6), 1105-1107. https://doi.org/10.1080/14680777.2014.975450

Jenkins, J., & Finneman, T. (2018). Gender trouble in the workplace: applying Judith Butler's theory of performativity to news organizations. *Feminist Media Studies*, 18(2), 157-172.

Melville, S., Eccles, K., & Yasseri, T. (2019). Topic Modeling of Everyday Sexism Project Entries. *Frontiers in Digital Humanities*, 5, 28. https://www.frontiersin.org/article/10.3389/fdigh.2018.00028

Mills, S. (2008). *language and sexism*. Cambridge University Press.

NAOVARATTHANAKORN, M. P. (2017). *A CRITICAL DISCOURSE ANALYSIS OF WOMEN'S LANGUAGE FEATURES USED IN AN ANIMATED MOVIE* (Doctoral dissertation, THAMMASAT UNIVERSITY).

Ricarda, D., & Elke, Z. (2016) Online feminist protest against sexism: the German-language hashtag #aufschrei, *Feminist Media Studies*, 16:1, 35-54. https://doi.org/10.1080/14680777.2015.1093071

Sarrasin, O., Gabriel, U., & Gygax, P. (2012). Sexism and attitudes toward gender-neutral language: The case of English, French, and German. Swiss Journal of Psychology, 71(3), 113–124. https://doi-org.ezproxy.lib.ouhk.edu.hk/10.1024/1421-0185/a000078

Talbot, M. (2019). *Language and gender*. John Wiley & Sons

Turley, E., & Fisher, J. (2018). Tweeting back while shouting back: Social media and feminist activism. *Feminism & psychology*, 28(1), 128-132.

Weatherall, A. (2016). Sexism in Language. *The Wiley Blackwell Encyclopedia of Gender and Sexuality Studies*, 1-2.

Contact email: amberw123@163.com