



TECHNICAL EXPLAINER: CORPUS LINGUISTICS ANALYSIS

Methodology

This briefing sets out in detail the methodology of the Corpus Linguistics Analysis (CLA) into how the British press typically presents issues related to trans people, and whose voices are given prominence.

CLA is the statistically driven analysis of large, representative collections of naturally occurring language data. CLA methods are particularly adept for analysing how different forms of media represent different groups and/or issues, because they allow an analyst to look across large collections of texts that would normally be impractical to do manually. Broadly speaking, CLA quantifies frequencies of occurrence of individual words and grammatical structures in order to analyse them within context.

While CLA comprises several methodologies, for this analysis the following were used: collocation, N-gram, and concordance line analysis.

Collocation

Collocation looks at how the words that appear with a particular term affect how that term is interpreted, revealing patterns that may not be visible with manual reading alone. In this analysis, a span of 5 words to the left and 5 words to the right of the terms 'gender' and 'sex' was used and collocates of the search terms explored in context. Given that a search term could have hundreds of collocates, the analysis considered the top 10 collocates in any given month of an outlet's reports.

N-grams

N-grams are similar to collocation, as they reflect what words occur directly next to each other. N-grams can take any number of words (where N = the number of words that occur together to analyse). For example, bigrams represent two words occurring with each other, while trigrams represent three, and so forth. While quantifying these instances of words occurring together is a useful window into what is being presented,

it is also important to examine how those words are used in contextⁱ, which is where concordance line analysis becomes particularly useful.

Concordance lines

Concordance lines are a way of viewing individual words within context. CLA software can be used to present these contextualised uses in context as a list, which allows an analyst to see the vast ways the same word might be used in each individual article. This meant it is possible to explore every contextual use of a particular N-gram within any news report.

The corpora analysed

To examine how the British press typically present trans people, and whose voices are given prominence, four specialised corporaⁱⁱ were collected using the website Lexis+, which contains a regularly updated archive of all major UK national newspaper reports. The four outlets considered are the Guardian, the Sun, the Telegraph and The Times & Sunday Times, which represent a wide range of political leanings and audiences. Lexis+ includes both print and online articles and controls for duplication.

The corpora cover the period between January 2020 and April 2025, the month of the Supreme Court judgement in the appeal of *For Women Scotland Ltd v The Scottish Ministers* on interpretation of the words ‘sex’, ‘man’ and woman’ in the Equality Act 2010 for the purpose of sex discriminationⁱⁱⁱ.

The corpora were built around the search terms: ‘transgender’, ‘trans’, ‘transgendered’, ‘transsexual’, ‘trannie’, ‘tranny’, ‘mtf’, and/or ‘ftm’. While some of these terms are inappropriate, previous research has shown that these are various labelling strategies within the British press and therefore should be considered in building the corpora^{iv}.

The analysis was configured with a stop list of words and phrases that reduced the number of false positives from the search. This included terms such as: ‘transmitter’, ‘trans fats’, ‘Trans-Pacific’, ‘translate’, ‘TransPennine’, and so forth. For each article, the use of the search terms in context was done manually to ensure that the term being used related to transgender people, as opposed to something else. These files were then analysed with a free CLA software called Antcont.

The searched looked at:

- i) what terms collocated with ‘gender’, and how this changed across time
- ii) what terms collocated with ‘sex’, and how this changed across time
- iii) who was named in N-grams.

Importantly, given that a search term could have hundreds of collocates, within each month of reporting only the top 10 collocates with a 5-left 5-right word span of both gender and sex were analysed.

The N-grams search focussed on proper noun bigrams (i.e., two-word units indicating a specific person) and set a minimum frequency of occurrence at 8 per month. The bigrams included anyone's name if they appeared either as a full name (e.g., 'Boris Johnson') or if part of their name appeared in an N-gram (e.g., 'Boris says'; 'Johnson has'). However, other references to these people were excluded (e.g., 'the doctor', 'the Prime Minister').

Amnesty International UK is grateful to Dr Frazer Heritage, Senior Lecturer in Linguistics at the School of English of Manchester Metropolitan University, who carried out the CLA.

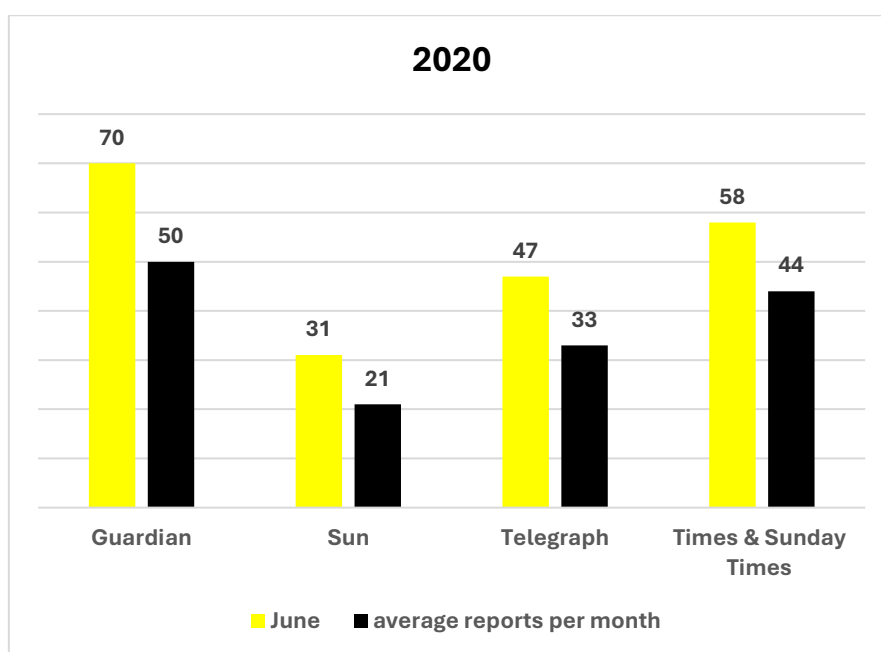
Spikes in coverage

The quantitative analysis of reports identified specific months when coverage spiked compared to the monthly average. The following tables show which events caused spikes in coverage. These are the most prominent relevant events identified for those months, however there may be other events that have caused spikes in coverage which are not as prominent.

2020

June

- J.K. Rowling authors an essay explaining ‘her reasons for Speaking out on Sex and Gender Issues’^v.



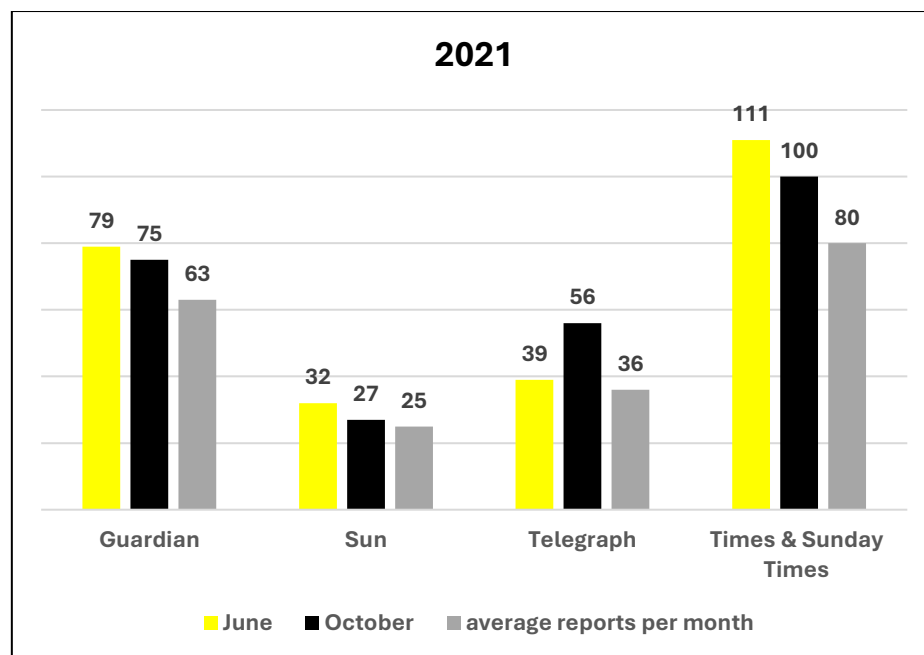
2021

June

- Maya Forstater wins her appeal tribunal^{vi}.
- Trans youth charity Mermaids challenges the Charity commission on its decision to grant LGB Alliance charity status^{vii}.

October

- Following accusation of transphobia Professor Kathleen Stock quits her job at Sussex University^{viii}.



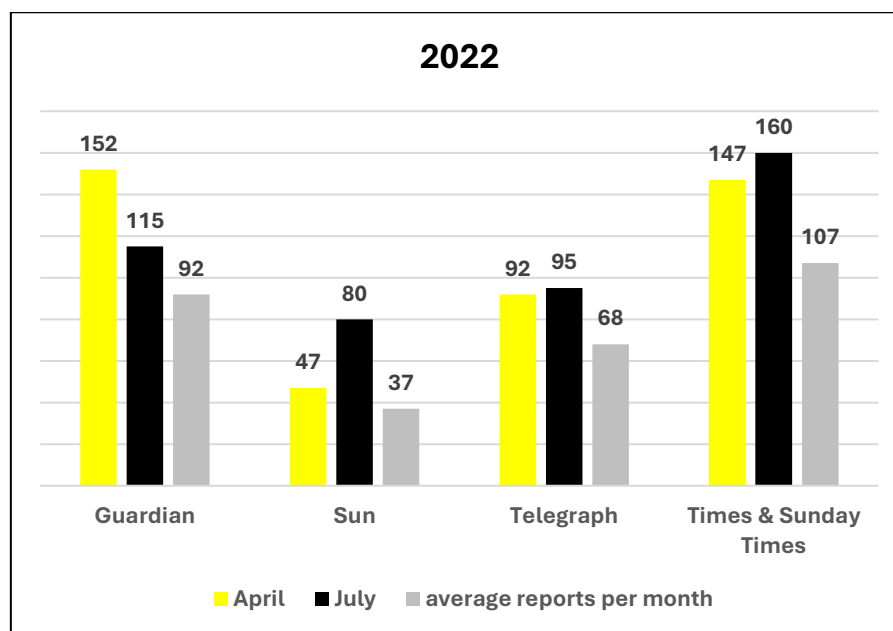
2022

April

- The government announces it will put forward a ban on conversion practices that excludes trans people^{ix}.
- British Cycling bans trans women from competing at elite level in its events^x.
- The EHRC publishes guidance on separate and single sex services^{xi}.

July

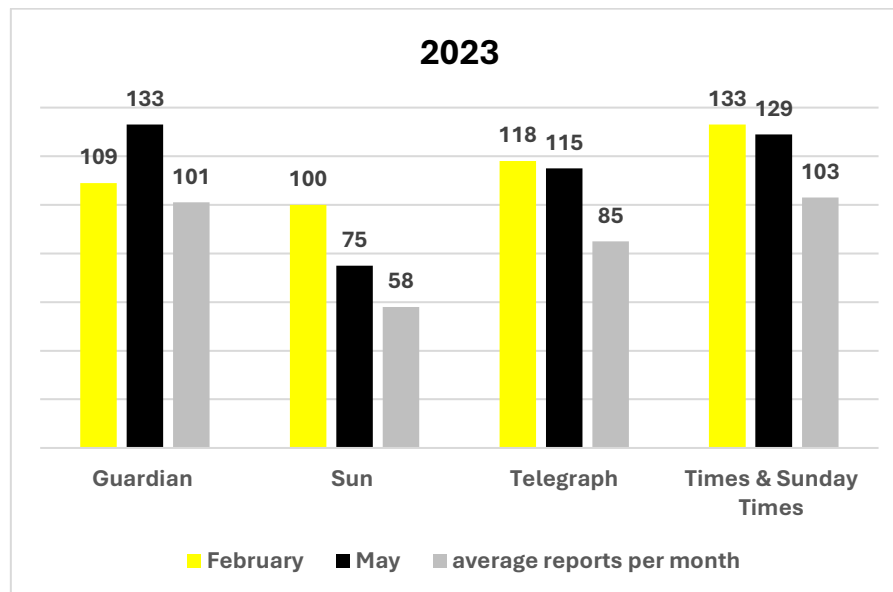
- For Women Scotland launches 2nd judicial review challenge against the Scottish Government^{xii}.
- Conservative party leadership elections^{xiii}.



2023

February: Brianna Ghey's murder.

May: Kathleen Stock gives a talk at Oxford University^{xiv}

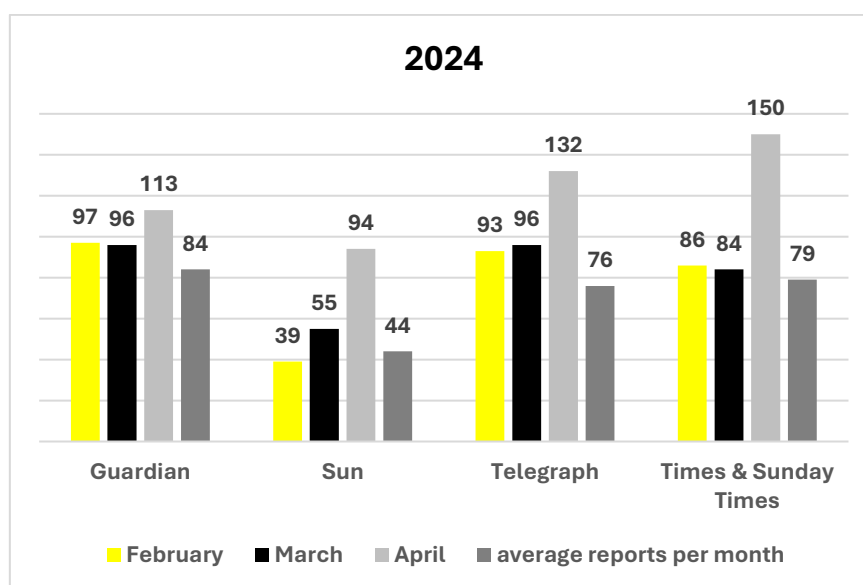


2024

February: For Women Scotland receives permission to appeal to the Supreme Court and J.K. Rowling supports the group with a £70,000 donation^{xv}.

March: The Department for Education's consultation on proposed 'Guidance for Schools and Colleges: Gender Questioning Children' closes^{xvi}.

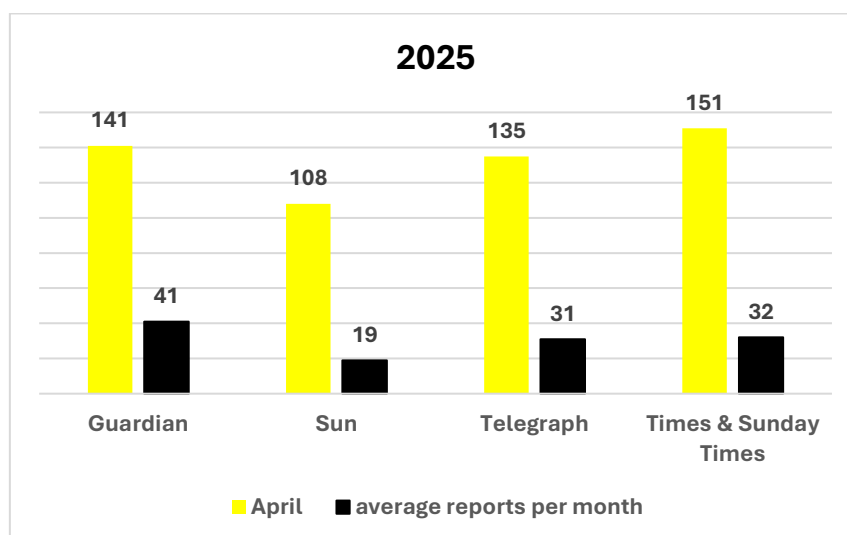
April: Final report of the Cass Review published^{xvii}.



2025

April

- Supreme Court Judgement *FWS v Scottish Ministers*^{xviii}
- EHRC's interim advice published.



List of GC groups identified

Name	Status	Type	Year established
Seen Network (Sex Equality and Equity Network Limited)	Registered company	Professional network	2022
Seen in the City	Informal	Professional network	2023
SEEN in HR	Informal	Professional network	2024
Parli SEEN	Informal	Professional network	2024
SEEN in STEM	Informal	Professional network	2024
SEEN in Journalism	Informal	Professional network	2024
SEEN in Schools	Informal	Professional network	2024
SEEN in Sport	Informal	Professional network	2024
SEEN in Health	Informal	Professional network	2024
SEEN in Retail	Informal	Professional network	2024
Local Authority SEEN	Informal	Professional network	2024
TU SEEN	Informal	Professional network	2024
SEEN in Publishing	Informal	Professional network	2024
Third Sector SEEN	Informal	Professional network	2024
Police SEEN UK	Informal	Professional network	2024
SEEN on Campus	Informal	Professional network	2025
FiLiA	Registered charity	Policy/advocacy/campaigns	2015
Bayswater support group	Registered company	Children/family	2019
Children of Transitioners	Informal	Children/family	2019
Get the L Out UK	Informal	LGB group	2018
Gender Critical Greens	Informal	Political party	2019
Labour Women's Declaration	Informal	Political party	2019

Legal Feminist	Informal	Policy/advocacy/campaigns	2018
Lesbian Strength	Informal	LGB group	2018
Merched Cymru	Informal	Policy/advocacy/campaigns	2020
OBJECT – Women Not Sex Objects!	Informal	Policy/advocacy/campaigns	2011
Older Feminist Network	Informal	Service provision	1982
Protect & Teach	Informal	Education	2024
Safe Schools Alliance UK (SSAUK)	Informal	Education	2018
Trans Widows Voices	Informal	Children/family	2018
Virago Women's Workshop	Registered company	Service provider	2021
LGB Alliance Cymru	Informal	LGB group	2019
Beira's Place	Registered company	Service provision	2022
Conservatives for Women	Registered company	Political party	2022
Fair Play for Women	Registered company	Policy/advocacy/campaigns	2019
For Women Scotland	Registered company	Policy/advocacy/campaigns	2020
Gay Men's Network UK	Registered company	LGB group	2022
Keep Prisons Single Sex	Dissolved	Policy/advocacy/campaigns	2022
Let Women Speak (formerly Standing for Women)	Registered company	Policy/advocacy/campaigns	2023
LGB Alliance	Registered charity	LGB group	2019
The Lesbian Project	Registered company	Policy/advocacy/campaigns	2022
Transgender Trend	Registered company	Education	2019
Woman's Place UK (WPUK)	Registered company	Policy/advocacy/campaigns	2021
Women's Declaration International (WDI)	Registered company	Policy/advocacy/campaigns	2019
Sex Matters	Registered charity	Policy/advocacy/campaigns	2024
Murray Blackburn and Mackenzie	Registered company	Policy/advocacy/campaigns	2022
Women's Rights Network (WRN)	Registered company	Policy/advocacy/campaigns	2022
LGB Christians	Informal	LGB group	2023
Liberal Voice for Women	Informal	Political party	2021
ScotPAG Scottish Professionals Advising on Gender	Informal	Professional network	2022

Women's Sports Union	Registered company	Policy/advocacy/campaigns	2025
----------------------	--------------------	---------------------------	------

-
- ⁱ Gillings, M.; Mautner, G. [Concordancing for CADS: Practical challenges and theoretical implications](#) International Journal of Corpus Linguistics, Volume 29, Issue 1, Feb 2024, p. 34 - 58
- ⁱⁱ These are large, representative samples of naturally occurring language - the singular of which is "corpus".
- ⁱⁱⁱ [For Women Scotland Ltd v The Scottish Ministers \[2025\] UKSC 16 \(16 April 2025\)](#).
- ^{iv} Baker, P. (2017). 'Bad Wigs and Screaming Mimis': Using Corpus-Assisted Techniques to Carry Out Critical Discourse Analysis of the Representation of Trans People in the British Press. In Cap. P & Hart. C, Contemporary Critical Discourse Studies (pp 211-236). *Bloomsbury*
- ^v J.K. Rowling [Writes about Her Reasons for Speaking out on Sex and Gender Issues](#), 10th June 2020
- ^{vi} [Forstater v CGD Europe & Ors \(RELIGION OR BELIEF DISCRIMINATION\) \[2021\] UKEAT 0105 20 1006 \(10 June 2021\)](#)
- ^{vii} [Mermaids v Charity Commission for England & Wales & Anor \[2023\] UKFTT 563 \(GRC\) \(06 July 2023\)](#)
- ^{viii} The Times, [Kathleen Stock resigns from University of Sussex after trans rights row](#), 29th October 2021
- ^{ix} BBC, [Conversion therapy: Ban to go ahead but not cover trans people](#), 1st April 2022
- ^x BBC Sport, [Transgender women no longer able to compete at elite female events run by British Cycling](#), 8th April 2022
- ^{xi} EHRC, [Guidance published for providers of single-sex services](#), 4th April 2022
- ^{xii} Scottish Legal, [Second judicial review brought against Scottish government over definition of 'woman'](#), 22nd July 2022
- ^{xiii} The Guardian, [Britain to have new PM by 5 September as Tory leadership rules announced](#), 11th July 2022
- ^{xiv} BBC, [Kathleen Stock: Protests at Oxford Union as talk goes ahead](#), 30th May 2023
- ^{xv} The Telegraph, [JK Rowling donates £70k to challenge ruling that men can become women](#), 18th February 2024
- ^{xvi} Department for Education, [Guidance for Schools and Colleges: Gender Questioning Children](#)
- ^{xvii} [Independent review of gender identity services for children and young people](#)
- ^{xviii} [For Women Scotland Ltd v The Scottish Ministers \[2025\] UKSC 16 \(16 April 2025\)](#)