



City Research Online

City, University of London Institutional Repository

Citation: Karaminis, T., Gabrielatos, C., Maden-Weinberger, U. & Beattie, G. (2023). Portrayals of autism in the British press: A corpus-based study. *Autism*, 27(4), pp. 1092-1114. doi: 10.1177/13623613221131752

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/32159/>

Link to published version: <https://doi.org/10.1177/13623613221131752>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Portrayals of autism in the British press: A corpus-based study

Themis Karaminis^{}, Costas Gabrielatos,
Ursula Maden-Weinberger and Geoffrey Beattie

Autism
2023, Vol. 27(4) 1092–1114
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/13623613221131752
journals.sagepub.com/home/aut



Abstract

Press representations of autism and autistic people both reflect and help shape public attitudes towards autism and neurodiversity and may establish critical barriers to social integration for autistic individuals. This study examined such representations in UK newspapers in the period 2011–2020 using a corpus-based approach. It also considered how press representations changed over time and differed with regard to reporting style (tabloids vs broadsheets) and political orientation (left- vs right-leaning). We created the *Autism UK Press Corpus*, which included all documents (~24K) referring to autism in 10 national newspapers. We used document counts (normalised by newspaper size) to assess the ‘newsworthiness’ of autism. We also employed a synergy of corpus-based and critical-discourse-analysis methodologies to study lexicogrammatical patterns and uncover explicit and implicit attitudes towards autism. Our results showed that the coverage of autism increased slightly over time, especially in broadsheets and left-leaning newspapers. Newspapers emphasised adversities associated with autism, often used negative language, and tended to focus on boys. These representations shifted gradually towards more difference-based descriptions and included more diverse age/gender groups, especially in broadsheets and left-leaning newspapers. We discuss the broad implications of these findings for the autism community and those interested in a more inclusive society.

Lay abstract

Any thriving society must recognise, accept and celebrate all of its diverse talent. But how accepting is British society towards autism and autistic people? This research addressed this question through the lens of the press since the press both reflects and helps shape public attitudes towards various social categories. We used specialised ‘corpus-based’ methods to carry out a large-scale study, which examined all articles referring to autism or autistic people in 10 national British newspapers in the period 2011–2020. We first investigated how often newspapers referred to autism. We found that the coverage of autism increased slightly over the years, suggesting that autism was becoming an increasingly newsworthy topic. Furthermore, the rise in autism coverage differed considerably between individual newspapers: it was more pronounced in the broadsheets than tabloids, and in left-leaning than right-leaning newspapers. But what was the focus of these articles? We found that newspapers emphasised the adversities associated with autism and portrayed autism with a lot of negative language. Newspapers also tended to focus on autistic children, and particularly on boys. There were some signs of change in more recent years, with some newspapers now representing autism as a difference and, in addition, referring to more diverse groups of autistic people. However, these changes tended to be confined to broadsheets and left-leaning newspapers. Our findings suggest that representations of autism in the contemporary British press are skewed towards stereotypically negative views, which may well hinder the acceptance of autism and the fostering of a more inclusive society.

Keywords

acceptance, autism spectrum disorders, corpus-based, critical discourse analysis, implicit attitudes, newspapers, representations

Introduction

Any thriving society must recognise, accept and celebrate all of its diverse talent. However, successful integration into society is a critical issue for autistic individuals

Edge Hill University, UK

Corresponding author:

Themis Karaminis, Department of Psychology, Edge Hill University,
St Helens Rd, Ormskirk, L39 4QP, UK.

Emails: themis.karaminis@edgehill.ac.uk; themkar@gmail.com

(Howlin & Magiati, 2017). Negative perceptions of autism (Cage et al., 2018) and the need for social camouflaging as a coping response to autism-related stigma (Hull et al., 2017, 2021; Perry et al., 2021) can significantly impact on the mental health of autistic people, who often present social anxiety and depression (Griffiths et al., 2019; Moss et al., 2015).

The acceptance of autism and autistic people is also central to the neurodiversity movement (Den Houting, 2019; Walker, 2012; Walker & Raymaker, 2020), which has gained increasing recognition in recent years and has been embraced by several leading autism researchers (e.g. Happé & Frith, 2020; Mottron, 2011; Nicolaidis, 2012; Pellicano & Stears, 2011; but see Jaarsma & Welin, 2012). The neurodiversity movement has shifted the understanding of autism from primarily deficit-based perspectives (cf. *Diagnostic and Statistical Manual of Mental Disorders* (5th ed; DSM-5; American Psychiatric Association, 2013)) to a more social understanding (Reindal, 2008; Shakespeare, 2006, 2010), which posits that disability arises from inadequate accommodations in physical, cognitive or social environments (Pellicano & den Houting, 2021). Thus, social difficulties in autism can be relational rather than a result of innate social impairments (Botha & Gillespie-Lynch, 2022; Davis & Crompton, 2021; Kapp, 2020; Kapp et al., 2013; Milton, 2012).

Importantly, the neurodiversity movement also advocates that human neurocognitive variation is the norm, and, although it can be challenging at times, it is essential for societal progress (Ballou, 2018; Kapp et al., 2013). This implies that autism ‘can be both an identity, and an embodied disability with aspects of impairment’ (Botha & Gillespie-Lynch, 2022, p. 94; see also Bailin, 2019; D. L. Baker, 2011; Ballou, 2018). Furthermore, the neurodiversity movement argues that autism and autistic people should be accepted and represented equally everywhere rather than cured or eradicated (D. L. Baker, 2011; Kapp et al., 2013).

However, the acceptance of autism and autistic people by society as a whole can be undermined by a variety of biases and stigma that are often perpetuated by the press and by the news media generally (Draaisma, 2009; Holton et al., 2014). The press reflects, constructs and indeed reconstructs public attitudes and beliefs towards a whole range of social categories (Corrigan et al., 2005; Edelman, 1988; Hall Jamieson & Waldman, 2004; Happer & Philo, 2013). This happens, first, by setting a news agenda (McCombs, 2004), which determines what is newsworthy about a topic (in this case, autism). Furthermore, news texts manipulate language to make certain pieces of information more salient than others. This process, referred to as framing (Entman, 1993; Hall Jamieson & Waldman, 2004; Lakoff, 2004), employs various linguistic means, such as ‘keywords, stock phrases, stereotypical images, sources of information, and sentences that provide

thematically reinforcing clusters of facts or judgments’ (Entman, 1993, p. 52).

A number of previous studies have examined the frames used in the press to portray autism and autistic people. These studies have employed content analysis methodologies, which typically involve some form of manual annotation of texts. For example, Holton et al. (2014) studied 473 articles from UK and US newspapers between 1978 and 2003. They found many stigmatisation cues, including an emphasis on social deficits and references to diagnostic labels and psychiatric conditions. In addition, so-called ‘loss frames’ (focusing on adverse outcomes of autism) were more frequent than ‘gain frames’ (focusing on treatments and cures). Finally, ‘episodic frames’, highlighting individual responsibility, outnumbered ‘thematic frames’ which, in contrast, emphasise societal issues around autism.

Other research has shown a high prevalence of negative narratives based on the disability-based model in articles from the *New York Times* published between 1973 and 2012 (Billawala & Wolbring, 2014; k=205 articles). Furthermore, Jones and Harwood (2009) suggested that in Australian newspapers (k=1228 articles, between 1996 and 2005), representations of autism lacked factual information and reflected a dual stereotype: autistic people are either dangerous and uncontrollable, or unloved and poorly treated. Huws and Jones (2011) also found stereotypical depictions of autism in the UK press between 1999 and 2008 (k=255 articles) and suggested that newspapers tended to offer the perspectives of parents, academics, health professionals and journalists rather than the perspectives of the autistic individuals themselves. Finally, a recent analysis of the Australian press found that articles tended to focus on the health and medical aspects of autism, and cure and treatments, rather than the more social dimensions (Baroutsis et al., 2021 k=1351 articles from 2016 to 2017).

A number of studies have examined both quantitative and qualitative changes in the representation of autism in the press over time. McKeever (2012) found that the coverage of autism increased steadily between 1996 and 2006 in the *New York Times* and the *Washington Post* (k=1517 articles). There was a decline in ‘science frames’ and a rise in ‘policy frames’ over this period, and treatments, interventions and policies were discussed consistently more frequently than the causes of autism. Yu and Farrell (2020) suggested that stigmatisation cues decreased over time in local and national newspapers in the United States published between 1998 and 2013 (k=982 articles), but they appeared in at least a quarter of articles published in 2013. Similarly, Lewin and Akhtar (2021) demonstrated a slow shift from disability-based perspectives to neurodiversity-based perspectives towards autism in the *Washington Post* between 2007 and 2016 (k=315 articles).

Other studies have examined the representation of autism across different newspapers and magazines. An early study (Robertson, 2009; $k=170$ articles) reported considerable variation in the number of autism-related articles published in five national UK newspapers in 2006, and in the attitudes of individual newspapers towards autism, including their stance regarding the possible link between autism and the measles, mumps, and rubella (MMR) vaccine. Furthermore, US magazines marketed as either ‘women’s’ or ‘gender-neutral’ offered different depictions of autism between 2000 and 2009, the former adopting subjective ‘tragedy frames’ and the latter more objective ‘medical and science frames’ (Clarke, 2012; $k=75$ articles). Moreover, local and national US-based newspapers employed similar frames in autism-related news, but to varying extents (Wendorf Muhamad & Yang, 2017; $k=413$ articles published in January 2014). Finally, with regard to non-English language newspapers, Tang and Bie (2016; $k=795$ articles, see also Bie & Tang, 2015) found that, between 2003 and 2012, the coverage of autism increased in five Chinese newspapers, with the articles depicting autistic people mainly as children, patients or savants. Pesonen et al. (2021) examined 210 articles from Finland’s largest daily newspaper between 1990 and 2016 and found that some articles offered balanced clinical perspectives aimed at educating the public, whereas others addressed societal aspects relevant to autism and included positive descriptions drawing upon Finnish cultural norms.

In this current study, we aim to contribute new results from a systematic investigation of press representations of autism in the United Kingdom over a 10-year period. We examined all documents referring to autism (~24K) in 10 major UK national newspapers between 2011 and 2020 to assess how portrayals of autism changed over time. We also examined the way autism was portrayed in different newspapers, focusing on reporting style (as a proxy of the readership’s educational level and socioeconomic status), comparing broadsheets and tabloids, and considering political orientation – examining and comparing left- and right-leaning newspapers.

We addressed these research aims using a different methodology than the content analysis employed in earlier research. Content analysis entails the risk of subjective, selective or impressionistic interpretations (see Breeze, 2011). To mitigate this risk, we used corpus-based critical discourse analysis (corpus-based CDA; P. Baker et al., 2013), a synergy of quantitative and qualitative methodologies from corpus linguistics and CDA. Broadly speaking, corpus-based CDA uses data-mining approaches to create large, representative and machine-readable datasets. These are analysed quantitatively to identify statistical trends in the occurrence (and co-occurrence) of particular words and grammatical structures. Such lexicogrammatical patterns index attitudes and the biases (either explicit or

implicit) of authors or speakers towards individuals, groups, actions, concepts or topics, even when the authors are unaware of their implicit biases or are consciously trying to obscure them. Qualitative analyses are also used in corpus-based CDA to consider broader textual contexts (e.g. paragraphs) and attitudes expressed in texts.

Our view is that corpus-based CDA can help uncover implicit attitudes and biases towards autism in the UK press. Our research should thus complement some of the findings from earlier studies, and offer fresh insights into the nature of these implicit attitudes and biases. Furthermore, we expect corpus-based CDA to reveal changes in the representation of autism in the British press over time (cf. Lewin & Akhtar, 2021; McKeever, 2012; Yu & Farrell, 2020), and differences in newspapers with different reporting styles and political orientation (cf. Robertson, 2009; also P. Baker et al., 2008, for representation of refugees and immigrants).

Method

The research design of this study is shown in Figure 1.

Corpus

We created the *Autism UK Press Corpus* by searching the *ProQuest European NewsStream*¹ database for documents related to autism from 1 January 2011 to 31 December 2020. The search included daily and Sunday, and print and online editions of 10 national UK newspapers: *Daily Express*, *Daily Mail*, *Daily Mirror*, *Daily Star*, *Daily Telegraph*, *Guardian*, *Independent*, *Observer*, *Sun* and *Times*.

We first designed an appropriate query for searching the *ProQuest European NewsStream* database. This query maximised the number of documents relevant to autism and minimised irrelevant documents (Chowdhury, 2004, p. 170; Gabrielatos, 2007, p. 6). It was constructed incrementally, starting from two clearly relevant terms: [autism*² OR autistic*]. More specifically, we employed a technique referred to as query expansion (Gabrielatos, 2007) to supplement the core query with candidate terms, selected from a pool of 11 terms: 8 terms referring to autistic people from Kenny et al. (2016) and 3 terms referring to neurodiversity. This technique used a Relative Query Term Relevance measure (RQTR; Gabrielatos, 2007, p. 14) to quantify the extent to which adding a new candidate term to our query resulted in the inclusion of new relevant documents in the corpus, while avoiding including irrelevant documents. This process yielded the following expanded query:

(autism OR autistic* OR Asperger* OR aspergic* OR Aspie? OR ASD OR neurodiver* OR ('on the spectrum' NOT 'on the spectrum of'))*.

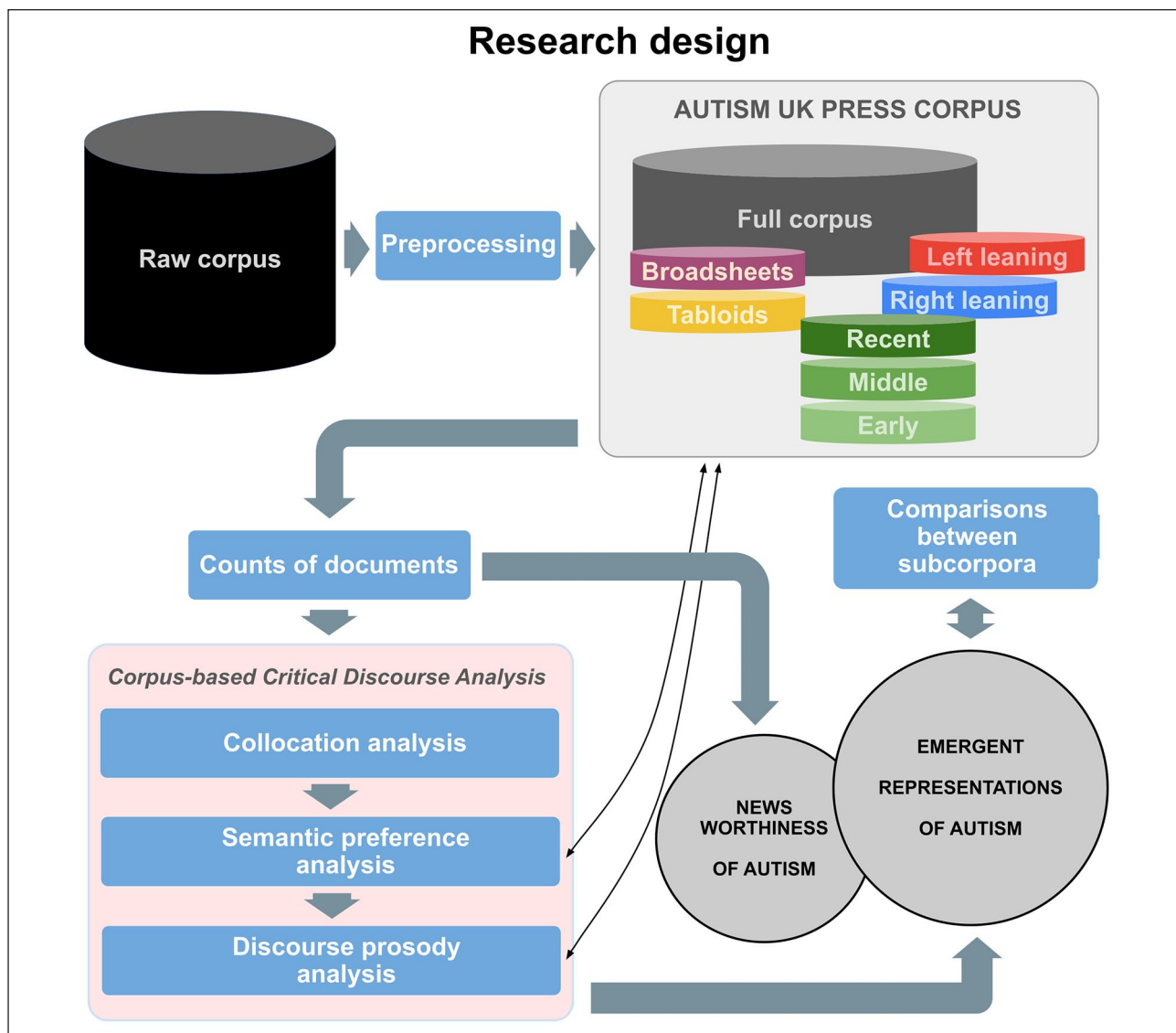


Figure 1. The research design of our study.

Downloading and preprocessing

Using the expanded query and *ProQuest's* duplicate removal filter, we downloaded documents from the 10 newspapers (in an ASCII text file format). Then a customised programming script (Python) removed all metadata (e.g. author, publisher) and created documents consisting of a title and main text. These were organised in a directory of files containing all documents from each newspaper over a 1-month period. These files were also arranged into subcorpora according to (1) publication period (Early: 2011–2014; Middle: 2015–2017; Recent: 2018–2020), (2) reporting style (Tabloids: *Daily Express*, *Daily Mail*, *Daily Mirror*, *Daily Star*, *Sun*; Broadsheets: *Daily Telegraph*, *Guardian*, *Independent*, *Observer*, *Times*) and (3) political orientation (left-leaning: *Daily Mirror*, *Guardian*, *Independent*, *Observer*; right-leaning: *Daily Express*, *Daily Mail*, *Daily Star*, *Daily Telegraph*, *Sun*, *Times*).

Counts

We derived *counts of documents* per newspaper per year. We also measured the *percentage of documents referring to autism* in each newspaper per year, as a proxy of the newsworthiness of autism. For a given newspaper, this measure was the raw corpus document counts divided by the total number of documents in *ProQuest*.

Reference corpus

We also created an additional 'reference' corpus consisting of general-content documents from the 10 newspapers. This corpus was used to establish whether certain lexicogrammatical patterns were pertinent to the *Autism UK Press Corpus* rather than a general characteristic of UK newspapers. The reference corpus consisted of 100 documents per month per newspaper for the period 2011–2020.

Corpus-based CDA

The analysis reported here focuses on the two core query terms, AUTISM* and AUTISTIC*, which correspond to 80% of all query-term instances (cf. Figure 3, subplot E below). We applied corpus-based CDA (P. Baker et al., 2008, 2013) to generate emergent representations of these terms from statistical patterns in their lexicogrammatical contexts. Our method involved three interconnected stages as follows.

Collocation analysis. First, we analysed collocations, that is, unusually frequent co-occurrences between two words in the corpus (Sinclair, 1991). One of these words ('node') was the term in focus (AUTISM or AUTISTIC), while the other word ('collocate') was any word co-occurring with the node within a window of five words before and after it. Collocations were valuable because they could be examined in the contexts in which they occurred so as to analyse the semantics of the terms in focus (Sinclair, 1991, pp. 115–116; see also Firth, 1957, pp. 195–196) and subsequently identify attitudes that these collocations expressed in context. This was done in the *semantic preference* and the *discourse prosody* stages of corpus-CDA (presented below in the corresponding subsections).

To derive collocations, we used a corpus-processing software, *WordSmith* (version 8) (Scott, 2021). We also allowed collocations to cross over sentence boundaries, in line with Sinclair (1991, pp. 111–112, 117) and Stubbs (1995, p. 246).

In addition, we constrained the collocations that were included in our final analyses based on three criteria: strength, reliability and frequency. First, we chose collocations representing a *strong* relationship, that is, one that does not apply to many other words (e.g. *raise awareness* is a stronger collocation than *develop awareness*). Strong collocations were identified using the *mutual information* metric, with a cutoff point of 3.0 (see Hunston, 2002). Second, we chose collocations that were also *statistically reliable* based on the *log-likelihood statistic* (G^2), with an alpha level of $p \leq 0.0001$ (see Rayson et al., 2004). Finally, from the collocations that were strong and reliable, we selected the 100 collocations with the highest *joint frequency*, that is, the number of occurrences in the corpus. Considering joint frequency mitigated a known bias of the mutual information metric to assign overly high scores to low-frequency words (Hoffmann et al., 2008, pp. 139–158).

Semantic preference analysis. Subsequently, in the semantic preference analysis, we examined the meanings of the collocates to establish the semantic preferences of the nodes AUTISM and AUTISTIC, that is, potential associations with groups of semantically related words (Stubbs, 2001, p. 65). For example, we examined whether AUTISM co-occurred with verbs expressing particular actions, or whether specific nouns or adjectives were used to describe autism. This

uncovered concepts, topics and issues (Phillips, 1989) routinely associated with autism and autistic people, and which were grouped into broader categories and subcategories (themes and subthemes, see also, P. Baker, 2004; Rayson, 2008). For example, a prominent category of collocations referred to age groups, gender and family roles, or combined two or three of these attributes (e.g. *boy, mother*). Within this category, collocations referring to autistic boys who were represented by their mothers were especially frequent. These prominent collocations drove the next stage of corpus-based CDA, namely the discourse prosody analysis (for a similar approach, see P. Baker et al., 2013).

Discourse prosody analysis. In the last stage of corpus-based CDA, we analysed manually broader textual contexts (e.g. paragraphs) which included one query term (AUTISM or AUTISTIC) and collocations from the most prominent category referring to age, gender and family roles. This analysis aimed to reveal *discourse prosodies*, that is, explicit and implicit attitudes (Stubbs, 2001, p. 66) towards autism, and enable a CDA of the press representations that emerged from corpus measurements. The situational context (in broad terms: real-world knowledge, beyond texts; Fromkin et al., 2011 p. 177) was also considered.

The discourse prosody analysis presented in this article focuses on a subcategory of the age, gender and family collocations referring to the prominent pattern of references to boys. We refer to these subcategory as BOY,³ including the collocation *boy(s)* and several semantically related words, for example, *son* and *lad*. The discourse prosody analysis was applied to eight sets of corpus samples, one from the full corpus and seven sets from the sub-corpora. Each set consisted of 100 randomly generated samples (using *WordSmith Tools 8*; Scott, 2021).

Four members of the research team were involved in the manual annotation of the sample sets, during which four different 'tags' referring to relationships, meanings, attitudes and discourses were set to appropriate values for each sample. The annotation scheme (Figure 2) consisted of Tag 1: General views towards autism; Tag 2: Diagnosis (Is the BOY referred to in the sample an autistic individual?); Tag 3: Agency (Does the autistic BOY have agency in the described action?); Tag 4: What are the underlying discourses? Following, Krishnamurthy (1996) and Gabrielatos and Baker (2008), this scheme was developed based on the definitions and descriptions of autism in dictionaries, diagnostic manuals and webpages of stakeholder organisations over multiple iterations of annotation and inter-annotator reliability meetings. When the annotation was completed, we counted alternative values within each tag.

Data analysis and addressing research aims

To examine the general representations of autism in the British press, we applied corpus-based CDA to the full corpus, using a combination of qualitative and quantitative

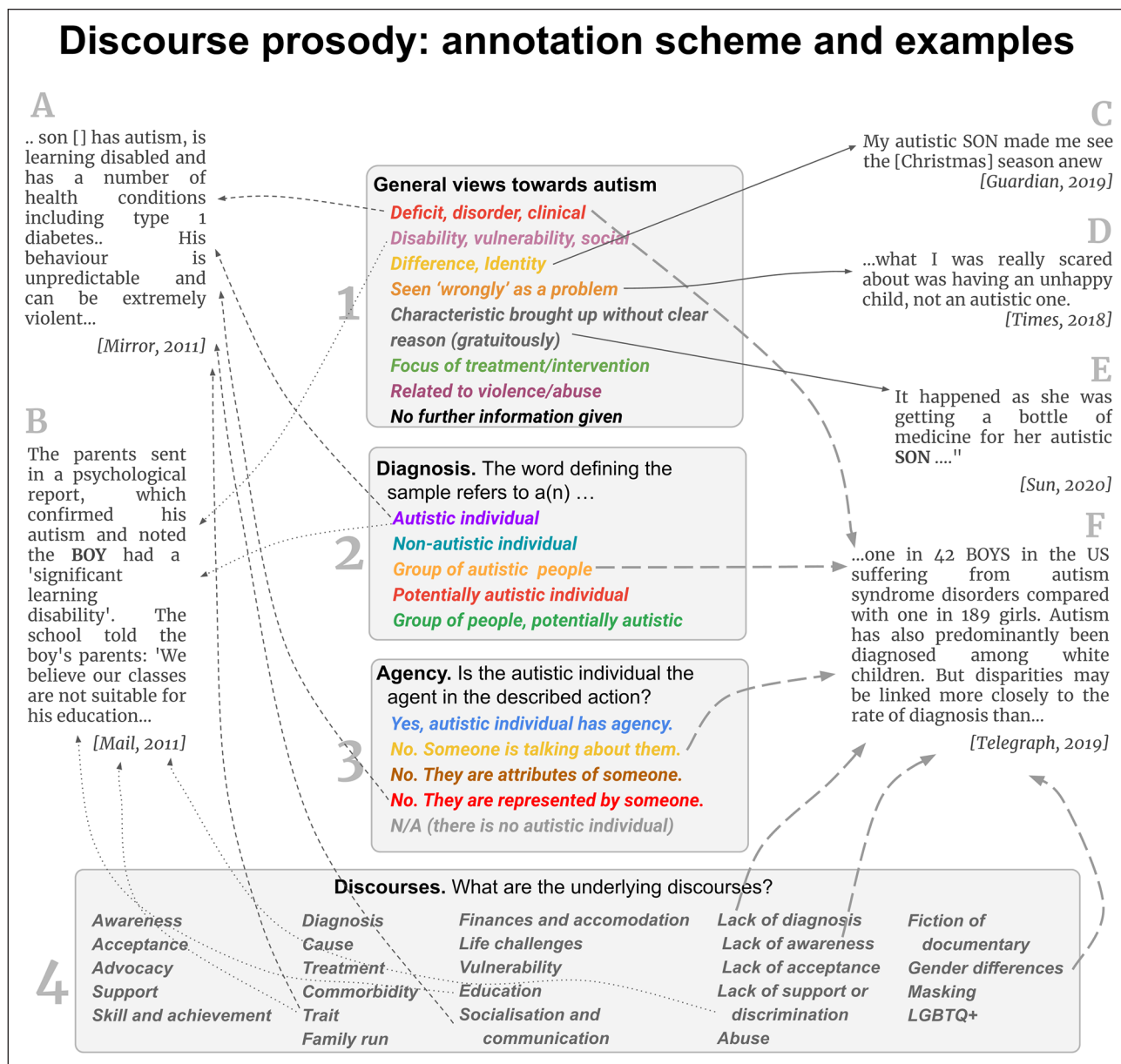


Figure 2. Annotation scheme and examples. The grey boxes show the five tags and the values they take. Corpus samples A, B and F are examples of annotation. Only one value was assigned to a sample for Tags 1–3; for Tag 4 (Discourses), multiple values could be assigned to the same sample. Corpus samples C, D and E are examples of the 'Difference', 'Seen wrongly as a problem' and the 'Mentioned gratuitously' values of Tag 1: General views towards autism.

procedures within each corpus-based CDA stage. In the semantic preference analysis, we complemented the qualitative analysis of collocations with counts of joint frequency, which quantified the salience of themes. In the discourse prosody analysis, we analysed the combined annotation results qualitatively, also performing quantitative comparisons of the proportion of a given value in two sample sets (e.g. the proportion of 'autistic individual' values in Tag 2-Diagnosis) using a *Bayesian Information Criterion* measure (BIC, Wilson, 2013, p. 8). A negative

BIC offered no reliable evidence for a difference; a positive BIC < 2.00 was taken as 'anecdotal' evidence for a reliable difference; a BIC ≥ 2.00 was 'positive evidence', while a BIC ≥ 6.00 or ≥ 10.00 was, respectively, 'strong' or 'very strong' evidence. For Tag 4 – Discourses, a correlation coefficient measure (calculated on counts for the 25 discourse values across two sets) was used to assess the overall similarity of discourses in two sample sets (e.g. whether discourses around autism were similar in Broadsheets and Tabloids).

We examined changes in the representation of autism over time, or according to reporting style and political orientation, by applying corpus-based CDA to appropriate subcorpora and adapting the semantic preference and the discourse prosody analyses so as to enable comparisons between subcorpora. In particular, the semantic preference analysis focused on *distinctive collocations*, that is, subcorpus-specific collocations (e.g. collocations in Tabloids, which were not among the collocations of the full corpus). Distinctive collocations showed how the emergent representation themes of the full corpus were differentiated between subcorpora. Similarly, the discourse prosody analysis contrasted how representations of autism differed across subcorpora (e.g. how annotation values differed in Tabloids and Broadsheets).

Ethical approval

The study was approved by the Science Research Ethics Committee of Edge Hill University (SREC: ETH2021-0008) and was conducted in accordance with its ethical procedures.

Community involvement statement

Two members of the autism community, who are parents of autistic children, were involved in the annotation of samples during the discourse prosody analysis. One of these individuals also contributed equal parts to all stages of the research process and co-authored this article.

Results

The autism UK press corpus

Figure 3 presents statistics of the *Autism UK Press Corpus*, which consists of approximately 24K documents, 900K sentences and 19 million words. There is roughly the same number of documents from tabloids and broadsheets (Panel B) in the corpus, but more documents from right- than left-leaning newspapers (~3:1 ratio, Panel D) (for statistics on the number of words, see Supplementary Figure S1). Furthermore, 98.2% of the occurrences of query terms are accounted for by *autism*, *autistic* and *Asperger's* (Panel E).

Figure 4 shows the percentage of documents referring to autism over time. Overall, this measure (range: 0.06%–0.97%) increased slightly over years (Panel B), with considerable differences between individual newspapers. The coverage of autism grew more rapidly in broadsheets than tabloids (Panel C), and in left- than right-leaning newspapers (Panel D), yielding a gradient pattern when the two factors were combined (Panel E).

Collocations and semantic preference analysis

Figure 5 shows the collocations of the terms *AUTISM* (subplot A) and *AUTISTIC* (subplot B), alongside their

joint frequency and mutual information. Three main themes emerged from these collocations.

Theme 1: adversities. Many of the collocations shown in Figure 5 refer to adversities associated with autism. This is a prominent theme, which emerged from several collocations (*AUTISM*: K=38, *AUTISTIC*: K=20), many of which were of high frequency. For example, 13 out of the 20 most frequent collocations of *AUTISM* were as follows: *spectrum*, *diagnosed*, *learning* [disabilities/difficulties], *disorder*, *MMR*, *disabilities*, *syndrome*, *form* [of autism], *risk*, *severe*, *diagnosis*, *conditions* and *vaccine*.

Collocations in this theme indicated a gradient of attitudes towards adversities associated with autism, including descriptions highlighting comorbidities and based on the medical model, descriptions using negative language, and perspectives highlighting disabilities and special needs. This variation is reflected in the (overlapping) sub-themes discussed below.

Subtheme 1A: comorbidities and medical model. This subtheme emerged from collocations referring to neurodevelopmental, physical, mental health and psychiatric conditions (e.g. *Attention-deficit/hyperactivity disorder* (*ADHD*), *cerebral palsy*, *epilepsy*, *schizophrenia*). These conditions were listed with autism in stories of individuals (e.g. ‘She helped her mother, single after her father walked out, to raise [. . .] who has epilepsy and autism’; published in Telegraph in September 2019), or when presenting research (‘It means that a 40 year-old father is about twice as likely to conceive a child that develops autism or schizophrenia as a 20 year-old father. . .’; published in Times, March 2014).

The collocations *link/linked* are also relevant to comorbidities (‘Herpes and autism link: Blood samples were taken from . . .’; Express, February 2017). Furthermore, *link/linked* were found in contexts discussing potential causes of autism (‘Could autism be linked to obesity in pregnancy?’; Mail, October 2015) – and often the disputed claims about a link between autism and vaccination (‘And there is still a leftover fear of autism linked to the MMR’; Sun, August 2019).

Subtheme 1B: negative language. Another emergent subtheme was the use of negative language. It arose from collocations overlapping with terminology from the medical model (e.g. *deficit*, *disorder*), and lay language words of negative connotation (e.g. *suffers*, *problems*, *partially/severely*, *symptoms*). For example, an autistic person was ‘[. . .], described by relatives as socially awkward and possibly suffering from autism and a personality disorder, then went to the school on Friday armed with a semi-automatic rifle and two handguns’ (Times, December 2012). A more recent example of negative language is: ‘Those who suffer from autism also don't have to wear a mask . . .’ (Telegraph, July 2020).

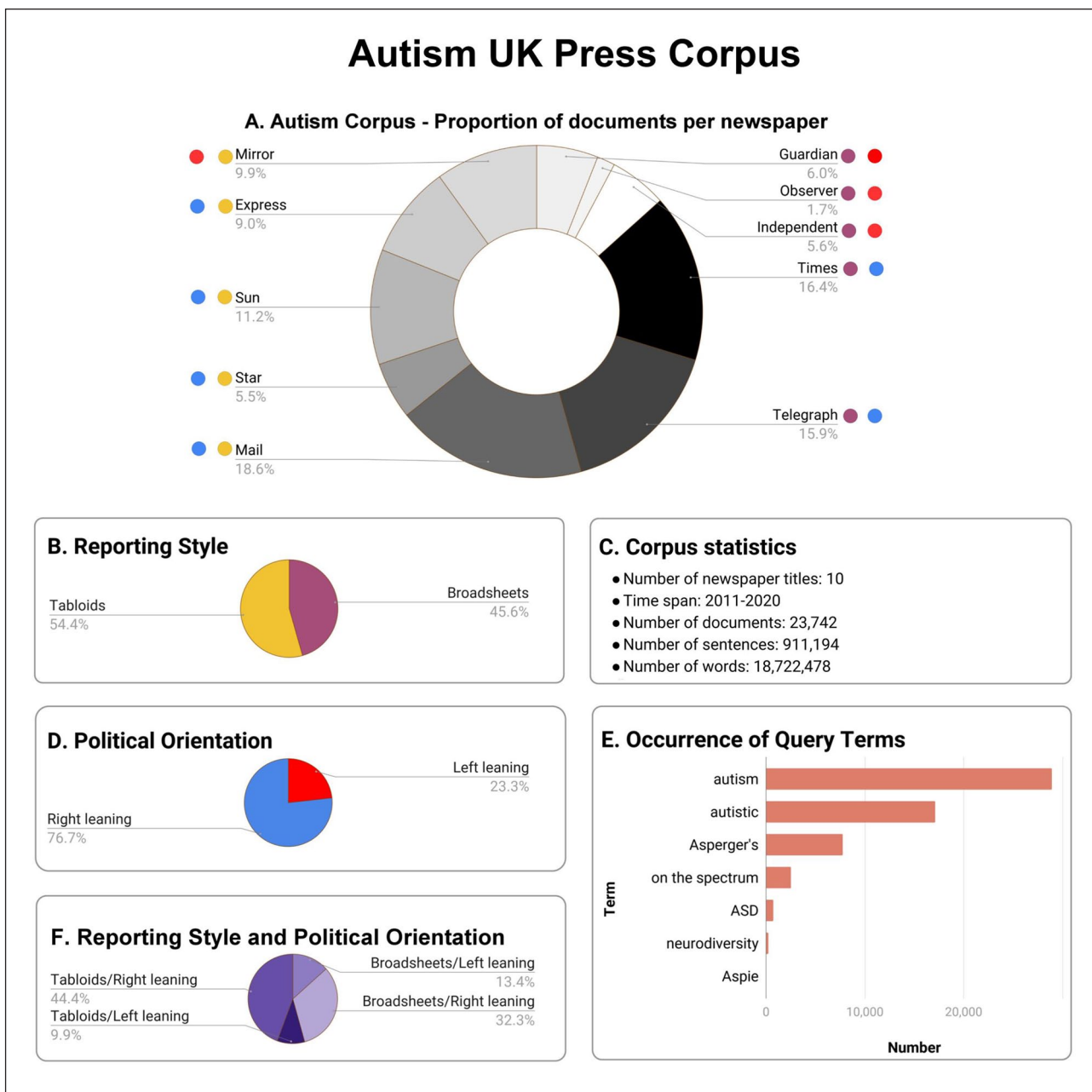


Figure 3. The *Autism in UK Press* corpus. Panel A shows the percentage of documents in the corpus per newspaper. The coloured dots next to newspaper titles indicate reporting style (purple: tabloids; yellow: broadsheets) and political orientation (red: left-leaning; blue: right-leaning). Panels B, D and F present counts of articles with newspapers grouped by reporting style, political orientation and combinations of these (correspondingly). Panel C presents corpus statistics and Panel E the frequency of the different query terms in the corpus.

Subtheme 1C: disabilities and needs. Another subtheme emerged from collocations identifying disabilities and special needs in autism (e.g. *disability/ies, conditions, special, needs*). These collocations were frequently found in contexts referring to autistic individuals with learning disabilities or special needs: ‘These are people like [. . .], who is 22, and has autism and a learning disability’ (Telegraph, February 2015). Notably, such collo-

cations could imply a variable sentiment towards autism. In some cases, they were used within overall negative descriptions, akin to Subtheme 1B (‘Because [. . .]’ autism is so extreme he needs constant care, which his parents can no longer give him’, Star October 2017). In other cases, they were used within more balanced accounts, identifying autism as a (learning) disability or, sometimes, suggesting an effort to offer an informed

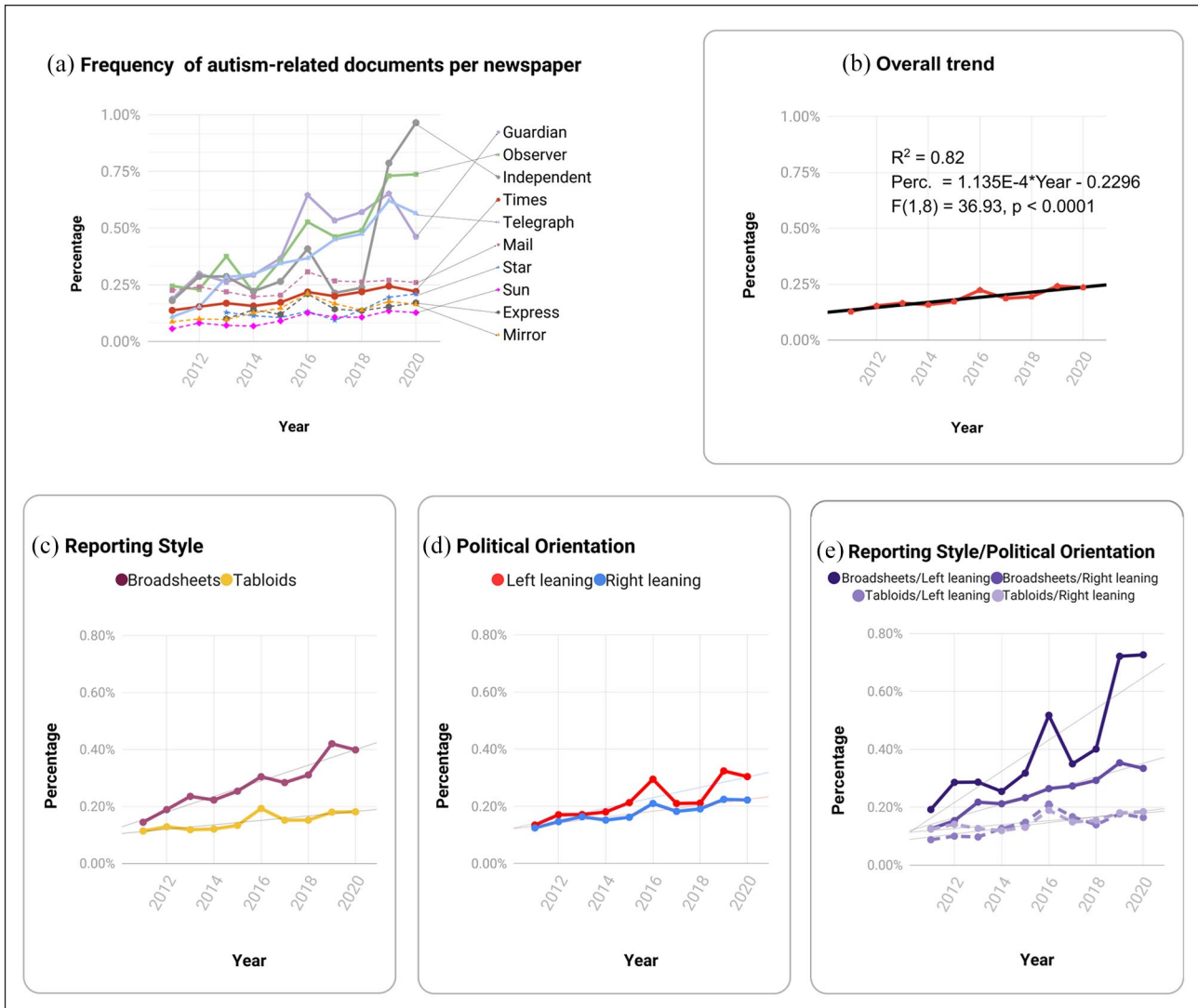


Figure 4. Coverage of autism over time. For a given newspaper, the percentage of documents referring to autism was calculated as the ratio of the number of articles in the Autism UK Press Corpus with the total number of documents (from the same newspaper) in the ProQuest database. Panel A shows this percentage for individual newspapers per year. Panel B shows the overall trend of a slight increase. Panels C, D and E group newspapers with respect to reporting style, political orientation and combinations of these.

description: ‘For the record, autism is a lifelong developmental disability that affects how people perceive the world and interact with others (I’m quoting here from the NAS)’ (Independent, March 2019).

Theme 2: age/gender groups and family roles. A second main theme was grounded on the prominent category of collocations (*AUTISM*: K = 10; *AUTISTIC*, K = 34), referring to age groups, genders and family roles. This theme arose from nouns, such as *son, child, young, mother, girls, father, parent, families, etc.*, and numbers suggesting age (e.g. *five, nine*) and reflected discussions about specific autistic individuals or individuals related to them (‘A 16 year-old autistic boy and his mother have won a

High Court case challenging his expulsion from school’; Mail, September 2019).

Subtheme 2A: emphasis on children, especially boys. The age/gender and family roles collocations referred mainly to children rather than adults. This held for both word types (child-related: K = 17, adult-related: K = 9) and word frequency. For example, *AUTISM* collocations referring to children (e.g. *child, pupil, boys*) were 8.6 times more frequent than collocations referring to adults (e.g. *adults, students*).

Furthermore, collocations revealed a strong gender bias towards boys. For both *AUTISM* and *AUTISTIC*, *son* and *boy* were the two most frequent gender-specific collocations.

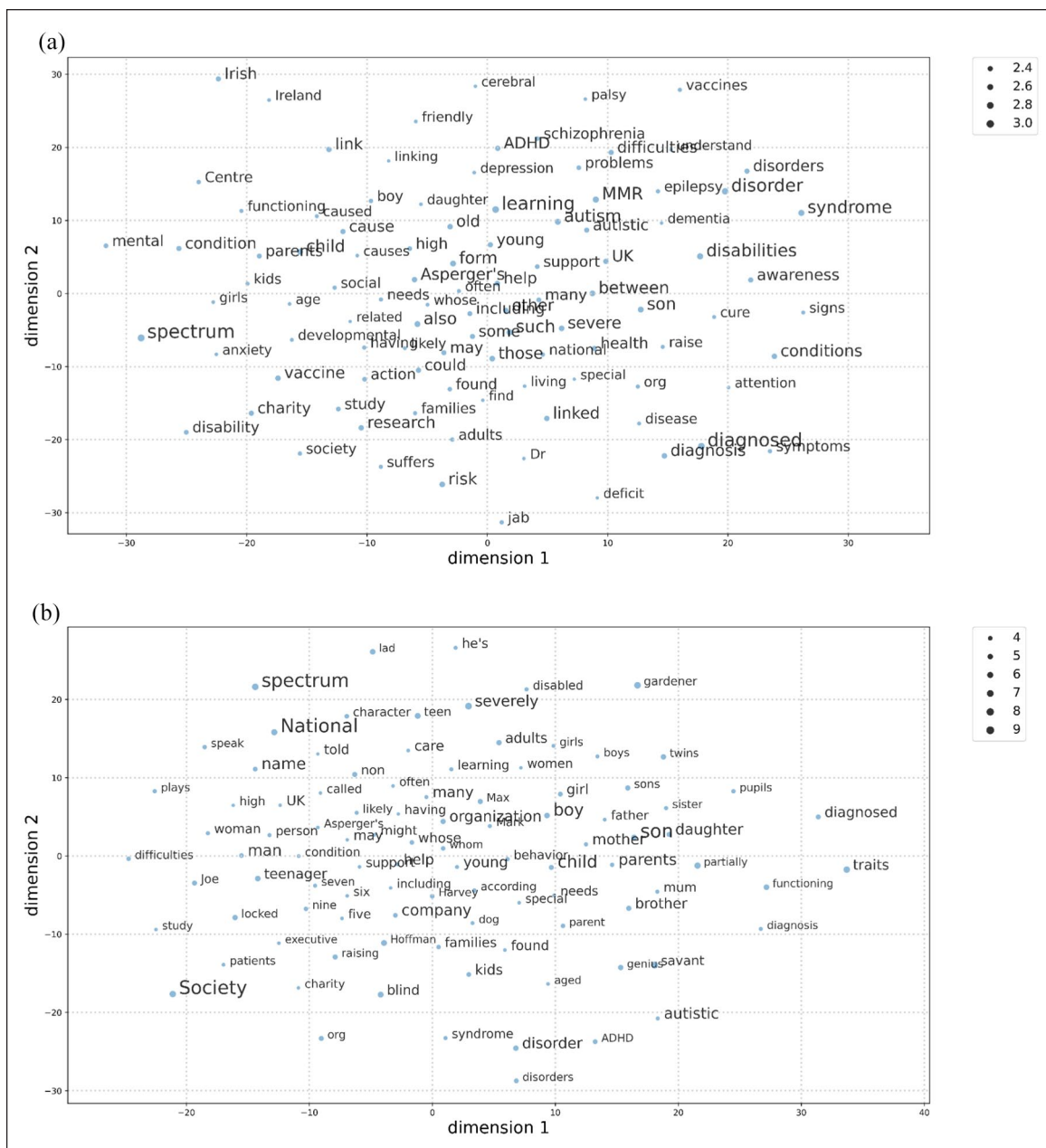


Figure 5. Collocations of the terms (a) AUTISM and (b) AUTISTIC in the full corpus. The collocations are placed on a semantic map, so as to show semantic relationships between words and clusters of semantic preferences. The semantic map was created using word vectors (w2v, word2vec-google-news-300; Mikolov et al., 2012), a computational linguistics resource that represents word meanings into a multidimensional semantic space. This space was visualised in two dimensions (1 and 2) using the t-SNE dimensionality-reduction technique (t-Distributed Stochastic Neighbour Embedding; van der Maaten & Hinton, 2012). A bigger font size in the collocates suggests higher joint frequency, that is, more co-occurrences between the collocate with the term in focus (font size is proportional to log10 joint frequency). The size of the blue dot on the lower left-hand side of the collocates shows mutual information.

For AUTISTIC, *son* was also the most frequent collocation (overall), two times more frequent than the next most frequent collocation (*child*, Figure 5(b)). A male-gender dominance was also observed in the plural number: *AUTISM* and *AUTISTIC* combined, *sons* were mentioned 4.6 times more

frequently than *daughters*, and *boys* were mentioned twice more frequently than *girls*.

Crucially, this bias towards boys did not arise from a general male bias in British newspapers, as in the Reference corpus, frequencies of gender-related words were relatively

balanced (e.g. *son: daughter*=1.28: 1.00, *boy/lad/school-boy: girl/lass/schoolgirl*=0.98: 1.00).

Theme 3: individuals and media representation. A final theme emerged from nouns and names referring to individuals associated with autism in the media. These were collocations of *AUTISTIC* ($K=7$) and referred mostly to TV series (e.g. Joe [Hughes] the main character in ‘The A Word’ TV series and *Max* [Vento]: the child protagonist, *character*). For example, ‘. . . , praise must be heaped on the young Max Vento who plays five-year-old Joe, the autistic boy at the heart of the series’ (Telegraph, April 2016).

Subtheme 3A: stereotypes. Two additional collocations, *savant* and *genius*, gave rise to a subtheme referring to the stereotypical presentation of autism in the media. In our corpus, these words were predominantly found in contexts referring to fiction: ‘. . . its titular hero, the paediatrician Shaun Murphy (played by Britain’s Freddie Highmore), is an autistic savant with the manner of a shy 12-year-old’ (Times, November 2017).

Distinctive collocations

We now turn to results on subcorpus-specific collocations, which, however, were not among the collocations of the full corpus. These distinctive collocations offer an insight into how Themes 1–3 were differentiated between subcorpora corresponding to different time periods, or with different reporting style and political orientation.

Changes over time. Figure 6 presents distinctive collocations for the Early, Middle and Recent subcorpora. There are distinctive collocations related to adversities associated with autism (Theme 1), comorbidities and medical model (Subtheme 1A) and negative language (Subtheme 1B) in all three subcorpora. Regarding Subtheme 1B, two distinctive collocations of *AUTISTIC* in the Early corpus, namely *tendencies* and *semi[-autistic]*, suggest poor awareness and especially negative attitudes towards autistic people. For example, ‘[A] respectable family which was hit with disaster last summer, after the semi-autistic 8-year-old son—who tends to make things up—had lashed out at his 13-year-old sister, leaving bruises . . .’ (Telegraph, February 2012). In more recent years, newspapers tended to use these words less frequently or in inverted commas.

Distinctive collocations also suggested that age and gender groups associated with autism (Theme 2) became more diverse over years. For example, *girl* was a distinctive collocation of *AUTISM* in the Middle (but not the Early) subcorpus. Furthermore, *transgender* was as a distinctive collocation of *AUTISTIC* in the Recent subcorpus, reflecting the recent interest of newspapers in gender diversity in autism. However, this issue was often discussed with criticism or doubt: ‘an expert in childhood

gender dysphoria, said youths diagnosed as transgender could in fact be suffering from autism, anxiety or depression’ (Mail, January 2017).

In addition, frequency counts suggested a slight decrease in gender bias in autism representation. There were at least four times more references to autistic males than females in the Early and the Middle subcorpora, however, this ratio was 2.5 in the Recent subcorpus.

Finally, there was a clear temporal pattern in the distinctive collocations regarding Individuals and Media Representation (Theme 3). For example, the Early corpus featured Christopher (the protagonist of *The Curious Incident of the Dog in the Night-time*, premiered at the National Theatre in August 2012); the Middle corpus featured [Max] Vento (*The A Word*, aired 2016) and the Recent corpus Shaun Highmore (protagonist of *The Good Doctor*, aired September 2017).

Reporting style. Figure 7 shows distinctive collocations in Broadsheets (subplots A and C) and Tabloids (subplots B and D). Five collocations falling into the subtheme of negative language (1A) – in particular, *affects*, *suffers*, *patient*, *severe*, *accused*, and *killed* – were specific to Tabloids. By contrast, the distinctive collocations *understanding* and *experience* were suggestive of a more positive and inclusive attitude in Broadsheets.

Furthermore, with regard to age and gender groups (Theme 2), distinctive collocations in Broadsheets suggested an indexing of more diverse groups of individuals associated with autism (*daughter’s*, *females*, *woman’s*, *mothers*, *schoolboy*, etc.). Finally, in relation to Individuals and media representation (Theme 3), distinctive collocations in Tabloids referred to celebrities, who often talked about their children’s diagnosis (‘The [. . .] star, who shares three autistic children with [. . .] opened up about the struggles she faces raising kids with the developmental disorder . . .’; Star, March 2020), or fictional characters. By contrast, broadsheet-specific collocations included a prominent researcher ([Baron-]Cohen), stakeholder organisations and individuals involved in these (*NAS*, *chief*, (Mark) *Lever*), and a leading figure of the anti-vaccine movement (*Wakefield*).

Political orientation. Turning to political orientation (Figure 8), distinctive collocations were more numerous in the left-leaning than the right-leaning subcorpus, probably due to the size difference between these two subcorpora (see Figure 3, Panel C). This would, of course, suggest that collocations of the full corpus were more representative of right-leaning newspapers. Nevertheless, many distinctive collocations in the left-leaning subcorpus referred to Individuals and media representation (Theme 3) and were related to fictional characters (e.g. *Dustin* and *Raymond*, referring to the Rain Man movie) or stakeholder organisations and related individuals (*chief*, *president*, (National Autistic) *Society’s*).

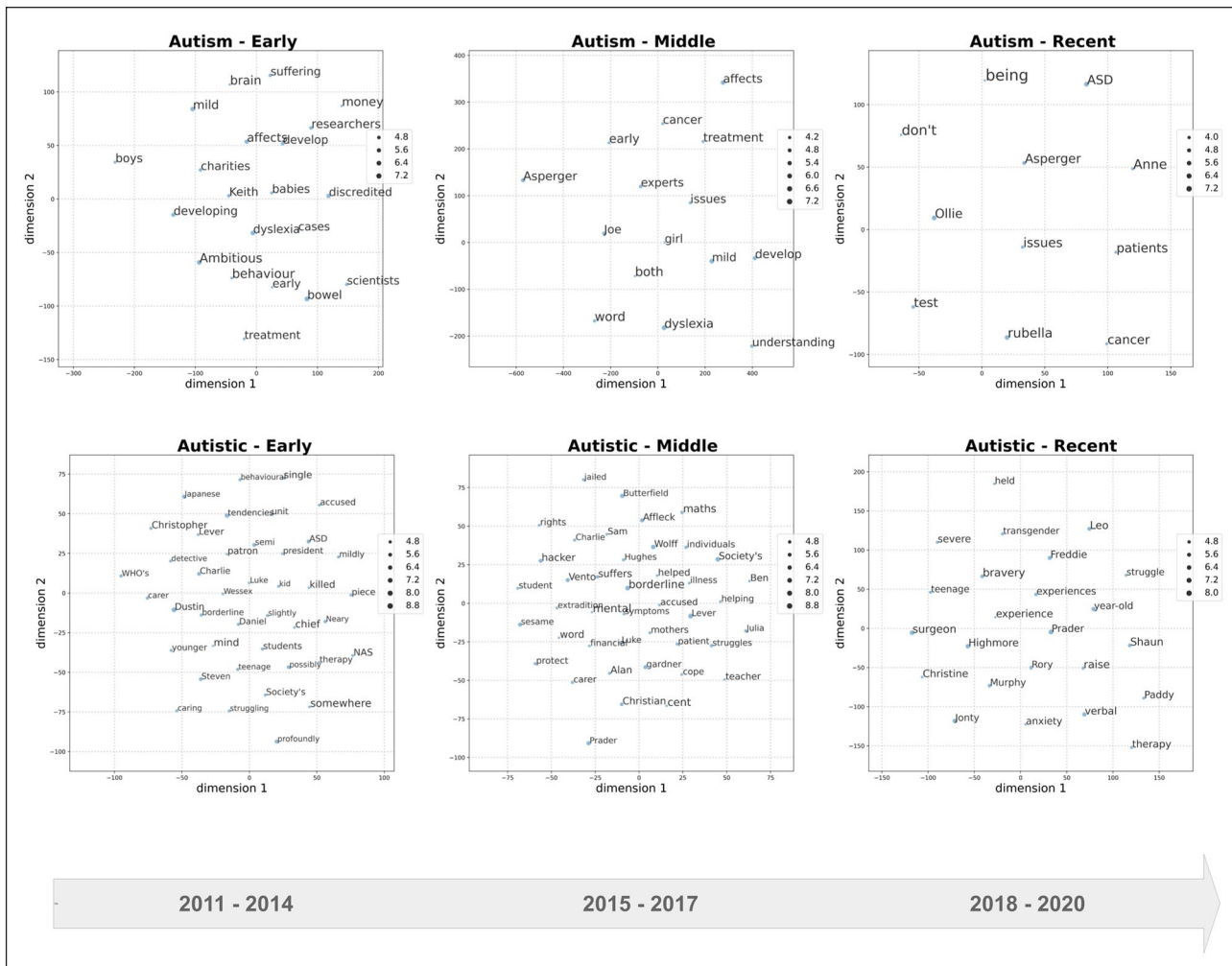


Figure 6. Distinctive collocations of the terms AUTISM (top row) and AUTISTIC (bottom row) in the Early (left column), Middle (middle column) and Recent (right column) subcorpora. Distinctive collocations were visualised similarly to Figure 5, that is, using a semantic map to group collocations according to their meaning, and with font size showing joint frequency and the size of the blue dot depicting mutual information. (The only difference is that the font size is now proportional to log joint frequency, that is, the log transformation used a different base from that used for the full corpus, natural vs 10).

Discourse prosody analysis

Full corpus. In this section, we report results from the discourse prosody analysis, in which we manually annotated representative sets of randomly chosen corpus samples based on an annotation scheme of four tags (Figure 2). This analysis aimed to consider the broader contexts in which collocations occurred and reveal underlying discourses and attitudes. For the full corpus, the results of the discourse prosody analysis are shown in Figure 9.

Tag 1 – general views towards autism. Counts of alternative values for Tag 1 (subplot A) suggested that autism was predominantly represented as a deficit or disorder (33.7% of corpus samples) or as a disability or vulnerability

(34.7%). An additional 14.3% of samples associated autism with a violent action, and an 11.2% discussed treatments or support. Only a tiny proportion of corpus samples (3%) presented autism as a difference. Overall, the results for Tag 1 were consistent with the semantic preference theme of adversities associated with autism (Theme 1).

Tag 2 – diagnosis. Turning to Tag 2 (subplot B), in the vast majority of corpus samples, the collocations (*boy, son, etc.*) referred to an autistic individual (96.9%), while in a small fraction of the samples, collocations referred to a group of autistic people (3.06%). This result was consistent with the second theme of the semantic preference analysis referring to the focus of newspapers on stories related to specific autistic individuals (Theme 2).

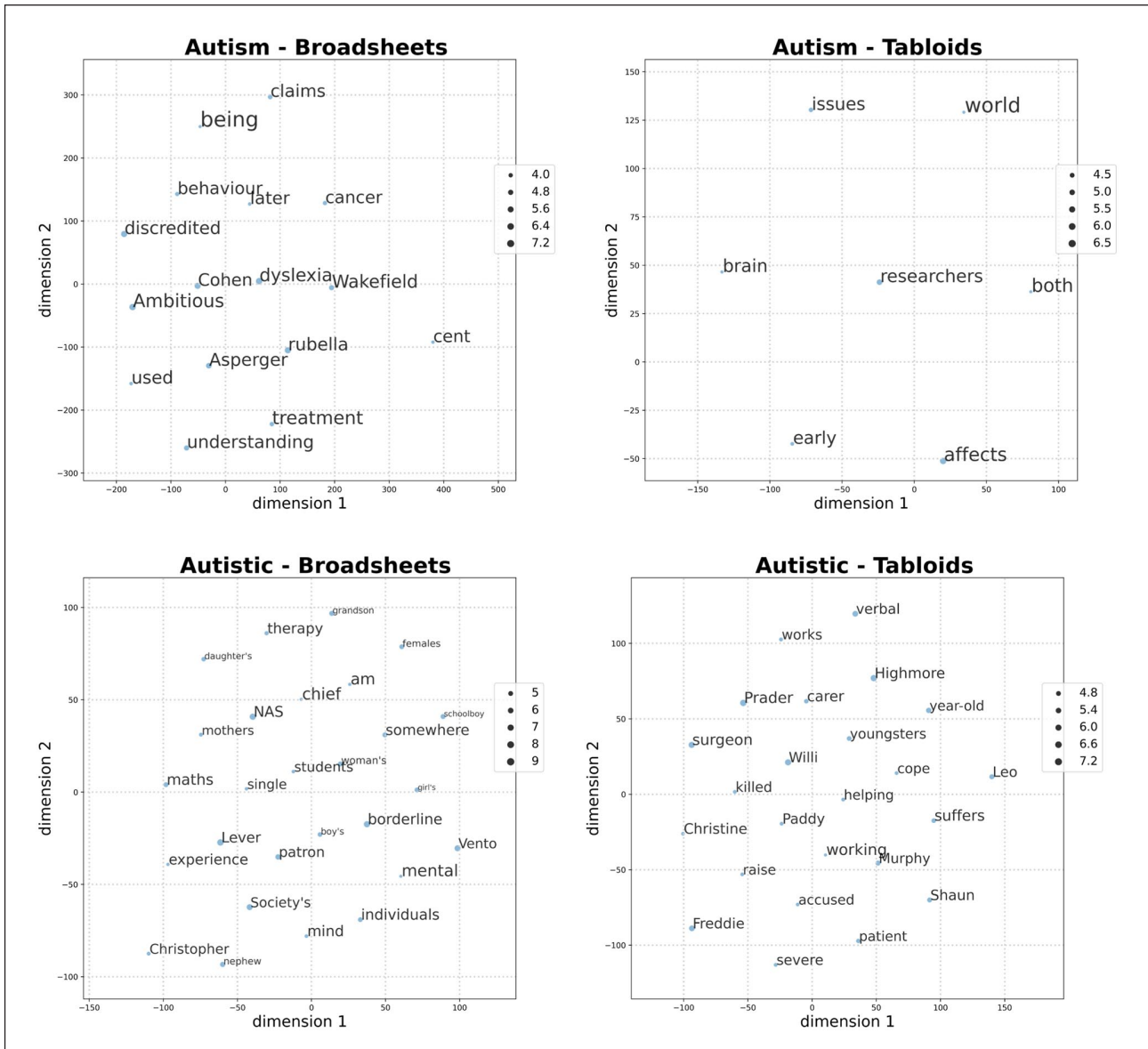


Figure 7. Distinctive collocations of the term AUTISM (top row) and AUTISTIC (bottom row) in the Broadsheets (left column) and Tabloids (right column) subcorpora. Distinctive collocations were visualised with the methods described in Figure 6.

Tag 3 – agency. With regard to whether the autistic individuals were portrayed as initiating or causing the described actions (Tag 3 – Agency, subplot C), in the great majority of corpus samples (90.8%), the autistic individuals were not presented as having agency. For example: ‘[S]he was left in tears by television presenter [. . .] after he posted an emotional video about a boy with autism’ (Express, August 2019). Furthermore, in more than half of the samples, they were represented by, or were an attribute of someone else. In two-thirds (66.6%) of the samples in which a carer was reported, this was the mother of the child (e.g. ‘a mum of an autistic boy’).

Crucially, a lack of agency was also observed in corpus samples that referred to adult autistic individuals (e.g. through the collocation *son*), suggesting that even in this case, autistic individuals were discussed mainly in relation to – or from the perspective of – a carer (usually a parent). For example, ‘But [. . .]’s dad says his son’s mild autism is hampering his attempts to find work’ (Star, January 2016).

Tag 4 – discourses. Turning to Discourses (Tag 4), subplot E shows the percentages for 25 discourse values using a histogram. Some of the frequent discourses (peaks in the

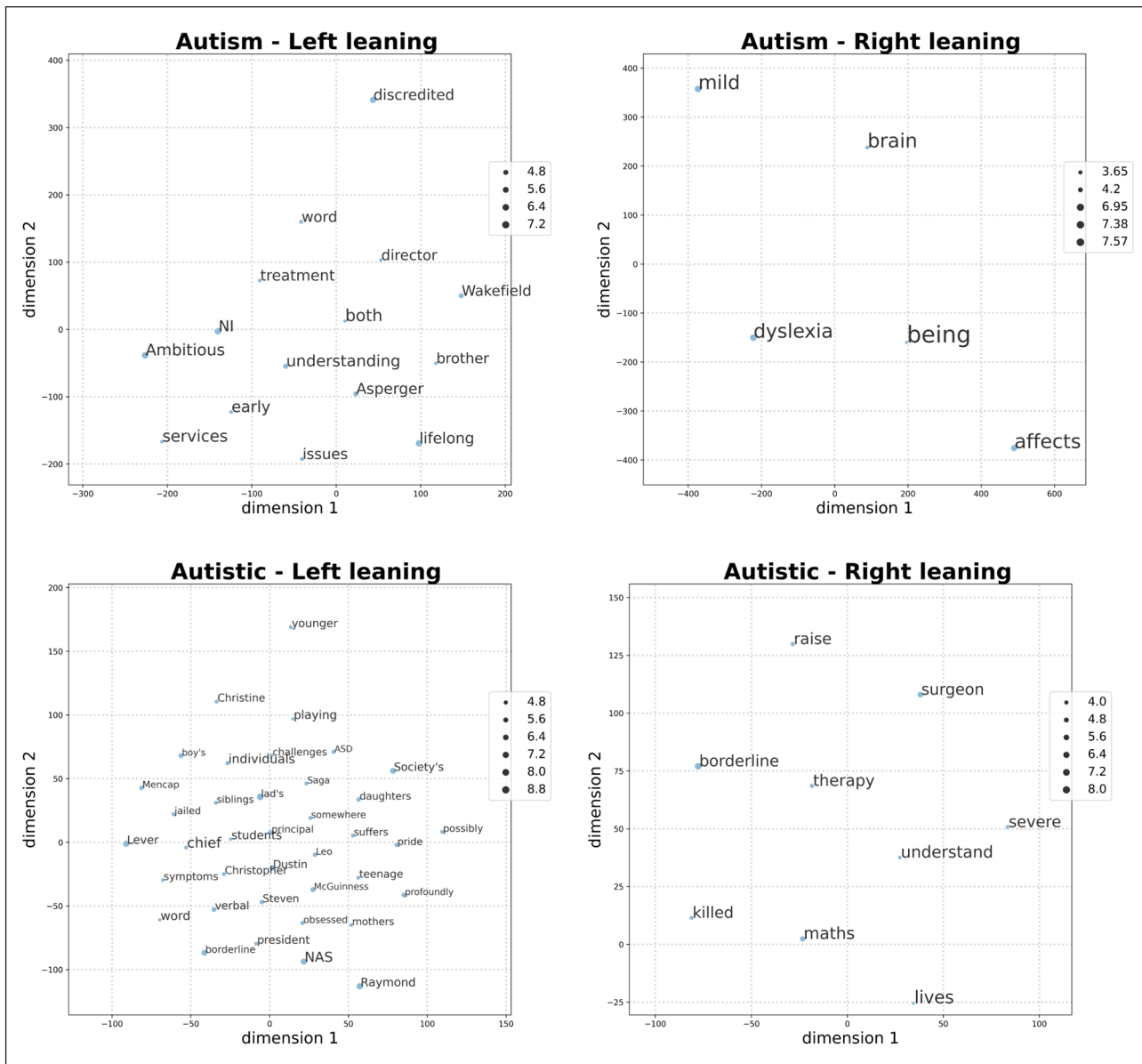


Figure 8. Distinctive collocations of the term AUTISM (top row) and AUTISM (bottom row) in left-leaning (left column) and right-leaning (right column) subcorpora. Distinctive collocations were visualised with the methods described in Figure 6.

histogram), were ‘comorbidity’, ‘abuse’, ‘lack of support/discrimination’, ‘advocacy’ and ‘fiction/documentary’. All these discourses were consistent with the themes identified in the semantic preference analysis and the annotation results for Tags 1–3.

Changes over time. Similarly to the semantic preference analysis, the discourse prosody analysis of samples from the subcorpora was useful for assessing changes over time and differences related to reporting style and political

orientation. Figure 10 summarises results from the annotation of the Early, Middle and Recent subcorpora.

With regard to Tag 1 – General views towards autism (subplot A) and Tag 3 – Agency (subplot B), both the proportion of difference-based descriptions of autism in Tag 1 and the proportion of agency values in Tag 3 presented some numerical increases over time. These changes would be consistent with the progression towards more positive and inclusive representations of autism shown in the semantic preference analysis; however, these numerical

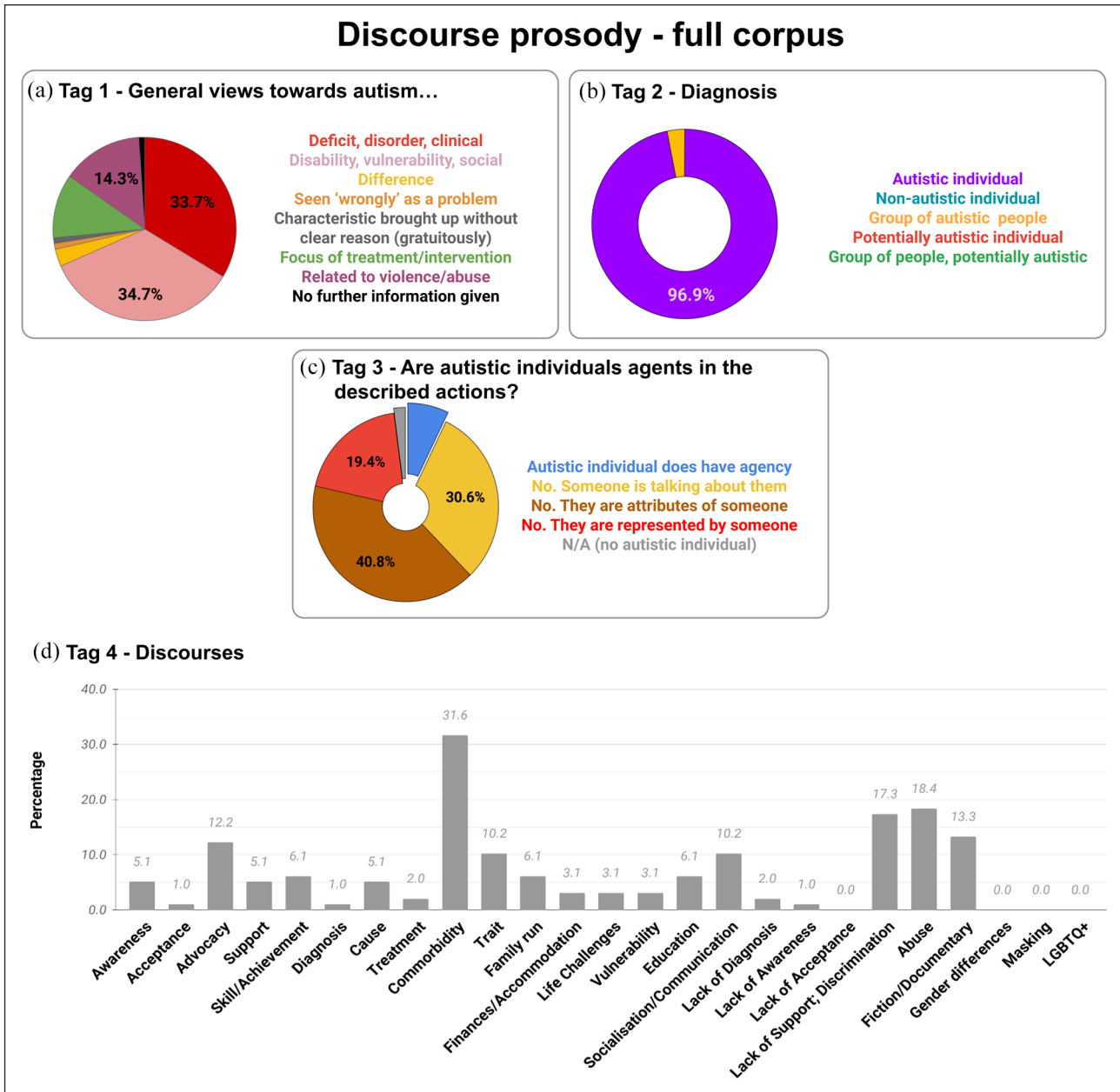


Figure 9. Results of the discourse prosody analysis: (a) General views towards autism (Tag 1); (b) Diagnosis (Tag 2); (c) Agency of the autistic individual in described actions (Tag 3); and (d) Histogram of discourse values for the sample sets for collocations in the category BOY (Tag 4).

differences were not corroborated statistically ('difference-based' in Early vs Recent: $BIC = -3.76$; agency in Early vs Recent: $BIC = -1.31$, 'no evidence').

With regard to Discourses (Tag 4, subplot C), the annotation results were, overall, highly similar for the three temporal subcorpora (correlation coefficients measuring similarity were in the range 0.67–0.73, all $ps < 0.001$). In terms of individual discourses, newspapers seemed to highlight the lack of support and discrimination against autistic people in more recent years ('lack of support/discrimination' in Early vs Recent: $BIC = 1.63$, 'anecdotal').

An example of this discourse is: '[. . .] have an autistic son, [. . .], five, and applied for an EHCP⁴ when he was four. This was rejected out of hand because . . .'; Times, December 2018).

Reporting style. Figure 11 presents annotation results for Broadsheets and Tabloids. Consistent with the results on distinctive collocations, there was a numerically higher proportion of difference-based views in Broadsheets than in Tabloids (Tag 1 – views towards autism; subplots A and B); however, this difference was not corroborated

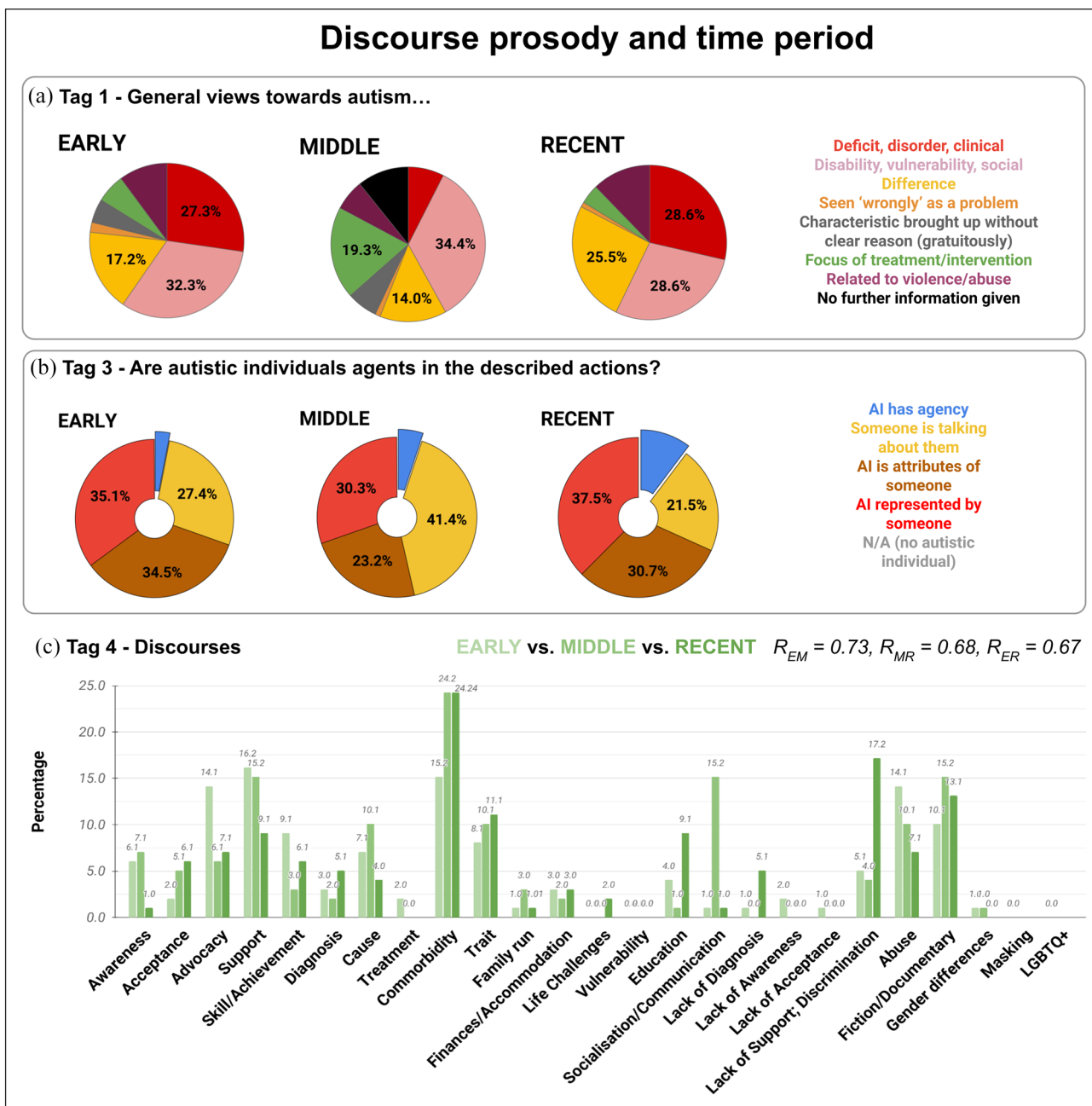


Figure 10. Results of the discourse prosody analysis in the Early, Middle and Recent subcorpora: (a) General views towards autism (Tag 1); (b) Agency (Tag 3); and (c) Discourse values (Tag 4).

statistically (BIC=-1.89, ‘no evidence’). Furthermore, tabloids made more gratuitous references to autism than broadsheets (BIC=3.73, ‘positive evidence’), suggestive of a sensationalist reporting style (see Figure 2, sample E for an example).

With regard to Agency (Tag 3 – subplot B), autistic boys were more often portrayed as being represented by someone else in Tabloids than in Broadsheets (BIC=3.69, ‘positive evidence’); however, cases in which boys had agency of the described actions appeared in a comparable

proportion of documents (BIC=-3.60) in Tabloids and Broadsheets. An example of an autistic boy having agency was: ‘They included the case of a 13 year-old boy with autism who complained of an attack by a 15 year-old friend . . .’ (Independent, May 2014).

The overall comparisons (correlation coefficients) of Discourses around autism suggested a high degree of similarity between Broadsheets and Tabloids ($R=0.73, p < 0.001$). In terms of individual discourses, Broadsheets tended to show a greater focus on the causes of autism than

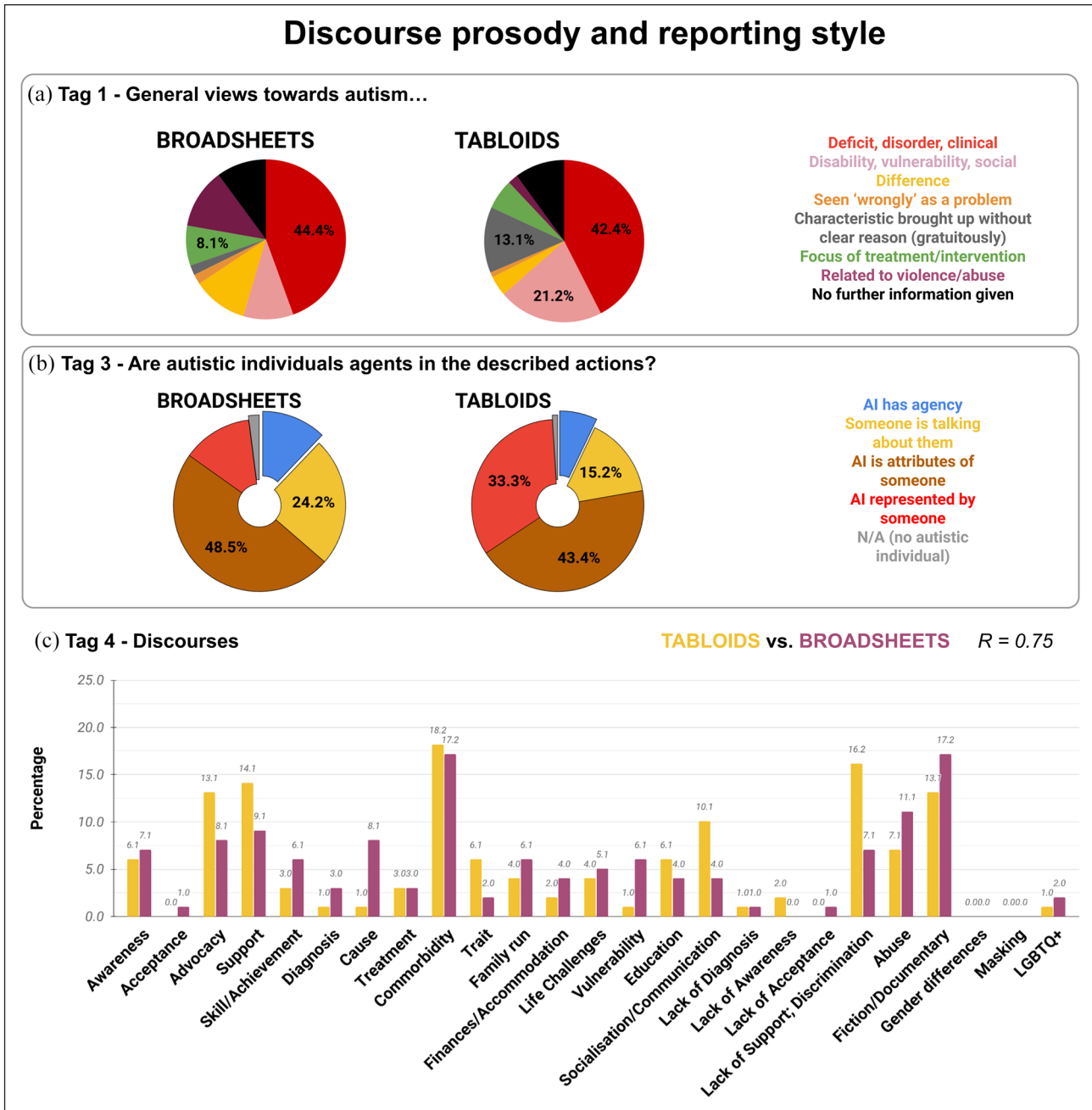


Figure 11. Results of the discourse prosody analysis in the BroadSheets and Tabloids subcorpora: (a) General views towards autism (Tag 1); (b) Agency (Tag 3); (c) Discourse values (Tag 4).

Tabloids (BIC=0.91, ‘anecdotal’); by contrast, Tabloids showed a tendency to emphasise the lack of support and discrimination, and issues related to socialisation and communication, but these results were not corroborated statistically (all BIC < 0).

Political orientation. Figure 12 shows annotation counts for left- and right-leaning newspapers. The proportion of deficit-based, disability-based and difference-based views towards autism (subplot A, Tag 1 – views towards autism)

did not differ significantly with political orientation (all BIC < 0). However, there was a higher proportion of samples referring to autism in the context of a treatment and cure in samples from right-leaning newspapers (BIC=0.5, ‘anecdotal’). Tag 3 – Agency also differed slightly as left-leaning newspapers tended to present a higher proportion of autistic boys as having agency of the described actions (subplot B; BIC=0.7, ‘anecdotal’).

Turning to Discourses (Tag 4; subplots C), the overall similarity of left-leaning and right-leaning newspapers

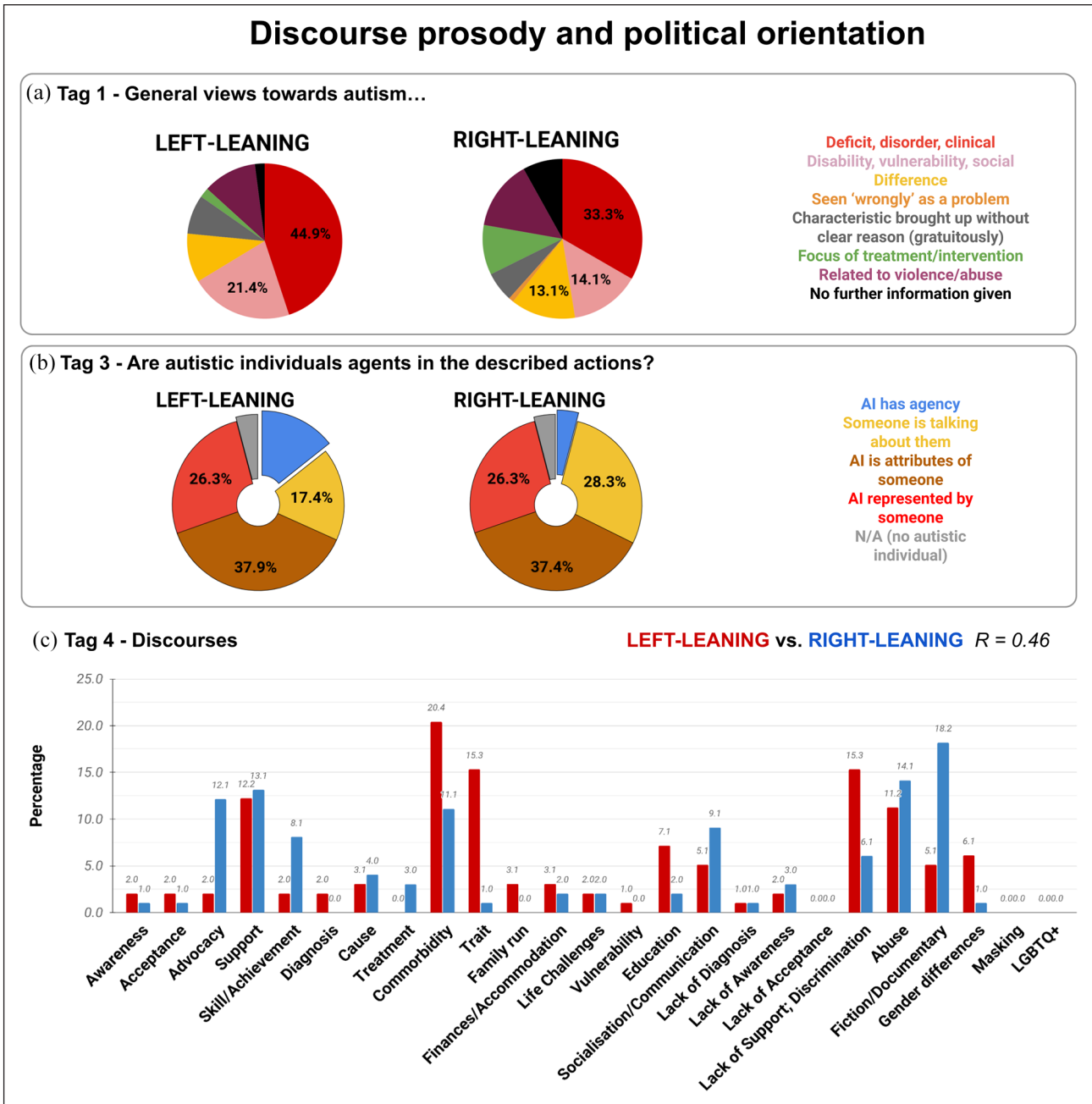


Figure 12. Results of the discourse prosody analysis in the left-leaning and right-leaning newspapers: (a) General views towards autism (Tag 1); (b) Agency (Tag 3); (c) Discourses (Tag 4).

was $R=0.42$, $p=0.021$. In terms of individual discourses, there was a much higher proportion of discourses on ‘autism traits’ in left-leaning than in right-leaning newspapers ($BIC=9.56$, ‘strong evidence’). This discourse value marked samples from left-leaning newspapers that expressed a range of attitudes towards autism. Some samples seemed to perpetuate negative stereotypes (many from the only left-leaning tabloid, e.g. ‘One lad aged 13 with autism used to sit by himself’; *Mirror*, August 2020) and some others put forward more positive views (mainly

in broadsheets, e.g. ‘... value autistic traits. My son doesn’t lie, he’s never sly or manipulative, and it would not occur to him to put anyone down’, *Guardian*, March 2018). There was also a higher proportion of ‘advocacy’ values ($BIC=2.54$, ‘positive evidence’) in left- compared to right-leaning newspapers, arising from stories on individuals (often from public life) who advocated for their children and autistic people in general, and more ‘fiction/documentary’ values ($BIC=2.39$, ‘positive evidence’). Left-leaning newspapers also tended to focus more on

'lack of support/discrimination' discourses, while right-leaning newspapers included more 'fiction/documentary' discourses not corroborated statistically (both $BIC < 0$).

Discussion

In this study, we examined portrayals of autism in the British press between 2011 and 2020. Our study was the first to examine such representations of autism in the press using corpus-based CDA (P. Baker et al., 2008), recognised as a rigorous and synergistic framework. Corpus-based CDA enabled us to analyse a larger (and hence more representative) collection of newspaper documents than previous studies and help uncover explicit and implicit attitudes towards autism encoded in lexicogrammatical patterns. We also examined how press representations changed over the time period, and similarities and differences with respect to both reporting style and political orientation.

Our analyses revealed that the proportion of newspaper documents referring to autism rose slightly over the time period, with a greater increase in broadsheets compared to tabloids, and left-leaning versus right-leaning newspapers. For example, in 2020, the proportion of documents referring to autism in the Independent (a 'left-leaning broadsheet') was more than seven times higher than in the Sun (considered the most prominent and popular 'right-leaning tabloid').

However, more important than quantity is what is reported about autism and autistic people and the language used to convey this information. The emergent representations of autism from the semantic preference analysis encompassed three themes: the prevalence of descriptions of adversities related to autism with associations with neurodevelopmental, physical and mental health conditions and many features of negative language; a prominence of individual-based accounts, with a strong focus on children and boys rather than adults and girls; and repeated references to figures in public life or well-known fictional characters. These themes were consistent with some findings from earlier studies, which suggested that newspapers offered biased and often inaccurate information about autism with misleadingly negative and ableist accounts, and promoted stereotypical views towards autistic people, which were not inclusive of many groups of autistic individuals (e.g. Baroutsis et al., 2021; Billawala & Wolbring, 2014; Holton et al., 2014; Jones & Harwood, 2009).

This study demonstrates that similar biases are to be found in the contemporary British press. It also shows how these biases can emerge from lexicogrammatical patterns distributed across documents with a variety of content (including articles, media coverage, regular columns, readers' letters, etc.) with relevance to autism. The repeated occurrence of these patterns establishes underlying associations between autism and other concepts, as outlined in theories of meaning (Firth, 1957; Landauer &

Dumais, 1997), and these associations, consolidated in the text, can then influence public perceptions towards autism and autistic people. For example, co-occurrences of the term *autism* with terms from the disability-based model and words referring to physical and mental health conditions (Theme 1A – comorbidities and medical model, and corresponding discourses) emphasise the shared characteristics between autism and such conditions, influencing how we interpret and think about autism itself (see Balfour, 2019, p. 58; and P. Baker, 2005, p. 70, for similar arguments on discourses around schizophrenia and homosexuality).

The discourse prosody analysis elaborated the emergent representations of autism based on the annotation of relationships, assumptions and attitudes. For example, with regard to the focus on children, the discourse prosody analysis showed how autistic children lacked agency with respect to the actions described and were discussed predominantly from the perspective of their parents (in line with Huws & Jones, 2011; see also Stevenson et al., 2011 on 'infantilization', and Botha & Cage, 2022, on 'objectification'). The annotation also revealed common discourses, such as associations with other conditions, abuse, lack of support and discrimination, and corroborated earlier findings that autistic people are rarely portrayed as living successful and independent lives (Holton et al., 2014).

With regard to changes in the representation of autism over time, our results suggested a slow move towards more positive descriptions. There was also partial evidence that newspapers tended to provide more difference-based accounts of autism over the time period, consistent with the observations of Lewin and Akhtar (2021), who were based exclusively on their analysis of the *Washington Post*. This slow movement may well reflect a degree of societal progress, and the growing impact and influence of the neurodiversity movement. However, it is worth noting that the term 'neurodiversity' itself only appeared in a negligible percentage of documents and was not a collocation (see Lewin and Akhtar, 2021, for a similar finding).

With regard to reporting style, our analysis revealed some important differences in the portrayals of autism in broadsheets and tabloids. The combined results support the idea that tabloids offer more negative and deficit-based portrayals of autism and autistic people than broadsheets. This occurs alongside more individual-based (episodic) stories, referring to individuals, celebrities or fictional characters, often with a sensationalistic reporting style, which maximises both readers' interest and circulation figures (Molek-Kozakowska, 2013).

With regard to political orientation, the differences in the representation of autism between left- and right-leaning newspapers appeared to interplay with reporting style, as for example, the emphasis of left-leaning newspapers on the 'traits' discourse was qualitatively different for left-leaning tabloids and broadsheets. One important confound that should be taken into account in the interpretation of

these findings is that the majority of the national UK tabloids are right-leaning, with only one left-leaning tabloid (see Figure 3).

Media representation and language around autism is very important for the autism community (Botha et al., 2021; Bury et al., 2020; Gernsbacher, 2017; Kenny et al., 2016; Lei et al., 2021; Sinclair, 2013; Vivanti, 2020). The findings from the detailed and corpus-driven examination of attitudes, biases, stereotypes and inequalities in the representations of autism in British newspapers in this study are a useful resource for autism advocacy groups and policymakers working on promoting the acceptance of autism and fostering a fair, inclusive and respectful representation of autistic people in the press and a positive autistic identity.

For example, one important contribution of our study is that it uncovers important asymmetries in the quantity and quality of the representations of autism in different parts of the British Press. Some autism advocates might be interested in prioritising work aiming to narrow this gap. Another contribution of our approach is that it has the potential to assess changes in attitudes towards autism over time. Given that our approach can be used with very large and representative datasets, it could be an important tool for measuring the impact of policies for promoting the acceptance of autism in and through the press.

Our study is not without its own shortcomings. The influence of the press and the media more generally on public attitudes and beliefs is widely acknowledged and has been demonstrated empirically in a number of domains (Corrigan et al., 2005; Edelman, 1988; Hall Jamieson & Waldman, 2004; Happer & Philo, 2013). However, the extent to which representations of autism in newspapers shape the readers' perceptions, or the way differences between individual newspapers translate to different public attitudes towards autism and autistic people, clearly warrants further investigation. We assume that the influence of the press on public opinion is likely to be exerted through complex bidirectional interactions between the reader and the newspaper context, rather than any simple causal unidirectional relationship. For example, people both explicitly and implicitly filter the information that they are presented with, and, of course, they are informed by a variety of media sources. In addition, readers might choose to be informed by media that matches their prior beliefs, and this process of influence may, on occasion, be one of the amplifications of existing beliefs rather than one of change.

Furthermore, although qualitative and quantitative methodologies are used synergistically and complementarily in corpus-based CDA, there is necessarily a degree of residual subjectivity, inherent in all qualitative methods, in the semantic preference and the discourse prosody analysis. Future studies should attempt to corroborate these analyses using participatory research frameworks (Fletcher-Watson et al., 2019; Pellicano et al., 2017). Participatory

frameworks could enable the meaningful involvement of the autism community in this research process and lead to an understanding of the language around autism in newspapers that is relevant to the autism community and their experiences and values. It could also lead to a better translation of findings into policies and better outcomes (Pickard et al., 2022).

In addition, since this study focused on British newspapers, its findings may only apply to the United Kingdom and not necessarily generalise to other countries. This may be especially true for low- and middle-income countries (De Vries, 2016), but it may also hold for some high-income countries (for example, Buijsman et al., 2022). It is essential to conduct corpus-based studies on other countries and cultures, especially those underrepresented in autism research, to gain a nuanced understanding of press representations of autism and their relevance to autistic people from a global perspective (De Leeuw et al., 2020).

Finally, the results presented in this article focused on representations of the terms *autism* and *autistic*, while the discourse prosody analysis considered samples referring to the prominent subcategory BOY. Our new research in this area is examining the representation of alternative terms (e.g. *Asperger's*; see Baron-Cohen, 2000) and how representations may be differentiated for particular groups of individuals from within the autism community, for example, age and gender groups. There are likely to be other interesting and significant patterns in the more general data in this area still waiting to be discovered, all no doubt exerting their own influence on people's everyday attitudes towards autism. This article outlines a new approach to these issues and hopefully illustrates how some of this work could be done.

Acknowledgements

The authors thank Mia Grice, Siobhan Hagan and Dave Menagh for their help in the analysis, and Shantanu Gupta for helping with supplementary visualisations. We are also grateful to Prof. Liz Pellicano for her support in various stages of the research project.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by the British Academy/Leverhulme Small Research Grant SRG20\201250, awarded to Themis Karaminis and Costas Gabrielatos.

ORCID iD

Themis Karaminis  <https://orcid.org/0000-0003-2977-5451>

Supplementary materials

Supplementary materials for this study can be found at <https://osf.io/5yf4r/>.

Notes

1. <https://proquest.libguides.com/Europeannewsstream>
2. The asterisk is the wildcard character, representing any number of characters. Effectively, asterisks broadened the core query to include *autisms*, *autism's*, *autistics*, etc.
3. BOY is in uppercase font to denote that it is a name of a subcategory of collocations. Note that we also conducted an extended semantic preference analysis in which we examined how representations of autism differed for particular age, gender and family words. This analysis considered the following four subcategories of age, gender and family roles collocations: BOY, GIRL, FATHER and MOTHER. The extended analysis is not presented in this article, for reasons of space.
4. Education, Health and Care Plan.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).
- Bailin, A. (2019, June 6). Clearing up some misconceptions about neurodiversity. *Scientific American*. <https://blogs.scientificamerican.com/observations/clearing-up-some-misconceptions-about-neurodiversity/>
- Baker, D. L. (2011). *The politics of neurodiversity: Why public policy matters*. Lynne Rienner.
- Baker, P. (2004). Querying keywords: Questions of difference, frequency and sense in keywords analysis. *Journal of English Linguistics*, 32(4), 346–359. <https://doi.org/10.1177/0075424204269894>
- Baker, P. (2005). *Public discourses of gay men*. Routledge.
- Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyzanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273–305. <https://doi.org/10.1177/0957926508088962>
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). Sketching Muslims: A corpus-driven analysis of representation around the word 'Muslim' in the British press, 1998–2009. *Applied Linguistics*, 34(3), 255–278. <https://doi.org/10.1093/applin/ams048>
- Balfour, J. (2019). 'The mythological marauding violent schizophrenic': Using the word sketch tool to examine representations of schizophrenic people as violent in the British press. *Journal of Corpora and Discourse Studies*, 2, 40–64. <https://doi.org/10.18573/jcads.10>
- Ballou, E. (2018, February 6). What the neurodiversity movement does: And doesn't – Offer. *Thinking Person's Guide to Autism*. <http://www.thinkingautismguide.com/2018/02/what-neurodiversity-movement-doesand.html>
- Baron-Cohen, S. (2000). Is Asperger syndrome/high-functioning autism necessarily a disability? *Development and Psychopathology*, 12, 489–500. <https://doi.org/10.1177/10883576020170030801>
- Baroutsis, A., Eckert, C., Newman, S., & Adams, D. (2021). How is autism portrayed in news media? A content analysis of Australian newspapers articles from 2016–2018. *Disability & Society*. Advance online publication. <https://doi.org/10.1080/09687599.2021.1971067>
- Bie, B., & Tang, L. (2015). Representation of autism in leading newspapers in China: A content analysis. *Health Communication*, 30(9), 884–893. <https://doi.org/10.1080/10410236.2014.889063>
- Billawala, A. S., & Wolbring, G. (2014). Analyzing the discourse surrounding Autism in the New York Times using an ableism lens. *Disability Studies Quarterly*, 34(1), Article 3348. <https://doi.org/10.18061/dsq.v34i1.3348>
- Botha, M., & Cage, E. (2022). 'Autism research is in crisis': A mixed method study of researcher's constructions of autistic people and autism research. https://www.researchgate.net/publication/361995985_Autism_Research_is_in_Crisis_A_mixed_method_study_of_researcher's_constructions_of_autistic_people_and_autism_research; <https://doi.org/10.31219/osf.io/w4389>
- Botha, M., & Gillespie-Lynch, K. (2022). Come as you are: Examining autistic identity development and the neurodiversity movement through an intersectional lens. *Human Development*, 66, 93–112. <https://doi.org/10.1159/000524123>
- Botha, M., Hanlon, J., & Williams, G. L. (2021). Does language matter? Identity-first versus person-first language use in autism research: A response to Vivanti. *Journal of Autism and Developmental Disorders*. Advance online publication. <https://doi.org/10.1007/s10803-020-04858-w>
- Breeze, R. (2011). Critical discourse analysis and its critics. *Pragmatics*, 21(4), 493–525.
- Buijsman, R., Begeer, S., & Scheeren, A. M. (2022). 'Autistic person' or 'person with autism'? Person-first language preference in Dutch adults with autism and parents. *Autism: The International Journal of Research and Practice*. Advance online publication. <https://doi.org/10.1177/13623613221117914>
- Bury, S. M., Jellett, R., Spoor, J. R., & Hedley, D. (2020). 'It defines who I am or 'it's something I have': What language do [autistic] Australian adults [on the autism spectrum] prefer? *Journal of Autism and Developmental Disorders*. Advance online publication. <https://doi.org/10.1007/s10803-020-04425-3>
- Cage, E., Di Monaco, J., & Newell, V. (2018). Experiences of autism acceptance and mental health in autistic adults. *Journal of Autism and Developmental Disorders*, 48(2), 473–484. <https://doi.org/10.1007/s10803-017-3342-7>
- Chowdhury, G. G. (2004). *Introduction to modern information retrieval* (2nd ed.). Facet Publishing.
- Clarke, J. N. (2012). Representations of autism in US magazines for women in comparison to the general audience. *Journal of Children and Media*, 6(2), 182–197. <https://doi.org/10.1080/17482798.2011.587143>
- Corrigan, P. W., Watson, A. C., Gracia, G., Slopen, N., Rasinski, K., & Hall, L. L. (2005). Newspaper stories as measures of structural stigma. *Psychiatric Services*, 56(5), 551–556. <https://doi.org/10.1176/appi.ps.56.5.551>
- Davis, R., & Crompton, C. J. (2021). What do new findings about social interaction in autistic adults mean for neurodevelopmental research? *Perspectives on Psychological Science*, 16(3), 649–653. <https://doi.org/10.1177/1745691620958010>
- De Leeuw, A. E., Happé, F., & Hoekstra, R. A. (2020). A conceptual framework for understanding the cultural and

- contextual factors on autism across the globe. *Autism Research*, 13, 1029–1050. <https://doi.org/10.1002/aur.2276>
- Den Houting, J. (2019). Neurodiversity: An insider's perspective. *Autism*, 23(2), 271–273. <https://doi.org/10.1177/1362361318820762>
- De Vries, P. J. (2016). Thinking globally to meet local needs: Autism spectrum disorders in Africa and other low-resource environments. *Current Opinion in Neurology*, 29(2), 130–136. <https://doi.org/10.1097/WCO.0000000000000297>
- Draaisma, D. (2009). Stereotypes of autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1522), 1475–1480. <https://doi.org/10.1098/rstb.2008.0324>
- Edelman, M. (1988). *Constructing the political spectacle*. University of Chicago Press.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58.
- Firth, J. R. (1957). *Modes of meaning: Papers in linguistics 1934–1951*. Oxford University Press.
- Fletcher-Watson, S., Adams, J., Brook, K., Charman, T., Crane, L., Cusack, J., Leekam, S., Milton, D., Parr, J. R., & Pellicano, E. (2019). Making the future together: Shaping autism research through meaningful participation. *Autism*, 23, 943–953. <https://doi.org/10.1177/1362361318786721>
- Fromkin, V., Rodman, R., & Hyams, N. (2011). *An introduction to language* (9th ed.). Cengage Learning Wadsworth.
- Gabrielatos, C. (2007). Selecting query terms to build a specialised corpus from a restricted-access database. *ICAME Journal*, 31, 5–44.
- Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK Press 1996–2005. *Journal of English Linguistics*, 36(1), 5–38. <https://doi.org/10.1177/0075424207311247>
- Gernsbacher, M. A. (2017). Editorial perspective: The use of person-first language in scholarly writing may Accentuate stigma. *Journal of Child Psychology and Psychiatry*, 58(7), 859–861. <https://doi.org/10.1111/jcpp.12706>
- Griffiths, S., Allison, C., Kenny, R., Holt, R., Smith, P., & Baron-Cohen, S. (2019). The Vulnerability Experiences Quotient (VEQ): A study of vulnerability, mental health and life satisfaction in autistic adults. *Autism Research*, 12(10), 1516–1528. <https://doi.org/10.1002/aur.2162>
- Hall Jamieson, K., & Waldman, P. (2004). *The press effect: Politicians, journalists, and the stories that shape the political world*. Oxford University Press.
- Happé, F., & Frith, U. (2020). Annual Research Review: Looking back to look forward – Changes in the concept of autism and implications for future research. *Journal of Child Psychology and Psychiatry*, 61, 218–232.
- Happer, C., & Philo, G. (2013). The role of the media in the construction of public belief and social change. *Journal of Social and Political Psychology*, 1(1), 321–336. <https://doi.org/10.5964/jspp.v1i1.96>
- Hoffmann, S., Evert, S., Smith, N., Lee, D., & Berglund Prytz, Y. (2008). *Corpus Linguistics with BNCweb: A Practical Guide*. Peter Lang.
- Holton, A. E., Farrell, L. C., & Fudge, J. L. (2014). A threatening space? Stigmatization and the framing of autism in the news. *Communication Studies*, 65(2), 189–207. <https://doi.org/10.1080/10510974.2013.855642>
- Howlin, P., & Magiati, I. (2017). Autism spectrum disorder: Outcomes in adulthood. *Current Opinion in Psychiatry*, 30(2), 69–76.
- Hull, L., Levy, L., Lai, M.-C., Petrides, K. V., Baron-Cohen, S., Allison, C., Smith, P., & Mandy, W. (2021). Is social camouflaging associated with anxiety and depression in autistic adults? *Molecular Autism*, 12(1), Article 13. <https://doi.org/10.1186/s13229-021-00421-1>
- Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., & Mandy, W. (2017). 'Putting on my best normal': Social camouflaging in adults with autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 47(8), 2519–2534. <https://doi.org/10.1007/s10803-017-3166>
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- Huws, J. C., & Jones, R. S. (2011). Missing voices: Representations of autism in British newspapers, 1999–2008. *British Journal of Learning Disabilities*, 39(2), 98–104. <https://doi.org/10.1111/j.1468-3156.2010.00624.x>
- Jaarsma, P., & Welin, S. (2012). Autism as a natural human variation: Reflections on the claims of the neurodiversity movement. *Health Care Analysis*, 20(1), 20–30. <https://doi.org/10.1007/s10728-011-0169-9>
- Jones, S. C., & Harwood, V. (2009). Representations of autism in Australian print media. *Disability & Society*, 24(1), 5–18. <https://doi.org/10.1080/09687590802535345>
- Kapp, S. K., Gillespie-Lynch, K., Sherman, L. E., & Hutman, T. (2013). Deficit, difference, or both? Autism and neurodiversity. *Developmental Psychology*, 49(1), 59–71. <https://doi.org/10.1037/a0028353>
- Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4), 442–462. <https://doi.org/10.1177/1362361315588200>
- Krishnamurthy, R. (1996). Ethnic, racial and tribal: The language of racism? In R. Caldas-Coulthard & M. Coulthard (Eds.), *Texts and practices: Readings in critical discourse analysis* (pp. 129–149). Routledge.
- Lakoff, G. (2004). *Don't think of an elephant! Know your values and frame the debate*. Chelsea Green Publishing.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, Article 211. <https://doi.org/10.1037/0033-295x.104.2.211>
- Lei, J., Jones, L., & Brosnan, M. (2021). Exploring an e-learning community's response to the language and terminology use in autism from two massive open online courses on autism education and technology use. *Autism*, 25(5), 1349–1367. <https://doi.org/10.1177/1362361320987963>
- Lewin, N., & Akhtar, N. (2021). Neurodiversity and deficit perspectives in The Washington Post's coverage of autism. *Disability & Society*, 36(5), 812–833. <https://doi.org/10.1080/09687599.2020.1751073>
- McCombs, M. E. (2004). *Setting the agenda: The mass media and public opinion*. Blackwell.
- McKeever, B. (2012). News framing of autism: Understanding media advocacy and the combating autism act. *Science Communication*, 35(2), 213–240. <https://doi.org/10.1177/1075547012450951>

- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2012). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26, 3111–3119.
- Milton, D. E. M. (2012). On the ontological status of autism: The 'double empathy problem'. *Disability & Society*, 27(6), 883–887. <https://doi.org/10.1080/09687599.2012.710008>
- Molek-Kozakowska, K. (2013). Towards a pragma-linguistic framework for the study of sensationalism in news headlines. *Discourse & Communication*, 7(2), 173–197. <https://doi.org/10.1177/1750481312471668>
- Moss, P., Howlin, P., Savage, S., Bolton, P., & Rutter, M. (2015). Self and informant reports of mental health difficulties among adults with autism findings from a long-term follow-up study. *Autism*, 19(7), 832–841. <https://doi.org/10.1177/1362361315585916>
- Mottron, L. (2011). The power of autism. *Nature*, 479, 33–35. <https://doi.org/10.1038/479033a>
- Nicolaidis, C. (2012). What can physicians learn from the neurodiversity movement? *American Medical Association Journal of Ethics*, 14(6), 503–510. <https://doi.org/10.1001/virtualmentor.2012.14.6.oped1-1206>
- Pellicano, E., Crane, L., & Gaudion, K. (2017). *Participatory autism research: A starterpack*. UCL Institute of Education.
- Pellicano, E., & den Houting, J. (2021). Annual research review: Shifting from 'normal science' to neurodiversity in autism science. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 63, 381–396. <https://doi.org/10.1111/jcpp.13534>
- Pellicano, E., & Stears, M. (2011). Bridging autism, science and society: Moving toward an ethically informed approach to autism research. *Autism Research*, 4(4), 271–282. <https://doi.org/10.1002/aur.201>
- Perry, E., Mandy, W., Hull, L., & Cage, E. (2021). Understanding camouflaging as a response to autism-related stigma: A social identity theory approach. *Journal of Autism and Developmental Disorders*, 52, 800–810. <https://doi.org/10.1007/s10803-021-04987-w>
- Pesonen, H., Itkonen, T., Saha, M., & Nordahl-Hansen, A. (2021). Framing autism in newspaper media: An example from Finland. *Advances in Autism*, 7, 2167–2177. <https://doi.org/10.1108/AIA-01-2020-0003>
- Phillips, M. (1989). *Lexical structure of text* (Discourse analysis monograph no. 12). English Language Research, University of Birmingham.
- Pickard, H., Pellicano, E., den Houting, J., & Crane, L. (2022). Participatory autism research: Early career and established researchers' views and experiences. *Autism: The International Journal of Research and Practice*, 26(1), 75–87. <https://doi.org/10.1177/13623613211019594>
- Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4), 519–549.
- Rayson, P., Berridge, D., & Francis, B. (2004). Extending the Cochran rule for the comparison of word frequencies between corpora. In G. Purnelle, C. Fairon, & A. Dister (Eds.), *Le poids des mots: Proceedings of the 7th International Conference on Statistical analysis of textual data (JADT 2004), Volume II: Louvain-la-Neuve, Belgium, March 10–12* (pp. 926–936). Presses universitaires de Louvain.
- Reindal, S. M. (2008). A social relational model of disability: A theoretical framework for special needs education? *European Journal of Special Needs Education*, 23(2), 135–146. <https://doi.org/10.1080/08856250801947812>
- Robertson, J. W. (2009). Informing the public? UK newspaper reporting of autism and Asperger's syndrome. *Journal of Research in Special Educational Needs*, 9(1), 12–26. <https://doi.org/10.1111/j.1471-3802.2009.01112.x>
- Scott, M. (2021). *WordSmith tools* (Version 8). Lexical Analysis Software.
- Shakespeare, T. (2006). *Disability rights and wrongs*. Routledge.
- Shakespeare, T. (2010). The social model of disability. In L. J. Davis (Ed.), *The disability studies reader* (pp. 266–273). Routledge.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- Sinclair, J. (2013). Why I dislike 'person first' language. *Autonomy: The Critical Journal of Interdisciplinary Autism Studies*, 1(2). <http://www.larry-arnold.net/Autonomy/index.php/autonomy/article/view/OP1>
- Stevenson, J. L., Harp, B., & Gernsbacher, M. A. (2011). Infantilizing autism. *Disability Studies Quarterly*, 31(3), Article 1675. <https://doi.org/10.18061/dsq.v31i3.1675>
- Stubbs, M. (1995). Corpus evidence for norms of lexical collocation. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 245–256). Oxford University Press.
- Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Blackwell.
- Tang, L., & Bie, B. (2016). The stigma of autism in China: An analysis of newspaper portrayals of autism between 2003 and 2012. *Health Communication*, 31(4), 445–452. <https://doi.org/10.1080/10410236.2014.965381>
- van der Maaten, L., & Hinton, G. (2012). Visualizing non-metric similarities in multiple maps. *Machine Learning*, 87(1), 33–55. <https://doi.org/10.1007/s10994-011-5273-4>
- Vivanti, G. (2020). Ask the editor: What is the most appropriate way to talk about individuals with a diagnosis of autism? *Journal of Autism and Developmental Disorders*, 50(2), 691–693. <https://doi.org/10.1007/s10803-019-04280-x>
- Walker, N. (2012). Throw away the master's tools: Liberating ourselves from the pathology paradigm. In J. Bascom (Ed.), *Loud hands: Autistic people, speaking* (pp. 225–237). The Autistic Press.
- Walker, N., & Raymaker, D. M. (2020). Toward a neuroqueer future: An interview with Nick Walker. *Autism in Adulthood*, 3, 5–10. <https://doi.org/10.1089/aut.2020.29014.njw>
- Wendorf Muhamad, J., & Yang, F. (2017). Framing autism: A content analysis of five major news frames in U.S.-based newspapers. *Journal of Health Communication*, 22(3), 190–197. <https://doi.org/10.1080/10810730.2016.1256453>
- Wilson, A. (2013). Embracing Bayes factors for key item analysis in corpus linguistics. In M. Bieswanger & A. Koll-Stobbe (Eds.), *New approaches to the study of linguistic variability* (Vol. 4, pp. 3–11). Peter Lang.
- Yu, N., & Farrell, L. C. (2020). Autism in the media: A longitudinal study of stigma cues and framing of attribution. *Newspaper Research Journal*, 41(4), 489–505. <https://doi.org/10.1177/0739532920969887>