



Corpora & Discourse

INTERNATIONAL CONFERENCE

2026

Book of Abstracts



Contents

Plenary speakers.....	3
Abstracts: Oral presentations	8
Abstracts: Poster presentations.....	170

Plenary speakers



**Corpora &
Discourse**
INTERNATIONAL CONFERENCE
2026



Matteo Fuoli

University of Birmingham, UK

For Matteo's bio, see [here](#).

A new methodological synergy? Mixing CADS and experiments to investigate audience effects

In one of the seminal papers in corpus-assisted discourse studies (CADS), Baker et al. (2008) advocated for the integration of corpus linguistics and discourse analysis as a “useful methodological synergy”. Today, CADS is a thriving field with many applications across a range of domains, including mass media, healthcare, gender and sexuality discourses. However, while a CADS approach provides valuable insights into semantic, rhetorical, and ideological patterns in discourse, it alone cannot reveal how linguistic choices affect text consumers. In this talk, I argue that to examine the impact of discursive practices on individuals and society and enhance the explanatory potential of CADS research, we should incorporate experiments into our methodology. I will support this argument by presenting the results of recent co-authored mixed-methods studies focusing on corporate discourse (Fuoli and Hart, 2018), online political blaming (Hansson et al., 2023), and computer-mediated interaction (Fuoli et al., in prep.).

References

- Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, 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), 273-306.
- Fuoli, M., & Hart, C. (2018). Trust-building strategies in corporate discourse: An experimental study. *Discourse & Society*, 29(5), 514-552.
- Fuoli, M., Wilding, E., & Wickert, A. (in prep.). Trusting strangers online: how language shapes behavior in an online trust game.
- Hansson, S., Fuoli, M., & Page, R. (2023). Strategies of Blaming on Social Media: An Experimental Study of Linguistic Framing and Retweetability. *Communication Research*, 51(5), 467-495.



Brian King

University of Hong Kong, Hong Kong

For Brian's bio, see [here](#).

Corpora and online discourse: The first vicennium

Twenty-odd years ago, at the dawn of a new millennium, a small, scattered and hopeful set of researchers began assembling corpora of online discourse and analysing them. Venturing into corpus study of the 'social media' of the time (though not yet widely called that), we were confronted by sticky challenges in ethics, privacy, and representativeness, and we were vexed in our corpus analysis efforts by the elevated (and novel) "noise" of computer-mediated interaction and expression. Some of these obstacles have since been tackled to varying degrees of success. More robust consent procedures and improved annotation software mean privacy and basic data handling are less problematic today. But difficulties persist – especially in defining and achieving true representativeness of social media discourse and managing multifaceted conversations that cross languages, modalities, and platforms. In hindsight those days represent a relative 'age of innocence' on the world wide web. New problems have arisen as online environments evolve for better or worse. Ephemeral communication, algorithmic curation, and multilingual multimodal exchanges all demand fresh approaches for corpus collection and analysis. Here I examine which issues have been (arguably) resolved, which seem to have resisted 'easy' solutions, and review what new hurdles have begun to vex those currently pushing the limits of possibility with corpora and online discourse. I will end by offering a critical summary for corpus-assisted discourse analysts seeking to document, interpret, and understand the fast-moving target of online discourse over the past two decades and into the next.



Michaela Mahlberg

Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

For Michaela's bio, see [here](#).

Discourses and Transdisciplinarity

Language is ubiquitous and discourses provide insights into every area of human experience. The study of corpus data – as evidence of human experience – opens up opportunities for working across disciplines. At the same time, engaging with research in other disciplines can challenge our own working practices and may take us out of our corpus linguistic comfort zone. In this talk, I will draw on examples from a number of case studies to consider how we can address discursive problems across disciplinary boundaries. The concept of ‘wicked problems’ that is used in environmental sciences, for instance, to describe extreme weather events such as floods and droughts, relates to the description of meanings across different discursive contexts. Also in digital humanities, approaches to literary and cultural history connect to work in corpus-assisted discourse studies, with the relationship sometimes being masked by different disciplinary terminology. The opportunities and challenges of working across disciplines are becoming even more complex through developments in AI that have put the spotlight on the ubiquity of language in an unprecedented way. Hence, it is even more important that we continue to develop our linguistic foundations, both theoretical and methodological foundations, to be able to address the range of discourses of human experience.



Elena Semino

Lancaster University, UK

For Elena's bio, see [here](#).

Using corpus-based discourse analysis to cross boundaries and build bridges

There is a long and growing tradition in the application of corpus linguistic methods to address questions and issues from outside linguistics and outside academia. In this talk I reflect on what I have learnt over the last decade by working with researchers from other disciplines and/or non-academic stakeholders on projects where corpus linguistics made an essential contribution. These projects concern specifically: public discourses on social care in the UK; the use of visual aids in specialist consultations about chronic pain; and the differences between clinical and non-clinical first-person accounts of voice-hearing. I discuss the challenges and frustrations, alongside the opportunities and rewards, of doing such work. Overall, I suggest that expertise in corpus linguistics puts us in an ideal position to cross boundaries and build bridges between disciplines, and between academia and the rest of society. I also hope to spark a discussion about how we can best continue to do this kind of work in the challenging times ahead, with the growth of generative AI and, at least in the UK and some other countries, the combination of impact agendas and financial crises in the University sector.

Abstracts:

Oral presentations



**Corpora &
Discourse**
INTERNATIONAL CONFERENCE
2026

Tech CEO Discourse on YouTube over a Decade: A Corpus-Based Discourse Analysis

Rakan Alibri (University of Tabuk, Saudi Arabia)

Chief executive officers (CEOs) of major technology companies have become among the most influential public figures worldwide. Through interviews and podcasts, they not only promote their corporations but also shape public understandings of technology, innovation, and other societal issues. YouTube, as one of the most powerful social media platforms and the second-largest search engine after Google, attracts billions of users (Charle 2025). In the United States alone, over 80% of adults use it daily (Pew Research Center 2025). Yet, podcasts and interviews on YouTube represent one of the “dusty corners” of discourse studies (Marchi and Taylor 2018), an emerging genre that remains underexplored despite its potential as a rich source of naturally occurring spoken data. This study draws on a corpus of such interviews featuring current and former CEOs of Google, OpenAI, Tesla / X, Nvidia, Meta, and Microsoft, published between 2016 and 2025. The videos were manually selected, downloaded via a Python script, and transcribed using AssemblyAI, a speech-to-text tool offering accurate and cost-free transcription.

The resulting CEO corpus was analysed in Wmatrix 7 (Rayson 2008), using the Spoken BNC1994 as a reference corpus, with statistically overused semantic domains identified using a log-likelihood threshold of 6.63 and LogRatio to indicate effect size. The analysis identified 221 items, several of which fell into thirteen broad semantic themes that characterise CEOs’ discourse, including Science and Technological Advancement; Intensification and Importance; and Temporality and Tense. These point to a discourse that is highly future-oriented, evaluative, and affectively charged, marked by frequent reference to technology, progress, agency, and achievement. The CEO corpus was divided into two sub-corpora based on the release of ChatGPT: Pre-ChatGPT (2016–2022) and Post-ChatGPT (2023–2025). The Pre-ChatGPT phase showed greater emphasis on themes such as Economy, Work, and Education; Intensification and Importance; Agency and Achievement; and Materiality and the Environment, reflecting a discourse centred on productivity, capability, and material progress. The Post-ChatGPT phase, however, displayed higher frequencies in Governance, Power, and the Global Order; Evaluation and Judgment; Media, Communication, and Sports; and Mortality and Life, indicating a shift towards more politicised and reflective discourse. Shared domains such as Emotions and Affect and Temporality and Tense exhibited notable lexical variation, with the later period showing stronger emotional intensity and a greater focus on future-oriented expressions.

This study contributes to corpus-assisted discourse research in two main ways. Empirically, it highlights the discursive distinctiveness of language used by technology CEOs in YouTube interviews and podcasts, showing how these media interactions construct authority, innovation, and responsibility in public communication. Methodologically, it demonstrates the opportunities and challenges of YouTube as an authentic source of spoken data and the value of AI-based transcription tools such as AssemblyAI for building and analysing corpora of emerging media genres.

References

- Charle Agency. 2025. "YouTube Statistics: 2025 Data on Usage, Demographics and Trends." Charle Agency. <https://www.charleagency.com/articles/youtube-statistics/>
- Pew Research Center. 2025. "5 Facts about Americans and YouTube." Pew Research Center. <https://www.pewresearch.org/short-reads/2025/02/28/5-facts-about-americans-and-youtube/>.
- Rayson, Paul. 2008. 'From Key Words to Key Semantic Domains'. *International Journal of Corpus Linguistics* 13 (4): 519–49. <https://doi.org/10.1075/ijcl.13.4.06ray>.
- Taylor, Charlotte, and Anna Marchi. 2018. *Corpus Approaches to Discourse: A Critical Review*. Taylor & Francis Group.

The “Southern Question” in Il Forum dei Brutti in Italy: a CADS-DHA investigation.

Selenia Anastasi (Sapienza, University of Rome, Italy)

Federica Formato (University of Brighton, UK)

It is historically proven that Southern Italy faces different economic and social struggles from the rest of the country, inextricably linked with gendered roles and ethno-national imaginaries. In this paper, we unfold constructions of gender and the South (vis-à-vis North) through an investigation of a diachronic corpus of Il Forum dei Brutti, the Italian main Incel forum (Anastasi 2025).

Gramsci (1966) argues that the historical imbalances between the meridione (South) and the settentrione (North) are due to a problematic fascist legacy and, more broadly, a politics that has not fostered financial and social unity. The geopolitical divide is still maintained nowadays through anti-southern feelings (anti-meridionalismo). The media have contributed to construct ways of seeing the two geographical areas, with the South constructed as “dark, dirty, uneducated, backward and parasitic” (Viola 2019, 46). For these reasons, we triangulate CADS and DHA; while CADS can offer key linguistic patterns to examine, DHA “demystif[ies]’ the hegemony of specific discourses” (Rieslig and Wodak 2015, 88).

Methodologically, the corpus we investigate (named IFDB_South) is made of 11 sub-corpora divided into years (from 2013 to 2023, total number of tokens 270,037; total number of posts 4,049) retrieved with specific seed words, e.g. terrone, sud, terronface, nord, initially identified by Anastasi (2025). Starting from a categorisation of keywords through SketchEngine in IFDB_South, we identify three dominant discourses: (i) Discourse of migration (immigrati/migrant; immigrato meridionale/migrant from the south; reietto/outcast), reconstructing the duality between space and people; (ii) Discourse of biological/somatic determinism using a pseudoscientific vocabulary, concerning the classification of people’s aesthetic based on alleged genetic influences of dominant populations (fenotipo/phenotype, aplogruppo/haplogroup, levantino/levantine). References to the white supremacist forum Stormfront are also present (Heritage 2023); and (iii) Discourse of North vs South stereotyping (camorrista/mafioso, nordico vero- puro/true northerner), where conventional slurs, such as terrone, used to summarise the anti-southern feeling (see Viola 2019, Orrù 2023) and neologisms also appear, e.g. terronface; vikingface; cerbiattina Nordica/northern deer-ess.

These discourses are gender-bound and present unique and distinct constructions about women and men. Specifically for Southern women, the discourse contains discussion of sexual objectification at the intersect of these two axes of discrimination (Hill-Collins and Bilge, 2020) - gender and class (this related to the geopolitical space). For instance, these women are described as sexually available as a means of emancipation from a condition of poverty. Conversely, men from the South are described as being more sexually attractive, through a racist perspective.

References

Anastasi, Selenia. "Misogyny beyond borders: A Cross-Linguistic Corpus Assisted Analysis of Transnational Incel Communities." PhD diss., University of Genoa, 2025

Gramsci, Antonio. *La Questione Meridionale*. Createspace Independent, 2012.

Heritage, Frazer. *Incels and ideologies: Exploring how incels use language to construct gender and race*. Palgrave, 2023.

Hill Collins, Patricia & Bilge, Sirma. *Intersectionality*. John Wiley & Sons, 2020.

Orrù, Paolo. "Terroni: la discriminazione territoriale nel discorso pubblico italiano." In *Lingua e discriminazione: studi diacronici, lessicali e discorsivi*. Edited by D. Pietrini. Peter Lang GmbH, 2023.

Viola, Lorella. "A discourse-historical analysis of media representation of Italian internal migration." In *Migration and Media*. Edited by Viola Lorella and Andreas Musolff. John Benhamins, 2019.

Viola, Lorella & Verheul, Jaap. "The Media Construction of Italian Identity: A Transatlantic, Digital Humanities Analysis of italianità, Ethnicity, and Whiteness, 1867-1920", *Identity* 19:4. 294-312, 2019.

A Study on Energy Policy Discourse in the International Energy Agency

Mengxuan Bai (University of Southampton, UK)

This paper presents a corpus-assisted critical discourse analysis of how the International Energy Agency (IEA) uses language to construct authority and urgency in global energy governance. It asks how IEA flagship reports legitimise particular energy transition pathways and how they position different state and institutional actors in relation to these preferred futures.

The analysis draws on a specialised corpus of IEA policy reports and scenario documents published between 2014 and 2024. Fairclough's three-dimensional model provides the overall analytical frame, and corpus tools are used to identify keywords, collocation patterns and recurrent evaluative and modal choices. Concordance lines are then examined qualitatively in order to explore how actors are represented and how specific transition pathways are justified in the texts and in their wider intertextual circulation in policy and media debates.

Findings indicate that deontic modals such as “must” and “need to”, together with clusters around terms such as “net zero”, “clean energy transition” and “advanced technologies”, present particular policy directions as necessary, urgent and grounded in technical expertise rather than in political preference. Nominalisations and passive constructions frequently background responsibility for decisions while reinforcing the IEA's role as an authoritative expert voice. Intertextual references to national strategies and multilateral agreements further embed these framings, and other actors are often depicted mainly as recipients of finance, technology or capacity building rather than as co-authors of transition pathways. The paper argues that these patterns support technocratic, capital-intensive models of transition and narrow the discursive space for decentralised or community-based alternatives. In doing so, it contributes to work on discourse, power and global governance by showing how corpus-assisted CDA can be used to interrogate the authority of expert energy scenarios.

Constructing Legitimate Violence: A Corpus-Assisted Analysis of Verbs in Russian War Telegram Blogs

Olga Bikkulova (University of Bern, Switzerland)

This paper forms part of an ongoing PhD project entitled “Russian Military Telegram Blogs as a Genre: A Corpus-Driven Study”, which investigates contemporary Russian war discourse through the analysis of a newly emerged genre — Telegram blogs produced by war journalists and military bloggers. These blogs gained significant prominence after the Russian–Ukrainian war entered its current full-scale phase. The broader project examines how authors construct their communicative identities and articulate their perspectives, and which linguistic strategies do they employ to persuade audiences to align with their viewpoints. The study adopts the framework of Corpus-Assisted Discourse Studies (CADS) as developed by Baker, Partington, McEnery, Taylor, Marchi, Friginal, Hardie, Ancarno, and others, as this approach enables a systematic investigation of emotionally charged discourse grounded in quantitative evidence. The present paper narrows its focus to verbal semantics, examining verbs as key linguistic resources for construing agency in contemporary war discourse. Special attention is paid to verbs denoting lethal action, including those that share the semantic component “to kill” or “to take life,” such as убивать (“to kill”), уничтожать (“to destroy”), поражать (“to strike”), устранять (“to eliminate”), among others. The analysis also considers typical collocational patterns involving these verbs, including subjects, objects, and adverbial modifiers. This comparative approach demonstrates how lexical choices encode differing conceptualisations of violence and varying degrees of responsibility attribution.

The quantitative analysis is based on a manually annotated corpus of Russian military Telegram blogs comprising texts published between 2017 and 2025 and encompassing both pre-war and wartime discourse. For reference and comparison, the study draws on a Russian media corpus and a large corpus of the Russian-language internet, which make it possible to identify genre-specific features of military blogs in contrast to other types of public discourse. All corpus queries and statistical analyses are conducted using the Sketch Engine platform.

The findings reveal a noticeable post-2022 shift from explicit verbs of killing towards more technicalised, impersonal, and institutionally framed verbal choices. Verbs such as уничтожать (“to destroy”) and поражать (“to strike”) show a significant increase in relative frequency, while direct references to killing become comparatively less salient. The analysis shows that destruction is predominantly represented through recurrent verb–adverb collocations (e.g. methodically, successfully, daily), which construe military action as routinised, technical, and depersonalised activity. This framing foregrounds procedures and results while reducing the visibility of individual human casualties. At the same time, several verbs that might be expected to increase in wartime discourse unexpectedly decline in frequency. For instance, истреблять (“to exterminate”), a verb commonly associated with the eradication of pests (e.g. insects or rodents) and often assumed to be productive in dehumanising representations of the enemy, becomes less frequent in the corpus. In other words, the lexical profile of military blogs reflects not merely a general increase in violence-related verbs but a patterned, discourse-specific choice of verbs whose meanings, connotations, and stylistic qualities align with how bloggers

frame and legitimate wartime action. These findings highlight the role of genre-specific verbal patterning in structuring representations of military action, suggesting that legitimisation operates through routine linguistic choices rather than overt ideological positioning.

Naming the Unheard: Bridging Tools to Analyse Representations of Indigenous Language Practices in the Canadian Parliament (2019)

Ann-Sophie Boily (Carleton University, Canada)

Indigenous languages in Canada have long faced threats that continue to affect their maintenance and intergenerational transmission. In 2019, in response to decades of advocacy by Indigenous communities (Collectif, 2018, United Nations, 2007), the Canadian government adopted Bill C-91, the Indigenous Languages Act—its first and only legislation aimed at supporting the reclamation, revitalization, and promotion of Indigenous languages. Despite an apparent consensus on the need to protect these languages, the Parliamentary debates leading to the Bill's adoption reveal deep ideological tensions, notably in identity construction and language naming practices.

This presentation examines how Indigenous languages were represented in these debates through a corpus-assisted discourse studies (CADS) and discourse-historical approach (Reisigl & Wodak, 2001). Building on work highlighting the potential of CADS to uncover ideologies in discourse (Vessey, 2017), this study addresses a gap in research on political discourse surrounding Indigenous languages in Canada. Using AntConc to identify word frequencies, keywords, collocations, and clusters, and NVivo for qualitative thematic coding and argumentation analysis, I explore how linguistic and discursive choices frame Indigenous languages representation at Parliament.

Results show that naming practices and identity construction serve to make long-erased language practices visible and heard. At the same time, processes of instrumentalization, appropriation, and resurgence contribute to both reproducing and transforming colonial relations of power.

Finally, the paper discusses methodological challenges in combining corpus-based and qualitative discourse approaches. Current tools offer limited integration between quantitative and in-depth critical analysis, and tasks such as systematically cataloguing language naming practices remain highly time-consuming. These reflections point to ongoing needs for methodological innovation in CADS research engaging with Indigenous language practices and political discourses.

References

- Collectif. (2018). *Atlas des peuples autochtones du Canada* (Canadian Geographic).
- Reisigl, M., & Wodak, R. (2001). *Discourse and Discrimination: Rhetorics of Racism and Antisemitism* (1st ed.). Routledge.
- United Nations. (2007). *United Nations Declaration on Rights of Indigenous Peoples*.
- Vessey, R. (2017). Corpus Approaches to Language Ideology. *Applied Linguistics*, 38(3), 277–296.

Filmic discourse: From corpus analysis to the English as a Foreign Language classroom

Paula Wood Borque and Pilar Mur Dueñas
(Universidad de Zaragoza, Spain)

Films offer a rich source of realistic input to be brought into the English as a Foreign Language (EFL) classroom, fostering language skills and communicative competence (e.g. Bednarek, 2018; Donaghy et al., 2023). Notwithstanding the multimodal nature of these audiovisual materials (e.g. Bateman & Schmidt, 2012; Herrero & Vanderschelden, 2024), this paper will focus exclusively on the analysis of the verbal mode of a corpus of films selected for their use in the Secondary Education EFL classroom. It is our aim to analyse filmic discourse based on the scripts of 95 scenes taken from 19 films covering a wide range of genres deemed appropriate for adolescent English L2 viewers. We seek to identify keywords, especially lexical and grammatical features which can help discourse analysts to interpret such scenes, and language educators to make the most of them in the EFL classroom. The analysis is driven by the following research questions: (1) what are the keywords in the scripts from specific film scenes selected to be used in EFL Secondary Education classrooms?; what insights do these keywords offer into the narrative and dialogic construction of meaning in filmic discourse?, (2) what are the most common lexical words?; how can these be related to particular topics which can be brought into the EFL classroom and/or focus on form (vocabulary)?; (3) what are the most common grammatical words?; how can these be related to particular situations and contexts and be targeted in the EFL classroom to focus on form (grammar)?

To answer these questions, a corpus-based two-fold analysis was conducted using the software AntConc 4.3.1 (Anthony, 2024). First, a keyword list of our FISEC corpus (Films Script for the English Classroom), totalling 25,526 words, was generated using as reference corpus a sample of the COCA (Corpus of Contemporary American English). A close look was taken at those words included among the first 50. Concordance analyses were also performed on these keywords to identify their use in context. Second, a 100-word list of the FISEC corpus was created and words classified according to their lexico-grammatical nature, followed by in-depth concordance analyses to derive implications for teaching EFL.

Results from the keyword list point out the dialogic and interactive nature of filmic discourse (e.g. Bednarek, 2018) in different contexts and situations in which the films selected take place - ranging from family reunions to product advertisements and job meetings. In addition, the narrative component of filmic discourse is also reflected in this keyword list, highlighting actions and states. Similar insights are gained from the word list. Concordance lines of the most common and relevant words in the FISEC corpus can help determine if and how the scenes can be brought into the EFL classroom not only for meaning (interpreting the message conveyed) but also for contextualised form-focused instruction (raising awareness on particular lexico-grammatical choices). Overall, the study reveals the usefulness to undertake corpus analysis to make the most of authentic materials to be exploited in the EFL classroom.

References

- Anthony, L. (2024). AntConc (Version 4.3.1) [Computer Software]. Tokyo, Japan: Waseda University. <https://www.laurenceanthony.net/software/AntConc/>
- Bateman, J. A. & Schmidt, K.H. (2012). *Multimodal Film Analysis: How Films Mean*. Routledge.
- Bednarek, M. (2018). *Language and Television Series: A Linguistic Approach to TV Dialogue*. Cambridge Applied Linguistics.
- Donaghy, K., Karastathi, S. and Peachey, N. (2023). *Multimodality in ELT: Communication Skills for Today's Generation*. Oxford University Press.
- Herrero, C. & Vanderschelden, I. (2024). *Transmediation: Pedagogical Applications of Film Adaptation in Modern Foreign Language Education*. In A. Bolaños García-Escribano and M. Oaknín (eds.). *Inclusion, Diversity and Innovation in Translation Education* (pp. 121-153). UCL Press.

How is National Debt Dramatized in Canadian Newspapers?

Pier-Pascale Boulanger (Concordia University, Canada)

Chantal Gagnon (Université de Montréal, Canada)

The traditional view in electoral democracies is that liberal governments argue for higher public spending, whereas conservative parties favour budgetary restraint backed by tax cuts. Competing fiscal narratives across the political spectrum are relayed by journalists delivering facts on the current state of the national debt, while offering forecasts on the issue. The framing by the press can vary, as the national debt can be measured and reported in different ways. For instance, it is more advantageous for the governing political party to foreground Canada's net debt (the gross debt minus the government's financial assets) or the federal debt (which excludes debts incurred by provinces), because it is of a lower amount, thus less alarming. An opposing view may bring the gross debt to the forefront, intent on raising concerns about the interests paid by taxpayers. The debate can be rational, calling on economic knowledge to explain to non-expert audiences that, over time, the debt shrinks as the gross domestic product grows and inflation rises. The narrative can be emotional, such as when policymakers argue that today's debt is a burden on the livelihoods of future generations. How are these views, voices and facts handled by reporters when debt is discussed, and is national debt always bad news?

Our paper delivers the results of our investigation. We studied how journalists covered the national debt in a bilingual corpus of two Canadian newspapers, *The Globe and Mail* (2.9 million words, English) and *La Presse* (1.2 million words, French). We crafted this comparable corpus using the query terms “debt” and “dette,” with a focus on 2009, under Conservative PM Stephen Harper, and 2023 under Liberal PM Justin Trudeau. These were years of economic challenges following the 2007-2008 financial crisis, and the pandemic. Combined with the quantitative approach afforded by frequency analysis, the qualitative tools of critical discourse analysis (CDA) were employed to gain insight into agency and intensifying strategies (Fairclough 2010; Reisigl and Wodak 2001). We paired CDA with a criterion indicative of dramatic coverage, which has proven effective in a study where style was surveyed and categorized as conveying curiosity or drama (Boulanger and Gagnon 2025). We also investigated positive/negative bias using a method developed in a study of the word “confidence” in a corpus of business news (Gagnon and Boulanger 2025).

We found that the prevailing discourse in *The Globe and Mail* and *La Presse* frames high levels of public debt as dramatic. This negative portrayal — intensified through metaphors and spectacular numbers — is consistent across 2009 and 2023. The language tends to be strongly argumentative, aimed at persuading the public of the risks associated with rising levels of debt. Politicians and advocacy groups are the most frequently cited voices across newspapers and years, reinforcing a discourse that is strategically constructed to influence public perception and policy priorities. In short, national debt is conveyed as bad news, which is surprising since both media outlets are centrist. Not one cue was found construing the national debt as a positive form of investment in civil society.

References

- Boulanger, Pier-Pascale and Chantal Gagnon. 2025. "How Constructive is the Coverage of Canadians in Debt?" In *Constructive News Across Languages and Cultures*, edited by Ashley Riggs and Lucile Davier. Routledge, 89-109.
- Fairclough, Norman. 2010. *Critical Discourse Analysis. The Critical Study of Language*. 2nd ed. Routledge.
- Gagnon, Chantal and Pier-Pascale Boulanger. 2025. "Investigating Positive/Negative Bias in Canadian Newspapers Through Translation: A Study of 'Confidence' in a Corpus of Business News." *Perspectives* 33(4): 743–760.
<https://doi.org/10.1080/0907676x.2023.2179930>.
- Reisigl, Martin and Ruth Wodak. 2001. *Discourse and Discrimination: Rhetorics of Racism and Antisemitism*. Routledge.

Detecting Discourses Around Sustainability in Social Media: A Lexical Multidimensional Analysis

Arianne Brogini (Pontifical Catholic University of São Paulo, Brazil)

‘Sustainability’ is usually defined through environmental concerns, with a focus on how societies attempt to reconcile environmental preservation with economic activity in capitalist systems. Researchers describe the concept as a negotiated balance among environmental protection, social well-being, and economic development, commonly referred to as the “three pillars” or the “triple bottom line” of planet, people and profit. These formulations treat sustainability as a response to the tension between ecological limits and the demands of capitalist expansion (Purvis, Mao & Robinson, 2019; Mensah, 2019). The United Nations introduced its Sustainable Development Goals (SDGs; United Nations, 2015), which contributed to the visibility of sustainability in policy and public debates. At the same time, the term has circulated across many different contexts, where it has been taken up for different purposes. To date, research has documented the incorporation of this concept in the public sphere, such as its blending with corporate social responsibility, but no comprehensive study of the spread of sustainability-driven discourses in social media seems to exist.

To fill this gap, we collected a corpus comprising 93,689 tweets from 42,503 users, spanning 2018 to 2022, each containing the Portuguese word ‘sustentabilidade’ (sustainability). We ran a Lexical Multi-Dimensional Analysis (Berber Sardinha & Fitzsimmons-Doolan, 2025), which detected eight dimensions across the texts. Each dimension comprised correlated sets of lemmas, expressing the major discourses on sustainability. For instance, Dimension 1, ‘Innovation and Sustainability vs. Economic Realism,’ contrasts the corporate culture of fostering innovation and actions towards sustainability and SDGs with the practical concerns of profit, capitalism, and the economic crisis impact on the population and workers. Dimension 4, ‘Renewable Energy and Efficiency vs. ESG Communication,’ compares renewable energies and low-carbon initiatives for savings against the communication strategies in governance related to ESG and SDGs. And Dimension 5, ‘Knowledge Domain vs. Tangible Credit’, presents sustainability as linked to its scientific community as an extensive knowledge domain. Conversely, the negative pole brings forward the certification market, which encompasses multimodal expressions of trust that reinforce the concept of sustainability. Overall, the dimensions show how the concept has been stretched to serve a wide range of discourses, which contributes to its dilution. The dimensions are ranked by statistical magnitude, which helps sort the discourses by prominence. In contemporary usage, the term has developed considerable semantic elasticity, particularly in institutional and corporate settings, where it is used to address diverse and sometimes competing aims. The dimensions capture the resulting discursive tensions associated with this elasticity, showing how groups with conflicting interests rely on the same term to position themselves ideologically. All dimensions will be discussed in the presentation.

References

Berber Sardinha, T., & Fitzsimmons-Doolan, S. (2025). *Lexical Multidimensional Analysis*. Cambridge University Press.

Mensah, J. (2019). Sustainable development: Meaning, history, principles, pillars, and implications for human action. *Cogent Social Sciences*, 5(1), 1653531. <https://doi.org/10.1080/23311886.2019.1653531>

Purvis, B., Mao, Y., & Robinson, D. (2019). Three pillars of sustainability: In search of conceptual origins. *Sustainability Science*, 14, 681–695. <https://doi.org/10.1007/s11625-018-0627-5>

United Nations. (n.d.). Sustainable Development Goals. <https://www.un.org/sustainabledevelopment/>

Coercive Control in the Courtroom: A corpus-based Critical Discourse Analysis on the representation of domestic abuse in Depp v. Heard

Charlie Brookes (Cardiff University, UK)

Domestic abuse is widely recognised as a pattern of power and control, with perpetrators employing a range of emotional, verbal, and psychological tactics to dominate and subjugate their victims (Stark 2007; Pitman 2017). While many victims describe these non-physical forms of abuse as equally, if not more, harmful than physical violence (Velonis 2016; Andrus 2021), legal responses often focus primarily on isolated incidents of physical harm (Barlow et al. 2020; Sheley 2021). This focus is reflected across both legal and media discourses, where linguistic analyses have shown a tendency to depict domestic abuse as a series of discrete physical assaults, often overlooking the broader, ongoing dynamics of coercive control (Cotterill 2003; Trinch 2003; Lee and Wong 2020). Moreover, language that perpetuates harmful stereotypes, particularly those that blame victims and mitigate perpetrator responsibility, has been consistently identified within institutional discourses related to domestic abuse and gender-based violence (Matoesian 2001; Lloyd and Ramon 2017). This study integrates corpus linguistics with Critical Discourse Analysis to examine the linguistic representation of domestic abuse in the Depp v. Heard (2019) defamation trial, a case centred on competing allegations of abuse. The presentation outlines how key semantic domain and concordance analyses were applied to identify lexical and grammatical patterns across four core genres of the trial: opening statements, direct examination, cross-examination, and closing arguments. Through these corpus-based methods, the research uncovers significant, and often problematic, themes regarding the nature of abuse, its victims, and its perpetrators. It further illustrates how the plaintiff (Depp) and the defence (Heard) construct contrasting narratives of domestic abuse at different stages of the trial, revealing how these linguistic representations align with their respective legal strategies.

References

- Andrus, Jennifer. 2021. *Narratives of Domestic Violence: Policing, Identity and Indexicality*. Cambridge: Cambridge University Press.
- Barlow, Charlotte, Kelly Johnson, Sandra Walklate, and Laura Humphreys. 2020. "Putting Coercive Control into Practice: Problems and Possibilities." *The British Journal of Criminology* 60 (1): 160–79.
- Cotterill, Janet. 2003. *Language and Power in Court: A Linguistic Analysis of the O.J. Simpson Trial*. Basingstoke: Palgrave Macmillan.
- Cross, Chelsea K. 2022. "Coercive Control and the Limits of Criminal Law." *UC Davis Law Review* 56 (1): 195–246.
- John C. Depp II v. Amber Laura Heard. 2019. Circuit Court of Fairfax County, Virginia, No. CL-2019-2911.
- Lee, Cynthia, and Jessica S. Wong. 2020. "99 Reasons and He Ain't One: A Context Analysis of Domestic Homicide News Coverage." *Violence Against Women* 26 (2): 213–32.

- Lloyd, Margaret, and Shulamit Ramon. 2017. "Smoke and Mirrors: UK Newspaper Representations of Intimate Partner Domestic Violence." *Violence Against Women* 23 (1): 114–39.
- Matoesian, Gregory M. 2001. *Law and the Language of Identity: Discourse in the William Kennedy Smith Rape Trial*. Oxford: Oxford University Press.
- Pitman, Theresa. 2017. "Living with Coercive Control: Trapped within a Complex Web of Double Standards, Double Binds and Boundary Violations." *The British Journal of Social Work* 47 (1): 143–61.
- Sheley, Erin. 2021. "Criminalizing Coercive Control Within the Limits of Due Process." *Duke Law Journal* 70.
- Stark, Evan. 2007. *Coercive Control: The Entrapment of Women in Personal Life*. Oxford and New York: Oxford University Press.
- Trinch, Shonna L. 2003. *Latinas' Narratives of Domestic Abuse: Discrepant Versions of Violence*. Amsterdam: John Benjamins.
- Velonis, Amanda J. 2016. "'He Never Did Anything You Typically Think of as Abuse': Experiences with Violence in Controlling and Non-Controlling Relationships in a Non-Agency Sample of Women." *Violence Against Women* 22 (9): 1031–54.

“A new meme, ‘overtourism’ has been unleashed”: Corpus-assisted explorations of the use of overtourism in UK national newspapers

Erik Castello (University of Padua, Italy)

Dario Del Fante (University of Ferrara, Italy)

The term overtourism is a buzzword used to talk about a phenomenon that is in fact not new in tourism scholarship. It refers to tourism-led overcrowding in areas previously used exclusively by residents and to how such situations are perceived by various stakeholders with different agendas (Yrigoya et al. 2024). The OED has not yet devoted an entry to it, which suggests that the definition of its denotation and connotation still needs scholarly attention. This study aims to contribute to filling this gap and sets out to explore the linguistic and discursive construction of overtourism in the British news media, with a special focus on its use over time and on the grammar of its representation in news discourse (van Leeuwen, 2008). Ultimately, it aims to inform the creation of corpus-assisted activities that can guide language learners through the exploration of patterns of use of specific terms and to engage them in discussions about tourism (and other related social phenomena) and its impact on the environment (Poole 2024). The study explores four main research questions: How is the term overtourism used diachronically in UK newspapers? Who are the main social participants (actors, places and concepts) revolving around it and how are they represented linguistically? What social actions do they perform and what roles are they given? What differences are there, if any, between the use of overtourism in broadsheets and in tabloids? It adopts corpus linguistic and CADS methods (Baker 2023, Partington et al. 2013), with a view to identifying keywords and lexico-grammatical patterns that encode representations of the phenomenon and of related participants and processes. It is based on a corpus of about 3,100,000 tokens, specially compiled from UK National newspapers through Nexis Uni, using the search terms overtourism/over-tourism. The software used is #LancsBox X 5.0.0 (Brezina and Platt 2024).

A preliminary analysis shows that its first attested uses date to May 2017. Its presence increased until 2019, decreased until 2022 due to the pandemic, and has been growing exponentially since 2024. A diachronic analysis of collocates suggests that in 2017 the very meaning of the term was often explained, in 2018-2019 the impacts of activities (e.g., cruising) were discussed, while lately journalists have often reported on protests and on proposals to tackle its negative effects. An inspection of a selection of concordance lines indicates that the main social agents are impersonal groups (e.g., locals, protesters), businesses (e.g., hoteliers, Airbnb), and names of places that are used metaphorically (e.g., destinations, Venice). These tend to take active roles (e.g., march against, complain about) and less frequently passive ones (e.g., affected by, suffer from). Finally, while broadsheets tend to depict overtourism metaphorically as an enemy that should be fought, tabloids are more inclined to discuss its effects and how they can be reduced.

References:

Baker, Paul. (2023). *Using Corpora in Discourse Analysis*. Bloomsbury.

Brezina, Vaclav, Platt, William. (2024). #LancsBox X 5.0.0 [software package], lancsbox.lancaster.ac.uk

Partington, Alan, Duguid, Alison, and Taylor, Charlotte. (2013). *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. John Benjamins.

Poole, Robert. (2024). A corpus-assisted ecolinguistic analysis of hurricanes and wildfires and the potential for corpus-assisted eco-pedagogy in ELT classrooms. In M. Bortoluzzi and E. Zurru (Eds.), *Environment as Lifescape: Ecoliteracy and Communication in Action*. Bloomsbury.

van Leeuwen, Theo. (2008). *Discourse and Practice: New Tools for Critical Discourse Analysis*. Oxford University Press.

Yrigoy, Ismael, Horrach, Patricia, Escudero, Luis, and Mulet, Carles. (2024). Co-opting overtourism: tourism stakeholders' use of the perceptions of overtourism in their power struggles, *Journal of Sustainable Tourism*, 32(4), 818-834.

The deconstruction of the climate crisis on anti-science websites

Isabelle Clarke and Beth Cook
(Lancaster University, UK)

In September 2018, the United Nations (UN) Secretary-General António Guterres emphasised that “climate change is the defining issue of our time... Let there be no doubt about the urgency of the crisis.” This sparked a clear shift in mainstream discourse on climate change with organisations like the UN and others, climate scientists, research units, politicians and various media outlets changing their terminology to refer to climate change as a ‘crisis’ and ‘emergency’ to be more scientifically precise (Carrington, 2019). But how has this discursive shift been addressed, if at all, within alternative, anti-scientific news, websites and blogs?

Previous research (e.g. Dunlap and McCright, 2015; Clarke, 2023) has found that websites and blogs known to promote pseudoscience and conspiracy theories often downplay the seriousness of climate change. This is achieved through subtle strategies like calling climate activists ‘alarmists’ and referring to climate policies as ‘extreme’. Rather than overtly denying climate change, which is hard to do in the face of growing scientific evidence, such strategies instead present climate change as being not as bad as climate scientists and activists are making it out to be (Clarke, 2023). Our analysis builds on this work, conducting a corpus-assisted discourse analysis on a corpus of 19,961 texts from 186 anti-science websites spanning 23 years (1999-2021) to uncover how such overt calls for urgency are discussed.

Our analysis finds that ‘the climate crisis’ is not often discussed on pseudoscience and conspiracy websites, occurring only 1,286 times out of 37,991,956 word tokens. However, when it is mentioned, the climate crisis is deconstructed – referred to only so as to diminish its legitimacy (i.e. delegitimation). For example, one way in which the climate crisis is delegitimised is by labelling it ‘the Climate Crisis Industry’ and ‘the Climate Crisis, Inc.’, which presents the climate crisis as a profitable fraudulent business/industry. Such texts often accuse activists and individuals as being paid actors and managers whose roles are to engineer hysteria. Moreover, the climate crisis is often presented as fake and unproven through terms like ‘manmade’, ‘humanmade’, ‘so-called’, ‘imaginary’, and ‘alleged’, all of which foster doubt in the seriousness and authenticity of the climate crisis.

References

- Carrington, D. (2019) Why the Guardian is changing the language it uses about the environment. <https://www.theguardian.com/environment/2019/may/17/why-the-guardian-is-changing-the-language-it-uses-about-the-environment>
- Clarke, I. (2023) The discourses of climate change denialism across conspiracy and pseudoscience websites. In S. M. Maci, M. Demata, M. McGlashan and P. Seargeant (eds.) *The Routledge Handbook of Discourse and Disinformation*, pp.325-341. Routledge.
- Dunlap, R. E. and McCright, A.M. (2015) Challenging climate change: The denial countermovement. In R. E. Dunlap and R. J. Brulle (eds.) *Climate Change and Society: Sociological Perspectives*, pp.300-332. Oxford University Press.

An investigation of catastrophic in climate-themed public-oriented research communication

Niall Curry (Manchester Metropolitan University, UK)

David Wright (Nottingham Trent University, UK)

Lorenzo Zannini (Università degli Studi di Napoli L'Orientale, Italy)

Fear appeals in climate communication have been shown to undermine sustained research engagement and to lead to feelings of denial, avoidance, defensiveness, anxiety, and fatalism among the public. Given the goal of public-oriented research communication to inform the public of climate knowledge and shape public responses to climate issues, a critical understanding of the relevance of fear appeals in such communication is needed. To this end, this talk examines the use of the term catastrophic in public-oriented research communication, using a corpus of climate-themed texts from The Conversation Australia, Canada, UK, and US. Adopting corpus-assisted discourse analysis approach, the talk presents a study of how the use of the term catastrophic has changed across time and region. Specifically, it highlights the polysemy of the term in public-oriented research communication, the complex contexts in which the term has evolved, and the implications of fear appeals for this emergent register.

‘As an American...’: A Corpus-Assisted Study of Nation-Related Self- and Other-Positioning in Reddit Crisis Discourse

Mark Dang-Anh (IDS Mannheim, Germany)

Giorgia Riboni (University of Turin, Italy)

In times of crisis (such as the Corona pandemic), individuals tend to resort to strong in-group identification and out-group derogation as a way of coping with uncertainty and insecurity (cf. Tajfel & Turner 1979). The closure of borders in the initial phase of the COVID-19 emergency triggered this dynamic, which was also observable on social media. One of the most widely used platforms, Reddit (Pfurtscheller 2023), enables users to initiate discussions, post and reply to comments as well as up/downvote them, thereby facilitating targeted interactions (such as those about border policies) in a sequential and hierarchical manner.

Against this backdrop, this paper explores discursive positionings on Reddit with regard to the issue of the closure of borders. The study combines stancetaking analysis (Du Bois 2007) with a Corpus-assisted Discourse Studies approach (Gillings et al. 2023) to investigate how recurrent border- and nation-related expressions are used to index self and other affiliations. The platform Reddit was chosen due to its techno-discursive makeup (cf. KhosraviNik 2018), which allows users to interact in community-like subreddits, signal their stance and (dis)align with one another with respect to the topic debated (cf. Kiesling et al. 2018). Data was taken from Pushshift's Reddit dump from the year 2020 (Baumgartner et al. 2020). The target corpus comprised threads with titles containing covid or corona in conjunction with border. Two reference corpora were created by stratified sampling from the same data, excluding border and excluding border, covid, and corona, respectively.

Results indicate that users tend to adopt a national stance in Reddit threads on border closure. Firstly, highly ranked lexical items and patterns—such as adjectival attributions (e.g., American) and demonyms (e.g., Canada)—were identified as markers of nation-related discourse. This was confirmed by the analysis of the keywords of the target corpus (using the two reference corpora as benchmarks). Secondly, self-identifications like as an American or Canadian here function as recurring linguistic means of conveying nation-related perspectives in debates over political issues amid COVID-19 and border closures. Finally, the micro-linguistic level of analysis revealed different positionings adopted by US-affiliated Redditors, who account for the vast majority of the platform, with respect to bordering countries. While talk about Canada showed no marked tendency, discourse about Mexico was characterized by a predominantly negative evaluative stance. However, it turned out that nation-related discourse is, by itself, not inherently nationalistic but depends on positionings that promote national unity, distinctiveness, or superiority.

References

- Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The Pushshift Reddit Dataset. arXiv.
- Du Bois, John W. 2007. "The Stance Triangle." In *Stancetaking in Discourse*, edited by Robert Englebretson, 139–182. Amsterdam: John Benjamins Publishing Company.

- Gillings, Mathew, Gerlinde Mautner, and Paul Baker. 2023. *Corpus-Assisted Discourse Studies. Elements in Corpus Linguistics*. Cambridge: Cambridge University Press.
- KhosraviNik, Majid. 2018. "'Social Media Techno-Discursive Design, Affective Communication and Contemporary Politics.'" *Fudan Journal of the Humanities and Social Sciences* 11(4): 427–442.
- Kiesling, Scott F., Umashanthi Pavalanathan, Jim Fitzpatrick, Xiaochuang Han, and Jacob Eisenstein. 2018. "'Interactional Stancetaking in Online Forums.'" *Computational Linguistics* 44(4): 683–718.
- Pfurtscheller, Daniel. 2023. "'Vom Fundus zum Korpus: Reddit als Medium und digitale Sprachressource.'" *Korpora Deutsch als Fremdsprache* 3(2): 137–159.
- Tajfel, Henry, and John C. Turner. 1979. "'An Integrative Theory of Intergroup Conflict.'" In *The Psychology of Intergroup Relations*, edited by William G. Austin and Stephen Worchel, 33–37. Monterrey, CA: Brooks Cole.

Evolving Discourses about “Influencers” on BlueSky

Daria Dayter (Tampere University, Finland)

This study investigates the evolving discourses surrounding the discourse keyword “influencer” within the context of migration to the new microblogging platform BlueSky. Using corpus-assisted discourse analysis of a 1,5 mio tokens corpus of BlueSky posts (collected with the BlueSky Scraper, Moncomble 2024, “influencer(s)” as the search query), I examine how this term is used and perceived during three waves of Anglophone user migration to BlueSky (primarily from Twitter/X). Wave 1 occurred after October 2022, Musk’s purchase of Twitter and the announcement of planned changes; Wave 2 started in February 2024, when an invite was no longer needed to join BlueSky; Wave 3 occurred post-November 2024, the US election of Trump. I investigate collocates in grammatical frames [verb] + influencer, influencer + [verb], and modifier [adjective|noun] + influencer to uncover shifting attitudes and associations with these roles over time.

The theoretical framework for the analysis is provided by the Positioning Theory, an approach to discursive negotiation of identity as “depending upon the positions made available within one’s own and others’ discursive practices and within those practices, the stories through which we make sense of our own and others’ lives” (Davies and Harré 1990: 46). The samples of 200 concordance lines for each grammatical frame collocate in each wave (or fewer, if the subcorpus does not contain sufficient number of collocates) are inductively coded for the storylines that BlueSky users assign to ‘influencers’. For each resulting storyline, I describe the language patterns (if any), the concepts and moral judgements made relevant and the subject positions made available within the storylines.

At this stage, the findings based on the analysis of the entirety of the corpus indicate the presence of three main storylines: aspirational, humorous, and critical. In the aspirational story, the subject position “legitimate worker” is made available to influencers, and the common linguistic patterns include first-person future-oriented verbs “I aspire / want /plan” and concessives. In the humorous story, the subject position is “comic foil”, and language includes playful labels like “corgi influencer”. In the critical story, the subject position is “the grifter” and the language patterns include negative qualifiers, quotation marks, and downtoners/minimizers. I plan to extend the analysis to each wave separately, in order to arrive at conclusions about the development of BlueSky user identity throughout migration stages.

It appears that the meaning of ‘influencer’ shifts from emblem of entrepreneurial success to shorthand for superficiality or even moral hazard, reflecting the recent influencer fatigue visible across social media (and manifested, anecdotally, in the “deinfluencing” trend). BlueSky users actively police this term and continually negotiate who is entitled to influence and on what moral grounds. Platform ideologies (BlueSky being distinctly left-leaning, less toxic, with emphasis on community curation rather than celebrity-driven engagement, cf. Pew Research 2025; Nogara et al. 2025) also undoubtedly shape how influencer roles are redefined and contested.

References

Davies, Bronwyn, and Rom Harré. 1990. "Positioning: The Discursive Production of Selves." *Journal for the Theory of Social Behavior* 20: 43–63.

Moncomble, Florian. 2024. BlueskyScraper (Version 0.4). Université d'Artois.
<https://fmoncomble.github.io/blueskyscraper/>

Nogara, Giovanni, Samin Abdali, Tommaso Radicioni, Robert West, and Fabian Flöck. 2025. "A Longitudinal Analysis of Misinformation, Polarization and Toxicity on Bluesky." arXiv.
<https://arxiv.org/abs/2505.02317>

Pew Research Center. 2025. "Bluesky Has Caught On with Many News Influencers, but X Remains Popular." Pew Research Center, May 29, 2025.
<https://www.pewresearch.org/short-reads/2025/05/29/bluesky-has-caught-on-with-many-news-influencers-but-x-remains-popular/>

Responsibility in times of instability: a cross-linguistic and cross-cultural analysis of parliamentary debates in Finland, Italy and the UK.

Dario Del Fante (University of Ferrara, Italy)

Jenni Räikkönen (University of Helsinki, Finland)

In this paper, our focus is on the framing of responsibility in the parliamentary debates of Finland, Italy and the United Kingdom. We examine the use of the noun responsibility and the adjective responsible (*vastuu** in Finnish, *responsibilit /responsabile* in Italian) in the debates held between February 2022 and September 2025. During this period, multiple crises have become entangled, creating a polycrisis (Lawrence et al. 2022: 2), which offers a fruitful context for studying how the concept of responsibility is used in parliamentary discourse.

Our data consist of the transcripts of all parliamentary sittings between February 2022 and September 2025 in the UK House of Commons, Parliament of Finland and the Italian Chamber of Deputies and the Senate of the Republic [~89 million words in total]. By employing methods of corpus-assisted discourse studies (e.g., Baker 2023), we analyse who are framed as being responsible and for what. Our aim is to better understand differences in the three datasets and whether challenges in the global political situation can be observed in the use and understanding of the concept.

By examining data from three different countries, we can avoid making too far-reaching conclusions about the concept based on one language and culture only. Furthermore, British, Finnish and Italian parliamentary debates are compared because of the differing roles and positions the respective countries have in the global political landscape. The UK, historically one of the most powerful European nations, left the EU in 2020. The UK has sought to re-define itself as a “global” trading and financial actor, indicating a shift in its global position and international relations (Egan & Webber 2023). In Finland, Russia’s attack on Ukraine in 2022 marked a drastic shift in public opinion and foreign policy, eventually leading to Finland joining NATO in 2023. Finally, Italy, one of the founding members of the EU and a central European country, occupies a strategic geographical position within the Mediterranean Sea. The country plays a delicate role in managing relations between North Africa and Europe regarding migration policy, in particular. These differences possibly influence how responsibility – particularly in relation to international conflicts – is understood.

Despite there being international conflicts to which responsibility is a crucial concept, national issues dominate the discussions. In particular, contributions focusing on whether the previous or the current government is responsible for the economic situation are common in the Finnish and British parliaments. Instead, Italian parliamentary discussions on responsibility generally focus on issues of cybersecurity, AI, and digital transition. Some of the differences between the three datasets regarding the use of responsibility are due to differences in the code of conduct of parliamentary sittings between the three countries. For instance, while British MPs should not address other MPs directly, Finnish and Italian MPs can, which creates differences in who are given responsibility. Furthermore, linguistic and cultural differences seem to affect the discursive portrayal of responsibility more than the global political situation. However, responsibility does have a role in discussions on international conflicts, as well.

References

Baker, P. (2023). Using corpora in discourse analysis (2nd edn.). Bloomsbury Academic.

Egan, M., & Webber, M. (2023). Brexit and 'Global Britain': Role adaptation and contestation in trade policy. *International Politics*. <https://doi.org/10.1057/s41311-023-00503-2>.

Lawrence, M., Janzwood, S., & Homer-Dixon, T. (2022). What Is a Global Polycrisis? Version 2.0. Discussion Paper 2022-4. Cascade Institute. <https://cascadeinstitute.org/technical-paper/what-is-a-global-polycrisis/>.

Figurative framings of Nation in British Conservative Party Parliamentary Discourse: an overview from 2015 to 2022

Dario Del Fante and Ilaria Iori (University of Ferrara, Italy)

Representations of nations and nationalism have long been central to the rhetoric of European right-wing populist parties. In the UK, the growing influence of the Conservative Party has likewise been shaped by nationalist narratives and metaphors in parliamentary discourse and manifestos, especially during crises (Hayton, 2012). These discourses typically frame nationalism through historical continuity, cultural heritage, and traditional values, reinforcing national identity (Daniel, 2022). Metaphors can shape political narratives and influence public perception (Musolff, 2016), especially in times of crisis, when metaphorical (re)conceptualisations can promote particular interpretations of ongoing issues (Entman, 1993).

This paper examines metaphorical representations of nation and nationalism in British Conservative Party parliamentary discourse during three major crises from 2015 to 2022: the migration crisis, Brexit, and the COVID-19 pandemic. These overlapping events—described as a “polycrisis” (Lawrence et al., 2024)—intensified anxieties around sovereignty, identity, and governance, offering a rich context for analysing nationalist metaphors. The 2015 migration crisis sharpened debates on borders, shaping the rhetoric of the 2016 Brexit referendum (Outhwaite, 2019). Brexit reflected decades of political discontent, while COVID-19 (2020–2022) deepened economic instability and social fragmentation, reinforcing nationalist narratives (Mohammadi, 2023).

We draw on the ParlaMint-GB 4.1 corpus (Erjavec et al., 2023), part of a multilingual, diachronically comparable collection of parliamentary corpora. Our dataset comprises three sub-corpora corresponding to each crisis: 2015–2016 (migration), 2017–2019 (Brexit), and 2020–2022 (COVID-19). Using a corpus-assisted metaphor analysis approach (Semino, 2008), we identify metaphorical mappings of the NATION target domain via source domains used by Conservative MPs. Our research addresses three questions:

- What conventional metaphors of the nation recur across the three crises?
- Which metaphors are unique to specific periods?
- Can a broader metaphorical mapping of the “nation in crisis” be hypothesised?

Our findings suggest a certain degree of continuity in how the UK is metaphorically represented in the three crises. Specifically, in all periods, the UK is personified in terms of a vulnerable person who needs to be protected and defended within a war-like and container frame. Such representations contribute to the discursive construction of a permanent state of emergency, which may justify exceptional political measures.

References

- Daniel, R. (2022). The liberal/conservative nationalism divide. *Nations and Nationalism*, 28(2), 523–538. <https://doi.org/10.1111/nana.12812>
- Entman, R. M. (1993). Framing: Toward clarification. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>

- Erjavec, T. et al. (2023). ParlaMint corpora.
- Hayton, R. (2012). *Reconstructing Conservatism?* Manchester University Press.
- Lawrence, M. et al. (2024). Global polycrisis. *Global Sustainability*, 7, e6.
<https://doi.org/10.1017/sus.2024.1>
- Mohammadi, A. N. (2023). Political Framing of COVID-19. *IJCL*, 29(2), 189–212.
<https://doi.org/10.1075/ijcl.22087.moh>
- Musolff, A. (2016). *Political metaphor analysis*. Bloomsbury.
- Outhwaite, W. (2019). Migration crisis and Brexit. In *Oxford Handbook of Migration Crises* (pp. 93–109). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780190856908.001.0001>
- Semino, E. (2008). *Metaphor in discourse*. Cambridge University Press.

Representations in Polish Parliamentary Debates on Abortion (1989-2024)

Tomasz Detlaf (University of Warsaw, Poland)

Abortion has been a topic of highly contentious public debate in Poland since the democratic transition of 1989, which opened the possibility of changing abortion law and therefore debating the issue.

This paper traces the major changes in Polish parliamentary discourse on abortion, focusing on the representations of the key objects of this discourse: abortion, women, and the fetus/child. Parliamentary debates were selected as they are directly related to the lawmaking process, are influenced by and at the same time have a significant impact on broader public debate, and allow for systematic comparisons over time.

The analysis presented in this paper covers the period from the 1989 semi-free parliamentary elections to 2024, when the abortion regulation was debated in the Sejm (lower chamber of the Polish parliament) for the last time. Abortion was a subject of the debate in the plenary session of the Sejm in as many as 21 of those 36 years.

Using the Polish Parliamentary Corpus (Ogrodniczuk 2018) and data from Sejm's website, a corpus of speeches from debates on bills and resolutions regarding abortion was created. To identify words used to represent analyzed objects of discourse, lists of keywords characterizing abortion debates were prepared using speeches from all other debates as a reference corpus. The lists were prepared for each parliamentary term separately and later merged to avoid omitting terms specific to particular periods. From these lists, words referring to abortion, women, and the fetus/child were selected, and their frequencies, concordances, and collocations were analyzed.

The word "abortion" itself, at the beginning rarely used, has consistently gained popularity and is now often used as a neutral term by supporters of abortion rights. At the same time, "abortion" has a strongly negative semantic prosody (through collocates and negatively charged derivative words such as "abortive" or "abortionist"), similarly to what Malory (2024) found in internet English.

Representations of women became more frequent and diverse after the protests of 2016 and 2020, which popularized the language of feminist dissent (cf. Chatupnik and Brookes 2022). Prior to this period the word "child" was more frequent in parliamentary debates on abortion than "woman." After the protests, "woman" is used more frequently, but also words that represent women as close persons (such as "sisters" or "daughters"), as well as "Polish women" relating to national identity are used more often. In addition, representations of specific women mentioned by their names became more common.

The representation of the "unborn" fetus/child as "innocent" and "defenseless" was not present in the discourse from the beginning, but emerged only after 2005 under the influence of John Paul II's rhetoric, commonly quoted after his death. Representations of impaired fetuses/children as "disabled" persons appeared almost exclusively in 2013-2019, during the period of the activity of social movements seeking to remove access to abortion in cases of foetal impairment.

References

- Chałupnik, Małgorzata, and Gavin Brookes. 2022. 'Discursive Acts of Resistance: A Multimodal Critical Discourse Analysis of All-Poland Women's Strike's Social Media'. *Gender and Language* 16 (3): 308–33. <https://doi.org/10.1558/genl.20148>.
- Malory, Beth. 2024. 'Polarized Discourses of Abortion in English: A Corpus-Based Study of Semantic Prosody and Discursive Salience'. *Applied Linguistics* 45 (3): 481–97. <https://doi.org/10.1093/applin/amad042>.
- Ogrodniczuk, Maciej. 2018. 'Polish Parliamentary Corpus'. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, edited by Darja Fišer, Maria Eskevich, and Franciska de Jong. European Language Resources Association (ELRA).

A linguistic exploration of SHAME and GUILT in English and Japanese parenting websites

Eugenia Diegoli (Ca' Foscari University of Venice, Italy)

SHAME and GUILT are self-conscious emotions that involve a negative evaluation of the self (SHAME) or one's behaviour (GUILT); they correlate with specific action tendencies (desire to hide, flee, repair the damage) and physiological manifestations (flushing, head lowering). SHAME and GUILT are also social emotions because they are generally experienced in front of (real or imagined) others and are functional to the maintenance of normative standards.

In this talk, I explore how speakers of English and Japanese describe and express SHAME and GUILT in research questionnaires and parenting websites, in the attempt to identify cross-linguistic similarities and differences in the semantic profiles of these two emotions. Within corpus linguistics, the challenge is one of operationalisation: how do we move from labels that denote emotion (shameful, guilty) to language that expresses or signals it? (Bednarek 2008; Marchi 2025).

In the first part of the study, I identified possible linguistic markers of SHAME and GUILT by looking for patterns in the English (n=845) and Japanese (n=1,340) responses to a scenario-based questionnaire that combines two measures of shame- and guilt-proneness used by psychologists (the TOSCA-3 [Tangney and Dearing 2002]; and the KA-JiKoKan12 [Kikuchi and Arimitsu 2002]). Specifically, for each language, I identified the words that occur statistically more often in the questionnaire responses, when compared to a general corpus of that language variety. I take these (key)words to be possible markers of SHAME and GUILT.

The two lists of keywords present areas of overlap: expressions of apology (mōshiwake 'apologise', shazai 'apology', gomen 'sorry', apologize, apologetic, sorry), emotion labels (hazukashii 'embarrassing', kōkai 'regret', okott(e) 'angry', guilty, embarrassed, angry), and intensifiers (zettaini 'absolutely', seishin 'wholeheartedly', very, profusely), among others.

The second part of the study builds on the keyword analysis to identify potential discourses of SHAME and GUILT in two relatively comparable web corpora (ca. 10 million tokens each):

- i. the Hatsugen Komachi tōkō data corpus (<https://www.nichigai.co.jp/dcs/index5.html>), representative of language use on the Japanese parenting website Hatsugen Komachi;
- ii. a corpus collected from the subreddits r/Parenting and r/UKParenting.

The two corpora were searched for textual units with at least three of the previously identified keywords, assuming that in such units we may see a cumulative effect that increases the chances that they actually relate to experiences of SHAME and GUILT.

Following a coding protocol, a sample was manually annotated for several variables, including: (a) emotion type(s); (b) emotion cues (the linguistic elements that motivate the annotation); (c) trigger(s) (what triggers the emotion). Each sentence was annotated for up to three emotion types to reflect the multi-compositionality of emotional experiences. The collocational analysis of emotion terms complemented the semantic annotation.

The findings suggest that, despite important differences, emotion co-occurrence is the norm in both the English and Japanese data, and that SHAME and GUILT tend to correlate with ANGER, FEAR and/or SADNESS.

En route, the study reflects on how far we can get in the identification and analysis of emotional experience using corpus methods alone, and is more generally relevant to scholars interested in (relative) ineffability and linguistic variation.

References

Bednarek, Monika. 2008. *Emotion Talk across Corpora*. Palgrave Macmillan.

Kikuchi, Akio, and Kōhki Arimitsu. 2002. KA-JiKoKan-12. <https://jspp.gr.jp/doc/KA-Jikokan-12.pdf>.

Marchi, Anna. 2025. 'Hypothesis-Testing in Corpus-Assisted Discourse Studies: A Methodological Exploration'. *International Journal of Corpus Linguistics*, ahead of print, September 9. <https://doi.org/10.1075/ijcl.24049.mar>.

Tangney, June Price, and Ronda L. Dearing. 2002. *Shame and Guilt*. Guilford Press

A Multivariate Analysis of Quotation in German Parliamentary Debates

Nathan Dykes (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)

Parliamentary debates constitute a hybrid genre: although delivered orally, they are highly planned, institutionally regulated, and in many cases, oriented toward public and media audiences rather than deliberative exchange (Burkhardt 2011). In this context, verbatim quotation, an otherwise highly atypical feature of spoken language, is strategically employed as a discursive resource through which speakers position themselves, invoke authority, and align parliamentary talk with external voices and texts. While reported speech has been studied in narrative and media discourse via corpus techniques (Čermáková & Mahlberg 2018, Taboada 2025), the analysis of parliamentary quotation with its highly specific functions has remained mostly qualitative (Gruber 2015, Atkins & Finlayson 2016) and focused on isolated functions, leaving open how different linguistic choices systematically interact at scale.

This study investigates verbatim quotation in GermaParl, a corpus of German parliamentary discourse (Blaette & Leonhardt 2023). Quotations were extracted using CQL queries and annotated for the type of actor being cited (e.g. contemporary politicians, experts, literary figures), syntactic form as operationalised in the query (e.g. QUOTE says actor vs. actor says QUOTE), reporting expressions, and metalinguistic marking via explicit self-insertion (ich zitiere 'I quote'). I applied multiple correspondence analysis, an exploratory technique to identify co-occurrence patterns in categorical data, to derive statistical dimensions, which were then interpreted qualitatively. The approach is conceptually similar to multidimensional analyses of register variation, but uses categorical variables rather than frequency correlations.

The statistical dimensions reveal systematic co-occurrence patterns that would be difficult to identify through isolated examples, pointing to distinct quotation practices.

One dimension foregrounds literary figures as the source of quotations (ein Zitat von Goethe 'a quote by Goethe'), which speakers use to foster alignment with external value systems. This dimension is associated with introducing the quotation via nouns rather than verbs. A different dimension combines media actors and metalinguistic verbs (schreiben 'write'), with speakers emphasising connectedness with contemporary matters. Stylistically, dimensions differentiate between agentive versus textualised quoting, spoken versus written sources, and varying degrees of speaker self-positioning. While some quotation strategies foreground institutional procedure and formality, others emphasise authority, or distance. In sum, co-occurring linguistic markers indicate distinct strategies through which MPs manage quotation to responsibility, legitimacy, and intertextual positioning.

References

- Atkins, Judi, and Alan Finlayson. "As Shakespeare so memorably said...": Quotation, rhetoric, and the performance of politics." *Political Studies* 64, no. 1 (2016): 164-181.
- Blaette, Andreas, & Leonhardt, Christoph (2023). GermaParl Corpus of Plenary Protocols (v2.0.1). Zenodo. <https://doi.org/10.5281/zenodo.10416536>.
- Burkhardt, Armin. *Das Parlament und seine Sprache: Studien zu Theorie und Geschichte parlamentarischer Kommunikation*. Vol. 241. De Gruyter, 2011.

- Čermáková, Anna, and Michaela Mahlberg. "Translating fictional characters: Alice and the Queen from the Wonderland in English and Czech." In *The Corpus Linguistics Discourse: In Honour of Wolfgang Teubert*, pp. 223-253. John Benjamins, 2018.
- Gruber, Helmut. "Policy-oriented argumentation or ironic evaluation: A study of verbal quoting and positioning in Austrian politicians' parliamentary debate contributions." *Discourse Studies* 17, no. 6 (2015): 682-702.
- Taboada, Maite. "Reported speech and gender in the news: Who is quoted, how are they quoted, and why it matters." *Discourse & Communication* 19, no. 1 (2025): 93-113.

Body, movement and sensation in fanfiction on Jane Austen

Nathan Dykes, Anastasia Glawion and Michaela Mahlberg
(Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)

Fanfiction is non-commercial literature written by fans in response to an existing work. Jane Austen fandom, in particular, has been long-standing, resulting in a body of texts with different narrative patterns and aesthetic conventions. While corpus linguistics has paid little attention to fanfiction, combining original and fan work offers a unique window into cultural reinterpretations of classical literature while contributing to the exploration of “overlooked text types” (Lischinsky 2018).

The present talk compares references to the body across both corpora: A Jane Austen corpus (six novels, ~900k tokens) and a collection of 77 Austen-based fanfiction stories from the platform AO3 (~400k tokens). We used WMatrix 7 (Rayson 2008) for USAS semantic tags. For our analysis, we used the Python library FlexiConc (<https://pypi.org/project/FlexiConc/>) to combine USAS tags with flexible concordance analysis. The tag B1 (Anatomy and physiology) serves as our primary window into embodiment, and we compare body reference patterns across the corpora in three steps: 1) frequent lexical realisations of B1 tags 2) USAS bigrams left of the node and 3) ranking concordance lines using KWICGrouper (O'Donnell et al. 2008; Mahlberg et al. 2020) by number of secondary body-related tags comprising sensory (X3) and movement terms (M1, M2, M8).

Frequent B1 realisations reveal contrasting vocabularies. While the corpora share over half of the top 20 terms, fanfiction favours concrete body parts (hand, eyes). Austen's dominant 'body' term heart is metaphorical, the second most-frequent item body is part of the pronominal expression every body; and her unique terms are indirect bodily references (roused, nerves).

The USAS bigrams show both continuity and divergence. Shared bigrams indicate conventional gestures (shook his head; took her hand). The tag sequence B1 Z5 appears in both corpora but with different realisations: fanfiction deploys it for body part coordination (neck and fingers), while Austen uses either routinised expressions (arm in arm) or non-bodily references (rouse and wake). Unique fanfiction bigrams indicate intimate physical contact (feeling and kissing), while Austen-specific patterns are largely non-corporeal (with all my heart).

KWICGrouping reveals further differences. In fanfiction, dense sensory passages fall into two patterns: sight vocabulary during emotionally charged exchanges between characters (looked, watching), and tactile or auditory language in intimate scenes. Austen's sensory passages are both fewer and less concentrated, favouring hypothetical and negative constructions around seeing (prevented from seeing; unseen). Movement patterns show a similar divergence: while fanfiction emphasises dynamic character positioning and physical engagement, Austen foregrounds stillness and metaphorical terms (sitting still; set her heart upon).

The results highlight two characteristic features of fanfiction. First, they confirm its well-documented focus on erotic content. Second, while often treated as a purely literary phenomenon, fanfiction shows clear traces of cinematic influence, as evidenced by the prominence of gaze dramaturgy as a narrative device in fan works (Cuntz-Leng 2015).

References

- Lischinsky, Alon. "Overlooked text types: From fictional texts to real-world discourses 1." In *Corpus Approaches to Discourse*, pp. 60-81. Routledge, 2018.
- Mahlberg, Michaela, Viola Wiegand, and Anthony Hennessey. "Eye language–body part collocations and textual contexts in the nineteenth-century novel." *Phraséologie et stylistique de la langue littéraire/Phraseology and Stylistics of Literary Language*. Berlin: Peter Lang (2020): 143-176.
- O'Donnell, Matthew Brook. "KWICgrouper–Designing a tool for corpus-driven concordance analysis." *International Journal of English Studies* 8, no. 1 (2008): 107-122.
- Rayson, Paul. "From key words to key semantic domains." *International Journal of Corpus Linguistics* 13, no. 4 (2008): 519-549.

What happens when a feature simply is not there? A short study of a gender-neutral pronoun in translation.

Irene Elmerot (University of Gothenburg, Sweden)

Tora Hedin and Jonáš Thál (Stockholm University, Sweden)

The Swedish gender-neutral, singular pronoun “hen” has entered public and private discourse as a tool for inclusive reference, as well as to denote individuals who do not align with the binary gender system (Wojahn 2024). Yet when translating “hen” into languages without a canonical gender-neutral equivalent, how is inclusivity carried over, altered, or lost? This study addresses this question through a discursive analysis of translations in the most recent edition of InterCorp (Cinková et al. 2024).

While the use of “hen” has become increasingly commonplace in Sweden, for example in official European Union documents in Swedish (Melzer 2011; Council of the European Union 2018), there is no equivalent gender-neutral pronoun in, for example, the Czech language. Despite the fact that this pronoun has been the focus of linguistic and sociocultural analyses (Gustafsson Sendén, Renström & Lindqvist 2021), the question of how it is translated into languages without corresponding gender-neutral pronouns remains relatively unexplored. Existing studies on gender-neutral language in Czech are also often contrastive in nature (Kolek 2019; Ivanová & Kyseřová 2022), but do not study pronominal options.

In this presentation, the questions are 1) What translation strategies are employed to render “hen” into Czech texts (fiction and non-fiction) and 2) How do these strategies affect the perceived gender in the target passages? There will also be a short discussion on what these effects reveal about broader mechanisms of cross-linguistic transfer of sociolinguistic norms.

Methodologically, this study combines contrastive corpus linguistics with discourse analysis, examining nine identified translation strategies. These translations have been observed to result in alterations to the connotations of the target texts. The findings of this study underscore the challenges associated with the transfer of gender-neutral language across languages characterised by divergent grammatical and cultural norms, albeit within the same language family. Practically, the findings offer insights into where inclusivity is most fragile in cross-language transfer and where it can be reinforced.

References

Cinková, Silvie, Lenka John, Martin Vavříň, and Adrian Jan Zasina. 2024. Korpus InterCorp – Švédština. V. 16ud. InterCorp. Institute of the Czech National Corpus, Charles University, released September 17. kontext.korpus.cz. <https://kontext.korpus.cz/>.

Council of the European Union. 2018. Inclusive Communication in the GSC. Publications Office. DOI: 10.2860/09975.

Gustafsson Sendén, Marie, Emma Renström, and Anna Lindqvist. 2021. ‘Pronouns Beyond the Binary: The Change of Attitudes and Use Over Time’. *Gender & Society* 35 (4): 588–615. DOI: 10.1177/08912432211029226.

- Ivanová, Martina, and Miroslava Kyselová. 2022. 'Attitudes Towards Gender-Inclusive Language Among Slovak, Czech, and Polish Speakers'. *Journal of Linguistics/Jazykovedný Casopis* 73 (3): 395–420. DOI: 10.2478/jazcas-2023-0015.
- Kolek, Vít. 2019. 'Options for Labelling Non-Heteronormative People: A German-Czech Comparison'. *Wiener Slawistischer Almanach*, no. 84: 303–14.
- Melzer, Sarah Maria. 2011. 'Att översätta könsneutralt: En empirisk undersökning med fokus på översättning till svenska och tyska på Europaparlamentet'. Master's thesis, Stockholms universitet.
https://www.tolk.su.se/polopoly_fs/1.57994.1321531966!/sarah_melzer.pdf.
- Wojahn, Daniel. 2024. 'Ten Years with Hen. Attitudes towards the Swedish Gender-Neutral Pronoun Hen in Public Online Comments and Official Language Recommendations (2012–2022)'. In *Public Attitudes Towards Gender-Inclusive Language*, edited by Falco Pfalzgraf. De Gruyter. DOI: 10.1515/9783111202280-014.

Taboo and Critical Discourse Analysis: A Linguistic Taxonomy of Transparency on Femicide

Sophie Eyssette (La Sapienza University of Rome, Italy)

In the United Kingdom, the successful Netflix series *Adolescence* launched a vivid debate on toxic masculinity and femicide (Comerford, 2025), bringing a sensitive issue into public eye. In 2023, it was estimated that 51,100 women and girls were killed by their partners or family members in the world, i.e., a woman is intentionally killed every 10 minutes (UN Women, 2024). The terms femicide and feminicide were respectively coined by the feminist researchers Russell (1976) and Largarde (2010) to refer to the killings of women due to their gender. This paper seeks to investigate whether these terms are transparently used or euphemised in media discourse.

For this purpose, this paper conducts a corpus-assisted discourse analysis, triangulating CL and CDA to determine the degrees of transparency on femicide in the UK press. The news coverage on femicide was explored over a five-year timespan: from 2020, the year when the first lockdown due to Covid began in Great Britain, up to 2025, the year of the data collection. This analysis frames femicide as a social taboo, insofar as killing any human being –among whom women– is prohibited in British law. Thus, as any social taboo, this prohibition leads to a tabooified discourse (Allan & Burridge, 2006), which this paper aims to analyse.

The results demonstrate that femicide is a social phenomenon increasingly acknowledged. However, they are overwhelmingly euphemised, thus damaging the transparency with which this taboo could be addressed. This paper classifies the linguistic avoidance strategies through which the taboo of femicide tends to be reported in the press. As such, degrees of transparency can be conceptualised on a continuum: articles can be transparent when they use accurate terminology, factual data and clear agency. On this continuum, some articles are much vaguer, mentioning femicide without the adequate terminology, establishing ambiguous agency and unclear, or even anecdotal, causality. Finally, there are non-transparent articles that display an absence of accurate terminology and clear agency, thus reporting this type of news as a social phenomenon statistically insignificant. The suggested taxonomy aims to be applied to other taboo topics and further expanded in CADs.

The Politicization of Climate Change: A Corpus-Assisted Ecolinguistic Analysis of Reddit

Mariasophia Falcone (University of Calabria, Italy)

Critical approaches remain underrepresented in climate change research (Pepermans and Maesele 2016). However, as climate change continues to be a reality, the central question is shifting from whether it exists to how societies should respond to it. This shift marks an ideological turning point towards the politicization of climate change, as from a policy and technocratic challenge it will become the competing cultural and conceptual battleground for diverging worldviews (Brown 2014; Marquardt and Lederer 2022; De Wilde 2011). Although not without risks, politicization may be both inevitable and necessary to imagine alternative socio-environmental futures (see Schmid-Petri 2017), where exclusionary views such as eco-fascism or climate apartheid may compete with approaches rooted in equity and justice, i.e., eco-socialist, eco-feminist, and postcolonial perspectives. So far, indicators of the politicization of climate change may have pointed to a rise in eco-fascism, a far-right ideology grounded in nativist beliefs that exploits environmental concerns to advance ethnonationalist and authoritarian objectives primarily by portraying immigration as the main cause for climate degradation (Forchtner 2019; Richards et al. 2022). This has been evidenced by acts of terrorism such as the 2011 Utøya and the 2019 Christchurch and El Paso shootings, as well as by the growing use of dehumanizing language such as illegal aliens (Hughes et al. 2022; Richards et al. 2022; Wang et al. 2025). In an attempt to investigate this from a discursive perspective, this study presents an analysis of climate-related subreddits using corpus-assisted ecolinguistics (Poole 2022). In particular, the corpus is analysed both quantitatively and qualitatively to evaluate whether exclusionary or nativist framings may be present, especially regarding the instrumental linking of immigration with ecological threats. The findings suggest that, as the politicization of climate change intensifies, climate-mitigation efforts become embedded in broader ideological conflicts, yet environmental action may be reframed as a vehicle for extremist ideology.

References

- Brown, Mark B. 2014. "Climate Science, Populism, and the Democracy of Rejection." In *Culture, Politics and Climate Change: How Information Shapes Our Common Future*, edited by Desera A. Crow and Maxwell T. Boykoff, 129–45. London: Routledge.
- De Wilde, Pieter. 2011. "No Polity for Old Politics? A Framework for Analyzing the Politicization of European Integration." *Journal of European Integration* 33 (5): 559–75.
- Forchtner, Bernhard. 2019. "Eco-Fascism: Justifications of Terrorist Violence in the Christchurch Mosque Shooting and the El Paso Shooting." *OpenDemocracy*, August 13.
- Hughes, Brian, Dave Jones, and Amarnath Amarasingam. 2022. "Ecofascism: An Examination of the Far-Right/Ecology Nexus in the Online Space." *Terrorism and Political Violence* 34 (5): 997–1023.
- Marquardt, Jens, and Markus Lederer. 2022. "Politicizing Climate Change in Times of Populism: An Introduction." *Environmental Politics* 31 (5): 735–54.

- Pepermans, Yves, and Pieter Maesele. 2016. "The Politicization of Climate Change: Problem or Solution?" *Wiley Interdisciplinary Reviews: Climate Change* 7 (4): 478–85.
- Poole, Robert. 2022. *Corpus-Assisted Ecolinguistics*. London: Bloomsbury Publishing.
- Richards, Imogen, Callum Jones, and Gearóid Brinn. 2022. "Eco-Fascism Online: Conceptualizing Far-Right Actors' Response to Climate Change on Stormfront." *Studies in Conflict & Terrorism* 48 (9): 982–1008.
- Schmid-Petri, Hannah. 2017. "Politicization of Science: How Climate Change Skeptics Use Experts and Scientific Evidence in Their Online Communication." *Climatic Change* 145 (3): 523–37.
- Wang, Yi Feng, Xinyi Zhang, and Yuwei Zhong. 2025. "Blaming 'Illegal Aliens' for Crisis?: Cultural Worldview Defenses in Effects of Death and Economic Threats on Anti-Undocumented Immigrant Attitudes." *International Journal of Intercultural Relations* 106 (May): 102150.

The Discourse of Authenticity: Constructing Commodity Value in Shanghai's High-end Sustainable Fashion

Qin Fan (Lancaster University, UK)

Introduction: This paper examines the discursive construction and commodification of authenticity within Shanghai's high-end sustainable fashion sector. As Harvey argues, "'claims to uniqueness, authenticity, particularity, and speciality...are as much an outcome of discursive constructions and struggles as they are grounded in material fact.'" The appropriation of linguistic and semiotic resources helps create authenticity and commodity value. This process constructs symbolic and economic worth by aligning with consumer expectations and cultural values, thereby reinforcing exclusivity and social distinction in elite consumption practices. While corpus-assisted discourse studies (CADS) have mapped lexical patterns in consumer texts and ethnography has contextualised consumption practices, a methodological gap remains in their combined application to complex questions of value. This research fills that gap with a novel, ethnography-led, corpus-validated mixed-methods approach that integrates multi-sited ethnography with CADS to triangulate between the breadth of recurrent linguistic patterns and the depth of their contextual, lived meanings.

Data and Methodology: The multi-modal dataset in this study concentrates on three high-end sustainable fashion brands based in Shanghai: klee klee, Norlha, and Recloning Bank. It comprises ethnographic field notes, interviews, media articles, and a dedicated corpus of 122 promotional articles from klee klee's official WeChat account. The corpus was compiled and analysed using Sketch Engine. The methodology involved four iterative phases: immersive ethnography, corpus analysis guided by field insights, a dialogic validation loop for triangulation, and final synthesis. This approach captures both overarching discursive patterns and their contextual implementation, providing a more comprehensive understanding of how value is strategically constructed within this elite market.

Key Findings: The analysis identifies six key rhetorical strategies that frame authenticity as an added value: (1) high quality, emphasising craftsmanship and rarity; (2) uniqueness, drawing on cultural and ethnic aesthetics; (3) locality, crafting narratives around ethical sourcing; (4) authorisation, building credibility through expert and media endorsements; (5) historicity, invoking tradition and timelessness; and (6) naturalness, advocating natural materials while selectively hiding technological dependencies. Collectively, they commodify authenticity, turning sustainability into a display of cultural capital aimed at elite consumers.

Discussion and Conclusion: These rhetorical strategies function as sophisticated mechanisms for perpetuating elitist narratives within Shanghai's sustainable fashion landscape. Authenticity emerges as a key discursive tool for elite positioning, operating as both a marketing strategy and a form of cultural capital that enables brands and consumers to co-construct privileged identities. By demanding specialised knowledge and cultivated taste to decode 'true' authenticity and sustainability, it creates new barriers to cultural participation. Consequently, the performative elitism expressed through this curated consumption reproduces, rather than challenges, conventional power structures. This study highlights the critical insight afforded by its integrated methodological approach, which captures how discursive constructions of value are systematically articulated and contextually enacted.

References

Bourdieu, Pierre. *Distinction: A Social Critique of the Judgement of Taste*. Cambridge, MA: Harvard University Press, 1984.

Harvey, David. *Rebel Cities: From the Right to the City to the Urban Revolution*. London: Verso, 2012.

Mapes, Gwynne. ""(De)constructing Distinction: Class Inequality and Elite Authenticity in Mediatized Food Discourse."" *Journal of Sociolinguistics* 22, no. 3 (2018): 265–287.

Linguistic and Conceptual Change in Disability Discourse 1950–2023: A Distributional-Semantic Analysis

Tim Feldmüller (IDS Mannheim, Germany)

Theresa Schweden (Johannes Gutenberg University Mainz, Germany)

Our study examines the linguistic construction of disability from 1950 to 2023 in German disability discourse, focusing on lexical and semantic change. Using distributional-semantic methods, we analyze how the vocabulary and concepts of disability have evolved over a span of 73 years. Our methodology combines word embeddings and collocation analysis, employing both semasiological and onomasiological perspectives. The findings are contextualized against disability models from conceptual history research (Waldschmidt 2020).

The analysis draws on a corpus of two German newspapers, *Der Spiegel* and *Die Zeit*, from the German Reference Corpus (DeReKo, Kupietz et al. 2018), comprising 777 million tokens divided into eight decadal subcorpora. We employ diachronically aligned word embedding models trained using a modified word2vec algorithm (Mikolov et al. 2013; Fankhauser and Kupietz 2017), combined with k-means clustering and network visualization to model semantic frames (Busse 2012; see Feldmüller accepted for methodological details).

Our macro-analysis of the network visualization of co-occurring word clusters reveals five recurring thematic domains of disability discourse: social belonging and group identities, medical aspects, legal dimensions, education, and crises/dynamics. The micro-analysis of nearest neighbors of specific anchor words with the stem *behinder** demonstrates, semasiologically, significant semantic broadening of these lexemes beginning around 1970. Initially describing process obstruction, these terms gradually encompass physical impairment, later expanding to mental and psychological disabilities. From 2000 onwards, disability increasingly emerges as a diversity dimension, co-occurring with other marginalized identities.

The onomasiological analysis of nearest neighbors to today's vector representations of the anchor words shows that, prior to 1970, words derived from *schädigen* ('to damage') served as discursive equivalents, particularly for war-related disabilities (*Kriegsgeschädigte*). These terms were embedded in socioeconomic contexts emphasizing financial dependency.

Our findings provide empirical evidence for the transformation from a deficit-oriented medical model toward more inclusive conceptualizations. The medical model, particularly prominent in the 1950s-60s, coexists with emerging elements of the social model recognizing societal barriers. By the 2000s, both models operate simultaneously—consistent with the relational model—while disability as a diversity dimension gains prominence.

References

- Busse, Dietrich. 2012. *Frame-Semantik: Ein Kompendium*. De Gruyter.
- Fankhauser, Peter, and Marc Kupietz. 2017. 'Visualizing Language Change in a Corpus of Contemporary German'. Paper presented at Corpus Linguistics International Conference 2017, Birmingham. Corpus Linguistics International Conference 2017.

Feldmüller, Tim. accepted. Das Framing von Extremismusvarianten Im Medialen Diskurs Der Jahre 1999 – 2021: Eine Corpus-Driven Methode Zur Erschließung Und Visualisierung Semantischer Frames. Frame-Based Approaches to Semantics. Language Science Press.

Kupietz, Marc, Harald Lungen, Paweł Kamocki, and Andreas Witt. 2018. 'The German Reference Corpus DeReKo: New Developments – New Opportunities'. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 7-12 May 2018, Miyazaki, Japan, edited by Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, et al. European language resources association (ELRA).

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. 'Efficient Estimation of Word Representations in Vector Space'. Version 3. Preprint, arXiv.
<https://doi.org/10.48550/ARXIV.1301.3781>.

Waldschmidt, Anne. 2020. Disability Studies zur Einführung. Junius.

Interdisciplinary applied corpus linguistics: A case study of a promising partnership

Shannon Fitzsimmons-Doolan and Jennifer Beseres Pollack
(Texas A&M University-Corpus Christi, USA)

This presentation will focus on the promise of interdisciplinary partnerships between corpus linguists and scholars from a range of other disciplines to generate meaningful applied scholarship rooted in discourse. It begins by introducing the field of interdisciplinarity and its tenets, and then moves on to highlighting key considerations of effective interdisciplinary collaboration generated by the field (Aboelela et al. 2017, Repko et al. 2020). Next, a case study—focusing on logistics—of the presenters’ collaborative research agenda integrating applied corpus linguistic and marine ecology serves as a central example.

The three interdisciplinary applied linguistics studies in this collaboration (Fitzsimmons-Doolan and Beseres Pollack, 2023, 2025, in preparation) investigate discourse related to natural resource management and restoration in the Gulf of Mexico. For each study, the presenters developed research questions through in-depth discussion of the relevance of texts, hypotheses informed by theory from both disciplines, and the analytical affordances of various corpus linguistic techniques. The results inform communication strategies and regional management efforts for coastal and estuarine ecosystems. Lessons learned from the collaboration include paying keen attention to project audience at the outset, engaging in explicit communication about many usually taken-for-granted decision points, and being proactive about publishing venues. These and other implications for discourse scholars interested in interdisciplinary work will be elaborated upon in the presentation.

This session will conclude by querying the attendees. What other interdisciplinary applied corpus linguistics studies of discourse are being conducted? How are their findings being disseminated? What information is needed to support more interdisciplinary applied corpus linguistics scholarship?

References

- Aboelela, S. W., E. Larson, S. Bakken, O. Carrasquillo, A. Formicola, S. A. Glied, J. Haas, and K. M. Gebbie. “Defining Interdisciplinary Research: Conclusions from a Critical Review of the Literature.” *Health Services Research* 42, no. 1 (2007): 329–46.
- Fitzsimmons-Doolan, S., and Beseres Pollack, J. (2023). “Shifting linguistic patterns in oyster restoration news articles surrounding the Deepwater Horizon disaster.” *Frontiers in Conservation Science: Human-Wildlife Interactions*: <https://doi.org/10.3389/fcosc.2023.1113844>.
- Fitzsimmons-Doolan, S., and Beseres Pollack, J. (2025). “Using corpus linguistics to investigate approaches to oyster fishery management across political boundaries.” *Environmental Management*: <https://doi.org/10.1007/s00267-025-02163-9>
- Fitzsimmons-Doolan, S., and Beseres Pollack, J. (In preparation). “Discourses and stakeholders in NOAA Oral History Archive.”
- Repko, A. F., R. Szostak, and M. P. Buchberger, M. P. (2020). *Introduction to Interdisciplinary Studies* (3rd ed). Los Angeles: Sage.

Engaging the public with health science: A corpus-based analysis of interactive strategies in spoken science popularization

Valeria Franceschi and Sara Corrizato
(University of Verona, Italy)

Science popularization, or parascientific communication (Luzón and Pérez-Llantada 2022; Mur-Dueñas and Lorés 2022), is a social process (Calsamiglia and van Dijk 2004) that plays a key role in bridging the gap between the scientific community and non-specialized audiences. This process is based on two macro-categories of strategies, both verbal and non-verbal, that have the purpose of making scientific content understandable and accessible to non-experts (Lorés 2023; Luzón 2013). The first category is recontextualization, which includes those strategies (e.g. definition, reformulation, figurative language, analogies with daily life, visuals, hyperlinks) that allow for the adaptation of scientific content to the target audience in different media. The second category is that of interactive strategies (e.g. humour, personal anecdotes, inclusive pronouns, rhetorical questions, evaluation of research), that are employed to create rapport with the audience and foster engagement with the material.

The focus of the present study is to investigate how experts in spoken science popularization genres create interest and engagement when discussing topics related to health and well-being through the use of interactive strategies.

The analysis draws on a purpose-built corpus of 104 spoken speech events on these topics, for a total of 338,533 words. The corpus is constituted of two macro-sections of interviews and monologues in different genres (i.e., TED talks, audio/video interviews and monologues, interviews broadcast on TV). Using the articulated xml speaker tag that allows for advanced queries within the corpus, a corpus-assisted methodology was adopted to extract and analyze concordance lines in order to identify engagement patterns and interpret their communicative function. The selected expressions include self-mention (when I) that frame expert experience and self-disclosure (Luzón 2013), and second-person pronouns (when you, if you) that invite audience identification and participation (Ye 2021). Preliminary findings show that experts talk about their past experiences as health practitioners, their patients or their own habits; generic “you” uses are also common, but other uses of the second-person pronoun are used to create real-life scenarios that can help the audience increase their health awareness and adopt healthy behaviors.

By examining these engagement strategies, the study aims to contribute to shedding light on how experts simultaneously establish authority and foster accessibility, offering insights into effective and inclusive approaches to science communication in the digital age.

References

- Calsamiglia, Helena, and Teun A. Van Dijk. 2004. “Popularization discourse and knowledge about the genome.” *Discourse & Society* 15 (4): 369-389.
- Lorés, Rosa. 2023. “Dual voices, hybrid identities: the recontextualization of research in digital dissemination scientific discourse.” *Círculo de Lingüística Aplicada a la Comunicación* 93: 69-84.

- Luzón, María José. 2013. "Public communication of science in blogs: Recontextualizing scientific discourse for a diversified audience." *Written Communication* 30 (4): 428-457.
- Luzón, María José, and Carmen Pérez-Llantada. 2022. Digital genres in academic knowledge production and communication: Perspectives and practices. *Multilingual Matters*.
- Mur-Dueñas, Pilar, and Rosa Lorés. 2022. "When Science Communication Becomes Parascience: Blurred Boundaries, Diffuse Roles." *Publications* 10 (2): 14.
- Ye, Yunping. 2021. "From abstracts to '60-second science' podcasts: Reformulation of scientific discourse." *Journal of English for Academic Purposes* 53: 101025.

Representation of blindness and blind individuals in the UK Press (2014-2023)

Costas Gabrielatos (Edge Hill University, UK)

Themis Karaminis (City St George's University of London, UK)

Ursula Maden-Weinberger (University of Cumbria, UK)

Valeria Occelli (Edge Hill University, UK)

Ocularcentrism is the belief that vision dominates other senses, positioning visual perception as the primary and standard way of understanding and acquiring knowledge (Jay, 2013). Everyday language reinforces this hierarchy through visual metaphors that equate seeing with understanding (e.g., “I see what you mean”), while “blind” is often associated with ignorance, misjudgment, or deception (Viberg, 1984). Ocularcentrism leads to stereotypical representations and contributes to discrimination against blind people (RNBI, 2022). Media representations shape how society understands disability and often reinforce harmful stereotypes (Zhang & Haller, 2013). This project aims to characterize ocularcentric language in the UK press to address the question: How do newspapers portray blindness and blind people?

To this aim, we created the Blindness UK Press Corpus by searching the Lexis+® database for news articles related to blindness, published between 1 January 2014 and 31 December 2023. The corpus includes ten national UK newspapers (Daily Express, Daily Mail, Daily Mirror, Daily Star, Daily Telegraph, Guardian, Independent, Observer, Sun, Times, and their Sunday editions) and contains 55,904 articles (46,086,373 words).

The analysis utilised two corpus tools, SketchEngine (Kilgarriff et al., 2014) and WordSmith Tools 8 (Scott, 2022), and was carried out in three stages (see Karaminis et al. 2023, 2025):

1. Collocation: co-occurrences of the lemmas BLIND and BLINDNESS
2. Semantic preference: identification of topics/issues related to the nodes via examination of their Word Sketches. This stage identified two general categories of reference: individuals and condition.
3. Discourse prosody: identification of attitudes via manual annotation of random samples relating to a) individuals and b) the condition. The multi-level annotation scheme evolved organically and iteratively based on the data.

The main results are:

- The vast majority of references (92%) relate to the condition rather than individuals.
- Blind individuals are presented as having agency in 34% and voice in only 14% of the instances. Only 9% of the instances show both agency and voice, while the most common portrayal (41%) grants blind individuals neither. Even when they have voice, they are often presented as lacking full agency (34%).
- References to the condition are predominantly metaphorical (88%), while metaphors appear far less often when referring to individuals (24%). These metaphors are largely negative (associated with ignorance or vulnerability) whether applied to individuals (86%) or the condition (63%). Positive metaphors mainly refer to fairness and impartiality.

- Individuals are framed within a small number of discourses: more than half (56%) negative (e.g. difficulty, victimhood), with one-third (34%) being positive (e.g. empowerment, inspiration).
- The discourses around the condition mainly focus on issues of pathology (64%) with the discourses identified in references to individuals being much less frequent (negative: 21%; positive: 9%).

References

- Jay, M. (2013). *Forces Fields: Between Intellectual History and Cultural Critique*. London, New York: Routledge.
- Karaminis, T., Gabrielatos, C., Maden-Weinberger, U. & Beattie, G. (2025). Gender and family-role portrayals of autism in British newspapers: An intersectional corpus-based study. *Autism*, 29(5), 1285-1302.
- Karaminis, T., Gabrielatos, C., Maden-Weinberger, U., & Beattie, G. (2023). Portrayals of autism in the British press: A corpus-based study. *Autism*, 27(4), 1092-1114.
- Kilgarrieff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014) The Sketch Engine: ten years on. *Lexicography*, 1, 7-36.
- RNBI (2022). *VI Lives: An in-depth understanding of the experiences of people living with vision impairment (VI) in the UK*.
- Scott, M. (2022) *WordSmith Tools version 8. Lexical Analysis Software*.
- Viberg, Å. (1984). The verbs of perception: a typological study. In B. Butterworth, B. Comrie & Ö. Dahl (Ed.), *Explanations for Language Universals* (pp. 123-162). De Gruyter Mouton.

Prudence and Power: Discursive Hierarchies in the Bank of Canada's Public Speeches

Chantal Gagnon (Université de Montréal, Canada)

This paper investigates how the Bank of Canada represents and translates key economic agents—households, financial institutions and businesses—across its bilingual public speeches. As Canada's central bank communicates in both English and French, its speeches provide a unique corpus for examining how translation shapes institutional discourse and how linguistic choices influence the way economic relationships and responsibilities are depicted within the national economy.

The research is based on a corpus of 376 speeches (188 in English and 188 in French) delivered by past Governors of the Bank of Canada between 2001 and 2022. Each speech was extracted from the Bank's website, totalling roughly one million words. Using WordSmith 9.0 and Sketch Engine, we conducted keyword, collocation, and concordance analyses within a critical discourse analysis framework (Fairclough 2010; van Dijk 1998). Our combination of quantitative and qualitative methods shows how linguistic patterns convey the institutional relationship between the central bank and its audiences.

Findings suggest that the Bank asserts a central role within the discourse. First-person plural forms, such as “we,” “nous”, occur more than 4,000 times, projecting collective expertise and control over the nation's economic narrative. Households, by contrast, appear mainly as affected agents—those who borrow, consume, and “add to their debt burden”. This repeated moral framing positions them as responsible for economic stability while downplaying the role of other forces, such as the impact of consumerism on households' behaviour (Velandia-Morales et al. 2022). In 2009-2010, in the aftermath of the financial crisis, the tone of obligation (“households need to be prudent...”) stands in stark contrast to the substantial losses incurred by international financial institutions and businesses in 2007-2008.

Cross-linguistic comparison reveals significant differences in interpersonal stance. English speeches tend to employ direct and inclusive formulations (“we need to ensure”), suggesting cooperation and shared purpose, whereas French versions sometimes favour impersonal expressions (“il faut s'assurer que...”), producing a more detached, bureaucratic voice. These differences align with cultural conventions of institutional discourse but also carry ideological consequences, subtly reinforcing communicative distance with francophone audiences.

The analysis further shows that businesses can be addressed as partners (“the way we run our businesses”), while households are the object of recommendation. This hierarchy of discursive agency reveals a latent struggle over voice and legitimacy within economic communication. The French speeches, often prepared as written texts rather than delivered orally, heighten this asymmetry: the governor's spoken interventions circulate primarily in English, leaving francophone publics at an additional remove.

By uncovering how bilingual institutional discourse encodes differing degrees of proximity, moral evaluation, and participation, the study demonstrates that the Bank of Canada's speeches are not merely vehicles of monetary policy but sites where economic power is articulated. In doing so, they shape how Canadians are invited to understand their roles and

responsibilities within an economic system marked by both complexity and inequality (Yassin et al., 2024).

References

Fairclough, N. 2010. Critical Discourse Analysis. The Critical Study of Language. 2nd ed. Routledge.

Van Dijk, T. A. 1998. Ideology : A Multidisciplinary Approach. London: Sage Publications.

Velandia-Morales, A., et al. 2022. Economic inequality increases the preference for status consumption. *Frontiers in Psychology* 12 (2022): 809101.

Yassin, S., et al. 2024. The troubling rise of income and wealth inequality in Canada, Policy Options, July 18, 2024. <https://policyoptions.irpp.org/2024/07/income-wealth-inequality/>

Historical Representations of Gay Male Identity in Contemporary Australian Legislation: A Corpus-Assisted Approach

Emma Genovese (University of Technology Sydney, Australia)

In this paper, I argue that contemporary Australian legislation continues to draw from harmful historical discourses to represent gay male identity. Specifically, provisions associated with schemes that aim to rectify past charges and/or convictions of homosexual offences instead (re)produce the figure of the ‘predatory homosexual’. As the homosexual subject of the law developed (Moran 1996), this predatory archetype appeared throughout history alongside other negative representations that positioned gay men as sinful (Robinson 2008; Altman 2012), mentally ill (Dalton 2016; Wilson 2008), or deviant (Russell 2020). While law reform has attempted to address the historical harms of criminalising gay men (George 2019), the amendments also obscured our understanding of how gay male identity is represented in the law. To address issues with identifying constructions, I use a corpus-assisted approach to uncover how gay male identity is generally represented in a corpus of contemporary Australian legislation (Genovese 2025). This approach draws from corpus-assisted legal linguistics (Vogel et al. 2018), and corpus-assisted critical discourse studies (Gillings et al. 2023, 24), to analyse provisions alongside external context related to the development of the legislation (Baker 2006; Mautner 2022). In analysing historical and modern terms, I reveal that despite associations with discourses of equality, certain provisions continue to draw from historical discourses of deviancy. By analysing these provisions alongside their development, I demonstrate that gay men continue to be represented as predatory or paedophilic through mechanisms that seemingly appear neutral. Ultimately, despite legislative reform, contemporary Australian legislation continues to reproduce homophobic constructions of gay male identity.

References

- Altman, Dennis. 2012. *Homosexual: Oppression & Liberation*. University of Queensland Press.
- Baker, Paul. 2006. *Using Corpora in Discourse Analysis*. Continuum Discourse Series. Continuum.
- Dalton, Derek. 2016. “A History of Repudiating Metaphors in Australian Parliamentary Debates Concerning the Decriminalisation of Homosexuality.” *Flinders Law Journal*, no. 18: 1–28.
- Genovese, Emma. 2025. Building the Australian Legislative Corpus 2023 — Combatting Issues and Highlighting Applications of General Legislative Corpora. February 21, 30-62 Pages. 30-62 Pages. <https://doi.org/10.14762/JLL.2025.030>.
- George, Allen. 2019. “Sex Offenders No More: Historical Homosexual Offences Expungement Legislation in Australia.” *Alternative Law Journal* 44 (4): 297–301.
- Gillings, Mathew, Gerlinde Mautner, and Paul Baker. 2023. *Corpus-Assisted Discourse Studies*. 1st ed. Cambridge University Press. <https://doi.org/10.1017/9781009168144>.
- Mautner, Gerlinde. 2022. “What Can a Corpus Tell Us about Discourse?” In *The Routledge Handbook of Corpus Linguistics*, 2nd ed., edited by Anne O’Keeffe and Michael J McCarthy. Routledge. <https://doi.org/10.4324/9780367076399>.

- Moran, Leslie J. 1996. *The Homosexual(ity) of Law*. Routledge.
- Robinson, Shirleene. 2008. "Introduction." In *Homophobia: An Australian History*, edited by Shirleene Robinson. The Federation Press.
- Russell, Emma K. 2020. *Queer Histories and the Politics of Policing. Queering Criminology and Criminal Justice*. Routledge.
- Vogel, Friedemann, Hanjo Hamann, and Isabelle Gauer. 2018. "Computer-Assisted Legal Linguistics: Corpus Analysis as a New Tool for Legal Studies." *Law & Social Inquiry* 43 (4): 1340–63. <https://doi.org/10.1111/lsi.12305>.
- Wilson, Emily. 2008. "Someone Who Is Sick and in Need of Help': Medical Attitudes to Homosexuality in Australia, 1960-1979." In *Homophobia: An Australian History*, edited by Shirleene Robinson. The Federation Press.

Unveiling structured antagonisms in leadership discourse: a corpus-assisted approach

Mathew Gillings (WU Vienna University of Economics and Business, Austria)

Gerlinde Mautner (WU Vienna University of Economics and Business, Austria)

Mark Learmonth (Nottingham Trent University, UK)

Andrea Whittle (Newcastle University, UK)

This paper illustrates how leadership discourse — both by and about leaders — often disguises power asymmetries through specific linguistic choices. Taking a CADS approach, we argue that corpus linguistics offers a useful lens to examine how leadership discourse conceals ‘structured antagonisms’ between leaders and followers. The paper unveils how patterns of language help to normalise power relations in leadership, and shows how corpus-assisted techniques can be integrated with critical approaches to leadership research (à la Critical Management Studies). While work in this area has long challenged the status quo of leadership studies, they have often relied on close textual reading of discourse. In contrast, CADS allows for systematic, large-scale analysis of linguistic patterns while retaining sensitivity to meaning and context (Gillings, Mautner, and Baker, 2023). To illustrate this, we present two case studies.

The first case study uses CADS to explore a 18,518-word corpus of interviews published in *The Times* (UK) — specifically where senior executives (n=733) responded to the prompt “What does leadership mean to you?”. Using Sketch Engine (Kilgarrieff et al., 2014), we identify frequent and salient lexical and grammatical patterns that effectively “edit out” hierarchy. Statistically significant keywords include positively loaded verbs like inspiring and motivating, which frames leadership as enabling and uplifting others to be the best that they can be. We also find that words implying some form of organisational hierarchy — such as subordinate or follower — are rarely used, and terms such as people and team are favoured instead, thus downplaying power asymmetries. The analysis demonstrates how leadership discourse covers-up unequal relations and instead favours humility and teamwork.

The second case study takes a more fine-grained approach, specifically examining how the term follower is used both in everyday English and within leadership scholarship. Using the 52-billion-word EnTenTen21 corpus (Jakubíček et al., 2013) accessible via Sketch Engine, we conduct a collocation analysis of the term, and look at their associated semantic prosodies. We find that, in everyday language, follower is predominantly associated with religion and social media. However, within leadership studies, follower is deployed as a neutral descriptor for employees. This linguistic recontextualization sanitises the unequal relations inherent in leadership. We ultimately argue that leadership scholars themselves continue to uphold this form of ‘structured antagonism’ in part due to disciplinary identity. Removing the term follower from professional academic discourse would undermine the conceptual foundation of the discipline, and so retaining the term is an attempt to preserve the distinctiveness of the field, even if it misrepresents actual workplace reality.

Our case studies highlight how leadership discourse obscures the inherent power within a leader-follower relationship. As such, our paper demonstrates how CADS, by identifying

linguistic patterns, challenges the ideological underpinnings of leadership studies. We conclude by advocating for a more reflexive use of corpus methods in leadership research — one that foregrounds the sociopolitical implications of language and harnesses linguistic methodology to do so.

References

Gillings, Mathew, Mautner, Gerlinde. and Baker, Paul. (2023). *Corpus-Assisted Discourse Studies*. Cambridge University Press.

Jakubíček, Miloš,, Kilgarriff, Adam, Kovář, Vojtěch, Rychlý Pavel, and Suchomel Vít (2013) The TenTen corpus family. Paper presented at 7th International Corpus Linguistics Conference, Lancaster, UK.

Kilgarriff, Adam, Baisa, Vít, , Bušta, Jan, Jakubíček, Miloš Kovář, Vojtěch, Michelfeit, Jan, Rychlý, Pavel and Suchomel, Vít (2014) The Sketch Engine: Ten years on. *Lexicography* 1: 7–36.

LLMs as Metadiscourse Annotators: A Three-Dimensional Framework for Context-Mediated Identification and Classification

Javad Haditaghi, Silvia Cavalieri and Marina Bondi
(University of Modena and Reggio Emilia, Italy)

The systematic annotation of metadiscourse in academic texts confronts a fundamental challenge: the absence of a principled framework for reproducible and consistent classification using LLMs. This contributes to what Törnberg (2024) calls an "academic Wild West" of inconsistent standards in LLM-based annotation. We address this gap by proposing a three-dimensional framework implemented through ensemble LLM annotation. Additionally, we consider metadiscourse as inherently context-mediated (Khedri et al., 2013; Pearson & Abdollahzadeh, 2023). The framework integrates three simultaneous dimensions: (1) Observable Realization, capturing objectively detectable properties including lexico-grammatical, syntax, punctuation, and crucially, reflexivity; (2) Functional Scope, determining how far metadiscursive functions extend across discourse structure to disambiguate structurally similar expressions that serve different rhetorical purposes at different discourse scales; therefore, we distinguish between micro-scope (single proposition), meso-scope (multiple propositions within coherent discourse segments), and macro-scope (section or text-level) operations; and (3) Hyland's taxonomy (Hyland, 2008) reconceptualized as an explicit three-level hierarchy: first distinguishing metadiscourse from propositional content (with borderline cases acknowledged), then classifying metadiscourse as interactive or interactional, and finally identifying specific functional types. Since metadiscourse boundaries are inherently fuzzy (Hyland, 2017), borderline cases in our framework indicate real linguistic complexity rather than annotation mistakes. Two types of borderline cases are classified: those situated between propositional and metadiscursive functions, and those positioned between specific functional types within the third level of Hyland's taxonomy. We implement this using three annotator LLMs (Claude, DeepSeek, Gemini) that independently identify and annotate each expression. Since "the nature of disagreements happens to be different for LLMs" (Bibal et al., 2025), we used a reasoner optimizer LLM (GPT-5) to resolve disagreements between annotators into consensus annotations. It systematically analyzes disagreements by identifying their sources (e.g., boundary interpretation, scope ambiguity, insufficient context) to produce the final annotations. The results include boundary corrections and explicit adjudication metadata documenting how each disagreement was resolved. This process synthesizes final annotations by weighing the linguistic and contextual evidence in each annotator's justification rather than counting votes. Annotations were drawn from Introduction and Methods sections of applied linguistics theses, with the full study expanding to over 1,000 expressions across different sections. Validation proceeded in two stages. For marker identification, the model achieved high accuracy (precision = 1.00, recall = .96, F1 = .98); the two undetected markers resulted from a per-sentence extraction limit. For classification of identified markers, agreement with expert annotations was excellent: reflexivity (94%, $\kappa = .88$), scope (98%, $\kappa = .97$), and Hyland's taxonomy ($\geq 98\%$, $\kappa \geq .85$). This pilot study (25 sentences; 52 expert-identified markers, 50 detected by the model) establishes the framework's viability. By separating observable features from interpretive scope judgments, this approach enables more granular, precise, and

replicable metadiscourse identification, offering a systematic methodology for large-scale corpus annotation.

Investigating Multimodal Meaning-Making in TV News Discourse about Immigration using the NewsScape Corpus

Christopher Hart (Lancaster University)

Corpus-assisted CDA has shown how refugees and migrants are constructed in online and print-news media, verbally, visually and multimodally (Gabrielatos & Baker 2008; Martínez Lirola 2017; Romano & Dolores Porto 2021). Owing to the difficulties associated with obtaining and analysing large quantities of televisual data, however, the discursive construction of refugees and migrants in TV news has not been subject to similar interrogation. This paper exploits the NewsScape Corpus – a massive multimodal corpus of broadcast news collated by the Distributed Red Hen lab (<https://www.redhenlab.org/>) – to investigate the multimodal representation of refugees and migrants in television news. Accessed via CQPWeb (Hardie 2012), the corpus is searched for target utterances representing four different constructions: refugees/*migrants have VERBed and refugees/*migrants are VERBing. With a focus on expressions of motion and following filtering of the data to exclude noise, the co-verbal images accompanying 474 utterances are analysed quantitatively and qualitatively. Results show that refugees/migrants are depicted in large rather than small groups, that they are depicted in transit somewhere along the migratory journey rather than in countries of origin or destination countries, that they are depicted on land more than at sea, that they are depicted in security contexts and that they are erased represented instead through abstract forms such as maps and silhouettes. Certain of these language-image combinations emerge as obtaining genre-specific multimodal constructional status (Steen & Turner 2013) and whose visual component is thus likely to be evoked even when not co-instantiated in discourse. The ideological implications of these patterns of representation are discussed from the perspective of multimodal CDA (Machin 2013) where, for example, large-group depictions are shown to have dehumanising effects (Azavedo et al. 2021). The paper presents an empirical case study but also serves to demonstrate the utility of the NewsScape Corpus as a resource for multimodal corpus-assisted CDA.

References

- Azavedo, R. T., De Beukelaer, S., Jones, I. L., Safra, L., and Tsakiris, M. (2021). When the lens is too wide: the political consequences of the visual dehumanization of refugees. *Human. Soc. Sci. Commun.* 8:115.
- Gabrielatos, C., and Baker, P. (2008). Felling, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press 1996-2005. *J. English Linguist.* 36, 5–38.
- Hardie, A. (2012). CQPweb – combining power, flexibility and usability in a corpus analysis tool. *Int. J. Corpus Linguist.* 17, 380–409.
- Machin, D. (2013). What is multimodal critical discourse studies? *Crit. Disc. Stud.* 10 (4), 347-355.
- Martínez Lirola, M. (2016). Linguistic and visual strategies for portraying immigrants as people deprived of human rights. *Soc. Semiot.* 27, 21–38.

Romano, M., and Dolores Porto, M. (2021). "Framing CONFLICT in the Syrian refugee crisis: multimodal representations in the Spanish and British press" in *Discursive approaches to sociopolitical polarization and conflict*. eds. L. Filardo-Llamas, E. Morales-López, and A. Floyd (London: Routledge), 153–173

Everyday Language Change and the Discursive Conceptualisation of Creative Practice: A Diachronic Correlation Analysis

Darryl Hocking (Auckland University of Technology, New Zealand)

A recent corpus-based study into the language of visual artists has demonstrated systematic diachronic shifts in how artists discursively conceptualise their creative practice and, crucially, has shown that these shifts statistically correlate with changes in the wider English lexicon (Hocking, 2022). Using a 235,000-word diachronic corpus of artists' interviews and statements (1950–2019), the analysis revealed that dominant discursive constructions of artistic practice, for example, as exploration, as problem-solving, or as capturing the essence of a phenomenon, rose and fell in tandem with broader patterns of everyday language use. These findings suggest a relationship between general linguistic change and the evolving conceptual frameworks through which creative practice is understood and articulated.

This presentation extends this line of inquiry by examining whether comparable patterns can be observed beyond the visual arts. To do so, it analyses a 433,281-word diachronic corpus of interviews with fiction writers, spanning the mid-twentieth century to the early twenty-first century. The study adopts the same two-stage analytical design. First, it identifies significant diachronic trends in the lexical and discursive resources fiction writers use to conceptualise their literary practices. Second, using the Corpus of Historical American English (COHA) as a reference corpus, it conducts correlation analyses to test whether these shifts also align with diachronic changes in the wider English lexicon.

The findings indicate that many of the discursive shifts observed in the language of fiction writers do correlate with broader patterns of everyday language change, reproducing key tendencies identified in the earlier study of visual artists. Although correlation is not treated as evidence of direct causation, when interpreted through discourse-theoretical perspectives that treat language as constitutive of social practice (e.g., Fairclough, 1992), the results support an interpretation in which changes in everyday language are implicated in shaping the aesthetic and conceptual development of fiction writers over time.

Together, the two studies provide converging evidence that everyday language may play a role in orienting the aesthetic and conceptual development of creative practices across domains. More broadly, the findings have implications for understanding how changes in general language use may be related to shifts in the specialist discourses through which different communities articulate, legitimise, and reconfigure their practices over time, as well as for the role of corpus-assisted discourse approaches in examining such relationships.

References

Fairclough, N. (1992). *Discourse and social change*. Cambridge: Polity Press.

Hocking, D. (2022). *The impact of everyday language change on the practices of visual artists*. Cambridge University Press. <https://doi.org/10.1017/9781108909693>

Laughed, cried, shouted: gender and emotional vocalisation in young adult fiction

Sally Hunt (Rhodes University, South Africa)

Maria Leedham and Sarah Jane Mukherjee (Open University, UK)

The representation of vocalisation is an important aspect of characterisation in fiction (Bednarek 2017, Sunderland 2011), enabling the reader to form impressions of the fictional people they encounter. This aspect of authorial choice reflects the range of pitch, volume, tone, emotion and power found in every context of human interaction, real or imagined. At the same time it sets up socially gendered expectations in readers. In this paper we present and explore the expression of emotion in the verbs laughed, cried and shouted, drawing out gendered patterns in representation in a 5-million word corpus of recent, best-selling Young Adult Fiction (YAF) (Mukherjee, Leedham and Hunt, 2025).

Verbs representing vocalisation include both speech reporting verbs, such as announced, promised or accused (cf. Caldas-Coulthard, 1994) and verbs which describe the manner in which the people speak or utter, with or without words, like yelled, cried, or giggled. Cutting across these are the verbs which include elements of emotion, for which we use the broader term ‘verbs of vocalisation’.

Gendered differences in the authors’ choice of verb have been found in several genres (cf. Johnson and Ensslin 2007: 233 on newspapers, Ruano 2016 on Dickens, Hunt 2025 on children’s fiction). Our focus here is on Young Adult Fiction in the DoRA corpus (Mukherjee, Leedham and Hunt, 2025).

In our presentation, we present the patterns evident in the use of three of the most frequent vocalisation verbs in DoRA: laughed, cried and shouted. These illustrate the powerful and subliminal role of authorial choice in the construction of gender through lexical priming. The first, laughed, is explored here.

In DoRA, laughed is more common amongst male characters (59%) than females (41%). While unmodified laughter is very common amongst female and male characters, with apparently happy, care-free amusement accounting for over 80% of tokens, the modifications reveal quite different constructions of female and male laughter, especially when negative. Overall, female laughter with negative modification is without malice, while male characters’ laughter frequently reflects hostility and wildness, as this selection of negative modifications shows:

Females

awkwardly cruelly nervously sharply; half laughed, half cried; tears in her eyes

Males

grimly sardonically wildly cruelly dementedly dismissively mean in scorn; in wild triumph; mirthlessly ruefully; his cackles bouncing off the walls; relishing my baffled horror; taunting

Males are unique in laughing at female characters, notably Edward, in *Midnight Sun*, at his girlfriend Bella. Edward laughs at many aspects of her existence, including her expressions, words, enthusiasm, requests for help and anger.

These patterns suggest that not only does laughter often reflect a much darker mood than simply having fun, it can be used in very gendered ways to maintain hierarchies of power in which the male laugher asserts his right to assess female characters' behaviour and feelings.

References

- Bednarek, M. 2017. The Role of Dialogue in Fiction. In *Pragmatics of Fiction*, edited by M. Locher and A. Jucker, 129–158. Berlin/New York: de Gruyter Mouton.
- Caldas-Coulthard, C. 1994. On Reporting Reporting: The Representation of Speech in Factual and Factional Narratives. In *Advances in Written Text Analysis*, ed. M. Coulthard, 305–306. London: Routledge.
- Hunt, S. 2025. *Linguistic Representations of Gender in Children’s Fiction: Feeling, Speaking, Doing*. Palgrave Macmillan. DOI <https://doi.org/10.1007/978-3-031-76662-6>
- Johnson, S. and A. Ensslin 2007. “‘But Her Language Skills Shifted the Family Dynamics Dramatically’. Language, Gender and the Construction of Publics in Two British Newspapers.” *Gender and Language* 1(2):229–254.
- Mukherjee, S.J., Leedham, M. and Hunt, S. 2025. Understanding Representation in Young Adult Literature Through the Voices of Young People. *New Review of Children’s Literature and Librarianship*, 31(1–2), 59–76. <https://doi.org/10.1080/13614541.2025.2575722>
- Ruano, P. 2016. An Analysis of Charles Dickens’s Gender-Based Use of Speech Verbs. *Gender and Language* 12: 192–217. <https://doi.org/10.1558/genl.31137>.
- Sunderland, J. 2011. *Language, Gender and Children’s Fiction*. London: Continuum.

Framing Chinese food in Western news media: A case study on Italian and American newspapers

Ilaria Iori (University of Ferrara, Italy)

Recent studies on Sinophobia in Western news media provided further support to Van Dijk's theories on the perpetuation of racism (e.g., Van Dijk, 1991) and found that COVID-19-related Sinophobic discourses often juxtaposed a positive 'Us' and a negative 'Them' (e.g., Costello et al., 2021; Kania, 2023; Mallapragada, 2021). Specifically, COVID-19-related Sinophobia seemed to be highly related to a biased representation of Chinese food (see Kania, 2023; King, 2020), with Western news media blaming Chinese food culture and 'exotic' food (e.g., bat soup) for the spread of COVID-19 (Jia & Lu, 2021). Despite the recurrence of such biased narratives in the representation of Chinese food in Western news media, little attention has been paid in the literature to the extent to which these narratives have changed after the COVID-19 pandemic. Hence, this paper explores representations of Chinese food in Western news discourse, comparing pandemic and post-pandemic periods in American and Italian newspapers. Methodologically, the study adopts a corpus-assisted approach (Partington et al., 2013) and combines critical discourse perspectives and corpus linguistics methods to find recurrent cross-cultural divergencies and similarities in how Chinese food was represented in the datasets. To support the analysis, a corpus of all news articles and editorials on China in the most read American and Italian newspapers was compiled for two periods (January-June 2020 and January-June 2025). The analysis uses #Lancsbox X (Brezina & Platt, 2025) to investigate collocates of words related to food (e.g., food/cibo, soup/zuppa) in the corpora. The results suggest that during the pandemic period, negative representations of Chinese food are shared between the corpora. Chinese food has often been represented as an 'exotic' delicacy, reinforcing negative stereotypes of China and recalling scenarios of danger, exoticism, and cultural otherness. In contrast, in the post-pandemic phase, especially in the United States, more diversified and at times positive representations emerged, with Chinese food increasingly framed as familiar, comforting, and embedded in everyday consumption.

References

- Brezina, V. & Platt, W. (2025) #LancsBox X [software], Lancaster University, <http://lancsbox.lancs.ac.uk>.
- Costello, M., Cheng, L., Luo, F., Hu, H., Liao, S., Vishwamitra, N., Li, M., & Okpala, E. (2021). COVID-19: A Pandemic of Anti-Asian Cyberhate. *Journal of Hate Studies*, 17(1), 108. <https://doi.org/10.33972/jhs.198>
- Jia, W., & Lu, F. (2021). US media's coverage of China's handling of COVID-19: Playing the role of the fourth branch of government or the fourth estate? *Global Media and China*, 6(1), 8–23. <https://doi.org/10.1177/2059436421994003>
- Kania, U. (2023). 'They really eat anything don't they?': Pronoun use in COVID-19-related Anti-Asian racism. In L. L. Paterson, *The Routledge Handbook of Pronouns* (1st edn, pp. 333–349). Routledge. <https://doi.org/10.4324/9781003349891-28>

King, M. T. (2020). Say no to bat fried rice: Changing the narrative of coronavirus and Chinese food. *Food and Foodways*, 28(3), 237–249.

<https://doi.org/10.1080/07409710.2020.1794182>

Mallapragada, M. (2021). Asian Americans as racial contagion. *Cultural Studies*, 35(2–3), 279–290. <https://doi.org/10.1080/09502386.2021.1905678>

Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and Meanings in Discourse: Theory and practice in corpus-assisted discourse studies (CADS)* (Vol. 55). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.55>

Van Dijk, T. A. (1991). *Racism and the press*. Routledge.

Insurance Contracting Discourses: Framing and Transparency

Katherine A Ireland, Tim R Samples and David Eckles
(University of Georgia, USA)

Insurance policies epitomize take-it-or-leave-it consumer contracting: long, standardized documents drafted unilaterally by insurers to govern high-stakes coverage that is often functionally mandatory (e.g., automobile liability). Courts and regulators have long recognized that many policyholders neither read nor understand these forms; the U.S. “plain language” movement led roughly 70% of states to introduce readability requirements in insurance statutes. Yet the relationship between these regimes, linguistic complexity, and the discursive construction of trustworthiness in insurance contracts remains under-described.

Corpus-based studies show how legal texts construct rights and institutional positioning through linguistic complexity (Samples, Ireland, and Kraczon 2024), while research on business registers offers tools for analyzing stance and credibility in organizational discourse (Fuoli 2018; Fuoli and Beelitz 2025; Kretzschmar et al. 2004; Cao and Zhao 2019). Critical discourse studies further emphasize how legal language can reproduce asymmetries of knowledge and power (Le and Machin 2023). Building on these strands, this paper applies corpus-assisted discourse analysis to insurance policies to examine how complexity and trust-related discourse vary across regulatory environments.

We introduce a new corpus of over 1,000 insurance contracts from all U.S. states and the District of Columbia (Florida excluded), comprising approximately 33 million tokens. Using a CADS approach, we ask: (1) which linguistic features are most strongly associated with complexity across jurisdictions, and (2) how do complexity patterns vary across state regulatory environments (e.g., presence/absence of readability requirements), and what do these patterns imply for transparency and perceived trustworthiness?

Analyses combine quantitative corpus methods with concordance-driven discourse analysis. We operationalize complexity using syntactic measures (e.g., dependency distance and embedding metrics) and keyness/collocation/dispersion to identify recurrent patterns relevant to risk allocation and consumer comprehensibility (e.g., definitional framing, agency attribution, procedural requirements). We then link these patterns to discourse functions (e.g., stance, responsibility allocation, “plainness” claims). Where appropriate, we model state variation with mixed-effects frameworks to separate jurisdictional differences from document-length and genre effects. The study contributes a replicable resource and a corpus-based account of how insurance contracts manage the tension between legal precision and comprehensibility, with implications for transparency, consumer protection, and the communicative consequences of “plain language” regulation (Zódi 2019; Samples, Ireland, and Kraczon 2024).

References

Cao, Miao, and Wei Zhao. “A Corpus-Based Analysis of Business Contract English Registers.” *International Journal of English Linguistics* 9, no. 3 (2019).

- Fuoli, Matteo. "Building a Trustworthy Corporate Identity: A Corpus-Based Analysis of Stance in Annual and Corporate Social Responsibility Reports." *Applied Linguistics* 39, no. 6 (2018): 846–885. <https://doi.org/10.1093/applin/amw058>. OUP Academic
- Fuoli, Matteo, and Annika Beelitz. "Corporate Buzzword or Genuine Commitment? A Corpus-Assisted Analysis of Corporate 'Net-Zero' Pledges by Major Global Corporations." *Applied Corpus Linguistics* 5, no. 3 (2025): 100142. <https://doi.org/10.1016/j.acorp.2025.100142>
- Le, Cheng, and Machin David. "The law and critical discourse studies." *Critical Discourse Studies* 20, no. 3 (2023): 243-255.
- Samples, Tim R., Katherine Ireland, and Caroline Kraczon. "TL;DR: The Law and Linguistics of Social Platform Terms-of-Use." *Berkeley Technology Law Journal* 39 (2024): 47–110.
- Stubbs, Michael. "Collocations and Cultural Connotations of Common Words." *Linguistics and Education* (1995): 370–390.
- Zódi, Zsolt. "The Limits of Plain Legal Language: Understanding the Comprehensible Style in Law."

From corpus to experiment and back again: What method and data triangulation can tell us about language and gender bias

Sylvia Jaworska (University of Reading, UK)

A central contribution of CADS research has been the empirical and systematic uncovering of linguistic patterns that reveal how language can perpetuate bias, marginalisation, and discrimination. However, an open question remains: do such patterns actually matter and if so, how, and for whom? This study addresses that question through a triangulated approach, focusing on the use of metaphorical language in workplace performance reviews (WPRs).

First, a corpus of 1,043 authentic WPRs (collected from the same level of senior leadership at a large professional services firm) was analysed to identify metaphorical expressions and their conceptual domains, with particular attention to how metaphors were used to evaluate male and female employees. The analysis revealed a diverse range of metaphors in WPRs and a clear gender pattern: men were significantly more often described using metaphorical language, particularly metaphors associated with SPORT, FORWARD MOVEMENT, BODY, WAR AND WIZARDRY.

To test the potential impact of such metaphorical framings on perception and decision-making in organisational contexts, an experiment was conducted using vignettes carefully constructed from authentic corpus data. A total of 280 participants rated the promotability of candidates described using either non-metaphorical or metaphorical language drawn from six conceptual metaphorical personae identified in the corpus (e.g., ATHLETE, SOLDIER, MASTER, DRIVER, DIVER, CONFIDANT). The results demonstrate the powerful role of metaphor in influencing perceptions of promotability but also a persistent gender bias: while certain metaphors appear to boost promotability ratings for both women and men, men consistently benefit more. Notably, the most promotability-boosting metaphors were sparsely used in the corpus and almost exclusively in evaluating male employees.

By combining corpus analysis with experimental design, this study offers novel analytical and methodological contributions. For CADS, it uncovers the metaphorical framing of workplace performance and shows how the use of metaphors reflects gendered constructions of leadership. It also demonstrates how corpus insights can be corroborated through experimental testing. For experimental research, it shows how corpus data can inform the design of more ecologically valid stimuli by adding real-life language use. Such a triangulated approach can offer a more comprehensive understanding of how language reflects and reinforces (gender) bias in society.

Mapping Trust in the Courtroom: A Corpus-Assisted Discourse Study of Jury Deliberations

Rebecca Jenkins and Nuria Lorenzo-Dus (Swansea University, UK)

This paper presents a corpus-assisted discourse study of how trust is linguistically constructed in jury deliberations. Drawing on a bespoke corpus of 11 mock juries (approximately 94,000 words and 10 hours of video-recorded talk), the study explores how jurors discursively construct the notion of trust vis-a-vis courtroom actors and evidence, particularly in the context of user-generated digital content. This dataset was collected as part of the overall project (TRUE) in which this work is embedded. Synergising theories and methods from psychology, legal studies and linguistics, the TRUE Project explores the impact of deepfakes on trust in user-generated evidence in accountability processes for human rights violations.

132 participants—diverse in age, gender, and ethnicity—watched a professionally staged fictional trial involving a real video of an airstrike in Yemen. The trial featured a former International Criminal Court judge (Sir Howard Morrison KC), eminent prosecution and defence counsel, and a digital forensics expert (Nick Waters). Following the trial, participants deliberated in groups as mock juries. These deliberations were video-recorded, transcribed, and compiled into a corpus for analysis.

Using SketchEngine’s wordlist and concordance tools, the study identified and analysed trust-related lexis, focusing on its frequency and dispersion, collocational patterns, and discursive alignment with key courtroom actors: the expert witness, the judge, the defendant, the prosecution and defence counsel, the jury, and the central piece of user-generated evidence. The analysis was guided by principles of CADS (Partington 2004), combining quantitative corpus methods with qualitative discourse interpretation to uncover how trust is constructed, negotiated, and attributed in deliberative talk.

To deepen the interpretive layer, the findings were mapped onto Young’s (2012) framework of narrative realms, conceptualizing courtroom actors as narrative “characters” that occupy distinct epistemic and moral positions. This alignment revealed that trust terms clustered most positively around the expert witness and the user-generated video evidence, while expressions of doubt or mistrust were more frequently associated with the defendant, and, in some cases, the legal counsel. Moreover, the jury often positioned itself as a subject of trust evaluation, revealing how jurors reflect on the legitimacy and efficacy of the jury as an institutional actor within the trial process.

These findings suggest that jurors are capable of attributing high levels of trust to authentic user-generated content and experts, even in an era of increasing concern over synthetic media. The study thus contributes to broader discussions about the epistemic authority of digital evidence, the role of narrative in legal discourse, and the interaction between lay and expert knowledge in institutional settings.

By integrating corpus tools with discourse theory and narrative analysis, this research demonstrates the value of CADS in unpacking complex social phenomena such as trust. It offers methodological insights for scholars working at the intersection of linguistics, law, and

media studies, and highlights the potential of corpus-assisted approaches to illuminate how language shapes—and is shaped by—emerging challenges in evidentiary practice.

References

Kilgarriff, Adam, and Pavel Rychlý. 2003. Sketch Engine. Lexical Computing CZ s.r.o. Accessed November 3, 2025. <https://www.sketchengine.eu..>

Partington, Alan. 2004. "From Frequency to Meaning: Using Corpora in Discourse Analysis." In *How to Use Corpora in Language Teaching*, edited by John Sinclair, 11–34. Amsterdam: John Benjamins.

Young, Katherine. 2012. *Taleworlds and Storyrealms: The Phenomenology of Narrative*. Vol. 16. Dordrecht: Springer Science & Business Media.

‘The takeaways doctors want to cancel’, ‘miracle foods’, and ‘culinary exotica’ for the ‘fearless eater’ - Representations of Chinese foodways in the UK press ‘post-COVID’

Ursula Kania and Lorna Ryan
(University of Liverpool, UK)

In May 2023, the World Health Organisation (WHO) declared that Covid-19 was no longer considered a ‘public health emergency of international concern’. However, the repercussions of this global crisis can still be felt on many levels. One socio-cultural legacy is the continued stigmatisation of (predominantly East & Southeast) Asians, triggered by the outbreak initially being associated with a ‘wet market’ in China, and exacerbated by media coverage drawing on longstanding stereotypes of (alleged) Chinese foodways (e.g., consuming bats; Kania 2023). As the 2023-24 and 2024-25 UK Home Office reports for England and Wales show, the ‘new normal’ includes well over 30,000 hate crimes targeting Asian/Asian British people per year.

Given the media’s role in reflecting, but also shaping public perceptions, this study – situated within corpus-assisted critical discourse analysis (e.g., Baker & McEnery 2015) – provides an analysis of UK press coverage of Chinese food culture ‘post-COVID’, with a particular focus on evaluating if/to which degree negative stereotypes of (alleged) Chinese foodways still persist, and if/how representations may have been affected by COVID-19.

The corpus, collected using LexisNexis, consists of 1,482 relevant articles from 12 UK broadsheets, tabloids, and regional newspapers, published in the 12-month-period following the WHO announcement (i.e., May 2023 - April 2024), totalling approx. 1.5 million words.

Quantitative and qualitative analyses are used to provide converging/complimentary evidence (keywords, collocations, using AntConc, Anthony 2024; identification of themes/discourses using NVivo; Fairclough 2010).

The results show that coverage of (British-)Chinese food culture is very multi-faceted, ranging from discussions of the popularity (but also ‘unhealthiness’) of takeaways and the promotion of alleged ‘miracle food[s]’ in traditional Chinese medicine to evaluations of (‘authentic’) Chinese food, which is still often construed as the exotic (and thus potentially inherently dangerous) other.

While the framing of Chinese cuisine as the ‘exotic other’ is nothing new (see e.g., King 2020), the current dataset demonstrates the continuing stigmatisation of (alleged) Chinese foodways ‘post-COVID’ through still drawing on (unsubstantiated) claims about the origins of COVID-19 (“Scientists in Thailand and Vietnam reckon that snake meat could replace beef, lamb and chicken. [...] Not so long ago, we learned that the Chinese in Wuhan were partial to the odd pangolin. And look how that turned out” (Richard Littlejohn, MailOnline, 18/03/2024).

Overall, the results of the study provide insights into the representation of Chinese foodways in the British press ‘post-COVID’, including the continuing use and hence normalisation of problematic stereotypes which may be (re)mobilised in harmful ways (not only) in times of crisis.

References:

Anthony, Laurence. 2024. AntConc, Version 4.2.4 Computer software.

<https://www.laurenceanthony.net/software>

Baker, Paul, and McEnery, Tony, eds. 2015. *Corpora and Discourse: Integrating Discourse and Corpora*. Palgrave.

Kania, Ursula. 2023. "They really eat anything don't they?": Pronoun use in COVID-19-related Anti-Asian racism." In *The Routledge Handbook of Pronouns*, edited by Laura L. Paterson. Routledge.

King, Michelle T. 2020. "Say no to bat fried rice: Changing the narrative of coronavirus and Chinese food." *Food And Foodways* 28(3): 237–249. doi: 10.1080/07409710.2020.1794182.

Lumivero. 2025. NVivo, Version 14.24.3 Computer Software.

<https://lumivero.com/products/nvivo/>

“AITA for not wanting to teach my child my mother tongue?”- A corpus-assisted critical discourse analysis of multilingual parenting discussions on Reddit

Ursula Kania, Sofia Lampropoulou, Annarita Magliacane and Paige Johnson
(University of Liverpool, UK)

Reddit is a social media platform with over 100,000 active communities and 57 million daily users (Reddit Press, 2021). Registered users (‘redditors’) post content in so-called ‘subreddits’ and upvote/downvote contributions made by others. Although Reddit has received some academic attention in relation to controversial communities (e.g., see Chang, 2020 on incels), subreddits containing less hostile content are still underexplored.

Taking a corpus-assisted critical discourse approach (Baker & McEnery, 2015), this study involves the analysis of 13 posts and 3,769 associated comments on multilingual parenting, taken from the AITA (=‘Am I the Asshole?’) subreddit, totalling 252,048 words. AITA is a forum for users to ask questions and discuss moral dilemmas (e.g., ‘AITA for refusing to make my bilingual daughter use an English term for my Brother?’).

The dataset was extracted using RedditExtractoR package for R (Rivera, 2022). Quantitative analyses (keywords, collocations) were combined with in-depth qualitative coding and analysis for dominant themes/discourses (Fairclough, 2010), using AntConc (Anthony, 2024) and NVivo, respectively. A particular focus was on the exploration of the stances users take vis-à-vis language learning and bilingual/multilingual identities, evaluating which (language) ideologies are reproduced and/or challenged.

We identified three prevalent themes: 1. Family/cultural heritage (e.g., “Learning [...] his other native tongue is about more than just conversing with his mother. It's about being connected to his roots.”); 2. Cognitive consequences of bi/multilingualism (e.g., “[His math teacher told] me off for speaking another language to him at home, blaming me for his “confusion” because “of course he’d be confused with all the languages flying about.”); 3. Prestige/stigma (e.g. “Learning two languages as a child is a privilege”).

Within all themes, there is a tendency for users to attribute negative evaluations of bi-/multilingualism to others (outside the subreddit), while themselves adopting a positive stance. For example, users may encourage others to pass on their heritage language(s) to their child(ren) while acknowledging the challenges posed by raciolinguistic ideologies (Rosa and Flores, 2017), within which the bilingualism of racialised speakers is viewed negatively (e.g., “Racism is still real. [...] It's very ‘cool’ when a white person is bilingual, it makes you smart. It's not cool when a minority is bilingual, they are still just a minority in America experiencing racism.”).

Overall, the results of this study contribute to and expand on existing research on social media discourse by providing a novel perspective on language ideologies and stances towards multilingualism in online discussion fora.

References

- Anthony, L. (2024). AntConc (Version 4.2.4) [Computer software]. Retrieved from <https://www.laurenceanthony.net/software>
- Baker, P. and McEnery, T. (eds.) (2015). *Corpora and Discourse: Integrating Discourse and Corpora*. London: Palgrave.
- Chang, W. (2020). "The monstrous-feminine in the incel imagination: Investigating the representation of women as 'femoids' on /r/braincels," *Feminist Media Studies*, 22(2), pp. 254–270. Available at: <https://doi.org/10.1080/14680777.2020.1804976>.
- Fairclough, N. (2010). *Critical Discourse Analysis: The Critical Study of Language*. 2nd edition. New York: Longman.
- Rivera, I. (2022). 'Package 'RedditExtractoR'. Available at: [RedditExtractoR: Reddit Data Extraction Toolkit \(r-project.org\)](https://github.com/RedditExtractoR/Reddit-Data-Extraction-Toolkit).
- Rosa, J. & Flores, N. (2017). "Unsettling race and language: Toward a raciolinguistic perspective." *Language in Society* 46, no. 5, pp. 621-647.

Objects of Feeling: A Comparative Study of 'Feel' in a Suicide and Anxiety Corpus

Charlotte-Rose Kennedy (Senso.cloud, UK)

Corrie MacMillan (University of Taipei & National Taiwan Normal University, Taiwan)

There is growing evidence for the role of anxiety disorders in suicide-related thoughts and behaviours in young people (e.g., Doering et al. 2019), and the need to improve suicide prevention among children and adolescents 'could not be higher' (Sleep et al. 2021: 4). One way in which UK schools identify the presence of anxiety and suicidal behaviours in children and adolescents is through implementing filtering and monitoring software on online school devices. Many filtering and monitoring systems use 'keyword monitoring' to track language use on devices to identify specific words or phrases (e.g. 'kill myself') that correlate with a specific form of risk (e.g. suicide) (McGlashan and Kennedy 2025). This enables schools to identify at-risk students and quickly intervene.

This study compares two corpora: a 573,471-word 'Suicide' corpus comprising posts from 'The Suicide Project' blog, and a 688,368-word 'Anxiety' corpus comprising posts from 'r/Anxiety' subreddit. The aim is to identify keywords and phrases that can be used to enhance the detection of mental health issues in filtering and monitoring software. Of particular interest is identifying points of similarity and difference between the two corpora with the goal of supporting UK schools in the earlier intervention of mental health issues in students.

To achieve this, we utilised CADS methods (Partington et al. 2013). Using Sketch Engine (Kilgarriff et al. 2014), we created a keyword list for each corpus using the enTenTen21 as the reference corpus. We focused on verbs in the keyword lists, using Halliday and Matthiessen's (2014) Systemic Functional Linguistic approach to group them by their ideational metafunction (i.e. material, mental, verbal, behavioural, relational, existential). We selected keywords denoting mental processes (e.g. 'feel') that appeared in both keyword lists. Conducting a 'word sketch' of feel found that both corpora included vague and indirect referents as grammatical objects (e.g. feel + thing/anything/nothing/something/way). Concordance analysis of collocational pairs revealed that these indirect referents were used to express the absence of emotion in the 'Suicide' corpus, but the presence of physical symptoms in the 'Anxiety' corpus.

The keywords and phrases we identified will inform Senso.cloud, a leading filtering and monitoring software, to help UK schools distinguish between forms of mental distress and provide students with timely and appropriate support. Incorporating the less explicit keywords into the software will help ensure that subtle and nuanced expressions of anxiety and suicidal ideation do not slip through the net, improving student safety.

References

- Doering, S., Lichtenstein, P., Gillberg, Ntr, C., Middeldorp, C.M., Bartels, M. and Lundstrom, S. (2019). Anxiety at age 15 predicts psychiatric diagnoses and suicidal ideation in late adolescence and young adulthood: results from two longitudinal studies. *BMC Psychiatry*, 19(1), 1-11.
- Halliday, M.A.K., Matthiessen, C.M.I.M. (2014). *Halliday's Introduction to Functional Grammar*, 4th edn. Routledge.

- McGlashan, M. and Kennedy, C-R. (2025). Corpus linguistics for safeguarding children online. *Applied Corpus Linguistics*, 5(3), 1-9.
- Partington, A., A. Duguid and C. Taylor. 2013. *Patterns and meanings in discourse: theory and practice in corpus-assisted discourse studies (CADS)*. Amsterdam: John Benjamins.
- Sleap, V., Williams, T., Stoianova, S., Odd, D., Gunnell, D., Chitsabesan, P., Irani, T., Rodway, C., Skelton, S., Tranter, S., King, A., McClymont, C., Fonagy, P., & Luyt. (2021). *Suicide in Children and Young People: National Child Mortality Database Programme Thematic Report—Data from April 2019 to March 2020*. Healthcare Quality Improvement Partnership (HQIP).
- Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý and V. Suchomel. 2014. "The Sketch Engine: ten years on." *Lexicography*, 1, 7-36.

Place as discourse: Employing Corpus-Assisted Discourse Studies (CADS) to study spatiality in texts

E. Dimitris Kitis (Xi'an Jiaotong-Liverpool University, China)

Over the past few decades, spatiality has emerged as a key concept in literary and cultural studies, with scholars examining how space and place are discursively constructed, represented, and imagined. This development has prompted increasing methodological interest in how textual analysis can engage with geographical imaginaries. One notable contribution is the field of Geographical Text Analysis (GTA) (Donaldson et al., 2017; McEnery et al., 2020; Paterson & Gregory, 2019; Porter et al., 2015), which applies corpus methods to uncover spatial patterns in large text collections.

Building on this trajectory, the present study introduces psychoscapes as a conceptual tool for analysing the psychological and discursive construction of urban spaces. In particular, I propose extending GTA by combining methods from Cognitive Linguistic Critical Discourse Studies (CL-CDS) and Corpus-Assisted Discourse Studies (CADS) to reveal the affective imaginaries and mental representations that shape perceptions of place. The approach is illustrated through examining media representations of two Johannesburg neighbourhoods, Hillbrow and Sandton, in a 141-million-word corpus of South African broadsheets. Using collocation analysis (Baker, 2023) alongside schema-based PoV modelling—angle, distance, and anchor (Hart, 2015), which construe different points of view (Langacker, 2008)—and conceptual blending producing evaluative metaphor (Fauconnier & Turner, 2002), the analysis reveals stark contrasts: Hillbrow is predominantly framed through crime, decay, and poverty, whereas Sandton is cast through affluence and consumerism, reinforcing its identity as a hub of luxury and financial power.

Foregrounding psychoscapes shows how corpus-based methods enrich the study of spatiality in texts, revealing fine-grained practices of micro-spection and micro-sensation that contribute to place-making and illuminate related phenomena such as socio-economic inequality.

References:

- Baker, P. (2023). *Using Corpora in Discourse Analysis* (2nd Edition). London: Bloomsbury.
- Donaldson, C., Gregory, I. N., & Taylor, J. E. (2017). Locating the beautiful, picturesque, sublime and majestic: Spatially analysing the application of aesthetic terminology in descriptions of the English Lake District. *Journal of Historical Geography*, 56, 43–60.
- Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.
- Hart, C. (2015). Viewpoint in linguistic discourse: Space and evaluation in news reports of political protests. *Critical Discourse Studies*, 12(3), 238–260.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- McEnery, T., Baker, H., & Dayrell, C. (2020). Working at the interface of hydrology and corpus linguistics: Using corpora to identify droughts in nineteenth-century Britain. In J. Egbert

& P. Baker (Eds.), *Using corpus methods to triangulate linguistic analysis* (pp. 52–84). London: Routledge.

Paterson, L. L., & Gregory, I. N. (2019). *Representations of poverty and place: Using geographical text analysis to understand discourse*. Cham, Switzerland: Palgrave Macmillan.

Porter, C., Atkinson, P., & Gregory, I. (2015). Geographical Text Analysis: A new approach to understanding nineteenth-century mortality. *Health & Place*, 36, 25–34.

Corporate censorship online: Vagueness and discursive imprecision in YouTube's advertiser-friendly content guidelines

Susanne Kopf (WU Vienna, Austria)

This paper deals with YouTube's advertiser-friendly content guidelines – the rules defining what YouTube deems advertiser (un)friendly and that YouTube creators seeking to monetise their content through advertising have to follow. Specifically, this study addresses the textual composition of YouTube's regulations with a focus on their specificity and discursive precision.

Adopting a corpus-assisted discourse analytical approach, the data comprises a corpus that encompasses YouTube's ad-friendly content policies and related explanatory material published on YouTube's 'Help' site (a total of 14,803 tokens). In the corpus, the term 'content' is the most frequent lexical item referring to creators' uploaded material, ranking fifth overall. Consequently, 'content' serves as the focal point of this analysis. In addition to examining the term's collocational profile (t-score, span of four), the study qualitatively investigates all 351 concordance lines of 'content' in their co-text. For this, I take a wide-angle view "equivalent to text extracts" (Partington et al. 2013, 18), examining at least forty tokens left and right of the node. This ensures a comprehensive understanding of how YouTube's discourse constructs notions of ad-(un)friendliness and allows for the identification of different forms of vagueness.

Findings suggest that at least 26% of the concordance lines containing 'content' in YouTube's ad-friendly guidelines exhibit at least one of eight forms of vagueness. Examples include indetermination in the representation of social actors (e.g. "some advertisers") (Van Leeuwen 1996, 51), vague quantification ("in most cases") (Drave 2002, 26) and the use of gradable adjectives without frames of reference ("strong") (Égré and Klinedinst 2011, 12). While such vagueness markers are readily identifiable on the text surface through concordance analysis and thus lend themselves to reliable quantification, the texts exhibit an additional recurrent form of vagueness that operates beneath the surface and resists straightforward quantification. Specifically, in the co-text surrounding 'content', YouTube frequently employs lexis that requires a "shared cultural frame of reference" (Evison et al. 2007, 139). For instance, interpreting what YouTube deems "controversial" varies dependent on cultural and local context.

Overall, the various forms of vagueness identified in the corpus leave content creators in the dark about the monetisability of specific material. Therefore, they may avoid pushing any boundaries in favour of producing noncontroversial and content that is safely monetisable. In doing so, monetising content creators may inadvertently contribute to reducing content plurality on YouTube.

References

Drave, N. "Vaguely Speaking: A Corpus Approach to Vague Language in Intercultural Conversations." In *New Frontiers of Corpus Research: Papers from the Twenty First International Conference on English Language Research on Computerized Corpora Sydney 2000 (Language and Computers)*, ed by P Peters, P Collins, and A Smith Cohen, vol. 36, 25–40. Brill, 2002.

Égré, P, and N Klinedinst. "Introduction: Vagueness and Language Use." In *Vagueness and Language Use* (Palgrave Studies in Pragmatics, Language and Cognition), ed by N Klinedinst and P Égré, 1–24. Palgrave, 2011.

Evison, J, M McCarthy, and A O’Keeffe. "Looking Out for Love and All the Rest of It’: Vague Category Markers as Shared Social Space." In *Vague Language Explored*, ed by J Cutting, 138–57. Palgrave, 2007.

Partington, A, A Duguid, and C Taylor. *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)* (Studies in Corpus Linguistics), vol. 55. Benjamins, 2013.

Van Leeuwen, T. "The Representation of Social Actors." In *Texts and Practices: Readings in Critical Discourse Analysis*, ed by C R Caldas-Coulthard and M Coulthard, 32–70. Routledge, 1996.

Discourses of democracy in the public domain: A comparative study of US and German news

Marc Kupietz (IDS Mannheim, Germany)

Tim Feldmüller (IDS Mannheim, Germany)

Tony Berber Sardinha (Pontifical Catholic University of Sao Paulo, Brazil)

Maria Claudia Nunes Delfino (Sao Paulo State Technology College, Brazil)

A portrayal of democracy as being in danger and in need of protection has gained prominence in public discourse (Levitsky and Ziblatt 2018). To explore the extent to which this and other discourses of democracy are pervasive, we conducted a corpus study of news texts from two major Western democracies, the United States and Germany. Previous computational work on the discourses of democracy relied on existing categories from political theory to determine the major representations (Dahlberg, Axelsson, and Holmberg 2020). In contrast, we employed an entirely bottom-up approach using two methods: Lexical Multi-Dimensional Analysis (Berber Sardinha and Fitzsimmons-Doolan 2025) and distributional language modeling (Fankhauser and Kupietz 2017). The paper therefore has two goals: a descriptive goal of identifying the major discourses, and a methodological goal of comparing corpus-driven methods for discourse identification. Our corpus comprises ca. 38 million words of news texts drawn from the NOW Corpus (for English) and DeReKo (for German). For the LMDA, we used keyword extraction to detect the salient lemmas across texts and extracted factors based on these lemmas (through Factor Analysis), which were interpreted as discourse dimensions. For distributional language modeling, we first focused on individual lexical items and used aligned word-embedding models (Fankhauser and Kupietz 2017) to identify expressions whose meanings in the democracy discourse diverge from those observed in the unfiltered reference corpora. Analyses of nearest neighbors in vector space and of collocates provided insights into the semantics of specific expressions. Second, we applied k-means clustering to the embeddings to derive word clusters to reveal discursive macro-structures (Feldmüller 2025). The LMDA revealed six dimensions for English and seven for German. To illustrate, English Dim. 1 centers on discourses in which democratic legitimacy depends on institutionally sanctioned journalism (pos. pole) and those that present democracy as a threatened procedural system (neg. pole). English Dim. 2 in turn distinguishes discourses of democracy as an institutional system that can deteriorate (pos. pole) from discourses in which democracy is employed as a slogan in subscription appeals (neg. pole). Meanwhile, German Dim. 1 includes a discourse that constructs democracy as a civilizational value (pos. pole) and a discourse that treats democracy as a governance mechanism that must be defended (neg. pole). And German Dim. 2 depicts democracy as being under threat by right-wing extremism (pos. pole) and a discourse that treats democracy as a governable technology (neg. pole). We compared the results of the two methods and will report on the findings in the paper presentation. In sum, the results confirm that a democracy-under-threat macro-discourse is indeed present in news, alongside other discourses in which democracy is construed as a form of heritage, an ideal whose loss would imply cultural regression.

References

- Berber Sardinha, Tony, and Shannon Fitzsimmons-Doolan. 2025. *Lexical Multidimensional Analysis: Identifying Discourses and Ideologies*. Cambridge: Cambridge University Press.
- Dahlberg, Stefan, Sofia Axelsson, and Sören Holmberg. 2020. "Democracy in context: using a distributional semantic model to study differences in the usage of democracy across languages and countries." *Zeitschrift für Vergleichende Politikwissenschaft* 14 (4): 425-459.
- Fankhauser, Peter, and Marc Kupietz. 2017. "Visualizing language change in a corpus of contemporary German." *Corpus Linguistics International Conference 2017*, Birmingham.
- Feldmüller, Tim. 2025. *Das Framing von Extremismusvarianten im medialen Diskurs der Jahre 1999–2021: Eine Corpus-driven Methode zur Erschließung und Visualisierung Semantischer Frames*. Vol. 2. Language Science Press.
- Levitsky, Steven, and Daniel Ziblatt. 2018. *How democracies die*. New York: Crown.

The discourse of the online place brand image of Hong Kong in an online travel forum

Phoenix Lam (The Hong Kong Polytechnic University, Hong Kong)

As a concept with its roots in marketing, place branding refers to long-term, holistic and systematic attempts to establish a location as unique and distinguishable from others. While place branding is essentially communication-based and discursive in nature, surprisingly little contribution has been made from linguistics and discourse studies amid the growing academic interdisciplinary interest in the field. Focusing on the key notion of place brand image which can be broadly summed up as the collective perceptions of a place brand, the present study examines how the place brand image of Hong Kong is discursively constructed in an online travel forum. Specifically, this corpus-assisted discourse analysis applies Sinclair's (1996) unit of meaning theory to investigate the lexical, grammatical, semantic and attitudinal patterns associated with the place brand "Hong Kong". Based on the collocational and colligational analysis of thousands of forum posts, the study shows that the brand image of the city is predominantly constructed through a discourse of movement, both internally and externally, across space and time. Further concordance analysis of the thematic construction of Hong Kong reveals mixed perceptions of the city covering both functional and experiential dimensions (Hankinson 2005), resulting in a multi-faceted brand image. The study thus demonstrates how a corpus-based discourse-oriented analysis can inform place branding research by highlighting not only the attributes of brand image of a place but also how such attributes are discursively realized in naturally-occurring digitally-mediated social interactions.

References

- Hankinson, Graham. 2005. "Destination Brand Images: A Business Tourism Perspective." *Journal of Services Marketing* 19 (1):24–32.
- Sinclair, John. 1996. "The Search for Units of Meaning." *Textus* IX: 75–106.

Representations of Ill Health in best-selling Young Adult fiction

Maria Leedham (The Open University, UK)

Sarah Jane Mukherjee (The Open University, UK)

Sally Hunt (Rhodes University, South Africa)

Stories are an important means for young people to shape their understandings of the world and their place within it. Books often serve as resources for exploring complex issues, including representations of ill health. This paper argues that Young Adult (YA) fiction enables adolescent readers to conceptualise illness, observe coping strategies, and potentially build resilience.

Drawing on our recently-compiled corpus of the 50 most commercially-successful YA books in the UK (2017-2022), this paper presents a corpus-assisted discourse analysis of representations of mental and physical ill health in YA fiction. We compare two sub-corpora: books featuring ill-health themes (n=10) and books without (n=40) to examine how illness and coping strategies are portrayed. Analyses included semantic categorisation (using WMatrix), and keyword analysis (AntConc). Expected USAS categories of Health, Medicines & medical treatment and Cleaning & personal care in the YA-Health Corpus were prominent, but semantic analysis also revealed an emphasis on Relationship: intimacy & sex and Size.

We hypothesize that health-related narratives foreground relationships as characters reassess their priorities, while the size category signals greater introspection and existential reflection (e.g. The universe is seeming really huge right now - *We Were Liars*; Letting myself get smaller and smaller - *Girl In Pieces*).

Keyword analysis highlights mental processes - thinking and feeling - as particularly salient. Nine out of ten YA-Health books use first-person narration, enabling deeper introspection and witness accounts of problems faced by characters (e.g. I feel like I'm living for these moments - *All the Bright Places*; I'm not sure why I told you the truth - *Thirteen Reasons Why*). Our paper thus demonstrates how the language of illness woven into stories serves to construct particular ways of being and coping for both mental and physical ill-health.

Our interpretation of linguistic patterns and their meanings is enriched by insights from focus groups with secondary school students (n=18). We argue that the language of illness and resilience in young people's reading repertoires is likely to have an impact on their understandings of their own lives and the ways in which they construct their own responses to ill health.

Moral panic over effeminacy: the representations of niangpao in Chinese newspapers

Run Li (Lancaster University, UK)

The study, using corpus-assisted critical discourse analysis, investigates the changing representations of ""niangpao (娘炮, niáng pào)""; a term often associated with pejorative connotations and used to describe individuals perceived as having feminine traits and lacking traditional masculinity, in Chinese-language newspapers from 2000 to 2023. The study identifies and analyses shifts in newspapers discourse surrounding ""niangpao"". The analysis, aided by the examination of keywords, collocates and concordances, focuses on five phases, each marked by significant changes in societal attitudes and regulatory responses. The findings reveal an initial emergence of ""niangpao"" through popular media, followed by a period of intense public scrutiny and moral panic, leading to educational and regulatory interventions aimed at curbing its visibility and influence. Particularly, education plays the role as a societal control mechanism in three phases examined, with an emphasis on gender-difference education and the promotion of traditional masculinity in youth. In addition, in recent years, the government's escalation in control measures, including official crackdowns on ""niangpao"" in media and entertainment industry, reflects broader societal anxieties about non-traditional gender expressions in Chinese society.

Multimodal Discourses of Voluntary Childlessness on Chinese Social Media: A Critical Study

Tianxin Li (Lancaster University, UK)

Since the official implementation of China's three-child policy in 2021, public discussions surrounding fertility intentions have intensified considerably (Zhu 2022; Jiang 2023), while studies on how voluntary childlessness is expressed, negotiated and circulated as a discourse practice in Chinese society remain limited. Based on this background, this study, grounded in Critical Discourse Studies (CDS), investigates both the reasons individuals give for voluntary childlessness and the discursive strategies they use to justify and negotiate these positions.

Empirically, the study uses the search terms “不生孩子(Not having children)”, “丁克 (DINK)”, “不想生(Not wanting to have children)”, and “不婚不育(No marriage, no child)” to collect relevant Weibo posts dating from May 2021 to May 2024, constructing a self-built corpus of 1,409,671 words, which was analysed using Sketch Engine (Kilgarriff et al. 2014). The analysis adopts corpus-based approaches, specifically keyword and frequency analyses (Baker 2023; Gillings et al. 2023), followed by concordance analysis informed by the appraisal framework in systemic functional linguistics (Halliday and Matthiessen 2014) and van Leeuwen's (2007) legitimation strategies, as well as visual grammar (Kress and van Leeuwen 2021).

The findings show that the cost of raising children is one of the most prominent reasons for voluntary childlessness. Lifestyle-related expressions also appear extensively, including preserving youth and maintaining an unambitious lifestyle. Moreover, time-related expressions are widely used to frame the present era as enabling childlessness (e.g., greater inclusiveness, gender equality, expanded female autonomy). Yet gendered beauty norms remain, as concerns about body shape and ageing continue to influence women's decisions.

Different reasons activate different discursive strategies. Celebrity references constitute a key means of constructing role-model authority; economic reasons typically employ rationalisation strategies; and reasons centred on quality of life often use mythopoesis and relational processes to frame childlessness as a desirable lifestyle. Furthermore, pronatalist discourses and counter discourses co-occur extensively and mirror, as both draw on social sanction (propriety) and rely on moral evaluation as a key legitimation strategy.

Multimodally, images and text co-construct meanings in two primary ways: (1) the image amplifies and makes more salient the emotional meanings conveyed in the text; and (2) the image provides lifestyle imaginaries that are not explicitly articulated in the text, making “not having children” not only understandable but also imaginable.

Overall, the study fills a gap in research on voluntary childlessness in Chinese social-media discourse and offers new insights into subjectivity, value negotiation and resistant discourse in the digital era.

References

Baker, Paul. *Using Corpora in Discourse Analysis*. 2nd ed. London: Bloomsbury Academic, 2023.

Gillings, Mathew, Gerlinde Mautner, and Paul Baker. *Corpus-Assisted Discourse Studies*. Cambridge Elements in Corpus Linguistics. Cambridge: Cambridge University Press, 2023. <https://doi.org/10.1017/9781009168144>.

Halliday, M. A. K., and Christian M. I. M. Matthiessen. *Halliday's Introduction to Functional Grammar*. 4th ed. London: Routledge, 2014.

Jiang, Quanbao, Chujun Zhang, Yunyun Zhuang, Yuyin Jiang, and Xiaoying Zhang. "Rising Trend of Childlessness in China: Analysis of Social and Regional Disparities with 2010 and 2020 Census Data." *BMJ Open* 13, no. 5 (2023): e070553.

Kilgarriff, Adam, Vít Baisa, and Jan Bušta. "The Sketch Engine: Ten Years On." *Lexicography ASIALEX* 1 (2014): 7–36. <https://doi.org/10.1007/s40607-014-0009-9>.

Kress, Gunther R., and Theo van Leeuwen. *Reading Images: The Grammar of Visual Design*. 3rd ed. London: Routledge, 2021.

van Leeuwen, Theo. "Legitimation in Discourse and Communication." *Discourse & Communication* 1, no. 1 (2007): 91–112. <https://doi.org/10.1>

The mistress, the sub, his hotwife and her lover: stock characters in erotic fiction

Alon Lischinsky (Oxford Brookes University, UK)

Kat Gupta (Royal Holloway University of London, UK)

In this paper, we offer an initial but systematic exploration of the types of characters found in erotic writing, their typical characterisation and their co-occurrence.

Stylisticians have used a variety of approaches (cf. McIntyre 2014) to gain insight into fictional characters (Stockwell & Mahlberg 2015; Evans 2023). Despite its fundamental role shaping the way people understand sexual desire, pornography remains understudied: characterisation is not often seen as an important dimension of porn, and has even been claimed to be detrimental to its enjoyment (Driscoll, 2006, p. 91).

To analyse the interpersonal relationships and social structures that characterise erotic writing, we examine the representation of the individuals in our corpus. Here we specifically explore the choice and linguistic context of common nouns with personal reference, one of the most direct indices of characterisation (cf. Čermáková & Mahlberg, 2021).

We analysed the top 1000 nouns unambiguously referring to people and groups of people (eg, boy, queens, guitarist). We identified common features across these terms through unsupervised annotation of semantic features (Moore & Rayson, 2022) and separately classified their functional role as social actors (van Leeuwen, 1996). Finally, we extracted the characteristic linguistic context of each term by computing its significant collocates, and used the Louvain method (Blondel et al., 2008) to detect communities of terms connected through shared collocates.

Results show that, unlike most other contemporary media genres, female characters are no less prominent in erotic fiction than male ones. Character and relationship types are drawn from a highly stylised set that features prominently relations of power, both purely sexual (sir, mistress, slave) and embedded in social relationships (doctor, boss, teacher). Characterisation emerges as a site in which power and gender are hotly contested; we find that this contestation and subversion is itself eroticised.

References

- Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 'Fast Unfolding of Communities in Large Networks'. *Journal of Statistical Mechanics: Theory and Experiment* 2008, no. 10 (2008): P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
- Čermáková, Anna, and Michaela Mahlberg. 'The Representation of Mothers and the Gendered Social Structure of Nineteenth-Century Children's Literature'. *English Text Construction* 14, no. 2 (2021): 119–49. <https://doi.org/10.1075/etc.00044.cer>.
- Driscoll, Catherine. 'One True Pairing: The Romance of Pornography and the Pornography of Romance'. In *Fan Fiction and Fan Communities in the Age of the Internet: New Essays*, edited by Karen Hellekson and Kristina Busse. McFarland & Company, 2006.

- Evans, Mel. 'Interjections and Individual Style: A Study of Restoration Dramatic Language'. *Language and Literature: International Journal of Stylistics* 32, no. 3 (2023): 297–328. <https://doi.org/10.1177/09639470231158695>.
- Leeuwen, Theo van. 'The Representation of Social Actors'. In *Texts and Practices: Readings in Critical Discourse Analysis*, edited by Carmen Rosa Caldas-Coulthard and Malcom Coulthard. Taylor & Francis, 1996.
- McIntyre, Dan. 'Characterisation'. In *The Cambridge Handbook of Stylistics*, edited by Peter Stockwell and Sara Whiteley. Cambridge Handbooks in Language and Linguistics. Cambridge University Press, 2014. <https://eprints.hud.ac.uk/id/eprint/20400/>, <https://www.hud.ac.uk/news/>.
- Moore, Andrew, and Paul Rayson. PyMUSAS: Python Multilingual UCREL Semantic Analysis System. V. 0.3.0. Lancaster University, released 4 April 2022. <https://ucrel.github.io/pymusas/>.
- Stockwell, Peter, and Michaela Mahlberg. 'Mind-Modelling with Corpus Stylistics in David Copperfield'. *Language and Literature: International Journal of Stylistics* 24, no. 2 (2015): 129–47. <https://doi.org/10.1177/0963947015576168>.

“5 skin cancer-care tools you should look out for”: A pragmalinguistic analysis of clickbaiting practices in digital scientific dissemination”

Rosa Lorés (Universidad de Zaragoza, Spain)

Within the context of digital science dissemination, knowledge brokers are reconsidering the way they “narrate” science, the way they project themselves as credible voices, and the relationship they establish with their audiences. As a result, new practices seem to be emerging to foster reader engagement, one of them being clickbaiting, which is designed to keep readers on the page for as long as possible and is typically associated with sensationalist discourse. Clickbaiting (Bloom and Hansen 2015; Scott 2023) has been widely investigated in online news headline contexts, frequently from a natural language processing approach or from the social perspective of reception studies. However, no research appears to have been conducted so far on digital science dissemination from a discourse-pragmatic, corpus-assisted perspective, as proposed in the present study.

Adopting a CADS approach to the analysis of clickbaiting in digital science dissemination, I constructed a corpus of headlines based on various practices retrieved from the database SciDis (Pascual and Sancho-Ortiz 2024): posts from The Conversation, feature articles from Popular Science and the Smithsonian Magazine and research digests from Science Daily. A collection of 630 headlines on topics related to health, natural sciences, and economy was evenly gathered, and the resulting corpus was subjected to quantitative and qualitative analyses supported by AntConc and NVivo. Two aspects were analysed: (i) headline structure (questions, listicles, the “you won’t believe” pattern, and informative ellipsis (use of suspensive dots or dashes)), and ii) pragmatic markers (deixis (the cataphoric this, and temporal and spatial deixis now and here)), hedging, evaluative language, emotional triggers, and direct appeals to the reader (i.e. directives and personal pronoun you).

Results reveal different degrees of use of clickbaiting tactics in digital science dissemination practices. Research digests, which popularise a single RA, show the lowest presence of these practices, with the use of question structures, high levels of hedging (as in academic discourse), and headlines that become summaries of the content provided in the text. Feature articles display a wider variety of structures, including questions, listicles, a more frequent use of cataphoric deixis (this/ these), and evaluative language, as well as other engagement markers such as the inclusive we. Finally, the platform The Conversation shows the highest presence and the widest range of clickbaiting tactics, including questions, listicles, ellipsis, and emotional triggers.

What can be inferred, then, is that clickbaiting practices occur to varying degrees depending on the type of digital text, and that several factors seem to account for this variation, one of them being the degree to which the digital text depends on the original source. Other practices pertaining to clickbaiting strategies are found to be absent from digital scientific dissemination (i.e., hype and hyperbolic language). This points towards a high level of interdiscursivity, which exploits the blend of discourses (scientific and journalistic / sensationalist) to maximise audience engagement, which might challenge the balance between ethos (credibility) and pathos (appeal to emotions) in digital science dissemination.

References

- Blom, Jonas Nygaard, and Kenneth Reinecke Hansen. 2015. "Click Bait: Forward-Reference as Lure in Online News Headlines." *Journal of Pragmatics* 76: 87–100. <https://doi.org/10.1016/j.pragma.2014.11.010>.
- Pascual, Daniel & Sancho-Ortiz, Ana Eugenia. (2024). Investigating recontextualisation processes in scientific digital practices: The SciDis Database. *Revista Electrónica de Lingüística Aplicada* 23, 101-118. <https://doi.org/10.58859/rael.v23i1.649>
- Scott, Kate. 2023. "Deceptive' Clickbait Headlines: Relevance, Intentions, and Lies." *Journal of Pragmatics* 218: 71-82. <https://doi.org/10.1016/j.pragma.2023.10.004>

Plural Identities and the Hermeneutics of Suspicion: A Corpus Assisted Discourse Analysis

Andrew Lustig (University of Toronto, Canada)

Oliver Harrison (University of Toronto, Canada)

Mark McGlashan (University of Liverpool, UK)

Gavin Brookes (Lancaster University, UK)

Sophie Simic-Lustig (Hudson College, Canada)

Benoit Mulsant (University of Toronto, Canada)

Online communities devoted to plural identity, where individuals experience multiple selves or “parts” within a single body, provide a unique window into how legitimacy and authenticity are linguistically negotiated. This paper presents a comparative corpus-assisted discourse analysis of four Reddit subreddits: r/DID, r/OSDD, r/tulpas, and r/plural. These spaces represent trauma-based, volitional, and inclusive understandings of plurality. Using a 1.3-million-word dataset comprising the top 200 most upvoted threads from each community, we identify “complement keywords,” or lexical types distinctive to one community and absent from the others. These are grouped into five recurrent discursive moves: authority-claiming, suspicion and gatekeeping, validation, boundary-realignment, and world-building or practice.

The analysis reveals divergent hermeneutic orientations across communities. r/DID anchors legitimacy in psychiatric authority while cultivating internal validation. r/OSDD discourse is saturated with suspicion, reflecting testimonial injustice and fragile legitimacy. r/tulpas constructs creative legitimacy through imaginative practice and aesthetic validation. r/plural redefines community boundaries through inclusivity and endogenic recognition. Together, these findings demonstrate how suspicion and empathy circulate unevenly across plural publics, generating distinct epistemic and affective economies.

Framed by Ricoeur’s hermeneutics of suspicion and empathy, Fricker’s epistemic injustice, and Warner’s concept of counterpublics, the study shows how platform affordances and moderation practices mediate recognition and doubt. By situating plural subreddits as counterpublics negotiating psychiatric authority and antagonistic publics such as r/DIDcringe, the paper advances understanding of how corpus-assisted discourse studies can illuminate the digital production of legitimacy, belief, and care.

“No hidden costs or sneaky fees”. Exploring the digital communication of financial institutions and payday lenders

Ursula Lutzky (WU Vienna University of Economics and Business, Austria)

The current cost-of-living crisis poses a significant threat to consumers’ financial well-being. Given the central importance of finance and money management to people’s day-to-day lives and recent financial pressures after the Covid-19 pandemic, consumer debt has attracted increased scholarly and public interest. Previous research, for example, explored the discourse of pawnbroking apps (Brookes and Harvey 2017), studied the representation of the cost-of-living crisis in UK newspapers (Parnell 2024), and investigated online discussions of consumer debt (Lawson et al. 2023). Few studies have, however, researched how financial institutions discuss lending, debt and the cost-of-living crisis through their own digital outlets. This paper addresses this gap by studying the digital communication of established financial institutions and payday lenders in the UK, exploring how they communicate with stakeholders through their websites.

It is based on the Financial Institutions and Payday Lenders Corpus (FINPAL) which comprises a total of 970,869 words. The corpus consists of website texts on the topics of loans, mortgages and financial advice from established financial institutions and payday lenders. For the former, these include UK banks such as Barclays and NatWest, governmental institutions such as UK Debt Advice, and charities such as StepChange (467,278 words). The payday lenders include companies such as Drafty, LoanPig and QuidMarket (503,591 words).

The FINPAL corpus is explored to uncover how established financial institutions and payday lenders communicate with their stakeholders when sharing advice and offering financial support. While a keyword analysis will reveal differences between the discourse of established financial institutions and payday lenders, these findings will then be explored further through the study of collocation and concordances. By combining different corpus linguistic methods, this study aims to provide new insights into the intricacies of digital financial discourse and offer new understanding of how debt and lending are discursively constructed online. In addition, its findings may inform tailored stakeholder recommendations with a view to improving their financial literacy and protecting the public from threats to their financial well-being.

References

- Brookes, Gavin, and Kevin Harvey. 2017. “The Discourse of Alternative Credit: A Multimodal Critical Examination of the *Cash Converters* Mobile App.” In *The Language of Money and Debt*, edited by Annabelle Mooney and Evi Sifaki. Palgrave Macmillan.
- Lawson, Robert, Ursula Lutzky, Andrew Kehoe, Matt Gee. 2023. “‘Sorry to hear you’re going through a difficult time’: Investigating Online Discussions of Consumer Debt.” *Applied Corpus Linguistics* 3(2), 100056. <https://doi.org/10.1016/j.acorp.2023.100056>
- Parnell, Tamsin. 2024. “‘Homeless’ and the Cost-of-Living Crisis in *The Guardian* and *MailOnline*: A Corpus-Assisted Analysis.” In *Critical Approaches to Polycrisis. Discourses of Conflict, Migration, Risk, and Climate*, edited by Tamsin Parnell, Tom Van Hout and Dario Del Fante. Springer.

“Il Mipaaft arriva su Instagram!” From food sovereignty to social media: An analysis of digital gastronationalism

Stefania M. Maci (University of Bergamo, Italy)

Gastronationalism is the phenomenon where nations use food to express national identity and differentiation, often supported by national policies such as geographical indications that protect certain foods legally and culturally. In this way, national identity and nationalist sentiments are reinforced, while influencing food markets (DeSoucey, 2010), enhancing national pride in traditional cuisine, and linking it closely to populist and nationalist political movements (Ichijo, 2020). In this regard, gastronationalism involves not just the preservation of tradition but also political exploitation of food traditions as a symbol of national pride and identity (Barca, 2024). As shown by Farina (2021), this supports domestic and international political strategy: UNESCO recognition of Italian food as intangible cultural heritage bolsters national identity and enhances the nation’s international image.

In Italy, the change of name of the Ministry of Agricultural, Food and Forestry Policies to ‘Ministry of Agriculture, Food Sovereignty and Forestry’ (2022) is in itself a clear example of institutional gastronationalism: it implies a strong reaffirmation of national sovereignty in the food sector against global and supranational dynamics. In the perspective of banal nationalism (Billig, 1995; Ichijo, 2020), everyday practices represented on social media, such as sharing recipes, images of typical products or even anchor-people, contribute to naturalising national identity.

Drawing on CDA (Wodak, 2021) and multimodality (van Leeuwen, 2014), this corpus-based contribution offers an analysis of the Ministry of Agriculture, Food Sovereignty and Forestry communication via Instagram from the new Minister’s mandate date (10 October 2022) to the date of my writing (613 posts altogether). The aim is to show the political implications, critical issues and contradictions, such as the risk of an excessive idealisation of food tradition that ignores contemporary multicultural reality or problems of sustainability and food safety (Ichijo, 2020; Farina, 2021).

References

- Barca, F. (2024). *Gastronationalisme: Derrière la fierté de la cuisine traditionnelle*. VoxEurope English. <https://voxeurop.eu/en/gastronationalism-cuisine-food-sovereignty/>
- Billig, M. (1995). *Banal Nationalism*. Sage.
- DeSoucey, M. (2010). *Gastronationalism*. *American Sociological Review*, 75(3), 432–455. <https://journals-sagepub-com.ezproxy.unibg.it/doi/10.1177/0003122410372226>
- Farina, F. (2021). *The politics of washoku: Japan’s gastronationalism and gastrodiplomacy*. In M. Castorina & D. Cucinelli (Eds.), *Studi e saggi*. Firenze University Press. <https://doi.org/10.36253/978-88-5518-506-6.09>
- Ichijo, A. (2020). *Food and Nationalism: Gastronationalism Revisited*. *Nationalities Papers*, 48(2), 215–223. <https://doi.org/10.1017/nps.2019.104>

- Martin, J. R., & White, P. R. R. (2005). Attitude: Ways of Feeling. In J. R. Martin & P. R. R. White (Eds.), *The Language of Evaluation: Appraisal in English* (pp. 42–91). Palgrave Macmillan UK. https://doi.org/10.1057/9780230511910_2
- van Leeuwen, T. (2014). Critical Discourse Analysis and Multimodality. In C. Hart & P. Cap (Eds.), *Contemporary Critical Discourse Studies* (pp. 281–297). Bloomsbury Academic.
- Wodak, R. (2021). *The Politics of Fear: The Shameless Normalization of Far-Right Discourse*. SAGE.

I feel like: Analyzing a key multi-word unit in an online anxiety support group context

Corrie MacMillan (University of Taipei & National Taiwan Normal University, Taiwan)

Alvin Cheng-Hsien Chen (National Taiwan Normal University, Taiwan)

Building upon the work of Collins and Baker (2023), this research examines the lived experiences of anxiety shared in the online support group r/Anxiety. The expression of feelings and emotions is a prevalent theme in mental health discourse, and the lemma ‘feel’ has been identified as central within anxiety discourse (Collins and Baker 2023). Focusing on the multi-word unit ‘I feel like’ extends this work by capturing first-person narration and offering insight into affective stance (Kiesling 2022).

The study draws on a 688,368-word corpus of posts from r/Anxiety and a Brown reference corpus. Sentence-bounded 3-grams were generated for both corpora, and multi-word units were treated as single tokens. Keyness was computed using KLDsigned (Gries 2024), a directional divergence measure comparing the distribution of each item across the two corpora. KLDsigned was computed for tri-grams present in both corpora. The results showed that ‘I feel like’ had the highest KLDsigned value, 0.0086. This motivated the selection of the target multi-word unit for concordance and collocational analysis.

Treating ‘I feel like’ as a single token enabled a precise examination of its co-text. The expression shows a clear asymmetry between its left and right contexts. The right-hand environment contributes most to meaning-making, clustering around somatic experiences (e.g. breathe, pass out), catastrophic interpretations (e.g. dying, going insane), and negatively evaluative self-descriptions (e.g. a failure, a burden). These patterns show how posters use the expression to report bodily sensations, escalate fears, and appraise themselves. The left-hand context, by contrast, performs more pragmatic work. Items such as ‘maybe’, ‘I guess’, ‘recently’, and ‘but’ signal uncertainty, introduce contrast, or set up the circumstances in which the feeling is expressed. Taken together, the patterns indicate that ‘I feel like’ functions as a key point in the sentence where posters shift from framing their situation to describing the experience of anxiety itself.

Rather than analysing a key word in context, this study adopts a key multi-word unit in context perspective, treating I feel like as the analytic node around which discourse patterns are examined. By extending the logic of word-based context analysis to multi-word units and modelling their left- and right-hand environments separately, the study demonstrates how such units function as stable discourse anchors in an online anxiety support community. Rather than analysing a key word in context, this study adopts a key multi-word unit in context perspective, treating ‘I feel like’ as the analytic node around which discourse patterns are examined. By extending the logic of word-based context analysis to multi-word units and modelling their left- and right-hand environments separately, the analysis demonstrates how such units function as stable discourse anchors in an online anxiety support community. This approach foregrounds the role of multi-word units in organising affective stance, showing how posters move from situational framing to the articulation of embodied, affective experience. More broadly, the study contributes to corpus-assisted discourse analysis by showing how multi-word-unit-

based, context-sensitive analysis can reveal patterns that are less accessible through single-word keyword or collocation approaches.

References

Collins, Luke, and Paul Baker. 2023. *Language, Discourse and Anxiety*. Cambridge: Cambridge University Press.

Gries, Stefan Th. 2024. *Frequency, Dispersion, Association, and Keyness: Revising and Tupleizing Corpus-Linguistic Measures*. Amsterdam and Philadelphia: John Benjamins.

Kiesling, Scott F. 2022. "'Stance and Stancetaking.'" *Annual Review of Linguistics* 8: 409-426.

There's no business like show business. A corpus-assisted analysis of the construction of celebrities' public self on Desert Island Discs

Anna Marchi (University of Bologna, Italy)

Niall Curry (Manchester Metropolitan University, UK)

This paper explores the construction of celebrities' public self in the iconic BBC radio program Desert Island Discs (henceforth, DID), using the DID Corpus – a 12 million word diachronic corpus of all the available episodes of DID between 1951 and 2024, annotated by occupation and speaker turn (2025). Using this database, we employ a corpus-assisted discourse analysis of the language produced by the facet of the entertainment industry that is most represented in the corpus: actors.

The analysis begins by comparing the speech of actors in the DID Corpus to castaways from other professional paths and by pursuing the key lexicon associated with work and fame in the actors' DID subcorpus (e.g. career, success, actor, acting, PERFORMANCE, PERFORM, talent). Using keywords as ways into the discourses surrounding the professional identities of people in “showbiz”, we interrogate extracts like the one presented below.

I did work in a pub in the East End for a short period when I was at drama school, in between terms, and I loved it, it was fun, but I couldn't have spent the rest of my life just doing that. You like being famous. I like being an actress, and I'm very proud that I've been recognised because of my ability, however small. I never wanted to be famous. I wanted to become well-known for being an actress, which was slightly different, because being famous, you can just be famous for no reason at all.

(Anita Dobson, Desert Island Discs 1988)

The close analysis of such extended concordance lines allows us to observe the piling up of narrative and explore what being an actor is about. Is it a job, a craft, a vocation? Is there a prevailing rhetoric of perseverance, of ambition, of talent, of privilege, of luck? The ultimate goal of our study is to develop a critical understanding of the meanings and the stories constructed around one's (or rather many's) professional and public self (or selves). Ultimately, the analysis contributes to the understanding of the construction of celebrity as product (Litter 2017) and demonstrates the value that CADS approaches can bring to discourse studies through the pursuit of patterns down screenfuls of concordance lines

References

Litter, J. 2017. Adrift or ashore? Desert Island Discs and celebrity culture. In J. Brown, N. Cook & S. Cottrell (Eds.) *Defining the discographic self. Desert Island Discs in context*. Oxford: Oxford University Press.

DID Corpus (2025). Desert Island Disks Corpus. No Sketch Engine. Available:

https://bellatrix.sslmit.unibo.it/noske/public/#dashboard?corpname=desert_island_disks [last accessed 4 Dec 2025]

Agency and involvement in the language of diabetes management. A corpus study of professional and lay discourses on type-1 and typ-2 diabetes

Georg Marko (University of Graz, Austria)

Type-1 and type-2 diabetes (T1D and T2D) are chronic conditions requiring continuous management. However, the wide availability and user-friendliness of advanced technical devices and medications, combined with the prominent role of lifestyle modification, have precipitated a significant shift from intervention by healthcare professionals towards greater agency and involvement of those directly (patients) or indirectly (family and friends) affected by diabetes (Schmeisl 2023). While this shift may be seen as a welcome development in chronic disease management, reflecting the desired transition from paternalistic to patient-centred models of healthcare (Nettleton 2006, 149), it risks placing sole responsibility for their well-being on patients themselves. This can result in overburdening them with self-management demands, while simultaneously granting at least figurative agency to non-human entities such as technical devices (e.g. insulin pumps, glucose monitors), medication (e.g. insulin, Metformin), and even food.

This study aims to investigate agency and involvement as constructed in the language of diabetes, specifically focusing on the representations of different people (healthcare professionals, patients, family and friends) and different objects (technical devices, medication, food) across professional and lay discourses on T1D and T2D.

Methodologically, the study draws upon a Corpus-Assisted Discourse Studies (inspired by, among others, Baker et al. 2006; Mautner 2008; Gillings et al. 2023) and Critical Discourse Analysis (Wodak and Meyer 2009), describing linguistic patterns in corpora in detail, relating them to the construction of socially relevant meanings, and critically reflecting upon potentially problematic but also positive sociocultural implications, such as those mentioned concerning agency and involvement in connection with diabetes management above.

Linguistically speaking, the research uses an adapted version of Halliday's (1994) concept of transitivity. It is based on the assumption that agency and involvement are scalar categories and that different morphosyntactic constructions affecting semantic roles can enhance or mitigate them to different degrees (e.g. the agent role in a clause (e.g. "the PATIENT does something") > goal role (e.g. "the PATIENT is treated") > circumstantial role (e.g. "someone talks about the PATIENT") > non-head in an noun phrase or a compound (e.g. "the PATIENT's parents", "PATIENT assistance"). I will trace these constructions in corpora, quantifying (how do frequencies affect the salience of agency and involvement of different entities?) and qualifying them (how do different semantic, lexical, and pragmatic aspects in the co-text modulate effects?).

The data to be examined are six self-compiled 1-million-word corpora, covering three scenarios – expert-to-expert discourse, represented by medical articles, expert-to-lay discourse, represented by self-help books, and lay-to-lay discourse, represented by reddit forums – and two types of diabetes (T1D and T2D). The tools used include Sketch Engine and Wordsmith 9.

References

- Baker, Paul, Costas Gabrielatos, Majid KhosraviNik, Michal Krzyżanowski, Tony McEnery, and Ruth. 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.
- Gillings, Mathew, Gerlinde Mautner, and Paul Baker. 2023. *Corpus-Assisted Discourse Studies*. CUP.
- Halliday, Michael A.K. 1994. *An Introduction to Functional Grammar* (2nd ed.). Edward Arnold.
- Mautner, Gerlinde. 2009. "Checks and Balances: How Corpus Linguistics can Contribute to CDA." In Wodak and Meyer (2009).
- Nettleton, Sarah. 2006. *The Sociology of Health and Illness* (2nd ed.). Polity.
- Schmeisl, Gerhard. 2023. *Schulungsbuch Diabetes*. Elsevier.
- Wodak, Ruth, and Michael Meyer, eds. 2009. *Methods of Critical Discourse Analysis* (2nd ed.). Sage.

Constructing regenerative travel: A corpus-driven, functionally motivated multimodal discourse analysis of Instagram hashtagging practices and eco-sensitive narratives

Elena Mattei (University of Bergamo, Italy)

This research explores how Instagram hashtagging practices construct regenerative travel as a pressing socio-ecological issue by integrating visuo-linguistic and metaphorical framing. Focusing on two multimodal sub-corpora—comprising both field-specific and general hashtags like #regenerativetravel, #travelwithimpact, and #travelwithpurpose—the study examines perceptions of environmental responsibility, community empowerment, and social transformation, also highlighting differences across emerging, ethically conscious travel forms (Mackenzie 2025; Rahmani and Carr 2021).

Using a mixed-method approach, the study combines Corpus-based Systemic Functional Linguistics (Halliday and Matthiessen 2014) with detailed multimodal analysis (Bateman et al. 2017). The corpus includes 2,200 Instagram posts and 250,000 tokens collected via API (Apify), alongside corresponding visual content. Keyword extraction, collocation analysis were conducted with Sketch Engine, complemented by semantic, Transitivity, and Appraisal manual coding using CATMA software. This helped identify key Indigenous and animist-regenerative metaphors, such as NATURE as NETWORK, KIN and LIVING, CONSCIOUS ENTITY (Rout and Reid 2020). Visual analysis of 200 images through a data-driven tagging system examined formal elements—(re)action type, participant roles, gaze, shots, angles, composition—and assessed how visuals reinforce or challenge ecological narratives (Pibiri and Mattei 2020).

Results show that regenerative tourism discourse contrasts with conventional neoliberal narratives, which depict landscapes and natural entities as static or aestheticized Phenomena (objects of observation of the romantic gaze, Mattei 2025; Urry and Larsen 2011). Instead, it emphasizes land as a nurturing, active agent engaged in reciprocal relationships with humans and ecosystems.

Such discourse strategies rely specifically on Transitivity and Appraisal patterns to encode agency through material processes of conservation and relational processes that describe, identify land, tourists and communities as both Actors and Carriers of Judgement and Appreciation resources that evoke empathy, ethical reflection, and protection toward ecological complexities. Such narratives also avoid emotionally charged Reactions and superficial Evaluations that reinforce the idealization, aestheticization of nature as a spectacle of human fantasy.

Concurrently, visual cues like direct gazes, eye-level angles and close-up shots of community active participation as well as reduced use of Reaction processes symbolizing consumption foster ethical awareness and encourage social activism, contesting dominant ideologies and advocating for ecological justice. In conclusion, this study shows how Instagram discourse may contribute to reframing the role of tourists, hosts and nature to support social change.

References

- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis – A problem-oriented introduction*. Mouton de Gruyter.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Introduction to functional grammar* (4th ed.). Routledge.
- Mackenzie, S. H. (2025). Enhancing community well-being via regenerative tourism?: Aspirations and approaches from Aotearoa New Zealand. In *The Routledge Handbook of Regenerative Tourism* (pp. 289–301). Routledge.
- Mattei, E. (2025). *The language of persuasion on Instagram: A systemic functional grammar of multimodal tourism discourse*. Routledge.
- Pibiri, G. E., & Mattei, E. (2020). *Statistically Reliable Image Tagging Software*.
<http://xor.isti.cnr.it:8000/login.html>
- Rahmani, Z., & Carr, A. (2021). Holistic tourism: A pathway towards eudaimonic well-being. In *Millennials, spirituality and tourism* (pp. 139–156). Routledge.
- Rout, M., & Reid, J. (2020). Embracing indigenous metaphors: a new/old way of thinking about sustainability. *Sustainability Science*, 15(3), 945-954.
- Urry, J., & Larsen, J. (2011). *The tourist gaze 3.0*.

Mainstreaming Misandry: examining the rise and spread of anti-feminist language in news media

Mark McGlashan (University of Liverpool, UK)

Jessica Aiston (Queen Mary University of London, UK)

Veronika Koller (Lancaster University, UK)

Alexandra Krendel (University of Southampton, UK)

This paper investigates the use of the term misandry – ‘the hatred of males’ (Oxford English Dictionary, 2023) – in English language news publications. This work builds on that of Marwick and Caplan (2018) – who used critical discourse analysis to explore usage of the term ‘misandry’ by men’s rights activists across various dedicated internet fora. Although semantically related to misogyny (‘the hatred of women’), Marwick and Caplan (2018, pp. 548–551) argue that misandry is functionally and discursively different in that it has been “used as a synonym for feminism and as a false equivalence to misogyny” and serves to “appropriate the language of leftist identity politics and claim a victimized stance”. Marwick and Caplan suggest, therefore, that misandry acts as a stand-in for the long-standing caricature of ‘the man-hating feminist’ as well as call-to-action against feminism.

In this paper, we examine whether misandry is gaining more widespread adoption outside of these dedicated spaces by exploring its use in mainstream publications. Specifically, the paper concentrates on a specialised corpus comprising 2,632 articles (~2.4million tokens) published in English language newspapers and web-based publications containing the terms ‘misandry OR misandrist OR misandric’ during 1983–2025. All data were retrieved from Lexis Nexis (<https://advance.lexis.com/>). The study details diachronic trends in the use of ‘misandry’, its international spread through worldwide publications, and the term’s linguistic associations and discourse prosodies over time.

In order to explore trends in use(s) and meaning(s) of misandry across our corpus, we combine methods for collocation analysis and key-keyword analysis so as to identify and examine relationships (emergent and evolving) between misandry and key-key-words appearing in all collocational windows for misandry. Rather than seeking to examine all collocates of misandry over time (which examines statistically salient relationships between words within a collocational window in contrast with all other contents of the same text or corpus), this approach focuses on only those relationships between misandry and words that are statistically salient from a keyness perspective within a defined span. This is achieved by identifying content-distinctive and content-generalisable key-key-words (Egbert & Biber, 2019; McGlashan & Krendel, 2024; Scott, 1997) for our corpus before exploring concordance windows for misandry containing these key-key-words. As such, this paper will concentrate on evolving meaning associations and discourse prosodies for misandry over time within mainstream English language news publications, thus contributing to our understanding of the mainstreaming of anti-feminist discourse.

References

- Egbert, J., & Biber, D. (2019). Incorporating text dispersion into keyword analyses. *Corpora*, 14(1), 77–104. <https://doi.org/10.3366/cor.2019.0162>
- Marwick, A. E., & Caplan, R. (2018). Drinking male tears: Language, the manosphere, and networked harassment. *Feminist Media Studies*, 18(4), 543–559. <https://doi.org/10.1080/14680777.2018.1450568>
- McGlashan, M., & Krendel, A. (2024). Keywords of the manosphere. *International Journal of Corpus Linguistics*, 29(1), 87–115. <https://doi.org/10.1075/ijcl.22053.mcg>
- Oxford English Dictionary. (2023). Misandry, n. Oxford University Press; Oxford English Dictionary. <https://doi.org/10.1093/OED/7147993009>
- Scott, M. (1997). PC analysis of key words - and key key words. *System*, 25(2), 233-45.

From European integration to contested values: a corpus-assisted discourse study of Europeanisation in the Croatian and Serbian broadsheet press (2007-2025)

Piotr Mirocha (Jagiellonian University in Kraków, Poland)

This longitudinal corpus-assisted discourse study (Baker et al. 2008; Baker 2023) investigates how Europeanisation has been constructed over the last two decades in the Croatian and Serbian broadsheet press by analysing more than 41,000 articles published between September 2007 and September 2025 in four ideologically diverse newspapers. Using collocation analysis, second-order collocates examination, and diachronic topic modelling to identify representative discourse samples (Mirocha 2023a; Bednarek 2024), the study traces how meanings associated with 'Europe' and 'Europeanness' have evolved in response to successive polycrises and shifting geopolitical factors.

The analysis identifies two major phases. The first is dominated by clusters such as 'European integration' and 'European standards', framing Europeanisation through economic crisis. While Serbian press focused on the global financial crisis and Croatian press on the Eurozone crisis, both emphasised 'fulfilling European standards' as objective necessities. The 2015 refugee crisis marked a turning point, raising the salience of crisis-related lexis in both countries (Mirocha 2023b).

The second phase marks a reorientation towards 'European values' and 'European way'. After 2022, security issues replace economic crises as the dominant frame. The Russian invasion of Ukraine intensifies crisis discourse in Serbian pro-government media, due to their reliance on Russian sources. Collocational patterns reveal divergent understanding of European values: liberal outlets foreground contradiction and defence, whereas pro-regime newspapers stress respect, sharing and commitment. Topic modelling indicates declining prominence of integration vocabulary.

The study shows how polycrises have continually reshaped the discursive landscape of Europeanisation, uncovering paradoxical trajectories: integration discourses diminish under contrasting conditions – Croatia's successful EU accession and Serbia's stalled progress. The findings demonstrate how quantitative tools can track long-term evolution in the meaning-making of Europe and reveal tensions accompanying disillusionment and uneven integration with the European Union (Jović 2023; Trošt 2023).

References

- Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, 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), 273-306.
- Baker, P. (2023). *Using Corpora in Discourse Analysis*. Bloomsbury Publishing.
- Bednarek, M. (2024). Topic modelling in corpus-based discourse analysis: Uses and critiques. *Discourse Studies*, 27(4), 659-671.

- Jović, D. (2023). Imagining Europe in a New and Small State: The Case of Croatia. In: Uvalić, M. (eds) Integrating the Western Balkans into the EU. New Perspectives on South-East Europe. Palgrave Macmillan, Cham.
- Mirocha P (2023a) The problem of migration in post-2007 Serbian media discourse about Europe: From topic modelling to qualitative discourse analysis. *Studia Linguistica Universitatis Jagellonicae Cracoviensis* 140: 255–266.
- Mirocha, P. (2023b). Europeanisation and the Dialectics of Crisis and Solidarity in the Croatian and Serbian Broadsheet Press between 2007 and 2017. *Southeastern Europe*, 46(3), 271-296.
- Trošt, T. P. (2023) Ambivalent Europeanization in the Western Balkans. *Current History* 122(842): 101–107.

Framing "limitations" as "bad news": A corpus-based comparative study of English and Spanish social science research articles

Ana I. Moreno (Universidad de León, Spain)

In research articles, communicating study limitations is a delicate rhetorical task that requires balancing transparency with the desire to present one's work persuasively. Spanish social scientists have reported particular difficulties when articulating limitations in English. This study investigates how "limitations" are discursively framed in English- and Spanish-medium social science journals, and how these framings may help explain such difficulties.

Working within intercultural rhetoric and genre-based approaches (Connor 2011; Moreno and Swales 2018), the study carries out a corpus-assisted discourse analysis of the local contexts in which limitations are presented in discussion and/or closing sections. The data consist of two comparable samples of ten discussion/closing sections each from pedagogy, sociology, psychology, business, and economics, drawn from the Exemplary Empirical Research Articles in English and Spanish (EXEMPRAES) Corpus. The sections had previously been segmented and annotated for communicative functions (Moreno 2021). Building on research into "bad news messages" in business communication (e.g. Lin 2020), limitations are conceptualised here as a form of bad news typically surrounded by text segments that serve distinct rhetorical purposes (e.g. preparation, mitigation, explication, reassurance).

The analysis identifies the sequences of rhetorical purposes surrounding each limitation and compares their frequency and ordering across the English and Spanish samples, complemented by interviews with ten authors. Preliminary results suggest that Spanish authors tend to embed limitations in extended explanations that highlight expertise and attribute constraints to external factors, often placing mitigation after the limitation. English authors more often foreground positive implications and future research, weave limitations into these takeaways, and deploy mitigation strategies before the "bad news". These patterned differences are interpreted with reference to writers' concerns with impression management and culturally shaped expectations of appropriate self-presentation in academic discourse.

The study demonstrates how corpus-assisted, functionally annotated data can illuminate cross-cultural variation in a key element of academic argumentation and open up new avenues for exploring cross-linguistic variation. It also points to implications for training intercultural rhetorical awareness in higher education.

References

- Connor, Ulla. 2011. *Intercultural Rhetoric in the Writing Classroom*. Ann Arbor: University of Michigan Press.
- Lin, Yan. 2020. "Communicating Bad News in Corporate Social Responsibility Reporting: A Genre-Based Analysis of Chinese Companies." *Discourse & Communication* 14 (1): 22–43.
- Moreno, Ana I. 2021. "Selling Research in RA Discussion Sections through English and Spanish: An Intercultural Rhetoric Approach." *English for Specific Purposes* 63: 1–17.
- Moreno, Ana I., and John M. Swales. 2018. "Strengthening Move Analysis Methodology towards Bridging the Function–Form Gap." *English for Specific Purposes* 50: 40–63.

New meanings of slow: a shift in time

Jessica Jane Nocella (University of Modena and Reggio Emilia, Italy)

The present study explores the changes in the meaning of the word slow in English, from the beginning of the Slow Food Movement (SFM) up to 2024.

Since 1989 and the early days of Carlo Petrini's SFM, technology has increasingly led people towards multi-tasking in an attempt to save time, while at the same time, back to the past trends are increasingly evident as seen in several other Slow Movements, for example, in catering and hand-made/retro chains. The concept of slowness is becoming a relevant and ethical topic that is often related to what is organic, local and sustainable, leading to the introduction of an anglicism in Italian, but also to a new whole re-semantization of the word slow in English.

The main aim of this study is to explore the semantic changes in the word slow diachronically in English. In order to examine such changes, a diachronic corpus from The Guardian newspaper was collected. The word slow was analysed within the different newspaper domains with a combination of both qualitative and quantitative methods (Sinclair 2004; Brezina 2018) exploring frequencies, collocates, and identifying significant "repeated events" (Sinclair 2004: 28) together with their semantic preferences.

Results from the diachronic analysis show that through the SFM, some specific new meanings of slow in Italian were then partly re-introduced in English. This is true for the dimensions related to ethics, wellbeing, environmental awareness. More specifically, the case of reborrowing of slow within the context of movements (slow fashion, slow tourism, etc.) has quite likely been influenced by the Italian SFM. As a matter of fact, this process of reborrowing involves both acquisition and loss in meaning in both languages. This seems to be a relevant phenomenon in a global context of constant linguistic contact where English appears to play a relevant role as a global language.

References

- Brezina, Vaclav. 2018. *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge University Press.
- Sinclair, John. 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.

Rebuild and Renew: Discourses of national health and nationhood under Starmer's Labour

Tamsin Parnell (University of Nottingham, UK)

In this presentation, I offer an empirical examination of the intertwined representations of the NHS and nationhood in the discourses of Keir Starmer's Labour government. Employing corpus-assisted critical discourse analysis, I analyse press releases, government-crafted news stories and speeches in two corpora: the Department of Health and Social Care (DHSC) corpus (~199,000 tokens) and the Cabinet Office (CO) corpus (~404,000 tokens) respectively. In the analysis of 5-grams, keywords, collocates, and concordances, I argue that the DHSC slogan that the NHS 'is broken, but not beaten' is a microcosm for the government's broader narrative of national reform and renewal. The discourses are interpreted as arising out of the specific polycrisis of COVID-19 and Brexit on the one hand, and long-standing political narratives about the NHS and (English/British) nationhood on the other. Within the presentation, I also reflect on some of the opportunities and challenges of working on politics-in-progress by taking up Darics et al.'s (2025) metaphor of "seeing the water". I use metaphors of rippled water and immersion; the 'rippled water' metaphor suggests working on politics-in-progress can be like looking through rippled water, with shifting visions of our reflection (our biases and identity) and the water itself (the topic). The metaphor of immersion recognises that working on the discourses of a government that is still in office can be difficult because we are immersed in the metaphorical water and do not have a complete sense of what the relevant context is, or how discourses might progress (in the face of, for example, political U-turns).

References

Darics, E., Staschen, O., Drury, M., Chojnicka, J., Mazzoli, M., Olthof, J., Serafis, D. and Wildfeuer, J. (2025). "Critical language awareness as a future imperative: seeing the 'water'." *Frontiers in Communication*, 10:1623193.
<https://doi.org/10.3389/fcomm.2025.1623193>

Explanatory processes in expert discourse: A view on Ask-an-Expert websites across disciplinary fields

Daniel Pascual (Universidad de Zaragoza, Spain)

Scientific knowledge is increasingly available to the wider public thanks to digital media. Careful efforts are made by experts to unpack their disciplinary fields and disclose complex topics that are socially relevant. The recontextualisation of expert discourse is key in this process, so that interaction between experts and diverse audiences may result in “communicative efficiency” (Roelcke 2002). The present study focuses on Ask-an-Expert websites as a digital practice where users pose questions answers by accredited professionals who aim to foster intelligibility, credibility, and trust. Relying on the SciDis Database, which compiles representative examples of scientific dissemination practices (Pascual and Sancho-Ortiz, 2024), three disciplinary fields are chosen to explore how potential “knowledge asymmetries” (Kastberg 2011) regarding expert discourse are bridged: nutrition, psychology, and climate. The corpus comprises two specific websites per field, and 25 texts per website, totalling 150 Ask-an-Expert texts. The main goal of this paper is to investigate the verbal explanatory processes through which expert discourse is adapted to accommodate lay users on these websites. The analysis follows a Corpus-Assisted Discourse Studies (CADS) approach, and presents a mixed-method design. Drawing on previous research (e.g., Cavalieri and Diani 2019; Mur-Dueñas 2024), the close exploration of the corpus led to the design of a taxonomy of ten salient explanatory processes: elaboration, explicitation, definition, denomination, enumeration, exemplification, justification, attribution, metaphor and comparison/analogy. The corpus was coded using the NVivo software. The frequency of use of these processes was retrieved to establish a comparison across disciplinary fields.

Results reveal shared communicative patterns and disciplinary differences in how experts tailor explanations for lay audiences. Across all six Ask-an-Expert sites, explicitation, exemplification, elaboration, and definition are central to catering for users’ knowledge gaps and recontextualizing specialised discourse. Explicitation and exemplification are particularly frequent, clarifying technical content and using concrete scenarios. Definition and elaboration are recurrently used to unpack terminology and expand initial answers. Clear discipline-specific profiles emerge as well. Ask-an-Expert texts on dietetics and nutrition emphasise practical justification and enumeration, reflecting the applied, action-oriented nature of nutritional guidance. Psychology and mental health sites rely more on metaphor, and analogy to contextualise emotional or behavioural issues, grounding advice in psychological theory and relatable experiences. Climate and environmental platforms favour exemplification and comparison with everyday processes to make complex scientific mechanisms more tangible. Overall, insights into verbal explanatory processes can help scientists and researchers facilitate disciplinary knowledge and conceptual complexity to the wide public, while bolstering credibility and trust in dialogic ways that contribute to making science accessible and comprehensible.

References

Cavalieri, S., and G. Diani. 2019. “Exploring Health Literacy: Web-Based Genres in Disseminating Specialized Knowledge to Caregivers. The Case of Paediatric

Neurological Disorders.” *Lingue Culture Mediazioni / Languages Cultures Mediation* 6 (1): 89–105. <https://doi.org/10.7358/lcm-2019-001-cadi>.

Kastberg, P. 2011. “Knowledge Asymmetries: Beyond ‘To Have and Have Not.’” *Fachsprache: International Journal of Specialized Communication* 34 (3–4): 137–51.

Mur-Dueñas, P. 2024. “Digital Dissemination Practices: An Analysis of Explanatory Strategies in the Process of Recontextualising Specialised Knowledge.” *Discourse and Interaction* 17 (1): 94–114. <https://doi.org/10.5817/DI2024-1-94>.

Pascual, D. and Sancho-Ortiz, A. E. 2024. "Investigating recontextualisation processes in scientific digital practices: The SciDis Database". *Revista Electrónica de Lingüística Aplicada* 23: 101-118. <https://doi.org/10.58859/rael.v23i1.649>

Roelcke, T. 2002. “Efficiency of Communication: A New Concept of Language Economy.” *Glottometrics* 4: 27–38.

A corpus-assisted multimodal discourse analysis of GLP-1 receptor agonist narratives on TikTok

Nicola Pelizzari (Università degli Studi di Brescia, Italy)

This study examines how GLP-1 receptor agonists are discursively constructed on TikTok. It focuses on how linguistic and multimodal resources shape public talk about pharmaceutical weight loss. GLP-1 medications were first designed to manage type 2 diabetes, but their capacity to induce weight loss has led to widespread popular discussion beyond clinical settings.

On TikTok, these drugs have achieved viral visibility, with users sharing personal stories of transformation. Yet research offers little linguistic or multimodal insight into how such narratives construct medical authority and selfhood in participatory media. With research access approved by TikTok (ID: 1723718073), a corpus-assisted multimodal discourse analysis (Bednarek, 2015) was conducted on a systematically sampled corpus of 100 English-language TikTok videos tagged #GLP1. The dataset totalled 137,032 words, 254 minutes of transcribed material, and a cumulative 114,193 likes, 36,886 comments, and 34,791 shares.

Transcribed speech was analysed as text using AntConc (Anthony, 2019), combining corpus-based keyword, collocation, and stance analysis (Baker et al., 2008; Du Bois, 2007) with multimodal annotation of gesture, gaze, camera framing, soundtrack, and text overlays (Kress & van Leeuwen, 2020). Metaphors were coded using MIPVU (Steen et al., 2010) and sentiment with VADER (Hutto & Gilbert, 2014). Quantitative tests assessed variation in lexical and evaluative patterns across creator types.

Findings reveal a predominantly weight-normative and biomedical framing, with recurring lexical fields of transformation, treatment, and personal success. Positive stance expressions such as life-changing, miracle, and game-changer strongly outnumber neutral or critical framings. Multimodal cues including direct gaze, smiling delivery, upbeat music, and before-and-after imagery reinforce verbal positivity and provide visual proof of efficacy. Professional creators adopt calmer delivery and neutral settings, producing more informational but less visible content. Statistical testing shows significant differences in stance and metaphor by creator type ($\chi^2 = 14.3$, $p < .01$), with lay users favouring miracle and battle metaphors and professionals preferring tool and journey frames.

Overall, TikTok's affect-driven style amplifies personal testimony and emotional framings of GLP-1 use, while clinically balanced or weight-inclusive perspectives remain marginal. These patterns suggest that algorithmic visibility and engagement practices help shape a discourse where pharmaceutical intervention is framed through affect and personal transformation.

References

- Anthony, L. (2019). AntConc (Version 3.5.8) [Computer software]. Waseda University.
<https://www.laurenceanthony.net/software/antconc/>
- Bednarek, M. (2015). Corpus-assisted multimodal discourse analysis of television and film narratives. In P. Baker & T. McEnery (Eds.), *Corpora and discourse studies: Integrating*

discourse and corpora (pp. 63–87). Palgrave Macmillan.

https://doi.org/10.1057/9781137431738_4

Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 139–182). John Benjamins.

<https://doi.org/10.1075/pbns.164.07du>

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Weblogs and Social Media* (Vol. 8, pp. 216–225).

<https://doi.org/10.1609/icwsm.v8i1.14550>

Kress, G., & Leeuwen, T.V. (2020). *Reading Images: The Grammar of Visual Design* (3rd ed.). Routledge. <https://doi.org/10.4324/9781003099857>

Steen, G. J., Dorst, A. G., Krennmayr, T., Kaal, A. A., & Herrmann, J. B. (2010). A method for linguistic metaphor identification: From MIP to MIPVU. John Benjamins.

<https://doi.org/10.1075/celcr.14>

From Recipes to Radicalization: A Corpus-linguistic Perspective on the Rhetoric of Far-right and Tradwife Influencers

Claudia Posch (University of Innsbruck, Austria)

This paper investigates the linguistic patterns used by influencers on social media use to normalise patriarchal ideologies. To achieve this, the study employs a Corpus-Assisted Discourse Studies (CADS) approach, focusing on identifying recurrent linguistic and argumentative patterns in two English corpora (approx. 3 million tokens), and two German corpora (215,000 tokens), extracted from YouTube and TikTok, including captions and user comments. The goal is to examine how overt political rhetoric and seemingly apolitical domestic content contribute to the "shameless normalisation" of far-right and anti-feminist discourses, as described by Wodak (2024, 361).

The analysis takes the form of an exploratory study and it centers on two influencers: Brittany Pettibone Sellner (BPS), a US-based far-right YouTube activist with close, personal ties to the far-right in Austria; and Simha Lily (SL), a German-speaking, Colombian-born Tradwife active on TikTok. Collocation and concordance analysis as well as large language models (Curry et al. 2025) were used to identify rhetorical and discursive patterns, and metaphors (Isentyeva 2019). BPS uses rhetorical techniques such as redefinition, straw-man fallacies, and appeals to emotion to reframe feminism as harmful. Her discourse is characterised by the strategic use of emotionally charged language, metaphorical mappings, and calculated ambivalence, which redefine feminism as a negative ideology while promoting patriarchal femininity as a solution. Collocational patterns (Kennedy et al. 2024) and concordance lines (Gillings & Mautner 2024) reveal how BPS's rhetoric interweaves feminist terminology with anti-feminist arguments, creating a discourse that appears relatable and credible to her audience. The user comment corpus echoes BPS' anti-feminist rhetoric, using emotionally charged language and negative metaphors to describe feminism, such as framing it as a "disease" or "toxic ideology". SL's content, on the other hand, appears apolitical at first glance but it aligns with conservative and far-right ideologies by presenting the Tradwife lifestyle as a natural and fulfilling alternative to feminist ideals. Similarly, comments on SL's videos reflect the emotional and nostalgic appeal of her content, explicitly conveying the implicit traditionalist and anti-modern messages.

This study demonstrates the value of CADS in analysing complex multimodal data, and in revealing how language, visuals, and ideology in online influencer culture interplay. By combining quantitative corpus analysis with qualitative discourse analysis, the research identifies how far-right influencers strategically blend personal narratives, emotional appeals, and aesthetic presentations to normalise reactionary ideologies. The findings show the importance of examining both overt and implicit discursive strategies to understand how social media content contributes to the "shameless normalization" of far-right and anti-feminist discourse. For further research larger, multilingual datasets should be investigated to study how corpus and discourse studies can contribute to understanding the role of social media in ideological radicalization.

References

- Curry, Niall, Tony McEnery & Gavin Brookes. 2025. A question of alignment – AI, GenAI and applied linguistics. *Annual Review of Applied Linguistics* 45. 315–336.
- Gillings, Mathew & Gerlinde Mautner. 2024. Concordance lines. What are they and how can they be used to explore representation? In Frazer Heritage & Charlotte Taylor (eds.), *Analysing representation: A corpus and discourse textbook*, 12–27. London, New York: Routledge Taylor & Francis Group.
- Islentyeva, Anna. 2019. The Europe of Scary Metaphors: The Voices of the British Right-Wing Press. *Zeitschrift für Anglistik und Amerikanistik* 67(3). 209–229.
- Kennedy, Charlotte-Rose, Gavin Brookes & Anna Cherniaeva. 2024. Collocates. What are they and how can they be used to explore representation? In Frazer Heritage & Charlotte Taylor (eds.) *Analysing Representation: A Corpus and Discourse Textbook*. London: Routledge.

‘Woman’ vs. ‘womb’: The construction of pregnant people in 50 years of amicus briefs from landmark abortion cases

Amanda Potts (Cardiff University, UK)

Jamie Abrams (American University, USA)

This talk presents an interdisciplinary legal-linguistic study of the amicus briefs that were filed in the milestone U.S. Supreme Court abortion cases of Roe, Doe, Casey, and Dobbs. Amicus briefs are intriguing discursive artefacts in which various constituencies construct abortion, women, fetuses, physicians, rights, and harms (Collins et al 2015). The resulting corpus comprises 1.1 million words of briefs spanning approximately 50 years. Applying corpus-based critical discourse analysis (Partington et al 2013), we systematically compare the rhetorical strategies across categories of amici (e.g., religious groups versus medical groups), analyze diachronic shifts in nomination strategies, and contrast argumentation in briefs seeking to restrict versus expand abortion access.

In this presentation, we will focus on ways that women have been represented over time in landmark abortion cases, and contrast the ways in which authors advocating to either restrict or expand abortion describe women. Through collocation analysis of adjectival attributes, we reveal how briefs arguing to expand abortion access were dedicated to intersectional narratives of harm and undue burdens that would be suffered by marginalized groups of women—even though these issues were not before the Court. By contrast, in nomination analysis, we show how authors wishing to restrict abortion depicted pregnant people as atomistic body parts (i.e. womb and cervix), or pregnant mothers. This represents the reproductive organs as the entirety of the pregnant person, eviscerates the relationship of the woman’s work and labor in pregnancy, and assigns the presumptive identity of mother regardless of pregnancy outcome. Intriguingly, the pregnant person is passive across all briefs and over time, indicating that the main party is being omitted from the discourse.

This study offers historical perspectives into evolving rhetorical strategies in abortion litigation, contemporaneous insights into the state of abortion politics, and future implications to amici activity and abortion advocacy. It charts a course forward for more effective engagement with the Court through (re-)incorporation of holistic nomination strategies, individual narratives, and agentive roles.

References

Paul M. Collins Jr. et al., The Influence of Amicus Curiae Briefs on U.S. Supreme Court Opinion Content, 49 *Law & Soc’y Rev.* 917 (2015).

Alan Partington, et. al., Patterns and Meanings in Discourse Theory and Practice in Corpus-Assisted Discourse Studies (CADS) 11 (2013).

Taking a stand on issues of gender and sexuality in Hong Kong: a corpus-driven analysis of letters to the editor

Benedict J.L. Rowlett (Hong Kong Baptist University, Hong Kong)

This talk presents findings from a corpus-driven analysis of letters to the editor addressing social issues of gender and sexuality in Hong Kong, China. The study combines keyword analysis with manual examination of concordance lines using Fairclough and Fairclough's (2011) framework of practical reasoning. The data derive from a larger discourse-analytical project based on a corpus of over 15,000 letters published in Hong Kong's English-language newspaper, the South China Morning Post, between 2014 and 2023 - a period of profound social and political upheaval in the city. Accordingly, debates on Hong Kong's sustainability as an "international" and "progressive" city amid increasing authoritarian restrictions on civil society intensified during this period. Keyword analysis of the corpus using AntConc (Anthony 2024) revealed a notable recurrence of words related to gender and sexuality, marking these as prominent areas of public concern. A sub-corpus of over 100 letters on gender/sexuality-related topics was therefore extracted to examine how such issues were presented and deliberated in this public forum.

The analysis focused on how social actors were positioned and how calls for action were articulated in relation to sexual politics in Hong Kong. Under Fairclough and Fairclough's (2011) practical reasoning framework, a concordance analysis identified recurring discourse patterns concerning circumstances, values, goals, and claims for action through which letter writers constructed their arguments. For example, concordances centered on the keyword "LGBT+" revealed that this group was generally represented as marginalized and deserving of "equal treatment" (circumstances/goals). Concordances on "Hong Kong" as a metonym demonstrated how the city (and its citizens) was represented as embodying values of diversity, openness, and cosmopolitanism, yet simultaneously as needing to "do better" to uphold these ideals (claims for action). Representations of the government (and its various offices) were particularly significant, where it was positioned both as part of the problem (circumstances) and as a responsible agent (claims) expected to enact equal rights legislation and thereby foster a modern, inclusive, and autonomous society consistent with Hong Kong's branding as "Asia's world city." As such, the analysis suggests that forms of homonationalist discourse - linking inclusive sexual citizenship to civic or national pride - underpin many of these arguments, though in ways that diverge from Western theoretical perspectives on homonationalism (cf. Lazar 2017). More broadly, the recurrent mobilization of effective and accountable governance as a value premise points to how engagement with gender and sexuality issues is just one example of how citizens have used the letters to the editor forum to address, confront and challenge the mechanisms of power.

The study concludes by demonstrating the potential of corpus-assisted discourse analysis to unpack the discursive strategies through which letter writers communicate civic values and identities, and to further investigate the South China Morning Post letters corpus as an archive of committed citizen voices - voices now increasingly constrained under tightening restrictions on Hong Kong's civic space and freedom of expression.

References

Anthony, Laurence. AntConc. Version 4.0.3. Tokyo: Waseda University, 2024.

<https://www.laurenceanthony.net/software/antconc/>.

Fairclough, Isabela, and Norman Fairclough. 2011. "Practical Reasoning in Political Discourse: The UK Government's Response to the Economic Crisis in the 2008 Pre-Budget Report." *Discourse & Society* 22 (3): 243–268.

Lazar, Michelle M. 2017. "Homonationalist Discourse as a Politics of Pragmatic Resistance in Singapore's Pink Dot Movement: Towards a Southern Praxis." *Journal of Sociolinguistics* 21 (3): 420–441.

Image-based sexual abuse in digital discourse: a critical analysis of YouTube commenters' discursive constructions of victim-survivors and perpetrators

Marta Sánchez-Cóccera (Complutense University of Madrid, Spain)

This study investigates image-based sexual abuse (henceforth, IBSA) through a discursive lens. IBSA refers to the range of abusive practices whereby perpetrators non-consensually create, distribute or threaten to distribute intimate images (Henry et al. 2021). Research on IBSA has been extensive, particularly from criminological, sociological, psychological and legal perspectives, revealing that IBSA is a technosocial, gendered form of sexual violence that is on a rising trend (Henry et al., 2021; Ringrose & Regehr 2025). However, to date, there is a paucity of linguistic incursions into this phenomenon, leaving important perspectives overlooked. To fill this gap, this study sets out to analyze a corpus of 7,000 user-generated YouTube comments (ca. 500,000 words) in response to three mediatized IBSA cases. YouTube, as a video-sharing platform, hosts numerous videos about IBSA, where users can express opinions about these wrongs, perpetrators and victim-survivors. Considering how embedded digital technologies are in everyday communication and how blurred the boundaries between the offline and online spheres are in today's society, YouTube enables users to participate in digital public discourse when commenting under video feeds, affecting both material and immaterial realities (Bou-Franch & Garcés-Conejos Blitvich 2016). Therefore, by examining how YouTube commenters publicly and discursively construct IBSA as well as victim-survivors and perpetrators online, I aim to identify the ideologies that frame IBSA offline.

The study draws on Feminist Critical Discourse Analysis (Lazar 2018) and Corpus-Assisted Discourse Studies (Gillings et al. 2023) to approach the analysis and interpretation of linguistic patterns and discourse strategies YouTube commenters (and broadly public citizens) rely on to construct IBSA. To this end, Keyword and Word Sketch analyses were conducted with Sketch Engine to identify salient lexical items and their most typical collocational patterns. Keywords were retrieved using the simple maths method, and only items with a minimum relative dispersion of 60% and a score threshold of 2.0 were retained. Thereafter, the most relevant 100 keywords were analyzed in context, alongside their collocations with a typicality score above 3. Results show polarizing tendencies in how victim-survivors and perpetrators are represented. While perpetrators are consistently condemned and negatively appraised, they are represented as norm-deviating individuals, dismissing the structural causes of IBSA. Victim-survivors' representations are, however, disparate. A segment of comments show support, representing victims as naive or godlike, whereas an overwhelming set of comments deems them responsible for their victimization through slut-shaming discourses. Importantly, the analysis finds that victim support is not at odds with victim blaming, as supportive comments may covertly reproduce rape-culture beliefs. This study highlights the ideological tensions underpinning public discourse around IBSA.

References

Bou-Franch, Patricia, and Pilar Garcés-Conejos Blitvich. 2016. "Gender Ideology and Social Identity Processes in Online Language Aggression against Women." In *Exploring Language Aggression against Women*, edited by Patricia Bou-Franch, 59–81. John Benjamins.

Gillings, Mathew, Gerlinde Mautner, and Paul Baker. 2023. *Corpus-Assisted Discourse Studies*. Cambridge University Press.

Henry, Nicola, Clare McGlynn, Asher Flynn, Kelly Johnson, Anastasia Powell, and Adrian J. Scott. 2021. *Image-Based Sexual Abuse: A Study on the Causes and Consequences of Non-Consensual Nude or Sexual Imagery*. Routledge.

Lazar, Michelle M. 2018. "Feminist Critical Discourse Analysis." In *The Routledge Handbook of Critical Discourse Studies*, edited by John Flowerdew and John E. Richardson, 372–387. Routledge.

Ringrose, Jessica, and Kaitlyn Regehr. 2025. *Teens, Social Media and Image Based Abuse*. Palgrave Macmillan.

The Use of Clinical Trial Studies to Support and Oppose Conspiracies during Pandemic Uncertainty: A Corpus Assisted Discourse Study of Reddit

Chris Sanderson (Lancaster University, UK)

Elena Semino (Lancaster University, UK)

Fiona Burns (University College London, UK)

Shema Tariq (University College London, UK)

Early in the COVID-19 pandemic, there was a paucity of information about the novel virus, resulting in a 'data void' (Golebiewski and Boyd 2019), and creating fertile ground for conspiracies. One particular conspiracy involved the antiparasitic drug Ivermectin as a miracle cure being kept from the public (Baker and Maddox 2022).

As part of the multidisciplinary PROACT EU-Response project, a European Commission funded project which seeks to better prepare Europe for future pandemics, we are analysing a corpus of 1562 Reddit submissions with 'Clinical Trial(s)' in the title, and their accompanying 87,972 comments. Submissions where the topic was an infectious disease, such as COVID, flu, or HIV were separated into an 'INFECT' subcorpus. A keyword analysis was then conducted using the rest of the corpus as a reference corpus.

'Ivermectin' was the top keyword for the INFECT subcorpus, closely followed by 'hydroxychloroquine'. These findings are consistent with known conspiracy theories about COVID-19 treatment. The antiparasitic drug Ivermectin emerged as a prominent theory early in the pandemic, positioned by key influencers as a miracle cure that was being kept from the public (Baker and Maddox 2022). This conspiracy has persisted, with US podcaster and influencer Joe Rogan claiming in January 2025 that Ivermectin cured his friends of cancer. Hydroxychloroquine was identified by Baker and Maddox (2022) as another prevalent conspiracy theory during the pandemic.

A concordance analysis revealed that discussions of Ivermectin frequently included intertextual references to academic articles of clinical trials. One linked article, which has since been retracted, attained 683,000 accesses (Zaidi and Dehgani-Mobaraki 2022). This paper lost the confidence of the Editor-in-Chief, as cited sources did not show clear clinical evidence of the effect of Ivermectin on COVID. None of the authors agreed to the retraction.

To investigate whether academic studies are used in our corpus more widely, all possible URL addresses were extracted. This resulted in 5803 links. Of these links, nature.com, sciencedirect.com, lancet.com and doi.org were all in the top 20 most frequent domains in our corpus, suggesting widespread intertextual references to academic articles in our corpus. This study looks at the use of these links in context, using concordance analysis to investigate how intertextual references to clinical trial studies are used to discuss potential treatments. The paper will go on to discuss implications for online conspiratorial health discourse, particularly in contexts of 'data voids' during emerging and novel infectious diseases.

References

- Baker, Stephanie Alice, and Alexia Maddox. 2022. 'From COVID-19 Treatment to Miracle Cure: The Role of Influencers and Public Figures in Amplifying the Hydroxychloroquine and Ivermectin Conspiracy Theories during the Pandemic.' *M/C Journal* 25 (1). <https://doi.org/10.5204/mcj.2872>.
- Golebiewski, Michael, and Danah Boyd. 2019. Data Voids: Where Missing Data Can Easily Be Exploited. *Data & Society*. <https://datasociety.net/wp-content/uploads/2019/11/Data-Voids-2.0-Final.pdf>.
- Zaidi, Asiya Kamber, and Puya Dehghani-Mobaraki. 2022. 'RETRACTED ARTICLE: The Mechanisms of Action of Ivermectin against SARS-CoV-2: An Evidence-Based Clinical Review Article'. *The Journal of Antibiotics* 75 (2): 122–122. <https://doi.org/10.1038/s41429-021-00430-5>.

A comparison of multivariate statistical approaches for corpus-based discourse analysis

Tony Berber Sardinha (Pontifical Catholic University of São Paulo, Brazil)

In the last few years, different multivariate approaches have been proposed that identify discourses in corpora using lexical variables. In this paper, we focus on three of these approaches, namely Lexical Multi-Dimensional Analysis (LMDA; Berber Sardinha and Fitzsimmons-Doolan 2025), Keyword Cooccurrence Analysis (KCA; Clarke, Brookes, and McEnery 2022), and clustering of word embeddings (BERT, Bidirectional Encoder Representations from Transformers; Devlin et al. 2019).

- (1) LMDA is a lexical extension of Biber's Multidimensional Analysis (Biber 1988). Correlations among the lexical variables capture the degree to which lexical items tend to occur together or avoid each other. Factor analysis is then applied to this correlation matrix to detect latent constructs (factors), which are interpreted as discourse dimensions. Each dimension represents a set of lexical features that co-occur across the corpus and is arrived at by interpreting the factor pattern and high-scoring texts.
- (2) KCA uses Multiple Correspondence Analysis (MCA) to identify patterns in keyword presence or absence. MCA produces a geometric representation in which each category receives a coordinate and a contribution value. The KCA dimensions are then interpreted by consulting the categories with the strongest contributions and the coordinates showing how these categories are distributed across texts.
- (3) Word embeddings are numerical vectors obtained from a neural language model. A vector represents a word through nonlinear mathematical transformations. The result is a long numerical string with hundreds of numbers, which do not correspond to identifiable linguistic units such as words or texts but to abstract numerical components created by the transformation process. To obtain groupings of embeddings, we applied k-means cluster analysis to the embeddings.

The three methods were applied to the same corpus of 4,800 immigration-related news texts (7.6 million words). The same keyword list was used for all three methods. To compare the methods, two analyses were conducted (quantitative and qualitative). The quantitative analysis showed low overlap across the dimensions and clusters. The results showed that the best mean match was for LMDA and KCA, with 7.88 overlap. In a closest-match comparison where each dimension pole or cluster was matched against all others, LMDA and MCA reached 59.56 percent overlap at best, and LMDA and embeddings, 24.71 percent. The qualitative analysis confirmed these differences. LMDA dimensions pointed to abstract and evaluative discourse, KCA dimensions signaled thematic strands reflecting the topical components, and the embedding clusters captured semantic and morphosyntactic groupings. Since these approaches rely on different statistical calculations, it is expected that their results will differ. However, the exact way in which they differed could not be predicted in advance. These will be illustrated and discussed in the paper presentation.

References

- Berber Sardinha, T., & Fitzsimmons-Doolan, S. (2025). *Lexical Multidimensional Analysis: Identifying Discourses and Ideologies*. Cambridge: Cambridge University Press.
- Biber, D. (1988). *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Clarke, I., Brookes, G., & McEnery, T. (2022). Keywords through time: Tracking changes in press discourse of Islam. *International Journal of Corpus Linguistics*, 27(4), 399-427.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Minneapolis, Minnesota.

Discourses in human and LLM-produced oral histories of the pandemic

Tony Berber Sardinha (Pontifical Catholic University of São Paulo [PUCSP], Brazil) Andrea Cristina Martins Santos (PUCSP), Amanda Chiarello Boldarine (PUCSP), Andressa Domingues Pedroso Salaru (PUCSP), Arianne Brogini (PUCSP), Carlos Kauffmann (PUCSP), Carolina Andrade (PUCSP), Cicero Soares da Silva (PUCSP), Leticia Gregorio Souza Rodrigues (PUCSP), Maria Claudia Nunes Delfino (Sao Paulo State Technical College, Brazil), Melina Maria Manasseh (PUCSP), Milena Bonfim (PUCSP), Renata Fernandes da Gama (PUCSP), Sara Elysa Dantas Souza (PUCSP), Simone Sarmento (Federal University of Rio Grande do Sul, Brazil) and Simone Vieira Resende (Cultura Inglesa College, Brazil)

Oral history is both a practice and a research method concerned with recording how people speak about events they consider noteworthy and with documenting their remembered experiences of the past (Abrams 2010). Oral histories are not literal reproductions of experience but retrospective constructions: interviewees assemble fragments of memory into a coherent account of their life trajectories for the purposes of the interview. In principle, LLMs might approximate such accounts when asked to fictionalize an oral history because they encode vast textual knowledge that reflects many kinds of lived experience. However, the generalities of human life found in mass produced texts cannot be easily reverse engineered into the specific worlds of unknown individuals. Therefore, whether a model can generate a plausible account of a single human life remains an open question. For that reason, a suitable test of LLM humanlikeness is whether a model can emulate one particular person's recollections as they appear in oral history accounts.

To do that, we built a corpus of (human) oral history based on an existing collection ('Journal of the Plague Year'), and tasked LLMs with assuming the identity of each individual who was interviewed. We implemented this in two ways. In the plain prompt condition, the model was simply told to speak as someone living in the USA during the pandemic who was being interviewed for an oral history project. In the extended persona prompt condition, the model received the same instruction but was also given biographical information about the actual interviewee (taken from project records) together with a summary of each of that person's answers, used as topic control for the responses. As a result, we had seven subcorpora, one for each LLM (GPT 5.1, Grok 4.1, and Gemini 2.5-Pro) with two prompt conditions (plain or persona), in addition to the human corpus. The final corpus amounted to 1974 texts and 7.8 million words.

The corpus was submitted to a Lexical Multidimensional Analysis (LMDA; Berber Sardinha and Fitzsimmons-Doolan 2025), where each text was POS tagged, content words were retrieved and counted, and a factor analysis was run, revealing seven factors, each with two poles. LMDA is an extension of Multi-Dimensional Analysis (Biber 1988) with which it shares constructs and procedures. The major goal of an LMDA is to detect discourses in corpora based on the principle that shared lexis (cross-text and within-text cooccurrence) signals shared discourses.

Discourses are understood as "socially shared, socially situated, representations of real-world phenomena conveyed implicitly through language use. Because they are shared, discourses and ideologies also constrain or limit how real-world phenomena are represented." (p.11). The results showed sharp differences between human histories and LLM histories. To illustrate,

Dim. 2, labeled as 'Emotional, family-centered pandemic storytelling vs Reflective, crisis and morale discourse,' showed human histories adopted emotional discourses whereas LLMs preferred morale discourses. And Dim. 4 (Adaptive pandemic lifestyle and silver linings vs Justice-oriented discourse) also saw a split: LLMs pushed lifestyle adaptation discourses whereas humans stressed justice-seeking discourses. In general, LLMs constructed different histories even when helped by the extended prompts. Plain prompts generally led to generic pandemic plots, whereas persona prompts usually exaggerated discourses. All dimensions will be detailed in the paper presentation.

References

- Abrams, Lynn. 2010. *Oral History Theory*. London: Routledge.
- Berber Sardinha, Tony, and Shannon Fitzsimmons-Doolan. 2025. *Lexical Multidimensional Analysis: Identifying Discourses and Ideologies*. Cambridge: Cambridge University Press.
- Biber, Douglas. 1988. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Investigating academics' discourses of writing: A Corpus-Assisted study of disciplinary expectations

Analisa Scerri (University of Malta, Malta)

Academic discourse depends on writing as its primary mode of knowledge communication. Given its social role in knowledge mediation, academic writing, like any form of social communication, is shaped by conventions (Curry, 2021). These conventions are iteratively defined and redefined by academics and can act as a form of gatekeeping in the context of academic publication. This reality has given rise to a large canon of work dedicated to unpacking and understanding academic discourse with a view to informing writing education. This practice is premised on the view that a clear understanding of what and how to write enables students to gain access to the discourse community of their disciplines (Lillis & Harrington, 2015) and to become core members of that community. Being gatekeepers to the discipline, academics' views and values towards what they deem to be appropriate influences how writing in the discipline is shaped and measured (Hyland, 2009).

This paper responds to this research context and presents a corpus-based analysis of discipline-specific academics' discourse surrounding writing expectations. By analysing interviews with academics, I identify what they consider to be features of good-quality student writing and how these expectations relate the textual and discursive features of actual students' texts. Specifically, the paper reports on analyses of semi-structured interviews with 12 discipline specific academics and exemplars of students' writing selected by the participants. The findings highlight that most of the academics express a strong sense of responsibility towards the students and their growth within the disciplinary community and are committed to respond positively to the emergent challenges, such as Generative AI and the rapid changes in students' identities. Interestingly, a mismatch in the attitudes and values of some academics is evident in their discursively constructed expectations, which are often characterised by deficit framings of students and their abilities. Notably, in such cases, the academics persistently refer to the surface features of language as being problematic in students' writing.

These and other findings, on which I elaborate in the talk, offer valuable insight into the ways academics construct and influence disciplinary norms. This insight is valuable as the practices of academics in such contexts are often sidelined in the literature, which tends to focus instead on students' writing and their engagement with the disciplinary genres and conventions (McGrath et al., 2019). By shifting the attention to academics as I do here, this study adds to our current understanding of the social and cultural views that shape undergraduate writing. In so doing, I also demonstrate the importance of multi-method corpus linguistics for unpacking and situating such complex discourses, by highlighting how the production of these findings required not only the adoption of a corpus-based approach informed by grounded theory, but also, the integration of multimodal textual analysis and corpus-assisted discourse analysis to explore and unpack academics' expectations.

References

- Curry, Niall. *Academic Writing and Reader Engagement*. New York: Routledge, 2021.
- Hyland, Ken. *Academic Discourse*. London: Continuum, 2009.

Lillis, Theresa, Kathy Harrington, Mary R. Lea, and Sally Mitchell, eds. *Working with Academic Literacies: Case Studies towards Transformative practice*. Fort Collins, CO: The WAC Clearinghouse; Anderson, SC: Parlor Press, 2015.

McGrath, Lisa, Raffaella Negretti, Kate Nicholls. "Hidden expectations: scaffolding subject specialists' genre knowledge of the assignments they set." *The International Journal of Higher Education Research*. 78, no.5 (2019): <https://doi.org/10.1007/s10734-019-00373-9>.

Discourses of immigration in human-authored and AI-generated news

Marilisa Shimazumi (Sao Paulo State Technology College, Brazil)

Tony Berber Sardinha (Pontifical Catholic University of Sao Paulo, Brazil)

With the growing use of AI in journalism (Reuters Institute for the Study of Journalism 2024), questions arise about which discourses are put into circulation when news texts are produced by AI agents. Given that these agents rely on large language models whose internal representations are opaque, and they will respond to meet different users' requests, the discourses that enter public circulation through AI-generated news are difficult to predict.

To investigate this, we conducted a comparative corpus-based study of the discourses around the issue of immigration in both human-authored and AI-generated news texts. We extracted 110,873 immigration-related articles published in the United States and the United Kingdom between 2016 and 2020 from the NOW (News on the Web) corpus (Davies 2010-). Articles were filtered to ensure immigration was the primary theme by using GPT to classify the texts in terms of whether the topic of immigration was a central topic to the text or not. The texts were classified by the model and those which were labeled as having immigration as a central topic were retained. From these, a random sample of 300 texts for each quarter of each year was drawn. These texts were classified in terms of their political orientation by GPT. Spot checks were performed which confirmed the reliability of the classification. From this pool, 1,200 texts were selected based on political orientation: 400 left-leaning, 400 right-leaning, and 400 without clear ideological marking. Each of these texts was then summarized by GPT in order to serve as input for the text generation phase. Finally, GPT was tasked with writing three versions of each input text through a prompt that instructed it to read the human text summary and assume one of three political orientations at a time: left-wing, right-wing, and non-partisan. This yielded 3,600 GPT-generated articles, which when added to the 1,200 human texts, resulted in a combined corpus of 4,800 texts and approximately 7.6 million words.

We applied Lexical Multi-Dimensional Analysis (LMDA; Berber Sardinha and Fitzsimmons-Doolan 2025) to identify the discourse dimensions in the corpus. LMDA is an extension of the Multi-Dimensional Analysis framework (Biber 1988) which uses factor analysis to detect recurring lexical sets across the corpus texts. The factors were interpreted as seven dimensions. For instance, Dim. 1 distinguished between discourses of immigration as an existential threat (positive pole) with portrayals of the hard lives of immigrants (negative pole). Dim. 2 centered on discourses of legal procedures to curb immigration (pos. pole) and discourses of geographic and geopolitical aspects of immigration (neg. pole). And Dim. 3 comprised a progressive moral defense of immigration (pos. pole) and a conservative discourse of immigration as a threat (neg. pole). ANOVAs were conducted on the dimension scores, which showed that the difference between AI and human discourses corresponded to up to 41.98% of the variation. Often, the AI reduced ideological asymmetry by producing more ideologically charged versions of the human left-wing texts. For instance, on Dim. 1, human right-wing texts had higher scores than left-wing ones, but the AI-generated texts raised left-wing scores to levels comparable to those of the right. In the paper presentation, we will present all the dimensions and discuss their implications.

References

- Berber Sardinha, Tony, and Shannon Fitzsimmons-Doolan. 2025. *Lexical Multidimensional Analysis: Identifying Discourses and Ideologies*. Cambridge: Cambridge University Press.
- Biber, Douglas. 1988. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Davies, Mark. 2010-. "'Corpus of News on the Web (NOW): 3+ billion words from 20 countries, updated every day.'" <https://www.english-corpora.org/now/>.
- Reuters Institute for the Study of Journalism. 2024. *AI and the future of news*. Reuters Institute for the Study of Journalism (Oxford).

Analyzing Xenophobic Discourses targeting China on X during the COVID-19 pandemic

Cícero Soares da Silva (São Paulo Catholic University, Brazil)

In this paper, we look at malicious representations of China on X (formerly Twitter) occurring in the context of the COVID-19 pandemic. In order to capture these detrimental representations, we scraped a corpus of ca. 100K tweets in Brazilian Portuguese containing ten highly xenophobic hashtags, which were used by right-wing followers to discredit China and spread hatred. The multimodal method followed in this study consisted of a combination of Lexical Multidimensional Analysis (LMDA; Berber Sardinha & Fitzsimmons-Doolan, 2025) and Visual Multidimensional Analysis (VMDA; Soares da Silva, 2025; Berber Sardinha et al., 2023). The LMDA used lexical units to detect traces of discourses across the texts, whereas the VMDA applied computer vision AI techniques to annotate the images posted along with the twitter messages. Two sets of dimensions were obtained (i.e. verbal and a visual). Six verbal dimensions were identified, the first two being: (1) Pandemic manipulation vs Pro-president hashtags: this contrasts the alleged manipulation involving the protection of corrupt officials and misleading pandemic data with the use of hashtags to ridicule China, support presidential policies, and target political adversaries. (2) System and media rejection hashtags vs Anti-China normalization: This captures the use of hashtags to reject the political system and its alleged ties with China, versus efforts to normalize anti-Chinese sentiments under the guise of common sense and cultural distortion. In turn, five visual dimensions were determined, the first two being: (1) China scam denouncement vs. Brazil Elite Accusations: This captures the contrasting positions of denouncing scams from China and accusing Brazilian intellectuals of siding with Chinese interests; (2) Weak local government pandemic response vs. Brave pandemic leadership: This contrasts the views on local governments' responses to the pandemic, either as weak and closure-promoting or brave and against closures seen as benefiting Chinese interests. All the dimensions will be discussed and illustrated in the paper presentation.

References

- Baker, P., & McEnery, T. (2015). Introduction. In P. Baker & T. McEnery (Eds.), *Corpora and discourse studies: Integrating discourse and corpora* (pp. 1–20). Basingstoke, England: Palgrave Macmillan.
- Berber Sardinha, T., & Fitzsimmons-Doolan, S. (2025). *Lexical Multidimensional Analysis*. Cambridge University Press.
- Berber Sardinha, T. (2024). Exploring multimodal corpora in the classroom from a multidimensional perspective In P. Crosthwaite (Ed.), *Corpora for Language Learning: Bridging the Research-Practice Divide* (pp. 25-36). Abingdon: Routledge.
- Collins, L., & Baker, P. (2023). Creating and analysing a multimodal corpus of news texts with Google Cloud Vision's automatic image tagger. *Applied Corpus Linguistics*, 3(1), 100043. <https://doi.org/10.1093/applin/xx>

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.

Soares da Silva, C. (2025, July). Exploring discursive multidimensionality and multimodality on Twitter: Analyzing xenophobic representations targeting China during the COVID-19 pandemic. The Corpus Linguistics 2025 International Conference (CL2025), Aston University, Birmingham, UK

How to kill your darlings: decision making (and decision breaking) in a participatory corpus study about teachers

Kathryn Spicksley (University of Glasgow, UK)

There is growing interest in participatory research methods in linguistics, as scholars work to decolonise their methods and ensure real-world outcomes (see, for example, Stirtz's 2024 edited collection *Participatory Linguistics: Methods and Case Studies from Around the World*). Within the field of corpus linguistics, increasing participation is often constructed as a matter of educating people to use corpus tools. This paper reflects on the limitations, challenges and benefits of heightening participant involvement in corpus-informed academic studies without the requirement that participants learn corpus methods.

The paper reflects on the methodological process of nesting a corpus-informed discourse study within a wider, participatory research project utilising a Delphi Panel method. The Delphi Panel approach involves engaging an anonymous group of experts within a specific field to provide information or feedback on a key issue, with the ultimate aim of achieving consensus. In this case, the focus of the study was to co-develop a mentoring Intervention to be used with new teachers in England, the aim of which would be to open up meaningful conversations about teacher professional identity in public discourse. The researcher and a panel of experts worked together to co-design a deck of cards to be used in mentoring situations with new teachers; each card illustrated a different professional 'identity' made evident through the discourse analysis undertaken in the research project. Expert panel members were recruited from teaching communities which were intended to benefit from the development of the mentoring cards. The project was initially inspired by the Metaphor and Cancer project, which utilised corpus methods to develop a real-world resource for health professionals.

The cards were informed by a corpus-assisted discourse analysis of the Teachers and Teaching in England Corpus (T&TEC). Utilising online questionnaires, expert panel members were involved in discussions around:

- corpus building;
- initial decision making concerning how to 'enter' the corpus;
- how to stratify and compare subcorpora;
- analytical focus (presence or absence of words, familiar or strange collocations, expected or unexpected linguistic patterns, foregrounded and backgrounded discourses);
- the utility of metaphorical constructions of teaching.

Panel members were then invited to comment on and discuss findings with each other through the Padlet online whiteboard platform. The paper evidences how this process heightened critical awareness amongst participants of the discursive practices which inform their professional identity.

The wider societal relevance of the study lies in its response to a deepening recruitment and retention crisis amongst teachers in England, and indeed internationally. Discourse-oriented research in education studies has repeatedly shown that constructions of teachers and

teaching within the public space produce unstable and contradictory professional identities, which new teachers struggle to effectively negotiate. The participatory nature of the project is located in a resistance to educational policy initiatives which position teachers as deficient, both intellectually and creatively. As such, the paper will reflect on the ethical imperative which drives participatory research, making the case for linguistic research which authentically and meaningfully engages the subjects of discourse in analysing, comparing and critiquing the language which positions them.

Measuring Division: The Language of Dissent in the Supreme Court and American Polarization

Rohan Srivatsa, Katherine Ireland, Sophia Nguyen, Jason Eappen, Meryem Acioz
(University of Georgia, USA)

Political polarization in the United States has been widely examined through legislative behavior, electoral politics, and mass political discourse. Far less attention, however, has been paid to how polarization is discursively constructed within elite judicial institutions. This paper examines how ideological polarization is enacted and negotiated through the language of dissenting opinions in the U.S. Supreme Court. Rather than treating polarization solely as a function of voting behavior or ideological scores, the study adopts a corpus-assisted discourse analytic approach to examine how disagreement is patterned linguistically over time.

Dissenting opinions are approached as a distinct institutional discourse genre. While they permit explicit disagreement, they are also constrained by norms of judicial legitimacy, professionalism, and institutional authority. This tension makes dissents a particularly revealing site for examining how ideological conflict is rhetorically managed within the Court. The analysis focuses on how stance, evaluation, and rhetorical framing in dissenting language shift across judicial eras, asking whether these changes align with, or diverge from, broader trends in political polarization.

The study draws on a diachronic corpus of Supreme Court dissenting opinions spanning the nineteenth century to the present, constructed using digitized judicial texts from the Caselaw Access Project. Dissents are systematically extracted, cleaned, and annotated with metadata including year of decision, Court era, justice identity, and ideological orientation. This enriched corpus enables longitudinal analysis of discursive change as well as comparative analysis across ideological groupings, allowing linguistic patterns to be interpreted in relation to institutional and historical context.

Methodologically, the project integrates quantitative corpus techniques with discourse-analytic interpretation. Corpus-driven keyword analysis, measures of evaluative stance and sentiment, lexical diversity, modality, and syntactic complexity are used to identify recurrent patterns of disagreement. These are complemented by semantic similarity modeling to explore how ideological distance is reflected in broader semantic framing. Dictionary-based measures are employed cautiously and interpreted alongside corpus-derived patterns in order to capture both overt and more strategic forms of evaluative positioning. Quantitative findings are treated as prompts for interpretation rather than endpoints, with emphasis placed on how linguistic patterns function within the institutional genre of dissent.

To ensure methodological transparency and robustness, the analytical pipeline was first tested on synthetic corpora designed to simulate ideological and rhetorical variation. This proof-of-concept stage confirmed that the corpus-based measures are sensitive to structured linguistic divergence prior to application to judicial texts. The validated framework is then applied to Supreme Court dissents to trace discursive change over time and to situate these patterns alongside external indicators of political polarization, such as congressional DW-NOMINATE scores.

By treating judicial writing as both data and discourse, this study contributes to research on institutional discourse, discursive change over time, and methodological practice in corpus-assisted discourse analysis. More broadly, it demonstrates how corpus methods can be productively integrated with interpretive analysis to examine ideological conflict within elite, high-stakes institutional settings.

Analysing stance in parliamentary briefings through semantic frames

Xinmei Sun (Lancaster University, UK)

This paper focuses on briefing reports produced by the UK House of Commons Library, an institution that explicitly presents its work as an impartial research and information service for MPs “regardless of political affiliation” (Commons Library, n.d.). Given the Commons Library’s central role in legal and political decision-making, it is important to assess whether they construct an institutional voice that is genuinely impartial, or whether particular actors and perspectives get foregrounded or backgrounded over time.

This study analyses a corpus of 235 briefing reports on immigration produced by the Commons Library between 1999-2024. The corpus is annotated with frame-semantic information using LOME (Large Ontology Multilingual Extraction) (Xia et al., 2021). Frame Semantics models the structured categories of experience that are evoked by lexical items (e.g., STATEMENT, POSSIBILITY, OBVIOUSNESS, DESIRABILITY) and the conceptual elements that they specify such as SPEAKER, MESSAGE, POSSIBLE_EVENT, EXPLANATION, etc. These frames and frame elements provide a principled basis for analysing stance-related categories such as attribution, modality, evidentiality and explicit evaluation across a large dataset.

Through data pre-processing and rule-based post-editing, the automatically generated annotations are refined to near-ceiling recall and accuracy, providing a reliable basis for downstream corpus analysis. Quantitative analyses of frame and frame-element frequencies, and dispersion patterns are integrated with qualitative concordance-based readings to investigate how immigrants, policymakers and institutions are represented; which social actors or policy positions are granted voice, credibility and/or responsibility; and how explicit and implicit evaluations are managed within a genre that presents itself as impartial.

The study also explores change over time by tracing longitudinal patterns in stance-related frames and frame elements relative to changes in the governing party. This allows an assessment of whether the Library’s institutional stance remains stable or shifts in line with broader political developments.

The paper thus offers, first, an empirical case study of parliamentary institutional discourse on immigration and its impartiality practices; and second, a methodological contribution to corpus-based discourse analysis by demonstrating how frame-semantic parsing can be used as a scalable, systematic resource for corpus-based discourse analysis, and how semantic frames facilitate the analysis of stance among other discourse phenomena.

References

- Coschignano, S., G. Minnema, and C. Zanchi. 2023. “Explaining the Distribution of Implicit Means of Misrepresentation: A Case Study on Italian Immigration Discourse.” *Pragmatics* 213: 107–125. <https://doi.org/10.1016/j.pragma.2023.06.002>.
- Fillmore, Charles J. 1982. “Frame Semantics.” In *Linguistics in the Morning Calm: Selected Papers from SICOL-1981*, edited by Linguistic Society of Korea, 111–137. Seoul: Hanshin.

House of Commons Library. n.d. "About Us." Accessed October 7, 2024.

<https://commonslibrary.parliament.uk/about-us/>.

Xia, Patrick, Guanghui Qin, Swarnima Vashishtha, Yanda Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron Steven White, and Benjamin van Durme. 2021. "LOME: Large Ontology Multilingual Extraction." In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations, 149–159. Association for Computational Linguistics.

<https://doi.org/10.18653/V1/2021.EACL-DEMOS.19>

Fake News, Cancel Culture, and the limits of the sayable: Towards a theory and typology of discourse on discourse

Isabelle Suremann (Zürcher Hochschule für Angewandte Wissenschaften, Switzerland)

Metalanguage is a well-researched topic in linguistics with different, partly overlapping frameworks, such as metapragmatics (Spitzmüller 2014; Silverstein 2021), folk linguistics (Niedzielski and Preston 2010), or language ideologies (Vessey 2016). A newer focus lies on statements about discourse and its conditions; that is, statements about the role of language in mediating social reality and constructing common sense, (perceived) limits of what can be said, and the power relations behind discourse. They include references to political correctness (Schroeter 2025), the limits of the sayable and their shifts (Niehr 2018), truth and deception (Schürgers 2025), and access to discourse and voice (Schröter and Jung 2024). Despite the clear relevance of these statements for discourse studies, a theory describing their linguistic characteristics and functions, as well as a comprehensive typology, have yet to be established. Instead, current analyses focus mainly on specific meta-discursive statements or formulas.

Crucially, discourse and discourse on discourse, or meta-discourse, can be understood to be in a reciprocal and co-constructive relationship, analogous to the relationship between language and metalanguage (Spitzmüller 2019). Therefore, these individual statements share an underlying function that contributes to the (re-)structuring of orders of discourse, discursive meaning-making, and common sense, forming a continuous cycle in which discursive statements elicit meta-discursive statements that, in turn, influence subsequent discursive statements.

The project presented here applies a mixed-methods approach combining logistic regression analysis with keyword and concordance analysis to identify linguistic features, functions, and categories of meta-discursive statements within the vaccination discourse in German-speaking Switzerland from 2000 to 2025. The presentation will focus on the keyword and concordance analysis of a specialised corpus consisting exclusively of texts with meta-discursive statements, allowing for the systematic identification of relevant lexis and categories of meta-discursive statements.

This project offers an initial step towards a theory and typology of meta-discourse. However, more empirical and theoretical work is needed to fully understand its characteristics and its central role in shaping orders of discourse and common sense.

References

Niedzielski, Nancy A., und Preston, Dennis R. 2010. *Folk Linguistics*. De Gruyter Mouton.

Niehr, Thomas. 2018. *Das Unsagbare sagbar machen. Immunisierungsstrategien im öffentlichen Diskurs*. De Gruyter Mouton. Print.

Schroeter, Melani. 2025. "Political correctness gone woke in polarised society: The emergence of a new keyword in an old 'culture war' discourse". *Journal of Corpora and Discourse Studies* 8 (Mai): 123–58.

- Schröter, Melani, und Theo Jung. 2024. "Speaking up and being heard: The changing metadiscourse about 'voice' in British parliamentary debates since 1800". *Language & Communication* 94 (Januar): 41–55.
- Schürgers, Elise. 2025. "Meta-discursive Formulas A Subtype of Keywords and its Pragmatic Effects: The Case of 'Fake News' in the Francophone Press". *Journal of Corpora and Discourse Studies* 8 (Mai): 217–40.
- Silverstein, Michael. 2021. "The Dialectics of Indexical Semiosis: Scaling up and out from the "Actual" to the "Virtual"". *International Journal of the Sociology of Language*. *International Journal of the Sociology of Language* 2021 (272): 13–45.
- Spitzmüller, Jürgen. 2014. "Metapragmatik, Indexikalität, soziale Registrierung". *Zeitschrift für Diskursforschung*, Nr. 3: 263–87.
- Spitzmüller, Jürgen. 2019. "'Sprache' – 'Metasprache' – 'Metapragmatik': Sprache und sprachliches Handeln als Gegenstand sozialer Reflexion". In Gerd Antos, Thomas Niehr und Jürgen Spitzmüller (Hgg.) *Handbuch Sprache im Urteil der Öffentlichkeit*. De Gruyter.
- Vessey, Rachelle. 2016. *Language and Canadian Media*. Palgrave Macmillan UK.

Explaining poverty: A discourse analysis of economic inequality in Swiss media discourse

Isabelle Suremann, Dario d'Agostino, Julia Krasselt and Philipp Dreesen
(Zurich University of Applied Sciences, Switzerland)

Even though social mobility is considered high in Switzerland (OECD 2018), economic inequality remains a persistent issue. In 2023, 1.4 million people in Switzerland were living in or close to poverty (Federal Statistical Office 2025). Poverty harms individuals in multifaceted ways beyond the immediate material deprivation, among others in respect to (mental) health (see, e.g., Mckenzie 2023). Moreover, how poverty and people living in poverty is represented in public discourse can lead to stigmatisation and social exclusion (Jensen and Tyler 2015, Tyler 2022, Kemper and Weinbach 2025). Despite its social significance, the issue of poverty in Switzerland remains under-researched, particularly from discourse-analytical perspectives (with few exceptions, e.g., Rosenberger Staub et al. 2019).

This paper explores how issues of economic inequality are represented and explained in German-, French-, and Italian-language media from Switzerland. Combining corpus linguistic methods with qualitative analyses, it addresses the following questions: 1) How is the semantic space around poverty and recipient of social assistance organised? 2) How are issues of economic inequality explained? 3) Which theories of class and economic inequality are thereby reproduced? 4) How does the discourse converge or diverge across the different language regions?

Using poverty and recipient of social assistance as seed words, the paper first maps the semantic space around these queries and compiles corpora from German-, French-, and Italian-language media from Switzerland. Fifty representative texts per corpus are analysed qualitatively to discover how economic inequality is explained, namely, which reasons are given (e.g., systematic vs. personal failures) and which economic or social theories are invoked. These findings are then deductively scaled to each corpus to assess whether they constitute recurrent language use patterns (Bubenhofer 2009) and thereby contribute to the “incremental effect of discourse” (Baker 2006: 13). Finally, the findings are compared across the corpora; particularly differences can be meaningful, as they may reveal discursive absences (Schröter and Taylor 2018).

Ultimately, this paper contributes to a broader understanding of how economic inequality is represented and explained in Swiss media discourse, and how the identified language patterns shape and solidify the legitimisation and naturalisation, or challenging and questioning, of economic inequality.

References

- Baker, Paul. 2006. *Using Corpora in Discourse Analysis*. First edition. Continuum Discourse Series. Zed Books.
- Bubenhofer, Noah. 2009. *Korpuslinguistik Als Methode Der Diskurs- Und Kulturanalyse*. Herausgegeben von Ekkehard Felder. Bd. 4. Sprache Und Wissen. Walter de Gruyter.

- Federal Statistical Office. 2025. „Poverty Rates - 2023 | Chart“. März 31.
<https://www.bfs.admin.ch/asset/en/34087848>.
- Jensen, Tracey, und Imogen Tyler. 2015. “Benefits Broods’: The Cultural and Political Crafting of Anti-Welfare Commonsense”. *Critical Social Policy* 35 (4): 470–91.
- Kemper, Andreas, und Heike Weinbach. 2025. *Klassismus*. Unrast Verlag.
- Mckenzie, Lisa. 2023. *Getting by: Estates, Class and Culture in Austerity Britain*. Policy Press.
- OECD. 2018. *A Broken Social Elevator? How to Promote Social Mobility*. OECD.
- Rosenberger Staub, Nicole, Philipp Dreesen, Julia Krasselt, u. a. 2019. *Linguistische Diskursanalyse zu «Sozialhilfe»*. Sozialdepartement der Stadt Zürich.
<https://www.stadt-zuerich.ch/sd-diskursanalyse>.
- Schröter, Melani, und Charlotte Taylor, Hrsg. 2018. *Exploring Silence and Absence in Discourse: Empirical Approaches*. *Postdisciplinary Studies in Discourse*. Palgrave Macmillan.
- Tyler, Imogen. 2022. *Stigma: The Machinery of Inequality*. Bloomsbury Academic.

Conveying the crisis: A corpus-assisted analysis of how left- and right-leaning UK newspapers frame the green energy transition

Yingnian Tao and Mark Ryan (Lancaster University, UK)

The green energy transition has recently become a hot topic in public and policy debates in the UK. This study employs corpus-assisted analysis to examine how the green energy transition is represented in left-leaning and right-leaning UK newspapers (2015-2024). Specifically, we identified the temporal changes in frequencies of emotional reporting and analysed and compared five groups of strong emotion words (represented by crisis, damage, greenwashing, catastrophe, and wealth) through collocation analysis and concordance analysis. Our findings indicate that although the sentiment of coverage is more positive than negative, negative sentiment is gradually increasing over time. Both sides cover environmental and economic aspects of the transition; however, they differ in their priority area and the way they engage with these areas. Left-leaning newspapers emphasise environmental degradation, blame fossil fuels for the climate crisis, and call for urgent action. Right-leaning newspapers focus more on the economic impacts of transition, often framing policies as threats to jobs or prosperity while downplaying accountability for emissions. Across the spectrum, reporting tends to adopt a techno-optimist, business-as-usual stance that rarely challenges the idea of continuous economic growth. Implications for media, readers, and policymakers are discussed.

Fear and pride: Emotion talk in UK migration debates

Charlotte Taylor (University of Sussex, UK)

Inward migration has been – and continues to be – extensively politicised in the UK context and beyond. In this paper, I focus on one element of public representations of migration; the way in which emotions are invoked persuasively in institutional discourses on migration. Following Bednarek (2008), I use the term ‘emotion talk’ to refer to talk about emotions, as opposed to ‘emotional talk’ which is talk in which emotions are displayed. Both ‘emotion talk’ and ‘emotional talk’ can perform persuasive appeals to emotions or discursive (de)legitimation (Reyes 2011). To date, ‘emotional talk’ has received more attention, given the centrality of emotion to conceptualisations of populism, and in parallel, the key role of securitisation in discursive framing of mobility – both of which lead to a key role of ‘fear’ in anti-immigration rhetorics. In this paper, I focus attention on the use of ‘emotion talk’ and identify which emotions are explicitly referenced in UK parliamentary debates on migration using the open-access Parlamin corpus which contains all transcripts over the period 2015-2021. The analysis addresses the discursive function of these mentioned emotions. Investigating function rather than just frequency means that a nuanced approach is required. The concordance lines of emotion talk (e.g. ‘proud’ or ‘grateful’) were coded according to a) who was speaking, b) who was described as experiencing this emotion, and c) what was positioned as the trigger or target of the emotion (e.g. what people are described as being ‘proud’ of). The findings show there is little difference along political lines in quantity of emotional talk collocates – despite the fact that affective legitimation has been strongly associated with populist discourse and, for part of the time period of this corpus, there was a populist prime minister. Furthermore, it shows a wide range of emotions are given prominence in this context; while there is a preponderance of negative emotions such as fear and shame, a key rhetorical role in framing the ‘(un)deserving migrant’ is played by talk of seemingly positive emotions such as pride, gratitude and tolerance.

The Grammar of Victimization Narratives in Incel Discourse: A Cognitive-Corpus Analysis

Antonio-Jesús Tinedo-Rodríguez (University of Córdoba, Spain)

Incel discourse represents one of the most extreme forms of misogyny in the digital “manosphere.” In their forums and communities, incels construct narratives in which they present themselves as sexually deprived men (“involuntary celibates”) while attributing responsibility for their suffering to women. This study focuses on the dehumanization strategies embedded in this discourse, with particular attention to the cognitive and emotional load of verbs co-occurring with specific incel terminology (e.g., femoid, chad, blackpill, Stacy, etc.) This term serves as a key resource for understanding how hatred toward women is constructed through incel rhetorics, projecting traits such as immorality, coldness, or lack of agency.

The research is grounded in cognitive and corpus linguistics and conceptual metaphor theory, as well as Haslam’s (2006) dehumanization taxonomy, which distinguishes between animalistic dehumanization (attributing lack of morality, self-control, or rationality) and mechanistic dehumanization (attributing passivity, emotional coldness, or fungibility). Verbs are central units of analysis because they structure events and agent roles, making them crucial for identifying patterns of victimization and implied violence in discourse.

The INCELCOR2 corpus, composed of publicly available forum posts, was compiled for analysis. A mixed-methods approach was employed. First, all verbs co-occurring with terms which are proper of incel discourse were automatically extracted and analysed in two syntactic configurations: (a) term as subject, and (b) term as object. Measures of frequency and association will be calculated to identify relevant patterns. Second, constructions were manually annotated according to Haslam’s (2006) taxonomy, classifying the dehumanization strategies activated in context. This combination of quantitative and qualitative analysis enables the identification of lexical-grammatical patterns alongside their conceptual projections and potential emotional triggers.

The main objective is to map how incel discourse constructs women as agents of harm or as objects of verbal violence, reinforcing the self-perceived victimhood of incels. The study also aims to demonstrate that verbs are a privileged resource for detecting and analysing hate speech, as they encode power dynamics and agency. This research contributes to understanding the interplay of language, cognition, and emotion in digital misogyny and provides conceptual and methodological tools that may support early detection systems for online hate speech, as well as interdisciplinary educational interventions designed to counteract narratives of violence and victimization.

References

Haslam, Nick. (2006). Dehumanization: An Integrative Review. *Personality and Social Psychology Review*, 10(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

The role of “sociopolitical keywords” in discursive depoliticization: the case of the terms fascism and responsibility

Erin Turneer (UCLouvain, Belgium)

This presentation focuses on the preliminary findings of a research project examining the role of “sociopolitical keywords”: terms such as populism, woke, or fake news that, in some of their uses, have the “capacity to ‘encompass’ or [...] ‘crystallize’ a whole set of ideas, values or political issues” (Hambye et al., 2025). The project seeks to deepen our understanding of these keywords by showing that they can play a crucial role in the language of persuasion (Partington, 2017) and, more specifically, in discursive depoliticization – “expression of a failure [...] in the discursive articulation of social conflicts” by actors tasked with “sustaining democracy” (Bonnet et al., 2024; our translation) – characterized by a set of strategies that contribute to “closing [...] debates” (Maesele & Ræijmaekers, 2020). The project is therefore structured around two research questions: (1) What roles do sociopolitical keywords play in political discourse? (2) What is their impact on processes of persuasion and discursive depoliticization?

To address these questions, we conduct a qualitative and collocational analysis of a corpus of authentic uses of eight sociopolitical keywords (responsibility, diversity, discrimination, fascism, extremism, inclusion, immobilism, transparency) extracted from political speeches delivered in the parliaments of French-speaking Belgium and France (2022-2025). To provide a diversified and generalizable overview of different types of keywords, these eight cases were selected according to several criteria established on the basis of previous research: their political connotation; their positive or negative evaluative valence; and their reference either to attitudes that political actors are expected to adopt or avoid in the conduct of political action (e.g. being responsible, transparent), or to ideologies or societal projects that political actors may defend or oppose (e.g. diversity, fascism).

The data are analyzed using an analytical framework inspired by Rondiat (2025). For each occurrence of one of the eight keywords, we identify its context, meaning, evaluative charge, argumentative function, weight and pattern, the targeted referent(s), and whether the use of the term—and the utterance in which it appears—can be characterized as depoliticized or not.

This presentation reports on the results of an initial, ongoing study comparing the uses of fascism and responsibility within our parliamentary corpus, which includes more than one hundred occurrences for each keyword. This study addresses research questions (1) and (2) for these two cases. We hypothesize that sociopolitical uses of these keywords recurrently follow argumentative patterns that contribute to forms of discursive depoliticization, notably through the naturalization of certain political positions or the disqualification of dissent. The comparative analysis aims to highlight these mechanisms and their variations, and to show that, despite their specificities, these terms belong to the same category of discursive phenomena, characterized by their function in the language of persuasion.

References

Bonnet, V., Marty, E., & Robert, C. (2024). Mécaniques de la dépolitisation. *Mots. Les langages du politique*, n° 134(1), 198. Cairn.

- Hambye, P., De Cock, B., & Rondiat, C. (2025). Studying keywords in discourse: Contrasting different theories and methods. *Journal of Corpora and Discourse Studies*, 8, 1-20. <https://doi.org/10.18573/jcads.149>
- Maesele, P., & Ræijmaekers, D. (2020). Nothing on the news but the establishment blues? Toward a framework of depoliticization and agonistic media pluralism. *Journalism*, 21(11), 1593-1610. <https://doi.org/10.1177/1464884917739476>
- Partington, A. (2017). *The Language of Persuasion in Politics: An Introduction*. Routledge. <https://doi.org/10.4324/9781315177342>
- Rondiat, C. (2025). 'Following the science'. Science* as a sociopolitical keyword during the pandemic | *Journal of Corpora and Discourse Studies*. <https://doi.org/10.18573/jcads.138>

The Evolution of Ecolanguage in Fashion Media Communication: A Corpus-Based Discourse Analysis

Agnese Ughetti (University of Ferrara, Italy)

Since the 19th century, the fashion industry has been transformed by the Industrial Revolution and later by the globalization of the supply chain. This evolution has led to the mass production of clothing, which has democratized fashion but, at the same time, has raised environmental dilemmas. As a result, sustainability has become a key theme in fashion communication, although it remains controversial, as vague green messages often provoke consumer skepticism (Pearson-Smith & Evans 2017).

The aim of this study is to shed light on the opacity of fashion ecolanguage by reconstructing its diachronic evolution in fashion media. Therefore, this research situates itself within the framework of Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) (Partington 2010), examining a specialized corpus to investigate how social and cultural changes are reflected in the language of fashion media. More specifically, this study addresses two main questions: (i) how has ecolanguage evolved from the late 19th century to the present day? (ii) what role do fashion media play in shaping green discourses in the fashion system?

Using the ProQuest database, an ad hoc corpus was compiled by selecting articles containing the terms *ecolog** and *sustainab** from the fashion magazine *American Vogue*, covering the period from 1965 to 2025. The linguistic analysis was carried out using #Lancsbox X (Brezina & Platt 2025) and was performed in two steps. First, a concordance and collocation analysis was conducted on a predetermined list of ecolanguage terminology (Thomas 2008), which was subsequently expanded to include more recent terms. The findings confirmed the actual relevance of these terms, although to varying degrees, and that *Vogue* generally tends to favor expressions with positive connotations. Furthermore, a detailed collocation analysis of fashion highlighted how sustainable fashion evolved from being loosely linked to ecology to becoming a ubiquitous term that often triggers consumer skepticism. Secondly, a keyword analysis was conducted to examine the diachronic evolution of topics associated with the query terms. The corpus was divided into six subcorpora, each spanning approximately ten years, and keywords were identified by comparing each subcorpus with the remainder of the corpus (Baker et al., 2020). These keywords were then grouped into macro-categories to facilitate comparison across subcorpora. This analysis showed that, prior to the 1990s, the query terms were associated with a broader range of macro-categories, including Travel, Art & Culture, and Food, whereas in more recent subcorpora this diversity has decreased. Focusing on the keywords of the Fashion & Beauty macro-category, the analysis revealed that sustainability, initially tied to cosmetics, gradually became increasingly associated with clothing.

In conclusion, the application of the MDCADS approach has helped bridge the research gap concerning the documentation of ecolanguage evolution in the fashion industry. The analysis also showed that this terminology has grown more complex and multidimensional over time. Therefore, it is crucial to continue monitoring how real-world developments shape and influence this evolving lexicon.

References

- Baker, Paul, Gavin Brookes, Dimitrinka Atanasova, and Stuart W. Flint. 2020. "Changing Frames of Obesity in the UK Press 2008–2017." *Social Science & Medicine* 264.
- Brezina, Vaclav, and William Platt. 2025. #LancsBox X [software]. Lancaster University.
- Partington, Alan. 2010. "Modern Diachronic Corpus-Assisted Discourse Studies (MD-CADS) on UK Newspapers: An Overview of the Project." *Corpora* 5 (2): 83–108.
- Peirson-Smith, Anne, and Susan Evans. 2017. "Fashioning Green Words and Eco Language: An Examination of the User Perception Gap for Fashion Brands Promoting Sustainable Practices." *Fashion Practice* 9 (3): 373–97.
- Thomas, Sue. 2008. "From 'Green Blur' to Ecofashion: Fashioning an Eco-Lexicon." *Fashion Theory* 12 (4): 525–39.

Disbelief, dismissal, denial, diversion: Using CADS to identify quantitative and qualitative absences in discourses relating to sexual violence

Rachelle Vessey (Carleton University, Canada)

In Partington's (2014) corpus linguistics-informed typology, the notion of "relative absence" refers to "what is entirely or relatively absent – what is either missing or rare – in one set of texts" (p. 129). This paper proposes two subtypes of "relative absence": quantitative relative absence in terms of words and phrases that are underused in a target corpus, and qualitative relative absence in terms of the strategies employed to discursively erase a social issue (or discursively construct a non-issue). Qualitative absence is examined via the words and phrases that are used in a target corpus instead of those that were expected but are absent (on expectations and absence, see Duguid and Partington, 2016, p. 40).

This paper focuses on the "relative absence" of sexual violence in a corpus of Canadian male military members' discourse (1.1 million words). Using SketchEngine, the procedure begins with the identification of quantitative absence via keyword and key cluster procedures, using two separate comparator corpora (311K words of female military members' discourse and the 1.2 billion-word Canadian subdomain of EnTenTen21). Although sexual violence is well-documented in the Canadian military, with females especially targeted (e.g., Duval-Lantoin, 2022), the keyword procedure reveals that males underuse sexual violence terminology in comparison with one or both comparator corpora. Thus, the focus on qualitative absence pivots on the question of what is being talked about instead of sexual violence. Themes are identified within the top 100 key words and clusters generated by each comparison, using collocates and concordance lines to establish meanings. Four main themes emerge, each enabling male military members to circumvent talk about sexual violence. (1) Keywords of quantification (e.g., 'average') and technical nominalizations (e.g., 'readiness') construe male expertise and pseudo-empiricism. Since the female data is where first-person accounts of sexual violence are provided, the male pseudo-empiricism facilitates disbelief in the veracity of female testimony, the style of which seems comparatively subjective and biased. (2) The keyword 'human' indexes how male respondents hail universal neutrality and equality as the answer to perceived problems of diversity and division plaguing the military workplace; however, a focus on humanism dismisses gender as a contributing factor in contexts where sexual violence occurs. (3) The keyword 'joke' raises the function of (sexist/sexualized) humour as a form of "indirect sexism" (Mills, 2008) whereby men contribute to female isolation in the male-dominated workplace while denying responsibility for this sexism. (4) The keywords 'harassment' and 'misconduct' occur in contexts where male respondents divert from discussions of sexual harassment and sexual misconduct towards other types of impropriety.

The identification of disbelief, dismissal, denial, and diversion helps to explain why sexual violence is relatively absent in the male data – i.e., it tends to be either erased or discursively constructed as a non-issue. While this paper focuses on the particulars of only one discursive context, the examination of qualitative absences may be relevant to other comparative CADS research.

References

- Duguid, A., and A. Partington. 2018. "Absence. You Don't Know What You're Missing. Or do you?" In *Corpus Approaches to Discourse*, ed. C. Taylor & A. Marchi, 38-51. Routledge.
- Duval-Lantoine, C. 2022. *The ones we let down: Toxic leadership culture and gender integration in the Canadian Forces*. McGill-Queen's University Press.
- Mills, S. (2008). *Language and sexism*. Cambridge University Press.
- Partington, A. 2014. "Mind the Gaps: The Role of Corpus Linguistics in Researching Absences." *International Journal of Corpus Linguistics* 19(1): 118–146.

The human-water relationship in 19th century fiction

Andreas Wagner (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany)

Michaela Mahlberg (Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany & University of Birmingham, UK)

The climate crisis confronts us with radical uncertainty (Johnson et al. 2023). Scientific predictions of weather extremes are increasingly challenging while anticipated changes in water scarcity, quality, and security (Lee et al. 2024) require urgent responses. From a societal point of view, it is becoming very complex to envision humanity's future. Perceptions that help people navigate this uncertainty are shaped by public narratives that align with their thoughts and beliefs (Beach and Wise 2022). Such 'base narratives' draw on different sources. In this paper, we argue that 19th-century fiction is one such source, to answer the question: What base narratives for the human-water relationship are presented in 19th century fiction?

We use a corpus of Charles Dickens' Novels and a general 19th-century reference corpus, overall 44 novels (Mahlberg et al. 2020). While Dickens is known for his depictions of human-environment relationships, the 19th-century corpus offers a broader perspective. In line with studies on corporate communication (Smith et al. 2025) and comparisons across different types of socially relevant texts (Mahlberg et al. forthcoming), we aim to identify textual patterns that provide insights into meanings of water. To focus on the portrayed human-water relationship, we semantically tagged the corpora for movement, weather and environmental issues (Rayson et al. 2004) and identified 'geonouns', such as rivers and boats (Steiner et al. 2023), pointing to water relations. In this way, we generated 'candidate sentences' for a two-step analysis. First, we qualitatively examined three novels to identify ad hoc categories of human-water relationships. Second, we used a random sample of 400 candidate sentences from the entire two corpora. To refine our categories of 'human-water relationships', we qualitatively studied concordance lines of the candidate sentences.

Our findings indicate that water extremes do not figure prominently in the corpora. Instead, references to waterbodies and domestic water are key components of base narratives. Waterbodies as spatial deixis and the distinction between man-made and natural features are crucial in positioning humans within their environment. Within the radical uncertainty of climate change, the base narratives of human-water relationships highlight the importance of water as a clearly perceivable feature of the environment of individuals, in contrast to large-scale water extremes. The absence of water extremes narratives aligns with research calls to connect these across disciplines (Kosow et al. 2022). Thus, our study forms the basis for a larger project that compares base narratives in classical novels and contemporary climate fiction.

References

Beach, Lee Roy, and James A. Wise. 2022. *The Theory of Narrative Thought*. Cambridge Scholars Publisher.

Johnson, Samuel G. B., Avri Bilovich, and David Tuckett. 2023. 'Conviction Narrative Theory: A Theory of Choice under Radical Uncertainty'. *Behavioral and Brain Sciences* 46: e82.

- Kosow, Hannah, Sabrina Kirschke, Dietrich Borchardt, et al. 2022. 'Scenarios of Water Extremes: Framing Ways Forward for Wicked Problems'. *Hydrological Processes* 36 (2): e14492.
- Lee, Hoesung, Katherine Calvin, Dipak Dasgupta, et al. 2024. 'Climate Change 2023 Synthesis Report Summary for Policymakers'. *CLIMATE CHANGE 2023 Synthesis Report: Summary for Policymakers*.
- Mahlberg, Michaela, Peter Stockwell, Viola Wiegand, and Jamie Lentin. 2020. 'CLiC 2.1 Corpus Linguistics in Context'. <https://clic-fiction.com/>.
- Rayson, Paul, Dawn Archer, Scott Piao, and Tony Mcenery. 2004. 'The UCREL Semantic Analysis System'. January.
- Smith, Justine, Thomas Bowerman, Becky Elliott, and Charlotte Taylor. 2025. Framing Water for Future Action: Corporate Communications. *Journal of Corpora and Discourse Studies*, no. 9.
- Steiner, Erik, Zephyr Frank, Ian Gregory, David Bodenhamer, and Ignatius Ezeani. 2023. Spatio-Textual Regions: Extracting Sense of Place from Spatial Narratives. August 26."

Right-wing pandemic crisis discourse: A case study of Bolsonaro's live broadcasts

Mirella Whiteman, Aline Zamboni, Arianne Brogini and Marcos Oliveira
(Pontifical Catholic University of São Paulo, Brazil)

During the COVID-19 pandemic, several right-wing political leaders adopted communication strategies that reduced direct engagement with the press and shifted public messaging toward social media. These typically combined skepticism toward public health restrictions with appeals to law and order and encouragement for citizens to maintain ordinary routines despite escalating health risks. Right-wing leaders such as Trump (US), Modi (India), and Johnson (UK) initially downplayed the severity of the pandemic, but their subsequent discourses diverged (Taraktaş, Esen, and Uskudarli 2024). Trump largely maintained a campaign-oriented, self-referential style, using Twitter to praise his own leadership and attack opponents. By contrast, Johnson and Modi revised their communication as the crisis intensified: Johnson reoriented his messaging by emphasizing public health measures and national solidarity, whereas Modi incorporated pandemic-related themes into a nationalist discourse. In general, right-wing populist leaders seem to share a preference for communication where domestic political contexts are mobilized for polarization and self-promotion (Taraktaş, Esen, and Uskudarli 2024). One prominent right-wing leader during the pandemic is Brazil's former President Jair Bolsonaro. As the pandemic progressed, he replaced traditional press briefings with weekly YouTube livestreams that became his primary channel of public communication. Through these broadcasts, he articulated a discourse centered on institutional authority and personal freedom and opposed measures he framed as ideological impositions, while downplaying the severity of the health crisis even as Brazil experienced very high COVID-19 mortality rates.

The goal of the current paper is to detect the discourses used by Bolsonaro during his YouTube livestreams in order to assess whether they conform to the general discourses observed among right-wing leaders during the COVID-19 pandemic as suggested in the literature. The study analyzes a corpus of 100 Bolsonaro's live broadcasts spanning 2020 to 2022 using Lexical Multidimensional Analysis (Berber Sardinha and Fitzsimmons-Doolan 2025). Transcriptions were made manually and tagged with TreeTagger; the counts of content word lemmas were normed and subjected to factorial analysis using SAS software. Three dimensions emerged: 1) Electoral polarization as moral conflict and conservative patriotism; 2) Infrastructure as political legitimacy and pragmatic governance; and 3) Economic hardship and blame shifting. The dimensions capture central features of contemporary conservatism such as nationalism, institutional skepticism, distrust of mainstream media, and moral polarization (van Dijk 2024). However, they also indicate that the discursive realization of right-wing ideologies varies according to domestic political conditions and electoral priorities. Bolsonaro's discourses closely resemble Trump's campaign-centered and confrontational discourses. At the same time, unlike Johnson and Modi, he did not recalibrate his messaging toward solidarity or collective risk management as the crisis intensified, despite escalating mortality. The Bolsonaro case shows right-wing crisis discourse to be flexible: leader-centered communication and moralized political oppositions remain constant, while the acknowledgment of public health risk and the use of the pandemic as an electoral instrument vary.

References

- Berber Sardinha, Tony, and Shannon Fitzsimmons-Doolan. 2025. *Lexical Multidimensional Analysis: Identifying Discourses and Ideologies*. Cambridge: Cambridge University Press.
- Taraktaş, Başak, Berk Esen, and Suzan Uskudarli. 2024. "Tweeting through a public health crisis: Communication strategies of right-wing populist leaders during the COVID-19 pandemic." *Government and Opposition* 59: 207-228.
- van Dijk, Teun A. 2024. *Discourse and ideologies of the radical right*. Cambridge: Cambridge University Press.

Disputed Museum Labels Through the Lens of CADS: Triangulating Corpus and Reception Analysis

Xioayu Xu (The Education University of Hong Kong, Hong Kong)

Cecilia Lazzeretti (Free University of Bolzano, Italy)

The aim of this study is to illustrate a methodological innovation in the analysis of museum texts by examining the linguistic features that support accessibility and comprehension in a corpus of English-language exhibit labels. Set in the M+ art museum in Hong Kong, the research employs a robust mixed-methods approach, integrating ethnographic reception studies with corpus-assisted discourse analysis (CADS).

The methodology involved several stages. First, a specialised corpus was constructed, comprising 68 authentic exhibit labels (34 pairs of 'disputed' versions) produced by museum professionals between 2023 and 2025. Each pair represents alternative drafts reflecting internal debates among curators about which version should be displayed, ensuring the corpus captures real-world communicative tensions and choices.

To assess audience reception, an on-site survey was conducted with museum visitors, stratified by educational background (undergraduate and postgraduate). Inspired by prior reception studies in museum contexts (Chmiel and Mazur 2012; Leung 2018), the survey instrument included sections on sociodemographic data, perceptions of exhibit labels as a genre, and direct preferences between paired label versions. This design enabled the identification of 'preferred' and 'less preferred' texts within each educational group. Subsequently, the corpus was divided into subsets according to these preferences. Using the UAM Corpus Tool (O'Donnell 2009), a comparative keyness analysis was performed between the preferred and less-preferred sets for each educational group. Concordance analysis was then applied to interpret keywords in context, revealing patterns in nominalisation, clause complexity, domain-specific vocabulary, and evaluative language.

Results reveal marked differences in linguistic preferences by educational background. Postgraduate visitors favour labels with higher nominalization, complex noun groups, and embedded clauses (e.g., frequent use of possessive markers like 'its', conjunctions like 'which', and prepositional phrases), reflecting a tolerance for abstraction and density of information. In contrast, undergraduates and visitors with less formal education prefer labels featuring tangible, domain-specific terms (e.g., 'design', 'craft'), contextual markers (e.g., 'Thai', 'Japanese'), and evaluative adjectives (e.g., 'new', 'traditional'). These labels employ more relational verbs in the present tense ('is'), demonstratives ('this'), and pronouns ('it'), supporting straightforward and factual communication.

This methodological triangulation, combining quantitative corpus analysis with qualitative audience reception, enables a nuanced understanding of how linguistic choices in exhibit labels affect accessibility for different visitor profiles. The approach demonstrates the value of small, principled corpora and experimental survey design for applied linguistic research in heritage communication.

References

- Chmiel, Agnieszka, and Iwona Mazur. "AD Reception Research: Some Methodological Considerations." In *Emerging Topics in Translation: Audio Description*, edited by Elisa Perego, 57–80. Trieste: EUT, 2012.
- Leung, Hoi Ching Dawning. "Audio Description of Audiovisual Programmes for the Visually Impaired in Hong Kong." PhD diss., University College London, 2018. UCL Discovery.
- O'Donnell, Mick. "Demonstration of the UAM CorpusTool for Text and Image Annotation." In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies Demo Session (HLT'08)*, 2008.

Gendered family narratives in a Japanese newspaper: A comparison of reader letters by women and men (1995–2024)

Keisuke Yoshimoto (Ryukoku University, Japan)

This paper examines gendered narratives in reader letter sections titled ‘Onna no Kimochi’ (Women’s Feelings) and ‘Otoko no Kimochi’ (Men’s Feelings) published in the liberal Japanese national newspaper *Mainichi* between 1995 and 2024. These sections invite readers to write about everyday feelings towards their families. While Women’s Feelings has existed since 1954, Men’s Feelings was introduced in 1995. This study examines how women and men negotiate gender roles within families, how their narratives differ by gender, and how these narratives have changed over three decades. In Japan, family relationships remain shaped by traditional gender norms. Women have long been constrained by fixed roles assigning them primary responsibility for housework, childcare, and family care (Ueno 1990). These norms are reinforced by legal and economic arrangements, including requirements that married couples share a surname and tax policies based on a male-breadwinner model. In contrast, men are expected to be the primary earners, and their identities are closely tied to paid work. When men deviate from this expectation, they often experience a loss of masculine recognition (Taga 2006). These asymmetric expectations provide an important backdrop for interpreting gendered family narratives.

Two corpora were constructed from all letters published in the Women’s and Men’s Feelings sections between 1995 and 2024: the WFeel corpus (27,897 letters; 7.1 million words) and the MFeel corpus (3,584 letters; 0.96 million words). The average age of contributors is 60.35 in WFeel and 66.23 in MFeel; therefore, the findings reflect tendencies within this age group. Keywords were obtained by comparing each corpus against the other using Sketch Engine. Collocates and concordance lines of keywords were examined and grouped into thematic categories to identify recurrent family-related narratives.

The results reveal a clear asymmetry in family narratives. Keywords in WFeel frequently refer to children and parents, including husbands’ parents, indicating that women largely bear responsibility for family care. WFeel also contains references to friends, hobbies, fashion, and emotional adjectives, suggesting that women’s narratives integrate family responsibilities with personal emotional experiences and reflect ongoing negotiation between caregiving roles and individual subjectivities. In contrast, MFeel keywords are dominated by work-related terms, several of which concern men’s anxieties about life after retirement, pointing to uncertainty and identity reconstruction in later life. To examine diachronic change, both corpora were divided into three decade-based subcorpora (1995–2004, 2005–2014, 2015–2024) and analysed using the remainder method (Brookes and Baker 2021). Keywords were obtained by comparing each subcorpus against the remainder and manually grouped into thematic categories based on concordance analysis. Changes in relative frequency were analysed using trend lines and the coefficient of determination. In both corpora, references to spouses in patriarchal terms steadily decline, indicating a shift away from hierarchical marital representations. References to past memories increase in the most recent decade. Close readings of concordance lines reveal ambivalent attitudes: while writers aspire for egalitarian family relationships, they also express nostalgia for traditional family models associated with their parents’ generation. These findings

demonstrate how gendered family narratives simultaneously reproduce and renegotiate dominant gender norms across generations.

References

Brookes, Gavin and Baker, Paul. (2021) *Obesity in the News*. Cambridge University Press.

Taga, Futoshi (2006). *Otokorashisa no Shakaigaku [Sociology of Masculinity] (Sekaishissha)*.

Ueno, Chizuko (1990). *Kafuchosei to Shihonseis [Patriarchy and Capitalism] (Iwanami)*.

Intersecting and (In/Ex)cluding Minorities in News and Rights Movements: A corpus-assisted integrated framework for analysing representations of LGBTQIA+ and Disability Prides in the UK and Italy (2020–2025)

Federico Zaupa (University of Modena and Reggio Emilia, Italy)

Pride manifestations have historically been associated with LGBT activism, but more recently they have embraced a broader range of minority identities facing discrimination, reflecting the growing awareness within contemporary rights movements that discrimination is intersectional (Crenshaw 1989). The celebration of LGBTQIA+ and disability Prides every June and July frames Prides as shared spaces of public mobilisation where minorities seek recognition and visibility, and construct identities through language and embodiment. At the same time, news coverage on Pride socially recontextualises these events and the issues for which activists call for political action, while foregrounding, marginalising, or erasing specific identities. Pride-related narratives therefore offer a relevant context for exploring discursive (in)visibilities of minority identities, particularly LGBTQIA+ and disabled communities. In discourse studies, corpus-assisted methods have been widely adopted to explore patterns of identity representation (Heritage/Taylor 2024), combining quantitative techniques of Corpus Linguistics (CL) with qualitative approaches of Critical Discourse Analysis (CDA). Typically, the quantitative phase informs the selection of concordance lines for subsequent qualitative analysis. However, for this stage, studies tend to rely on single focal CDA approaches, with no methodological toolkits integrating CL with the diverse analytical dimensions from different CDA traditions. Moreover, while CADS research on news discourse has extensively focused on LGBTQ+ communities (e.g., Wilkinson 2020) or, more recently, on disability representations (Karaminis et al. 2025), there is a scarcity of intersectional research or specific contributions on Pride discourses.

This study addresses these gaps by proposing an ad-hoc framework bringing together multiple CDA dimensions (e.g., Van Dijk 1998; Van Leeuwen 2008) to enable an intersectional and cross-generic analysis of the (self-)representation of LGBTQIA+ and disabled identities during Prides in the UK and Italy, across news and the websites of LGBTQIA+ and disability-supporting organisations. The research examines two exploratory corpora of articles from British and Italian broadsheet and tabloid newspapers published between 2020 and 2025 in June and July (approx. 500,000 tokens), and Pride-related webpages and press releases from LGBTQIA+ and disability-supporting organisations in the two countries (approx. 50,000 tokens). Findings reveal increased visibility of transgender identities across communication channels. However, both newspapers and websites represent LGBTQIA+ identities through collectivisation and passivisation, suggesting their lack of agency. LGBTQIA+ claims are also frequently conflated with trans and intersex debates, with discursive polarisations through in-group and out-group representations: while activists and some broadsheets legitimise their claims by negatively construing opposing viewpoints, conservative broadsheet and tabloid newspapers deploy similar evaluative strategies to undermine activists' demands by appealing to common-sense arguments. Following the outbreak of the Israel–Palestine conflict, intersectionality becomes more salient within activist discourse; by contrast, news discourse increasingly conflates LGBTQIA+ activism with pro-Palestinian movements, and challenges the legitimacy of Pride events. The prominence of trans and pro-Palestinian activism in the news reflects the marginalisation of disabled people in Pride narratives, who are predominantly represented through genericization,

contributing to the erasure of specific disabilities, including physical ones and neurodivergences.

Abstracts:

Poster presentations



**Corpora &
Discourse**
INTERNATIONAL CONFERENCE
2026

Subgenre Categorization in Finance Advice Videos: Triangulating register variation and semantic domains

Ylva Biri (University of Helsinki, Finland)

The ongoing project investigates subgenres and register variation within the broader genre of finance advice videos by so-called ‘finfluencers’, influencers specialising in finance. While general communicative goals of finance advice video include instruction, self-branding, and socialization with the audience, there is variation in which goal an individual video prioritises (see Bhatia 2023, 46). A qualitative analysis of self-presentation in financial advice videos (Author, in review) found that some videos are focused around the content creator’s autobiographical narrative whereas others focus on personal finance instruction more explicitly. This variation in communicative purposes likely corresponds to variation in register. The research explores what registers finfluencers draw on and how these choices in style may be associated with video topic or more specific communicative purposes.

The data is an English-language text corpus of 812,000 words, comprising 264 video transcripts extracted using the YouTube API. The corpus is balanced across channels by male and female content creators and channels with different subscriber counts (from 100,000 to over 1 million). The overarching goal is to categorise the videos based on their subgenres (or hybrid genres) based on their register and topic. Register variation is understood as variation in the frequency of lexico-grammatical features (Biber et al. 2021). Video topics are examined by combining keyword analysis with identifying semantic domains using the UCREL semantic annotation system (USAS) (Rayson 2008).

Preliminary quantitative findings on the data have identified clear variation for individual lexico-grammatical features: e.g., the frequency of 1st and 2nd person pronouns – features associated with a more interpersonal register – are associated with audience-engaging text types reminiscent of a self-help text. Although gender was not a significant explanatory variable for these features, previous research suggests that male and female influencers differ in the topics covered and the advice they provide (Ben-Shmuel et al. 2024).

Overall, the project sheds light on the communicative purposes and rhetorical strategies used within this understudied genre. It explores the appeal of the finfluencers, emphasising the genre’s potential impact on its audiences’ financial literacy.

References:

- Ben-Shmuel, Ambreen Tour, Adam Hayes, and Vanessa Drach. 2024. ‘The Gendered Language of Financial Advice: Finfluencers, Framing, and Subconscious Preferences’. *Socius: Sociological Research for a Dynamic World* 10 (January): 23780231241267131. <https://doi.org/10.1177/23780231241267131>.
- Bhatia, Aditi. 2023. *Digital Influencers and Online Expertise: The Linguistic Power of Beauty Vloggers*. 1st edn. Routledge. <https://doi.org/10.4324/9781003011705>.
- Biber, Douglas, Jesse Egbert, Daniel Keller, and Stacey Wizner. 2021. ‘Towards a Taxonomy of Conversational Discourse Types: An Empirical Corpus-Based Analysis’. *Journal of Pragmatics* 171: 20–35. <https://doi.org/10.1016/j.pragma.2020.09.018>.

Rayson, Paul 2008. 'From key words to key semantic domains', *International Journal of Corpus Linguistics* 13 (4), pp. 519–49.

YouTube Debates on Colourblind Casting: A Corpus-Assisted Discourse Analysis of Reactions to HBO's Severus Snape

Clara Cantos-Delgado (Autonomous University of Madrid, Spain)

YouTube has over 122 million active users daily (GMI Research Team, 2025) and is extensively known as a place where users consume audiovisual content and also where they interact with other users. Although polarization in this platform is not as extended (Levy, 2021; Liu et al., 2023) as in other social networking sites such as Twitter, it is still found, especially concerning (potentially) controversial topics (Garcés-Conejos Blitvich, 2010; Maślankowski et al., 2023; Zhang et al., 2025). This is the case, for instance, of the practice known as 'colourblind casting'. This practice consists of choosing an actor or actress to play a role independently of their ethnic background. Sometimes, this practice may be considered 'black-washing' when a character which was originally white is interpreted by a black actor or actress (Smith, 2021), but decisions behind these casting choices are multiple, including political, marketing strategies or simply instances of 'colourblind casting.' There has been many such cases recently including, among others, the character of Hermione in the 2016 London's West End play 'Harry Potter and the Cursed Child', the character of MJ in the 2017 'Spider-Man: Homecoming' and the character of Ariel in the 2023 life-action 'The Little Mermaid'. These casting choices have given rise to much controversy across the fandom, with the most negative viewers starting online movements such as #NotMy[Character's name]. Recently, HBO announced the casting of the black actor, Paapa Essiedua, for the character of Severus Snape in their 2027 new TV series of the Harry Potter franchise. This character is described in the novels as "a teacher with greasy black hair, a hooked nose, and sallow skin" (1997, p.126), a description fitting the interpretation by Alan Rickman, the actor who formerly played the character in the Harry Potter movie franchise (2001-2011). Thus, the 'race-swapping' decision has caused some commotion among fans and has led to some polarized opinions. This considered, this paper's main aim is to assess audience's opinion on the casting decision through a Corpus-Assisted Discourse Studies (CADS) approach (Partington et al., 2013). Sketch Engine was used to analyze the corpus of YouTube comments (N=3.609) posted to two YouTube videos (in Spanish) on the topic, by two Mexican female influencers. The results indicate significant differences across subcorpora considering the nature of the debates. The findings reveal statistically significant differences between subcorpora: commenters in Video 1 engaged more frequently with topics related to racism, as indicated by higher frequencies of ethnic-related adjectives in the Wordlists (e.g., 'negro' [black]). However, no notable differences emerged after examining the concordance lines regarding concerns about fidelity to the novels' storyline or character descriptions, showcasing a general concern over this issue across the subcorpora. Future developments will determine how the series' producers handle potential narrative adjustments and whether public opinion remains polarized or shifts as the release approaches.

Representing Corporate Climate Responsibility: A Corpus-Assisted Critical Discourse Analysis of UK and Chinese Companies

Hui Chen (Lancaster University, UK)

Climate change is one of the most urgent environmental challenges confronting humanity, and companies are under increasing pressure to demonstrate climate responsibility. Existing research (e.g., Jaworska 2018; Fuoli and Beelitz 2023) has focused on the energy sector and developed countries, leaving gaps in understanding how corporations represent climate change across socio-economic contexts and sectors with varying climate impact. This study addresses this gap by analyzing companies in the energy, technology, and finance sectors in China and the UK, two major emitters but with distinct contexts.

Using corpus-assisted critical discourse analysis, this study asks: How do UK and Chinese companies represent climate change in sustainability disclosures, and what ideologies underpin their similarities and differences?

The corpus comprises climate-related sections from annual, CSR/ESG, sustainability, and climate reports (2022–2024) of the top five companies per sector in China and the UK, selected from Forbes Global 2000 2025 (Disfold 2025 for UK technology firms). Only sections with explicit climate focus were extracted: standalone climate reports in full, or sections from other reports with climate-related titles or dense climate terminology. The resulting corpus comprises 1,021,055 tokens (UK) and 479,624 tokens (China). Corpus methods include keyword, collocation, and concordance analysis. Concordance analysis of significant keywords specifically focuses on transitivity and modality, social actor representation, and metaphor analysis.

Analysis reveals both shared strategies and distinct patterns. Both UK and Chinese companies employ obfuscation strategies through vague material processes (promote, assist) and journey metaphors ('transition to', 'advance towards'), creating dynamic scenarios while avoiding commitment. Responsibility is displaced through mental/verbal processes and deontic modality ('required') with backgrounded social actors. While social actors are functionalized and collectivized, UK firms emphasize stakeholder partnerships, whereas Chinese corporations foreground top-down management.

Divergences emerge in hedging and aspiration. UK companies use epistemic modality ('potentially', 'unlikely') and weak dynamic modality ('can'), portraying climate action as non-binding. Their discourse also features future tense, aspirational nouns ('ambition', 'commitment'), and mental processes ('aim to'), creating impressions of ethical intention without promise. Chinese firms employ self-activation strategies through positive evaluative lexis ('actively', 'steadily') and attributive relational processes, presenting action as accomplished states rather than future objectives.

These differences reflect distinct ideologies. UK discourse embodies market-based climate governance, constructing responsibility as compatible with stakeholder value, whereas Chinese discourse reflects state-led governance, legitimizing action through national policy alignment. Yet both share a belief in corporate continuity: climate action is framed as compatible with business-as-usual rather than requiring transformative change, enabling symbolic compliance

across contexts. These findings show that regulatory context influences legitimation but not obfuscation strategies, challenging the view that corporate climate discourse is uniformly driven by neoliberal rationality while stressing how distinct ideologies can lead to similar patterns of responsibility management.

References

- Fuoli, Matteo, and Anne Beelitz. 2023. "Framing the Path to Net Zero: Corporate Climate Change Discourse and Responsibility." *International Journal of Corpus Linguistics* 28 (4): 473–505. <https://doi.org/10.1075/ijcl.22123.fuo>.
- Jaworska, Sylvia. 2018. "Change but No Climate Change: Discourses of Climate Change in Corporate Social Responsibility Reporting in the Oil Industry." *International Journal of Business Communication* 55 (2): 194–219.

Legal Translation and International Legal Discourse: A Parallel Corpus of English–Arabic UN Texts

Hanem El-Farahaty (University of Leeds, UK)

Nouran Khallaf (University of Leeds, UK)

Asmaa Alduhaim (The Gulf University for Science & Technology, Kuwait)

Fathi Ahmed (Centre for International Law Research and Policy)

International legal translation operates as a central mechanism through which institutional authority, legal certainty, and ideological positioning are discursively produced and maintained across languages. Within the United Nations, legal discourse is not merely transferred but systematically recontextualised through translation, shaping how key legal concepts are stabilised, mitigated, or foregrounded in multilingual governance. Arabic, as one of the UN's six official languages, presents particular discursive challenges due to regional variation, legal hybridity, and the absence of a fully standardised supranational legal register (El-Farahaty 2015).

This project presents the first stage of a corpus-based discourse analytical study of English–Arabic legal translation in the UN General Assembly's Sixth Committee (Legal Committee), a primary site for the production of international legal discourse. The study introduces a parallel, open-access English–Arabic parallel corpus of Sixth Committee documents, available on GitHub and accessible via Sketch Engine (Kilgarriff et al. 2014). The corpus comprises resolutions, treaties, draft conventions, reports, and legal memoranda retrieved from the UN Official Document System and the UN Digital Library, all sentence-aligned and enriched with detailed metadata, including document type, session, date, subject area, and institutional source. Semi-automated alignment combined with rigorous manual verification ensures the reliable representation of discursively and legally complex constructions.

To evaluate the corpus as a resource for discourse analysis, we will present a pilot study that focuses on a set of politically sensitive legal concepts, such as self-determination, sovereignty, state immunity, territorial integrity, and use of force, treated as discursive nodes within UN legal argumentation. Drawing on corpus-based discourse analysis (CBDA) (Baker, 2023), which integrates quantitative corpus techniques with qualitative, context-sensitive interpretation, the study conducts a comparative analysis of English UN legal texts and their Arabic translations to investigate how recurrent patterns of evaluative and stance-bearing terms are re-articulated in translation. Through collocational and concordance-based evidence, the analysis explores shifts that contribute to mitigation in the treatment of politically sensitive legal concepts. In doing so, the study provides a discourse-analytic account of how translation participates in shaping institutional authority and neutrality in Arabic UN legal discourse.

References

Baker, P. (2023). *Using Corpora in Discourse Analysis* (2nd Ed.), London/New York: Bloomsbury Publishing

El-Farahaty, H. (2015). *Arabic–English–Arabic legal translation*. London/New York: Routledge.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36.
<https://doi.org/10.1007/s40607-014-0009-9>

Diplomatic Discourse on the New Silk Road: A cross-linguistic sentiment analysis of Europe's representation of China

Sophie Eyssette (La Sapienza University of Rome, Italy)

In 1877, the German geographer Ferdinand von Richthofen coined the term “Silk Road” to name the thousand-year-old trade routes connecting East and West (Bull et al., 2024). The dubbing of this ancient road by a European scholar illustrates the cultural construal of the East by the West (Said et al., 2005). In 2013, Xi Jinping, the President of People’s Republic of China’s, reactivated the former trade roads connecting the East to the West by launching the project One Belt, One Road, later renamed Belt and Road Initiative (Frankopan, 2018; Islam, 2019).

Research in linguistics has focused on the communication surrounding the New Silk Road (Van Noort, 2021), and its representation in diplomatic discourses from Asian and U.S. perspectives (Afzaal, 2021; Afzaal et al., 2022). However, the European perspective has not been examined linguistically. For this reason, this study seeks to address this research gap by analysing the linguistic patterns framing the NSR in European diplomacy.

The NSR’s multilingual and multicultural scope makes it an ideal case study for a novel methodological approach. This study incorporates diplomatic data from public websites from four European countries: the United Kingdom, Poland, Italy, and France. The major websites affiliated with the European Union (E.U.) are also screened. Consequently, a large-scale corpus of one trillion tokens was built, entitled DCBRI (Diplomatic Discourse of Belt and Road Initiative). Our method adopts a multi-layered framework that synthesizes lexicon-based emotion detection, sentiment analysis, and the extraction of diplomatic language features. By bridging computational sentiment analysis with critical discourse analysis, this study aims to reveal how language shapes—and is shaped by—the NSR’s global narrative. As such, this study aims to reveal how diplomatic discourses frame the relationships between Europe and China. The results demonstrate a complex perception of China, which is seen both as a hundred-year-old trade partner and as a new economic threat.

References

- Afzaal, M. (2021). Exploring Image of China in the Diplomatic Discourse: A Critical Discourse Analysis. *Journal of Development and Social Sciences*, 2(IV), 120–133. [https://doi.org/10.47205/jdss.2021\(2-IV\)11](https://doi.org/10.47205/jdss.2021(2-IV)11).
- Afzaal, M., Zhang, C., & Chishti, M. I. (2022). Comrades or contenders: A corpus-based study of China’s Belt and Road in US diplomatic discourse. *Asian Journal of Comparative Politics*, 7(3), 684–702. <https://doi.org/10.1177/20578911211069709>.
- Bull, S., Dacombe, S., Downes, A., Duguid, B., Greig, R., Lewin, R., Mumby, S., & Newman, L. (2024). *Story of the Silk Road*. All About History.
- Frankopan, P. (2018). *The new silk roads: The present and future of the world*. Bloomsbury Publishing.
- Islam, Md. N. (Ed.). (2019). *Silk Road to Belt Road: Reinventing the Past and Shaping the Future*. Springer Nature Singapore. <https://doi.org/10.1007/978-981-13-2998-2>.

Said, E. W., Todorov, C., & Said, E. W. (2005). *L'Orientalisme: L'Orient créé par l'Occident* (C. Malamoud & C. Wautier, Trans.). Éd. du Seuil.

Van Noort, C. (2021). *China's Communication of the Belt and Road Initiative: Silk Road and Infrastructure Narratives* (1st ed.). Routledge. <https://doi.org/10.4324/9781003184713>.

UK Parliament discourse on immigration (2007-2024) in a time of populist and post truth politics

Kenneth Fordyce, Lucia Michielin and Jessica Witte
(University of Edinburgh, UK)

The past two decades have seen the increased displacement of people following sociopolitical events such as the 2008 global financial crisis, regime changes in the Middle East and North Africa (Barry, 2012), war in Europe, and the Coronavirus pandemic. In parallel with increased levels of migration, there has been an increase in political rhetoric and disinformation around immigration, and a rise of populist right-wing political parties portraying immigrants and asylum seekers as a threat (Campani et al, 2022; Boeynaems et al, 2023).

This project, which investigates changes in political discourse during this turbulent period, is guided by the following research questions:

- In what ways did UK parliamentary immigration discourse change between 2007 and 2024?
- What links (if any) exist between changes in immigration discourse and specific sociopolitical events (2007-2024)?
- What patterns in immigration discourse can be found in relation to individual politicians, political parties and specific governments?

A corpus of all parliamentary debates in the UK House of Commons (2007-2024) focusing on immigration or asylum was collected from the Hansard online database using data extraction techniques alongside the Hansard database's Search API. This data was imported into OpenRefine (2025) alongside key metadata: debate date; debate title; name of politician; political party of speaker; government role (if applicable).

The dataset comprises over 13000 contributions from 374 separate debates. This poster will present the following steps in the process of data analysis:

- The refinement of the corpus data using OpenRefine (2025).
- Development of an Immigration Discourse Lexis based on word and n-gram lists generated using Antconc (Anthony, n.d.).
- Statistical analysis on the relationship between the use of the Immigration Discourse Lexis and: (1) sociopolitical events; (2) individual language use; (3) party affiliation; (4) government affiliation (if applicable).

It will also present on the value of establishing an interdisciplinary research team involving expertise in corpus linguistics and language data extraction techniques.

References

- Anthony, L. (n.d.) AntConc [Computer Software]. Tokyo, Japan: Waseda University.
<https://www.laurenceanthony.net/software/antconc/>
- Barry, A. (2012). Europe's Dilemma: Immigration and the Arab Spring
https://fpif.org/europes_dilemma_immigration_and_the_arab_spring/

Boeynaems, A., Burgers, C., A. Konijn, E., & J. Steen, G. (2023). Attractive or repellent? How right-wing populist voters respond to figuratively framed anti-immigration rhetoric. *Communications*, 48(4), 502-522.

Campani, G., Fabelo Concepción, S., Rodríguez Soler, A., & Sánchez Savín, C. (2022). The rise of Donald Trump right-wing populism in the United States: Middle American radicalism and anti-immigration discourse. *Societies*, 12(6), 154.

OpenRefine (2025) <https://openrefine.org/>

Would've, Could've, Should've: Conditionality and Hypothetical Modality in Populist Political Statements

Victoria Hogstad (University of Copenhagen, Denmark)

Modal auxiliaries and conditional constructions are central to populist discourse where they may function as persuasive tools for projecting authority and assigning blame - particularly in the context of negated reversed conditionals, as seen in the following:

“It’s probably why Putin decided to go into Ukraine, something he never would’ve done if I were president”

“We wouldn’t have had the problems of October 7th if I were president”

“We shouldn’t have any wars if the UN’s really doing its job.”

Unlike canonical conditionals introduced by ‘if’ (e.g., Dancygier 1998), reversed conditional structures foreground the consequence before the condition, enhancing rhetorical force by emphasising outcome and inevitability first. Drawing on cognitive rhetoric (Browse 2018, 2021) and, more broadly, cognitive linguistics (e.g. Croft & Cruse 2004; Langacker 1987; Lakoff and Johnson 1980), the analysis situates these constructions within broader processes of framing and narrative positioning.

This study examines counterfactual statements from a corpus of utterances by President Donald Trump from 2025. The examples above illustrate how for instance WOULD signals hypothetical certainty and SHOULD encodes normative obligation (Hoye 1997, p. 110), framing alternative realities and moral expectations.

The study adopts a mixed-method, corpus-driven approach (McEnery and Hardie 2012), applying quantitative corpus analyses such as (keyness analysis, collocation analysis, and collexeme analysis (Stefanowitsch and Gries 2003) combined with qualitative interpretation. Data include Donald Trump’s speeches and social media posts during his campaign and second presidency.

Collocation analysis shows that Trump employs a variety of modal and conditional constructions. Most notably, his use of negated reversed conditionals serves a clear rhetorical function: authority framing, achieved by foregrounding the consequence before the condition, typically expressed as his own authority. These collocational patterns construct hypothetical worlds of stability under his leadership, and, in turn, shift blame onto others. This may serve to highlight the implicit dimensions of populist language, where grammatical form becomes a means of projecting control and shaping perception.

References

- Browse, Sam. 2018. *Cognitive Rhetoric: The Cognitive Poetics of Political Discourse*. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/lal.31>.
- Browse, Sam. 2021. “Towards an Empirical Stylistics of Critical Reception: The Oppositional Reader in Political Discourse.” In *Style and Reader Response: Minds, Media, Methods*,

edited by A. Bell, S. Browse, A. Gibbons, and D. Peplow, 61–80. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/lal.36.04bro>.

Croft, William, and D. A. Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Hoye, Leo. 1997. *Adverbs and Modality in English*. London: Longman.

Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Vol. 1. Stanford: Stanford University Press.

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar: Descriptive Application*. Vol. 2. Stanford: Stanford University Press.

McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.

Stefanowitsch, Anatol, and Stefan Th. Gries. 2003. "Collostructions: Investigating the Interaction of Words and Constructions." *International Journal of Corpus Linguistics* 8 (2): 209–243. <https://doi.org/10.1075/ijcl.8.2.03ste>.

‘Skinny jabs’, surgery and stigma: Linguistic representations of medical weight loss treatments in the news

Abigail Kinsella (Aston University, UK)

Over a quarter of the UK adult population is estimated to be living with obesity (House of Commons 2023). Where behavioural interventions are ineffective (Dumbrowski et al. 2014), the long-established medical treatment is bariatric surgery (Boyers et al. 2021); however, this treatment is costly. The recent approval of cheaper injectable medications, such as GLP-1 agonists, has resulted in an uptake of weight loss drugs, with over 1.5 million estimated UK users as of April 2025 (NPA 2025).

The surge in demand for so-called ‘skinny jabs’ has received considerable attention in the news. Previous research on the representations of obesity (e.g. Bednarek et al. 2024; Brookes & Baker 2021) has found a prevalence of language practices that perpetuate weight stigma (discriminatory ideologies about obesity (Bednarek et al. 2024)), which can have negative impacts on the mental and physical health of those targeted (Brown et al. 2022).

As both the clinical use and cultural visibility of weight loss drugs increase, this work-in-progress study seeks to investigate reporting about medical weight loss treatments in the UK press and evaluate how/to what extent reporting draws upon discourses associated with weight stigma (e.g. shame). Methods from corpus-assisted discourse studies (Partington et al. 2013) and corpus-based critical discourse analysis (Baker 2023) will be used to analyse a corpus of UK news reporting from 2021-2026, comprising over 10,000 news articles.

This poster describes the corpus design and planned analytical procedure, which seeks to identify the topics and societal domains (e.g. healthcare, politics, entertainment) associated with weight loss treatments, how users of weight loss treatments are evaluated; represented voices (Fahs and Swank, 2025) and how these voices linguistically frame weight based issues (Nutter et al., 2024). The study hypothesises that a) the emergence of weight loss drugs has led to an increase in reporting about obesity and b) there is an emerging stigma associated with the use of weight loss drugs.

References

- Baker, P. (2023). *Using Corpora for Discourse Analysis* (2nd ed.). Bloomsbury.
- Bednarek, M., Bray, C., Vanichkina, D., Brookes, G., Bonfiglioli, C., Coltman-Patel, T., Lee, K., & Baker, P. (2024). Weight stigma: towards a language-informed analytical framework. *Applied Linguistics*, 45(3): 424-448. <https://doi.org/10.1093/applin/amad033>
- Boyers, D., Retat, L., Jacobsen, E., Avenell, A., Aveyard, P., Corbould, E., Jaccard, A., Cooper, D., Robertson, C., Aceves-Martins, M., Xu, B., Skea, Z., de Bruin, M.; and the REBALANCE team. (2021). Cost-effectiveness of bariatric surgery and non-surgical weight management programmes for adults with severe obesity: a decision analysis model. *Int J Obes (Lond)*, 45(10): 2179-2190. doi: 10.1038/s41366-021-00849-8.
- Brookes, G. & Baker, P. (2021). *Obesity in the News: Language and Representation in the Press*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/9781108864732>

- Brown, A., Flint, S.W., & Batterham, R.L. (2022). Pervasiveness, impact and implications of weight stigma. *eClinicalMedicine*, 47(101408).
<https://doi.org/10.1016/j.eclinm.2022.101408>
- Dombrowski, S. U., Knittle, K., Avenell, A., Araujo-Soares, V., & Sniehotta, F. (2014). Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. *BMJ* 348:g2646
doi:10.1136/bmj.g2646
- House of Commons. (2023). Obesity statistics.
<https://researchbriefings.files.parliament.uk/documents/SN03336/SN03336.pdf>
- National Pharmacy Association. (2025). <https://www.npa.co.uk/news/2025/july/around-half-of-adults-support-the-nhs-making-weight-loss-injections-free-through-their-community-pharmacy/>
- Partington, A., Duguid, A., & Taylor, C. (2013). Patterns and meanings in discourse: theory and practice in corpus-assisted discourse studies (CADS). (55 ed.) (Studies in corpus linguistics; No. 55). John Benjamins.

Theological and ideological positioning toward same-sex relationships, and how semantic domains emerge in denominational discourse.

Gareth Johns (Manchester Metropolitan University, UK)

This research focuses on the appointment of Sarah Mullally as the new Archbishop of Canterbury, the first female to do so in history, which was announced on 3rd October 2025. Many sources of journalistic discourse commented on her appointment as she is an advocate for same-sex marriage, which links the themes of gender and sexuality. Language constructs authority and legitimacy. WMatrix data revealed semantic fields like “in power” (Anglican Global South), “allowed” (Catholic texts) and “law and order” (Baptist texts). These fields show how discourse frames same sex marriage not just as a relational issue but as a matter of authority, permission and institutional control. Language is used to legitimise or delegitimise social practices, reinforcing power structures within religious communities. Language evolves according to social change. Diachronic contexts (e.g. UK Methodists affirming same sex marriage since 2021 compared to US Methodists lifting bans in 2024) show how shifts in policy are accompanied by shifts in discourse. This corpus evidence demonstrates that as society changes, language adapts: new collocates emerged (unions, blessings), which reflects evolving norms and values; semantic fields in the Anglican Western Liberal article identified ‘discernment’ and ‘bravery’ as statistically significant semantic fields when reporting on the new Archbishop of Canterbury. Language reflects the ideological divisions within society. My collocate analysis identifies that different denominations foreground, background and completely omit certain lexical items (e.g. marriage, blessings, relationships, and unions). This demonstrates how language choices determine theological and cultural positions. For example, Baptists emphasise legality and restriction, Methodists prioritise participation and belonging, and Anglican Western Liberals promote agency through decision and considered speech acts.

Language and Ideology in Crisis: A Critical Discourse Analysis of Hong Kong's COVID-19 News across Different Phases

Amber, Congyao Liu (SOAS, University of London, UK)

Since late 2019, the world has been enveloped in a pandemic caused by a novel coronavirus, SARS-CoV-2. Until now, COVID-19 has been around for five years, and humanity has grown accustomed to coexisting with it. However, it is worth noting that the pandemic has highlighted certain societal issues. The pandemic not only exposes various social and political issues but also amplifies the underlying ideological tensions that have been latent within society. This article aims to explore how power relations and ideological stances are embedded in the language of Hong Kong newspaper coverage. News reports from Oriental Daily and Headline Daily were compiled to build the corpus, with a focus on analysing keywords and their co-occurrence patterns. The theoretical framework combines Critical Discourse Analysis and Systemic Functional Linguistics, while the research is methodologically grounded in corpus pragmatics. AntConc is employed as the main analytical tool. Vaccination discourse is selected as a case study to provide a focused lens for investigation. To examine shifts in ideology and policy over different stages of the pandemic, selected periods from 2021 and 2022 are compared. This study finds that there has been a discursive shift from externalising the virus as a foreign threat to internalising pandemic governance within domestic frameworks from 2021 to 2022. And in 2022, the discourse increasingly stressed the governance on vulnerable groups, with discussions extending to issues of morality and human rights. Additionally, the reporting patterns of COVID-19 cases in Hong Kong media followed a distinct logic that categorised individuals into vaccinated and unvaccinated groups, which can be seen as a strategy to legitimise governmental vaccination policies. Moreover, the choice between BioNTech and Sinovac vaccines by public figures and ordinary citizens to some extent reflects their political stances as well as their considerations concerning vaccine safety and efficacy.

Metaphors and China's COVID-19 policies: a diachronic analysis of Chinese-English bilingual news editorials

Yufeng Liu (Lancaster University, UK)

China implemented stringent COVID restrictions in 2020 but eased them in December 2022. Since metaphors can both reflect and shape people's views and stances, it is relevant to study how the state-backed newspaper Global Times used metaphors to frame the policy shift for different target readers. It extends our previous research on the newspaper's 2020 editorials, revealing the fluidity of metaphors used by the same text producer during a period of significant transition. Despite changes in COVID-19 policies, the newspaper continued to use WAR metaphors between January 2022 and February 2023. However, the gradual increase in JOURNEY metaphors and decrease in WAR metaphors over time suggest a transition from a "militaristic" to a "smoother" portrayal of the pandemic's progression. While the newspaper consistently maintained pro-China and anti-US stances, it used fewer metaphors to describe the US position. The metaphor shift for China from "a person who does not lie at" to "a person who rolls up sleeves and works hard" aligns with the government's change in focus from the pandemic situation to economic recovery. It should be noted that translation practices further affect metaphor perception, with source-text readers exposed to more negative depictions of the pandemic and the US compared to target-text readers. Future research could explore these metaphors' impact on audiences.

Talking green in Vietnam: A corpus-assisted critical discourse study of news framing of environmental issues

Minh Hieu Nguyen (University of Birmingham, UK)

Three decades after the “Đổi Mới” reform in 1986, Vietnam has witnessed remarkable economic transformation, albeit at the expense of the environment. With environmental issues becoming more prominent, the news media have shown increasing interest in environmental reporting. However, little is known about how the Vietnamese press communicates these issues to the public. To date, most existing scholarship has focused on English-language news discourse from Western countries and major economic powers such as the United States, the United Kingdom, and China. Consequently, the diverse realities of environmental challenges in other contexts, particularly in the Global South, remain largely underrepresented. To address this research gap, I employ a corpus-assisted critical discourse analysis approach, informed by framing theory, to examine Vietnamese-language environmental news discourse.

To construct the corpus, I used Sketch Engine (Kilgarriff et al. 2014) to retrieve relevant articles from both private and state-owned online newspapers, including 24h, Dân Trí, Kênh14, Thanh Niên, Thời Báo VTV, Tiền Phong, Tuổi Trẻ, VietNamNet, and VnExpress. This process resulted in a corpus of 1,615 articles, totalling approximately 1.1 million words. I then performed a key n-gram analysis using viTenTen17 as the reference corpus to identify the environmental issues with the most press coverage. Moreover, I selected the three highest-frequency key n-grams, namely “ô nhiễm không khí” [air pollution], “rác thải nhựa” [plastic waste] and “biến đổi khí hậu” [climate change], for collocation and concordance analyses to explore how they are framed in terms of problem definition, causes, consequences, and solutions/actions (Entman 1993).

The findings show that air pollution, plastic waste, climate change, and ocean pollution are the most frequently reported issues, with air pollution receiving the most coverage. Additionally, air pollution is defined as a local problem caused by transportation and human activities that lead to health-related effects. In contrast, plastic waste receives relatively limited press coverage, particularly regarding its causes and impacts. Climate change is framed as a global challenge, often linked to natural phenomena and portrayed as affecting both Vietnamese citizens and local areas. Across these issues, solution-oriented framing is prominent, with a wide range of solutions reported for plastic waste and climate change. The identified framing patterns result from a complex interplay of factors at the micro level (journalists’ styles, experiences, and education), the meso level (newsroom norms and practices), and the macro level (Vietnam’s socio-economic, cultural, and political contexts).

By redressing the Western-centric scholarship on environmental news discourse, my study sheds light on how environmental issues are framed across national contexts. It also demonstrates the rigour of corpus linguistic techniques for analysing Vietnamese-language discourse, thus laying the groundwork for future research. Additionally, the insights gained can help enhance audiences’ critical media literacy, as well as inform journalists, media outlets, and policymakers in developing and implementing appropriate reporting strategies, guidelines, and actions. Future studies could examine how environmental news discourse varies across different types of newspapers or evolves over time. It would also be valuable to extend the

analysis beyond text to include multimodal resources, and to conduct surveys and interviews to better understand journalistic practices and audience perceptions.

References

Entman, Robert M. 1993. "Framing: Toward Clarification of a Fractured Paradigm." *Journal of Communication* 43 (4): 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, et al. 2014. "The Sketch Engine: Ten Years On." *Lexicography*, no. 1: 7–36. <https://doi.org/10.1007/s40607-014-0009-9>.

Playing the climate crisis: Climate emotions in environmental games

Janno Reincke and Michaela Mahlberg
(Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)

Video games are discursive environments that contribute to processes of individual and cultural meaning-making. Their narrative qualities lie not solely in the stories they tell, but in the ways these stories are mediated through the specific affordances of the medium. This perspective becomes particularly relevant in the context of environmental and climate-related issues, where video games offer distinctive possibilities to shape how environmental issues are narratively framed and made meaningful. While emotional engagement has become a central concern in climate change communication, particularly with regard to its influence on individual meaning-making and behaviour (Roeser 2012), the discursive potential of video game narratives to evoke such emotional experiences has received little systematic attention. Although a small number of corpus-based studies have begun to examine social discourses in video game texts (e.g. Heritage 2020), corpus-driven analyses of emotional meaning-making in environmental game narratives remain scarce.

Building on this perspective, the study explores how corpus-based approaches can be used to examine the emotional dimensions of environmental storytelling in video games. It draws on a small, purpose-built corpus of textual data from environmental games, including dialogues, narration, and descriptive texts. Focusing on textual components allows for a systematic corpus-based analysis of narrative discourse while acknowledging the broader multimodal context of video games. The analysis positions climate-related emotions as integral components of environmental storytelling and examines how they are embedded in the game texts. Using the Inventory of Climate Emotions (ICE; Marczak et al. 2023), which differentiates eight climate-related emotions, as a conceptual framework, the study develops an LLM-based emotion classification pipeline. Two general-purpose LLMs are tasked with identifying textual articulations of these predefined emotion categories in the game texts. This approach enables an examination of the distribution and recurrence of climate-related emotions and their discursive patterns across the corpus. It provides an empirical basis for interpreting how environmental games communicate climate change through emotional cues. Drawing on LLMs is not meant to replace the detailed analysis of the textual patterns. As we will show, it provides a step for the identification of patterns that then require further analysis.

The findings suggest that LLM-based emotion classification can reveal recurring patterns in the textual articulation of climate-related emotions in games, with emotions such as anxiety and sorrow appearing prominently, occasionally complemented by the more engagement-oriented emotion of climate enthusiasm. While exploratory in nature, the approach demonstrates the potential of LLM-based methods to support the analysis of emotional meaning-making in narrative media. Overall, the study positions video games as discursive environments in the broader context of climate discourse. It highlights emotions as key discursive resources in narrative meaning-making and demonstrates how corpus-based computational methods can be integrated into discourse-oriented analyses of video games.

References

Heritage, Frazer. 2021. *Language, Gender and Videogames*. 10.1007/978-3-030-74398-7.

Marczak, Michalina, Małgorzata Wierzba, Dominika Zaremba, Maria Kulesza, Jan Szczypiński, Bartosz Kossowski, Magdalena Budziszewska, Jarosław M. Michałowski, Christian A. Klöckner, and Artur Marchewka. 2023. "Beyond Climate Anxiety: Development and Validation of the Inventory of Climate Emotions (ICE): A Measure of Multiple Emotions Experienced in Relation to Climate Change." *Global Environmental Change* 83 (October): 102764. 10.1016/j.gloenvcha.2023.102764

Roeser, Sabine. 2012. "Risk Communication, Public Engagement, and Climate Change: A Role for Emotions." *Risk Analysis* 32 (6): 1033–40. 10.1111/j.1539-6924.2012.01812.x.

Constructing Climate Sceptic Identities: A Corpus-Assisted Discourse Analysis of Climate Sceptic Blogs

Izabella Schleier (Lancaster University, UK)

Climate sceptic discourses function as identity resources, with Bliuc et al. (2015) conceptualising the climate debate as conflict between opinion-based groups with distinct identities forming the basis for collective action. This work-in-progress study examines how climate sceptic blogs construct ingroup and outgroup identities while legitimising rejection of scientific consensus.

The project is grounded in the Discourse-Historical Approach (Wodak and Rheindorf 2022), using methodological triangulation to combine qualitative discourse analysis of blog About pages – which serve as focal points for bloggers' identity construction (Dennen 2009) – with corpus linguistic analysis of posts and comments from the same prominent climate sceptic blogs (NoTricksZone, Watts Up With That, Principia Scientific International, Climate Depot), spanning 2017-2025 (estimated 56 million words). Combining these approaches allows their strengths to complement one another, with corpus methods benefiting from CDA's theoretical categories and contextual attention while CDA is strengthened through corpus-based evidence of pattern generality (Baker et al. 2008). Drawing on theories of identity construction (Bucholtz and Hall 2005) and legitimation (van Leeuwen 2007; Reyes 2011), I examine how climate sceptics legitimise their own stance, with preliminary qualitative findings showing that bloggers authenticate themselves as alternative experts and delegitimise mainstream climate science as dogmatic.

As climate sceptic bloggers frequently represent their blogs as counterbalances against 'alarmist' mainstream media, a reference corpus of mainstream news articles will be compiled, enabling comparative keyword analysis across ideological positions and genres. Additionally, as some terms – such as 'climate science/scientist(s)' – may occur at similar rates but be used in different contexts or with different evaluative stances, a second comparison with EnTenTen will reveal words more frequent in climate sceptic blogs compared to general online English. Collocation analysis will focus on how social actors (e.g. scientists, activists, politicians) are described in climate sceptic blogs compared to mainstream news, examining predication strategies and attributed actions by analysing collocates using van Leeuwen's (2008) framework for representing social actors and social actions. These patterns can be further investigated using concordance analysis.

References

- Baker, Paul, Costas Gabrielatos, Majid KhosraviNik, Michał Krzyżanowski, Tony McEnery, and Ruth Wodak. 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): 273–306.
- Bliuc, Ana-Maria, Craig McGarty, Emma F. Thomas, Girish Lala, Marianne Berndsen, and Renata Misajon. 2015. "Public Division about Climate Change Rooted in Conflicting Socio-Political Identities." *Nature Climate Change* 5 (3): 226–29.

- Bucholtz, Mary, and Kira Hall. 2005. "Identity and Interaction: A Sociocultural Linguistic Approach." *Discourse Studies* 7 (4-5): 585–614.
- Dennen, Vanessa P. 2009. "Constructing Academic Alter-Egos: Identity Issues in a Blog-Based Community." *Identity in the Information Society* 2 (1): 23–38.
- Reyes, Antonio. 2011. "Strategies of Legitimization in Political Discourse." *Discourse & Society* 22 (6): 781–807.
- van Leeuwen, Theo. 2007. "Legitimation in Discourse and Communication." *Discourse & Communication* 1 (1): 91–112.
- van Leeuwen, Theo. 2008. *Discourse and Practice*. Oxford: Oxford University Press.
- Wodak, Ruth, and Markus Rheindorf. 2022. *Identity Politics Past and Present: Political Discourses from Post-War Austria to the Covid Crisis*. Exeter: University of Exeter Press.

Who (Should) Get to be Pregnant? A CADs Analysis of Gender and Pregnancy in UK Newspapers

Matilda Vokes (Queen Mary University of London, UK)

Pregnancy features across a wide range of UK newspaper genres: from showbiz and lifestyle, to politics and healthcare. Discourse around pregnancy is heavily gendered (Freed, 1999) and recently there has been more academic interest in discourses of LGBT+ pregnancies (Carpenter and Niesen, 2021). Attention is rightly bestowed on queer people's own experience with pregnancy but understanding how the dominant discourses of gendered language around pregnancy is important because of its ability to influence public opinion.

Newspapers offer a unique insight into influential national discourse around pregnancy and texts can represent current discourse on both pregnant straight cis women and queer/trans pregnancies. Therefore, this research asks what gendered or non-gendered categories of person are legitimised when discussing pregnancy? And how are gendered or non-gendered actors represented in UK news?

The corpus is comprised of the 11 UK national papers, including their Sunday and online equivalents. Due to a high volume of articles, each text had to mention pregnancy at least twice to be included giving a total of 49,524 texts over a five year period of 2020-24. Focussing on the construction of gendered actors, the person-related nouns that co-occurred with 'pregnant' were identified. These were categorised as gendered female (woman/en, girlfriend, wife), gendered neutral (person/people, worker/s, MP/s) and the texts were then grouped into two sub-corpora. A smaller number of texts were categorised as gendered male, but these will not be discussed here. Following a down-sampling method described by Baker (2020), 40 texts for each group were chosen for close analysis using a random number generator.

'Pregnant women' was, unsurprisingly, the most common co-occurrent, and was used with verbs such as feel, need, and consume in the subject positions, and affect, advise and induce in the object. A close analysis suggests women's pregnancies are framed in terms medical and dietary advice, as well as individual narratives of hardship. Overall, the impression of vulnerability was created in these texts.

Of the gender neutral terms, 'pregnant people' is the most common at 5.07 frequency per million, although far less than pregnant women/an which is 143.2/34.8 respectively. 'Pregnant people' is overwhelmingly the object of the sentence and occurs with verbs such as replace, refer, and include, with the preliminary CADs analysis highlighting that this is typically in a meta-linguistic context about the language of official guidance. Uses of 'pregnant people' is rarely to represent the experiences of pregnancy and texts tend to frame the term as an illogical and hypothetical linguistic choice. This is in contrast to other gender neutral terms like 'workers'.

In UK newspapers, pregnancy is constructed as legitimate for specific gendered actors, by only allowing attributing a subjective and felt experience to cis women's pregnancies. Gender neutral language is almost exclusively reserved for hypothetical and meta-linguistic discourse. Arguably, this discourse pushes a narrow narrative on to women, while excluding queer pregnancy entirely, ultimately contributing to a binary and essentialist discourse of pregnancy and gender.

References

- Freed A, (1999), Communities of practice and pregnant women: Is there a connection?,
Language in Society, 28, 257-271
- Carpenter E, Niesen R, (2021), “It’s Just Constantly Having to Make a Ton of Decisions That
Other People Take for Granted”: Pregnancy and Parenting Desires for Queer Cisgender
Women and Non-binary Individuals Assigned Female at Birth, Journal of GLBT Families
Studies, 17(2), 87-101
- Baker P (2020), “Analysing Representations of Obesity in the Daily Mail via Corpus and Down-
Sampling Methods” in Using Corpus Methods to Triangulate Linguistic Analysis, eds.
Baker P, Egbert J, London: Routledge

Tourism Chinese in Macau's Entertainment and Sports Industries

Shan Wang (University of Macau, China)

In recent years, the development of Chinese for Specific Purposes (CSP) has highlighted the need for linguistic resources tailored to specific industries and application contexts. Macau, one of China's two Special Administrative Regions, is working to build a "World Centre of Tourism and Leisure." This effort has created a growing need for professional Chinese language resources that reflect local characteristics and cultural orientations in tourism communication, service interaction, and cultural dissemination. This study focuses on building a specialized Chinese tourism corpus for Macau's entertainment and sports industries. This research contributes to the development of CSP tourism language services and pedagogical practices.

A user-friendly tool for corpus-based analysis of discourse using Lexical Multidimensional Analysis

Rogério Yamada and Tony Berber Sardinha
(Pontifical Catholic University of Sao Paulo, Brazil)

Lexical Multidimensional Analysis (LMDA; Berber Sardinha and Fitzsimmons-Doolan 2025) is an extension of Multidimensional Analysis (MDA; Biber, 1988, 1995) that was developed to detect lexical co-occurrence patterns that point to underlying discourses. As a development of the so-called Flagstaff School of Corpus Linguistics (cf. Cortes and Csomay 2015), LMDA shares principles with MDA, such as linguistic cooccurrence signaling a common underlying trait, multidimensionality, and a text-linguistic focus (Veirano Pinto, Berber Sardinha, and Delfino 2026). MDA has been criticized for being unnecessarily complicated (Xiao and McEnery 2005; Egbert, Larsson, and Biber 2020), and so LMDA inherits these criticisms by extension. We recognize the complexity of LMDA. The method relies on multivariate statistics and involves several processing steps, including part-of-speech tagging, lemmatization, feature selection, and feature counting on a text-by-text basis. It may also require converting continuous frequency measures to nominal values or applying frequency normalization. At the analytical stage, correlations are computed, and factor analysis is applied to identify latent sets of co-occurring features, followed by analysis of variation across texts or groups. Consequently, a successful implementation of LMDA presupposes familiarity with statistical reasoning and with statistical software such as SPSS, SAS, or R; although programming is not required, writing one's own scripts for automating parts of the workflow is an advantage.

In this paper, we introduce a new application that simplifies running an LMDA by offloading the burden of corpus processing and statistical analysis, requiring only that the user provide a corpus and adjust a few settings. The app is free to use and runs on different operating systems. To conduct an analysis, the user uploads a corpus, specifies the language, sets the minimum factor loading, and defines whether POS categories should be filtered out and if keywords need to be computed. The user also toggles the data type as categorical, ordinal, or interval. The app then tags the texts, performs POS filtering if needed, applies the appropriate correlation method based on the data type, performs an initial factor analysis, and presents an eigenvalue list to help the user determine the number of factors. Once the user sets this number, the app completes the factor extraction, removes low communality features and scores each text on each factor. The app output includes, among other elements, the factorial pattern, a spreadsheet containing feature counts and factor scores, and a sample of high-scoring texts from each pole of each dimension to help users interpret the factors as dimensions. It is hoped that this tool will help make LMDA more widely adopted by researchers interested in describing discourse patterns.

References

Berber Sardinha, Tony, and Shannon Fitzsimmons-Doolan. 2025. *Lexical Multidimensional Analysis: Identifying Discourses and Ideologies*. Cambridge: Cambridge University Press.

- Cortes, Viviana, and Eniko Csomay. 2015. "Douglas Biber and the Flagstaff School of corpus-based research: an introduction." In *Corpus-based research in Applied Linguistics: Studies in honor of Doug Biber*, edited by Viviana Cortes and Eniko Csomay, xv-xix. Amsterdam; Philadelphia: John Benjamins.
- Egbert, Jesse, Tove Larsson, and Douglas Biber. 2020. *Doing linguistics with a corpus: Methodological considerations for the everyday user*. Cambridge: Cambridge University Press.
- Veirano Pinto, Marcia, T. Berber Sardinha, and Maria Claudia Nunes Delfino. 2026. "Multi-Dimensional Analysis of Corpora." In *The Encyclopedia of Applied Linguistics*, edited by Carol Chapelle. Malden, MA: Wiley-Blackwell.
- Xiao, Zhonghua, and Tony McEnery. 2005. "Two approaches to Genre Analysis: Three genres in modern American English." *Journal of English Linguistics* 33 (1): 62-82.

How the Representation of Covid-19 in British News Media has Changed: A Corpus-assisted Critical Discourse Analysis

Yaan Yu (Lancaster University, UK)

Covid-19 is an acute, contagious disease which has spread globally since 12/2019 and was declared a pandemic by the WHO in 03/2020. The issues surrounding the public health emergencies like the Covid-19 pandemic are challenging and complex, because they have severe effects not only in medical, but also in political, social and economic spheres. Newspapers wield enormous influence over public opinion (van Dijk, 1991) and rather than just mirroring reality, newspapers have the function of constructing other versions of reality that are ideologically motivated (Baker, Gabrielatos, & McEnery, 2013). Thus, media representations of Covid-19 have been examined extensively since the outbreak of the pandemic, some investigate the news of covid from a general perspective (e.g., Basch et al., 2020), while others focus on specific aspects of the pandemic (e.g., vaccines or quarantine) (Nerlich & Jaspal, 2021; Ahmad Kamboh et al., 2022) or representations of particular groups (e.g., people with obesity or older people) in the context of the pandemic (Brookes, 2021; Morgan et al., 2021). Most research has focused on media discourse during the pandemic. However, so far, under the situation that though people are still contracting covid, they are talking about it in a different way and there is the public sense of ‘the pandemic is over’, little research has explored how the representation of covid-19 in the news media has changed over time.

This study addresses this gap by investigating how the representation of Covid-19 in the British press has evolved from the pandemic period to the post-pandemic era, and what the similarities and differences are in the evolution of the discourses around Covid-19 between the broadsheets and tabloids as well as the possible reasons. To answer the research questions, a corpus-assisted CDA of British news about Covid-19 is conducted, in which AntConc (version 4.3.1) is used as the analytic tool and Fairclough’s three-dimensional model is adopted as the theoretical framework. Two purpose-built corpora which respectively contain British broadsheet and tabloid news mentioning Covid-19 published between 01/01/2020 and 31/12/2023 are conducted. Totally, 9,600 reports were collected from 4 broadsheets and 4 tabloids, each newspaper contributed 1,200 articles collected evenly across months. Each corpus was divided into 16 three-month sub-corpora to enable diachronic analysis. Sequential comparisons were conducted within each corpus, with each sub-corpus serving as the reference for the next, generating keywords that occur significantly more frequently in the later period. These keywords were then thematically/semantically categorized and examined through concordance and context analysis to identify discourse patterns in broadsheets and tabloids across different pandemic phases.

Finally, similarities and differences in Covid-19 discourses between the two genres will be analyzed considering sociocultural, institutional, and political contexts. By adopting a longitudinal and comparative perspective, this study moves beyond snapshot analyses of Covid-19 discourse and seeks to capture how media representations develop over time. These insights can help health communication professionals understand how the discourse of global health emergency can evolve, informing the design of messaging for current and future public health issues.

References

- Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press* (1st ed.). Cambridge University Press.
- Brookes, G. (2022). "Lose weight, save the NHS": Discourses of obesity in press coverage of COVID-19. *Critical Discourse Studies*, 19(6), 629–647.
- Dijk, T. A. van. (1991). *Racism and the press*. Routledge.
- Nerlich, B., & Jaspal, R. (2021). Social representations of 'social distancing' in response to COVID-19 in the UK media. *Current Sociology*, 69(4), 566–583.



Corpora & Discourse

INTERNATIONAL CONFERENCE
2026