



Book of Abstracts

Contents

Committees.....	2
Plenary Speakers.....	3
Journal presentation.....	4
Full-length paper presentations	5
Short paper presentations	74

Committees

Organising committee

Charlotte Taylor (University of Sussex)
Anna Marchi (University of Bologna)
Dario del Fante (University of Padova / University of Sussex)
Jasmin Kidgell (University of Sussex)

Programme committee

Paul Baker (Lancaster University)
María Calzada Pérez (Universitat Jaume I, Castellón)
Alison Duguid (University of Siena)
Shannon Fitzsimmons-Doolan (Texas A&M University Corpus Christi)
Federica Formato (University of Brighton)
Nicholas Groom (University of Birmingham)
Sylvia Jaworska (University of Reading)
Anna Marchi (University of Bologna)
Gerlinde Mautner (Vienna University of Economics and Business)
Tony McEnery (Lancaster University)
Mark Nartey (Hong Kong Polytechnic University)
Alan Partington (University of Bologna)
Amanda Potts (Cardiff University)
Charlotte Taylor (University of Sussex)
Marie Veniard (University of Paris, EDA, fellow at Institut Convergences Migrations)
Rachelle Vessey (Birkbeck, University of London)

Plenary Speakers

If artificial intelligence is making our students dumb, can CADS make them smarter? Learning from corpora in the time of big data

Silvia Bernardini
University of Bologna

Applying Corpus Linguistics and Discourse Studies: Impact, Engagement and Interdisciplinarity

Gavin Brookes
Lancaster University

Journal presentation

Introducing APPLIED CORPUS LINGUISTICS (ACORP) JOURNAL

*Paul Thompson & Eric Friginal, Founding Co-Editors-in-Chief
University of Birmingham / Georgia State University*

In this presentation, we introduce Applied Corpus Linguistics (ACORP), a new international peer-reviewed journal published by Elsevier for the dissemination of research that reports or clearly supports the applications of corpus linguistics (CL) methods, theories, and findings to a wide variety of real-world contexts. This new journal is aimed at stimulating further development and dispersion of corpus tools, techniques, approaches and resources to research in diverse areas. In the coming years, we hope that the journal will develop into a repository for accounts of how corpus linguistics has been applied, a platform for discussion of major issues in applied corpus linguistics, and an inspiration to further innovations in applications of CL. We will briefly discuss the Aims & Scope statement of the journal and answer questions about the types of submission we would like to receive.

Full-length paper presentations

Challenges in corpus-based study of language-related discourses and ideologies

Adnan Ajšić
American University of Sharjah

aajsic@aus.edu
@AdnanAjsic
<https://www.aus.edu/faculty/dr-adnan-ajsic>

Recent lexical approaches to identification of (language-related) discourses, and language ideologies in particular, focus on the application of quantitative corpus-linguistic techniques to large data sets as a way to ensure more objective sampling methods and replicability of analytical procedures, as well as to minimize researcher inference (Ajšić, 2015, 2021; Fitzsimmons-Doolan, 2014; Subtirelu, 2013, 2015; Vessey, 2017). As with all research, however, there are challenges and limitations associated with this method, as well as specific data sets and research questions (see Baker & Egbert, 2016; Brookes & McEnery, 2019; Taylor & Marchi, 2018; Vessey, 2013). Based on extensive research conducted on two comprehensive, specialized research (11.6 million words from 16K articles) and comparator (22.4 million words from 37K articles) corpora comprising texts from Serbian newspapers and magazines from the period between 2003 and 2008, this paper discusses the challenges in using corpus linguistics generally, and exploratory factor analysis (and associated statistical techniques such as cluster analysis and analysis of variance) specifically, in mixed-methods research on language-related discourses and ideologies. The issues faced include: limited representativeness of data sets; sampling dilemmas; limited usefulness of popular software (e.g., WordSmith Tools) for use with inflectional languages; choice of appropriate and effective statistical techniques; discrepancy between abstract mathematical concepts such as outlier cases in factor analysis and concrete discourse patterns in 'naturally' occurring language; potential conflation of diachronic patterns inherent in some of the statistical techniques employed; and the ultimate inevitability of researcher inference.

The Role of Multilingualism Patterns and Functions on Branding and Self-Branding: An Investigation of Practices by Saudi Users on Twitter

Reem Al Madani
Cardiff University

Eligible for nomination to a student researcher award

The online presence of corporations and social media influencers plays a significant role in the way they communicate online. The increasing number of internet users means that corporations and social media influencers build creative social media strategies and interactive posts in order to engage with their followers and encourage them to share the content related to them and their brand. The current study examines the digital communication of corporations and social media influencers with a focus on the multilingual practices used to communicate with the existing and potential virtual audience on Twitter in Saudi Arabia. The study aims at investigating the use of the different forms of Modern Standard Arabic, Colloquial Arabic, English, in addition to the use of punctuations to express emotions and emoji in more than 13,000 corporations' and social media influencers' tweets.

The study combines the observation of online discourse in addition to contact with its social actors. The presentation will discuss the findings of the analysis of a corpus which was collected from the official Twitter accounts of 50 corporations and 30 social media influencers. It will explain how the different language varieties and functions that emerged from the data were coded in relation to previous studies such as Androutsopoulos (2013) and Halim & Maros (2014). The corporations' accounts represent different businesses such as oil industries, retail, sports, automotive, and education. On the other hand, the social media influencers represent different interests such as sports, media, technology, and science. Furthermore, the presentation will give an overview of the interview results with a sample of marketing executives from the corporations and social media influencers to reveal the motivations for the multilingual practices and its relation to the branding of the corporations and the self-branding of the social media influencers.

From “acoger refugiados” to “expulsar refugiados”: Exploring the victimization of refugees in a corpus of Spanish news articles

Gema Alcaraz-Mármol & Jorge Soto-Almela

Universidad de Castilla-La Mancha (UCLM) / Universidad Politécnica de Madrid (UPM)

gema.alcaraz@uclm.es

https://www.researchgate.net/profile/Gema_Alcaraz-Marmol

jorge.soto@upm.es

https://www.researchgate.net/profile/Jorge_Soto-Almela

Refugees have recently gained special attention due to the latest armed conflicts in the Middle East, forcing many people to leave their country of origin. The Spanish media have echoed this situation, publishing news articles about immigrants, refugees and asylum seekers almost daily. These migratory groups have already been explored in the print media from various linguistic perspectives such as Corpus Linguistics (Baker et al., 2008; Blinder & Allen, 2016), Critical Discourse Analysis (Baker et al., 2008; KhosraviNik, 2008, 2010) or mixed approaches combining both (Taylor, 2009; Fotopoulos & Kaimaklioti, 2016). Although we can find some studies on this area within the Spanish context (Alcaraz-Mármol and Soto-Almela, 2018; Soto-Almela and Alcaraz-Mármol, 2019), research is still limited as far as the Spanish media portrait of refugees is concerned.

This study focuses on how refugees are victimized in the Spanish written press. In particular, we will examine the semantic preference of the lemma *refugiado* (refugee) from 2010 to 2016. A corpus of 1,808,400 words collected from the digital libraries of *El Mundo* and *El País* - the two most widely read newspapers in Spain - is analysed. The analysis hinges on the quantification of *refugiado* and its categorization in terms of semantic preference, using first corpus analytic methods and then labelling or classifying the different collocates semantically. According to the Online Oxford Dictionary of English (2018), victimization is defined as “the action of singling someone out for cruel or unjust treatment”. Yet, the dictionary REDES, *Diccionario Combinatorio del Español Contemporáneo*, (Bosque, 2004) shows that not only does victimization include words of misfortune and mistreatment, but also words from the semantic fields of support, help and humanitarianism. Preliminary results show that victimization collocates represent almost 25% of the total number of collocates with the lemma *refugiado*. Three semantic subsets were identified: (a) refugees as unfairly treated, (b) as sufferers of tragic circumstances or (c) in need of help, oscillating between a paternalistic attitude under a false humanitarian positivity, and the idea of being unwanted and consequently mistreated.

News Translation and Power Negotiation: Corpus-assisted CDA Investigation of Saudi Arabia's Representations in the BBC's Coverage, 2013-2015

Asma Alqunayir

Qassim University, Saudi Arabia

Asma_A_Q@hotmail.com

@Asma_Alqunayir

Translators in newsrooms routinely apply what Bielsa and Bassnett (2009, p.10) refer to as 'absolute domestication' in which the source text (ST) is 'transedited' (Stetting, 1989, pp.371-82). This can lead to power 'abuse' and 'ethnocentric violence' against the ST, language and culture (Venuti, 1995). By incorporating corpus tools and critical discourse analysis (CDA) (Wodak and Meyer, 2016), this research investigates, first, the key themes that the BBC Monitoring Middle East (BBCM-M) service tended to focus on when reporting on Saudi Arabia from Arabic news output from 2013 to 2015 in relation to other British news sources reporting in English, second, the representations of Saudi Arabia disseminated by the BBCM-M and the extent to which it contributed to such representations and, third, the power dynamics between the Arabic source and English target texts through transediting.

The study reveals three key themes that characterise the BBCM-M's coverage of Saudi Arabia: identity, action and status and relations. There are four main representations of Saudi Arabia: 1) as filled with men of authority, who are unlike women, with real agency, 2) its rivalry with Iran and reliance on the US, 3) its three-dimensional image in relation to terrorism and 4) its paradoxical portrayal in relation to: power, policies and development. Importantly, these are 'anchored' to stereotypical 'social representations' (Moscovici, 2000) that fit into the 'system of representations' of Arabs and Muslims in Western media and literature (Said, 1978). The study also exposes a power imbalance in favour of English both prior to and during the 'transediting' process, which enabled the active contributions of BBCM M professionals to these representations. This research demonstrates how translation in a cross-cultural context such as news translation can be an apparatus of 'coloniality of power and knowledge' (Quijano, 2000). It also shows how a certain 'system of representations' can be sustained across time, languages and cultures via the constant reproduction of certain images that 'anchor' the same 'social representations' that exist in that system.

Analysing Intersectionality in Discourse: A Corpus-informed Methodology

Zainah Alshahrani & Michael Handford
Cardiff University

AlshahraniZT@cardiff.ac.uk
[@ZainahGhareeb](https://twitter.com/ZainahGhareeb)

HandfordM@cardiff.ac.uk

Eligible for nomination to a student researcher award

Intersectionality, which argues that to understand discrimination and privilege we need to see identity as an interaction of various social categories and structures, has been discussed in different fields, including law and feminism (Crenshaw 1998, 1991; Nash 2008, 2011), economics (McCall 2000, 2001; Davis 2006; Knapp 2005) and language and discourse (Baker and Levon 2016; Jaworska and Hunt 2017; Hunt and Jaworska 2019). However, while popular as a theoretical paradigm in research, intersectionality has received relatively little attention as a method of analysis (Hancock 2007). The present study, therefore, proposes a methodological framework that operationalises intersectionality as an analytical tool, in combination with methods from discourse and corpus linguistics. In so doing, it demonstrates how identities that intersect, such as gender, ethnicity, class and profession, can be analysed in discourse. It also exposes their hierarchical positions in discourse as privileged, marginalised or disadvantaged interactions in discursive contexts. The study thus aims to answer the main question: How can Intersectional identities be investigated through a corpus-assisted discourse analysis?

A corpus of BBC news articles on the topic of Saudi Women is used to demonstrate the methodology. The research question is unpacked and answered through different levels of analysis, involving macro-level 'wider context', meso level 'social structures' and micro level 'social actors' representations (Collins 1990). Identifying the social structures and pinpointing the social actors is achieved through the analysis of frequency lists, keyness, word sketching and clusters, from which semantic categories can be inferred. References to social structures and categories of identity, along with their collocates, are then analysed within concordance lines and longer extracts using discourse analytical tools. The methodology demonstrates how identities are constructed as intersectional entities and accordingly are located in the created social structure.

East, West and Westminster: A Corpus-Based Study of UK Parliamentary Discourse about the Unification of Germany

Stephen Appleton
University of Birmingham

saa777@student.bham.ac.uk

Eligible for nomination to a student researcher award

October 2020 marks thirty years since the unification of Germany. As this anniversary approaches, this paper offers a linguistic analysis of UK government ministers' discourse relating to Germany at the time of the change. Using a bespoke 1.5 million word corpus assembled from Hansard, the Official Report of proceedings in the UK Houses of Parliament, it explores correlations between changes of policy and changes of language in ministers' discourse.

Historical sources indicate that, although the government of the United Kingdom had long supported the unification of Germany, the sudden fall of the Berlin Wall in November 1989 brought reservations to the surface. It was not until February 1990 that the UK settled on a policy of accepting unification.

Using both quantitative and qualitative techniques, the paper presents evidence that a correlation between changing policy and changing language exists in three areas:

1. The prominence of Germany in ministers' discourse. This is measured both in absolute terms and relative to other countries, by frequency analysis of country names and other relevant lexical items.
2. Ministers' lexical choices to describe the evolving situation in Germany, including their use of 'unification' and 'reunification'.
3. The seniority of the ministers delivering the discourse, i.e. the role of the Prime Minister, Foreign Secretary and junior ministers in articulating the UK government's views.

The research is in the empirical tradition of Political Discourse Analysis, and the paper argues that this type of linguistic analysis can make a specialist contribution to other disciplines. In the same way as a political scientist may benefit from the specialist insights of an economist, sociologist or historian, it contends that a discourse analyst can equally enrich the understanding of politics. Specifically, it demonstrates the interplay of political and linguistic changes over a short but historically significant period of time.

Voices in Cross-lingual Global News: Same logos, but variable viewpoints

Banan Assiri
Cardiff University

AssiriBA@cardiff.ac.uk
@bananassiri

Eligible for nomination to a student researcher award

Besides their production and reception variation (Clausen 2003), international news media may generate different emphases of the same event and assert different forms of narration to support particular political, ideological and social perspectives. In this research, I employ both cross-linguistic corpus linguistics (Taylor, 2014; Vessey, 2013) and critical discourse analysis to discuss how reporting of the same events by the same news outlet in two languages may lead to varied and contradictory representations of particular political and social voices. This is of interest because it may problematize the notion that an individual news source has a coherent news story or stance and adds empirical support for the notion of 'representation' in the reporting of events (Hall, 1999). Using News Values approach (Bednarek & Caple, 2014), the voices representation (Bednarek, 2016) of political and social actors in the 'Arab Spring' in the Arabic- and English-language articles of two international news outlets, Aljazeera and the BBC, are investigated to reveal underlying mediated ideological meanings. This evaluation-analysis of the cross-linguistic/cultural corpora reveals a downscaling of news audience cultures. Also, in both Aljazeera and the BBC, there was evidence of how discourse features change to index legitimacy in the representation of political and social actors' voices in the 'Arab Spring' from contrasting cross-linguistic/cultural perspectives.

A bilingual corpus-assisted discourse study: Framing the outcomes of the Egyptian revolution

Safa Attia

Universidad Autonoma de Madrid

safattia01@gmail.com

@safattia

Eligible for nomination to a student researcher award

The latent resonance of the autocratic regimes in the North African and Middle East countries has transformed the region into what is called “Arab Spring”. Yet, each nation followed a unique path. Tunisia is the only country that stepped forward to establish democracy (Eskjær, 2012). However, none of the other nations came to a happy end; Egypt’s revolution ended up with more power and abuse in the government. Libya fell into a civil war and Yemen into a sectarian war. And Syria descended into an abyss with no end in sight (Galander, 2013). These uprisings were covered extensively and thoroughly in the world media, which helped a plethora of linguists to analyze each of them (Al-Abed Al-Haq & Hussein, 2011; Attia & Romero-Trillo, 2016; Dağtaş, 2013; Eskjær, 2012; Haider, 2016; Galander, 2013; etc.). This study scrutinizes the aftermath of the Egyptian revolution (2011-2015) through different media lenses; Al-Jazeera and Al-Arabiya –in Arabic and English versions- and BBC and CNN -in their English versions. This investigation combines both corpus linguistics and critical discourse analysis, by using AntConc 3.4.4w database (Anthony, 2014). The critical issue of this research is how the discourse of the Arab and Western media resonated in the mentioned media contexts through the analysis of collocations, concordances, and CDA. Moreover, a comparison will be made in terms of frequency distribution and story content.

The use of jihad in Extreme, Fringe and Moderate discourse

Paul Baker

Lancaster University

p.baker@lancaster.ac.uk

[@_paulbaker_](https://www.instagram.com/_paulbaker_)

<https://www.lancaster.ac.uk/linguistics/about/people/Paul-Baker>

This paper examines use of the word the word jihad in English texts that had been found in the possession of British citizens who had been successfully convicted, under counter-terrorism legislation, of crimes or of planning crimes (Holbrook, 2019:2). Based on expert close reading of the texts, they had been sorted into one of three categories (ibid 4-5). One hundred and seventy Extreme texts (1,775,340 words) endorsed or glorified violence in a contemporary context and/or stark dehumanisation. Fifty-four Fringe texts (486,650 words) were religiously or ideologically conservative and isolationist as well as politically radical and confrontational, but without any justifications conveyed for violence in present-day scenarios. Finally, 51 Moderate texts (1,721,442 words) contained no endorsement of violence or hatred towards identified communities.

I examine use of the word jihad (and alternative forms like jihaad) across these three corpora. As the most frequent word in this corpus which can be used to refer to violence, it was analysed as part of a larger study examining linguistic differences between the three corpora.

Using Sketch Engine, I examine significant gramrels containing jihad to obtain a sense of the different forms of representation around it. I also consider frequencies of semantically-related words like fight and strive, and collocates of jihad like wage, make and perform, considering how such verbs are used in reference corpora in order to understand how they imbue jihad with particular senses. The analysis identifies four main representations around jihad, which have different frequencies across the three corpora, indicating how the term's meaning is contested and reformulated in different contexts, contributing towards an exhortation towards violence in Extremist discourse.

A genuine fantasy? Comparing distinctive language used to represent people with schizophrenia in the UK tabloid and broadsheet press

James Balfour
Lancaster University

Eligible for nomination to a student researcher award

Studies in a variety of disciplines have demonstrated that media portrayals of people with schizophrenia are typically negative and stereotypical (e.g. Clement and Foster, 2008). A small number of studies examining newspaper portrayals have identified representations in the British tabloids as of particular concern. Here, people with schizophrenia are typically represented more negatively than the broadsheets, where they are typically represented in the context of violent crime, frequently using intolerant and dismissive labels, such as 'maniac' and 'psycho' (Bowen et al, 2019; Clement and Foster, 2008). These studies, while informative, are typically narrow in scope, typically focussing on the same set of language features at the expense of the broader picture. These studies also focus on features in the tabloids at the expense of the broadsheets.

In this talk, I discuss the findings of a keyword analysis carried out as part of a broader project examining representations of people with schizophrenia in the UK press between 2000 and 2015. I conducted a contrastive keywords analysis of the tabloid and broadsheet subcorpora. The top 100 salient keywords were then examined qualitatively in order to determine patterns in usage. The distinctive language in the tabloids suggests a tendency to represent people with schizophrenia in terms of a 'undesired differentness' (Goffman, 1963). This is illustrated by a close analysis of the verb phrase 'released into'. The broadsheet keywords suggest a more marked tendency to represent people with schizophrenia more positively, engaging in academic debates around the aetiology of the disorder. However, the broadsheets also represented people with schizophrenia in terms of a 'desired differentness', as inherently creative artists, which is also identified as potentially problematic.

Using CADS to uncover how right-wing extremists generate an alternative rape culture in online narratives

Kate Barber
Cardiff University

barberk@cardiff.ac.uk
@katebarber2015

Eligible for nomination to a student researcher award

Right-wing extremism, in its many guises, is spreading and expanding internationally (Caiani & Kröll 2015, p 331), particularly through online platforms. Linguistic analyses of far-right discourses have traditionally focused on nationalist rhetoric or racist and ethnoreligious-based invective (see, for example, Brindle 2016). Recently, academic focus has been on group identity within the far-right (Gray 2018, Lorenzo-Dus & Nouri 2020). However, the explicit anti-feminist stance held by some far-right groups, and how this contributes to their collective identities and ideologies, remains underexplored. Related to this, the growing men's rights activist movement known as the Manosphere is appearing to move closer to the far right online (Romano 2016, Marwick & Caplan 2018). The overall aim of this study is to highlight the extent to which ascribed and inhabited identities in both the Alternative Right (Alt-Right) and the Manosphere converge in the way they portray victims and perpetrators of sexual violence, particularly through the manipulation of narrative discourses.

This paper outlines findings from an ongoing corpus-assisted critical discourse analysis of 100 blog posts from Alt-Right and Manosphere sites. Following a brief overview of the data selection process, which involved building a network analysis of right-wing websites, and the XML mark up of lexical and clausal elements of the posts, the paper examines two aspects of the study: 1) the similarities and differences between the two factions in the way identities are indexed; and 2) how identities shift between narrative components to generate a shared alternative rape culture which excuses sexual violence against women while promoting white male victimhood. It is hoped that this study will contribute to research on the relationship between identities and narrative in extremist rhetoric and, eventually, to work on countering violent extremism online.

Correlation, collocation and cohesion: A corpus-based critical analysis of violent jihadist discourse

*Gavin Brookes & Tony McEnery
Lancaster University*

This paper explores the language of violent jihad, focussing in upon lexis encoding concepts from Islam. Through the use of correlation statistics, we will demonstrate that the words encoding such concepts distribute in dependent relationships across different types of texts. We will show that the correlation between the words cannot be simply explained in terms of collocation. Rather the correlation is evidence of other forms of cohesion at work in the texts. The variation in patterns of cohesion across a spectrum of texts from those advocating violence to those which do not promote violence demonstrates how these concepts are contested and redefined by violent jihadists and the role that collocation and other forms of lexical cohesion can play in the process. The paper concludes that the terms, and their redefinition, are a key part of the symbolic capital used by groups to create identities which licence violence.

How to be an expert online? A corpus-based study of expert identity construction on a Spanish medical forum

Barbara De Cock & Carolina Figueras Bates
Université catholique de Louvain / Universitat de Barcelona

barbara.decock@uclouvain.be
@barbaradecock
<http://www.uclouvain.be/barbara.decock>

The use of online communication tool is rapidly increasing in the medical sector, in spite of parts of society and of the medical profession being rather wary as to this evolution. In this paper, we want to contribute to this field by showing how the status of 'expert' is constructed in a corpus of interactions of a Spanish medical online forum. The forum under scrutiny contains both threads where only laypersons contribute and threads which are moderated by an actively participating qualified healthcare professional (doctors, nurses, dieticians...). In our study, we wish to focus on how the identity as an 'expert' is constructed by those asking questions and answering them. We will compare the case of healthcare professionals as experts with the case of non-qualified laypersons who have a certain expertise (e.g. because they have a long experience as patient of a certain illness) and, as such, (try to) intervene as an expert in the forum.

In order to do so, we will carry out a corpus analysis focusing on explicit references to the fact of being an expert (e.g. mentions of the profession or title) as well as on other ways of constructing expert status (e.g. references to experience). Second, we will look into the strategies used by both professionals and laypersons to legitimize knowledge as expert knowledge (e.g. reference to a scientific source). Finally, we will compare the strategies put in place in the discourse by and towards the qualified healthcare professionals vs. those in the discourse by and towards the non-qualified laypersons, in order to shed light on whether the expert status of these two groups of participants is discursively constructed in a similar way or not.

Understanding Populism: A Linguistic Analysis of Czech Parliamentary Discourse

Martina Berrocal, Václav Cvrček & David Lukeš
Institut für Slawistik und Kaukasusstudien

@MartuBerro

Populism has become a prominent trait of current political action in a great number of countries in Europe and beyond. A great part of the electoral success of populist politicians and parties is often attributed to a simple and effective communication with the electorate. A thorough linguistic analysis of discourse structure, that is, discursive strategies and linguistic means populist politicians use, will help us to understand this phenomenon that is “pathological of democracy” (Taggart, 2002) and that poses a potential threat to the pluralistic democracy (Müller 2016, Mudde and Kaltwasser 2014, 2017). On that account, we examine all the speeches of the Lower Parliamentary Chamber (Poslanecká sněmovna) Czech parliamentary discourse within three mandate periods (2010-2013, 2013-2017, 2017-2019), accounting for a spectrum of populist and non-populist discourses.

Methodically, the study applies the Multi-Dimensional Analysis (MDA) of Czech registers, which maps the repertoire of the alternating linguistic means. Mainly, it determines the invariant and the scope of the variation of Czech registers, where linguistic variation is perceived as a materialization of the linguistic choices made by the speakers (Cvrček et al. 2018b, 293).

The analysis studies linguistic profiles of the parliamentary speakers and aims to determine how populist language (and its varieties) differ from other parliamentary register varieties. Moreover, it examines the impact of extra-linguistic factors (e.g. party affiliation) or sociolinguistic variables (sex, education, age) on the variation of parliamentary register.

A comparative study of the discourse of Voice Output Communication Aids users' with and without their aids in the workplace.

Julie Bouchard & Laura Di Ferrante

Université du Québec à Chicoutimi / Sapienza Università di Roma

@juliebouchard02

<http://www.uqac.ca/portfolio/julie2bouchard/>

laura.diferrante@uniroma1.it

@ladiferro

In this comparative study, we investigate the talk of Voice Output Communication Aids (VOCA) users when they use their devices compared to when they choose not to and to produce intelligible vocalizations instead. VOCA are used to communicate more effectively by people who have a speech impairment. These devices produce synthesized speech from users' input. The input-to output process is time consuming, and the talk may be ready too late in the conversation. Thus, to be able to participate in the conversation, VOCA users may rely on vocalizations (Bouchard 2016; Di Ferrante and Bouchard 2020; Dominowska 2002); with vocalizations the talk is then in the same "time stream" as the other interlocutors but the problem of intelligibility is not resolved. This duality has an effect on the choices VOCA users make in terms of mode (VOCA or vocalization) when they interact with their peers. Using corpus linguistics methodology, we used the AAC and non-AAC workplace corpus (Pickering et al. 2019, Pickering & Bruce 2009), specifically the data collected from the four participants who used VOCAs. The talk produced using the VOCA device and the intelligible natural talk made through vocalization were separated into two collections that were analyzed using the AntConc software (Anthony, 2019). A word count analysis showed that the participants did not use the modes in similar proportions. The ratio of intelligible vs unintelligible vocalizations was used to determine the individual participant's ability to produce intelligible vocalizations. Interestingly, this result was found to correlate with the use of both modes in three of the four cases. Finally, word type-token ratios show the participants used more varied vocabulary with their devices than when they vocalized. This can be explained by the actions accomplished using the different modes.

Examining the use of reported speech in the PrEPUK corpus

Luke Collins
Lancaster University

l.collins3@lancaster.ac.uk
@LukeCCollins

Eligible for nomination to a student researcher award

This study examines the use of reported speech in a corpus of news articles relating to pre-exposure prophylaxis (PrEP) published in the UK in the period 2015-2019. Pre-Exposure Prophylaxis (PrEP) refers to treatment options designed to prevent the transmission of the Human Immunodeficiency Virus (HIV), specifically variants of a combination pill (tenofovir and emtricitabine) that has been reported to reduce the risk of infection by upwards of 86% (McCormack et al., 2016). Jaspal and Nerlich (2017: 489) report that in the UK press (2008-2015), although PrEP was represented as inspiring 'hope' and as a 'weapon in the battle against HIV/AIDS', "strategic expert quoting was employed to construct PrEP as a threat". The 'polarised' coverage of PrEP is considered to be one of the barriers to its wider provision in the UK.

A keyness analysis of the data presented here identified the broader patterns in the use of reported speech throughout 2015-2019, including which commentators were frequently quoted in the debates around PrEP provision. Focusing on the oft-cited quotes from key commentators in the debate in the context of the article, I refer to forms of attribution and style (Bednarek, 2006) to consider how comments are variously recontextualised as argumentation across different news outlets in the UK, including how they are positioned alongside reports of contrasting perspectives.

How critical is critical discourse analysis of advertising: Corpus-driven findings and methodological implications

Yixiong Chen & Csilla Weninger

National Institute of Education, Nanyang Technological University

chen023@e.ntu.edu.sg

Eligible for nomination to a student researcher award

Critical discourse analysis (CDA), which aims to explicate the genesis and maintenance of hegemony through discursive analysis, has been hailed for its potential of understanding social progresses. However, CDA has also been a target of critiques from its very beginning. On such recent critique is on its descriptive components: too much CDA, it is claimed, has been describing discursive practices rather than explaining social ideological repercussions of such practices. To address this critique, a corpus-driven genre analysis of CDA discourses on advertising was conducted. Specifically, a corpus was built from 98 published CDA research articles on advertising and was subsequently tagged with Swales' (1990) framework of academic genre and discursive moves. Preliminary analysis of the moves reveals uneven distribution of analytical efforts in the research literature: while more than 95% of the moves congregate on the descriptive and interpretive dimensions (Fairclough, 1989; 2010), few (n=9) fall in the explanatory dimensions. The finding suggests that CDA of advertising is not as critical as it aims to be, because few analytical efforts are directed to the explanatory dimensions that distinguishes CDA from other non-critical forms of discourse studies. It is argued that the findings can be attributed to methodological insufficiencies of SFG related frameworks used in the research literature, and that empirically oriented approaches to advertising offer more opportunities for CDA's critical purposes.

A Corpus-based Study of Discursive Representation of LGBT Identities and Movement in the Bangkok Post

*Duangkamol Chuaikun & Raksangob Wijitsopon
Chulalongkorn University*

Eligible for nomination to a student researcher award

The present study examines how each group of LGBT (lesbian, gay, bisexual and transgender) is represented discursively in the Bangkok Post, a major English-language newspaper in Thailand. The data comprise news articles from 2012 to 2019, which contain the words LGBT, lesbian(s), gay(s), bisexual, and transgender, making up a corpus of 217 texts with 88,054 tokens. Statistically significant collocates of each term were extracted via AntConc. The collocates were then categorised according to their semantic relations as observed from their concordance lines, hence forming different groups of semantic preference associated with each search term. Findings from the Bangkok Post corpus show that the different terms share a common semantic preference of “political movement”, consisting of such collocates as rights and community. However, the terms gay and transgender, two most frequent terms of the LGBT group, were represented differently; their significant collocates are also associated with “crimes”, “disease” and “beauty”. To explore to what extent these semantic preference patterns are idiosyncratic of Bangkok Post news about LGBT, the newspaper sub-corpus of COCA and the SiBol English broadsheet newspaper corpus were also examined. It is found that only the “political movement” collocational group is shared between the Bangkok Post and the other two newspaper corpora. This suggests that while the Bangkok Post echoes and contributes to the international dominance of the LGBT movement discourse, it plays a role in promoting distinctive discourses surrounding LGBT identities in Thailand as being associated with crime and health concerns of the country. This raises questions about Thailand’s governmental organizations’ claim that the country is LGBT-friendly.

A corpus-based study of representations of social care during the 2019 UK General Election campaign

Carmen Dayrell & Elena Semino
Lancaster University

c.dayrell@lancaster.ac.uk

@carmendayrell

<https://www.lancaster.ac.uk/people-profiles/carmen-dayrell>

e.semino@lancaster.ac.uk

@elenasemino

<https://www.lancaster.ac.uk/linguistics/about/people/Elena-Semino>

This paper employs corpus-based discourse analysis to examine how social care was talked about by the three main UK political parties and by Twitter users during the 2019 election campaign. This research was commissioned by representatives of activist organisations in the charity sector who seek to bring about major positive changes in discourse, policy and practice around social care.

Three corpora were analysed:

Manifestos Corpus: Election manifestos for the Labour Party, the Conservative Party and the Liberal Democrats; Speeches Corpus: Campaign and manifesto launch speeches delivered by the leaders of the above parties; Twitter Corpus: 897 tweets (41,081 words) that contain one of the following: social care OR free personal care OR #socialcare OR #freepersonalcare, posted between 27th Oct 2019 and 15th Dec 2019.

The analysis proceeded through a combination of corpus linguistics techniques (collocations and concordance lines) and, when evidence was scarce, traditional 'manual' discourse analysis. Across all three sets of data, we first examined the frequency distributions and discourse patterns of the phrases 'social care' and 'free personal care'. We then focused on references to social groups in the context of discussions of social care, focussing on: (i) people with physical and/or mental disabilities; (ii) the elderly; (iii) adults; and (iv) children. To complement the analysis, we explored the narratives around care providers and the phrase 'mental health'.

We report our key findings, including particularly key differences among three corpora in how social care is talked about. The paper concludes with some reflections on the implications of our findings, and the challenges and rewards of conducting and disseminating this kind of research.

Evaluation of medication safety oversight programs in the United States using corpus linguistics methods

*Carmen Dayrell, Darren M. Triller, Youssef Bessada & Elena Semino
Lancaster University / WellScriptEd Consulting / Albany College of Pharmacy and Health Science /
Lancaster University*

c.dayrell@lancaster.ac.uk
@carmendayrell
<https://www.lancaster.ac.uk/people-profiles/carmen-dayrell>

dtriller@WellScriptEdConsulting.com

youssef.bessada@acphs.edu

e.semino@lancaster.ac.uk
@elenasemino
<https://www.lancaster.ac.uk/linguistics/about/people/Elena-Semino>

Adverse drug events are a major cause of preventable patient harm and in 2014 the United States National Action Plan for Adverse Drug Event Prevention (NAP-ADEP) identified three drug classes (opioids, anticoagulants, and diabetes agents) as warranting coordinated Federal action to improve patient safety. Since then, the U.S. Department of Health and Human Services has supported state and local entities in providing prevention, treatment, and recovery services by thus reducing harms associated with opioids. However, it is unclear whether substantive efforts have been undertaken to reduce patient harms associated with the other drug classes: anticoagulants and diabetes agents.

Seeking to evaluate the current status of oversight of these drugs across care settings, we examine the extent to which the language used within current authoritative oversight documents addresses the key drug safety concerns identified in the NAP-ADEP. This study thus uses four corpora:

- i. The NAP-ADEP corpus (49,489 words);
- ii. Current surveyor guidance for hospitals (192,165 words);
- iii. Current surveyor guidance for skilled nursing facilities (308,458 words);
- iv. Guidance for Medicare drug plans (2011-2019; 407,901 words), which describe modifications to drug formulary and quality measurement criteria.

Mentions of each drug class were identified through Wmatrix semantic tagging. Relevant terms were selected from the domains of Medicines and Medical Treatment, Health and Disease, Substances and Materials Generally, and Unmatched Terms. Frequency distribution of terms and, as frequencies allowed, collocates were determined using LancsBox. Dominant discourse patterns were identified through close reading of concordance lines.

This analysis shows that the NAP-ADEP addresses safety concerns associated with opioids,

anticoagulants, and diabetes agents as being of similar priority and importance. The other three documents also address safety concerns with opioids, yet such concerns are not clearly articulated for anticoagulants or diabetes agents. This paper concludes with a discussion of the implications of these findings.

A corpus-based study of the boundary between opinion and hatespeech in Belgian French-speaking online political discourse

Barbara De Cock, Pauline Dupret, Philippe Hambye & Andrea Pizarro Pedraza
Université catholique de Louvain

barbara.decock@uclouvain.be
[@barbaradecock](#)
<http://www.uclouvain.be/barbara.decock>

In this study, we will present an analysis of discourses on the boundary between opinion expression and hatespeech, based on a corpus of online discourse of Belgian French-speaking politicians. Our corpus consists of the Facebook and Twitter productions from leading politicians from all major parties, which were collected one month right before and one month during the 2019 electoral campaign for the regional, national and European elections on May 26th 2019. While discourses falling within the legal qualification of hatespeech are extremely rare in this corpus, we will analyze the productions that are on the boundary between opinion expression and hatespeech, engaging in polarizing and stereotyping representations of certain communities based on criteria related to ethnicity, religion and sexual orientation.

We look into the linguistic strategies that are used to construe these polarizing representations, such as the use of deictics, generalization, metaphorical language use, and the construction of agentivity and intentionality (i.e. the representation of certain actions as intentional deeds, thus construing e.g. an immigrant group as willingly harming the local population). We both analyze the use of these individual strategies in our corpus and show how they are combined in creating a discourse that represents these communities as a threat. Finally, we will also discuss to which extent there are differences between discourse produced before and during the electoral campaign.

In doing so, we show how a detailed discursive analysis based on a corpus of contemporary online data can contribute to a better understanding of the general public debate and of people's conceptualisations concerning these communities.

Metaphors over time and spaces. A cross-linguistic corpus-assisted newspaper discourse analysis of the metaphorical representation of migration. The case of two New York newspapers and an Italian Newspaper between 20th and 21st Century.

Dario Del Fante

University of Padova / University of Sussex

The proposed research study focuses on two essential qualities of humanity: migratory movements and metaphors. A large body of work concerning the metaphorical representation of migrants (Santa Ana 1999; Semino 2008; Dervinytė 2009; Taylor 2014) have shown how newspaper and public discourse have represented immigrants negatively. The only remarkable exception is presented by Salahshour (2016) and Taylor (2020) who show that immigrants are positively represented as “force” which is “gained” by the country in, respectively, a contemporary newspaper corpus in New Zealand and in 2018 British parliamentary debates. Studies migration discourse seen from both a diachronic perspective and cross-linguistic perspective seems to be, however, under-researched. For this reason, I intend to investigate the weight of time and cultural factors in their relations with the use of metaphors. This study approaches metaphors through the theoretical framework of Cross-Linguistic Corpus Assisted Discourse Study (Partington et al. 2013; Vessey 2013; Taylor & Del Fante forthcoming). The dataset consists of two Italian corpora (ItMig1900 and ItMig2000) and two American English corpora (AmMig1900 and AmMig2000). These cover the periods 1900-10 and 2000-2010 respectively, two significant moments for the migratory history (Daniels 2002, Migration Policy Institute 2016). A corpus-driven (Steen et al. 2010), approach has been adopted to retrieve, with the aid of the corpus tool Lancsbox (Brezina, Timperley and McEnery 2018), all instances of metaphors, subsequently discussed through a qualitative analysis. Results show that migration in both languages is mainly represented in terms of quantity through liquid metaphors. The older corpora present similar results. The most used metaphors in the AmMig1900 and ItMig1900, are, respectively, “tide of”, “stream of”, “wave of and “flusso di” and “afflusso di” (influx), which, nevertheless, show a very low frequency of usage. The same metaphorical expressions have a higher frequency in the contemporary data. It would be possible that this type of metaphor firstly appeared during the 1900s period in both countries and then it has been conventionalized, losing its novelty as creative metaphor (Charteris-Black 2004).

The discursive construction of contemporary Spanish emigrants in the press

Marisol del-Teso-Craviotto

Miami University

My investigation explores media discourse on the Spanish emigration trend that followed the 2008 economic crisis. While immigration to Spain has received much attention, recent emigration of Spaniards has gone mostly unnoticed among discourse scholars, even though Spanish emigrants present linguistic and ideological challenges to the current configuration of the migration imaginary in Spain. Evidence from political discourse showed concerted attempts to separate Spaniards from other migrants conceptually, and thus my research aims at explaining the linguistic and ideological strategies involved in such attempts. Specifically, my study sought to identify the semantic features of the lemma *emigra**, establish the degree of membership of contemporary Spanish emigrants in the semantic prototype of the lemma, and explore the ideological implications derived from the semantic configuration of this lemma and its use in discourse.

I used a targeted corpus of news on Spanish emigrants from 2011 and 2015 from three mainstream Spanish newspapers, *El País*, *El Mundo*, and *ABC*. A total of 78 texts were analyzed in terms of concordances, collocates and clusters of the lemma *emigra**, word frequency, and lexical alternatives to the lemma. The results were then compared with a larger corpus (*Corpus del Español*, 2016) to ascertain whether the conclusions could be generalized.

Results show the lemma *emigra** consists of a few core semantic features, including “moving out of one’s birth country”, “economic reasons”, “victimization”, “poverty”, and “unskilled/uneducated people”. When referring to contemporary Spanish emigrants, however, only the first three semantic features are relevant. Additionally, lexical alternatives such as *exiliados* (‘exiles’), situate contemporary Spanish emigrants as peripheral members of the lemma, and thus as conceptually different from other emigrants. Therefore, it could be argued that *emigra** cannot simply be defined as an outward population displacement, but has additional semantic traits that allow for ideological uses of the lemma in discourse.

US State Departments press briefings and China Foreign Affairs Ministry. Differences and similarities in what they say and how they say it.

Alison Duguid & Alan Scott Partington
Siena University / Bologna University

alanscott.partington@unibo.it
<http://www.lilec.it/clb/>
<https://www.facebook.com/corpuslinguisticsSiBol/>

In this paper, we analyse two sets of transcripts of press briefings on international affairs both from 2018, namely, US State Departments press briefings (USSD-18) and the English-language translation transcriptions of the China Foreign Affairs Ministry press briefings (CMFA-18). The CMFA-2018 corpus consists of 223 briefings containing 356,385 tokens. The USSD-2018 corpus consists of 74 briefings containing 495,635 tokens. We have also collected similar sets for 2019, but found a sudden change, US press briefings discontinued after August and preference given to one to interviews with friendly media hosts.

We compare and contrast these datasets from the perspectives of content analysis, and linguistically, how they are discussed as two variations of a discourse type intended and perceived as a complex communicative event by all participants: the podiums (acting variously as principles, authors and vehicles), the journalists (addressees but also co-principles) and the real beneficiaries (other 'over-hearing' powers, the media and the public).

We analyse sets of key items and clusters, in particular key items which are completely absent from the other corpus, and concordances of these items. The relationship between the podiums and the press is complex; their goals are very different as are the risks they confront – to their own face, the administrations' faces and even to national faces. We discuss the negotiation of meanings including via the pervasive employment of forced priming (flooding the discourse with strategic messaging; Duguid 2009) performed by the officials and the degree of resistance to such forcing offered by their interlocutors. The most apparent difference between the two datasets is the formality and familiarity between podium and press corps.

Our analyses also reveal both differences and similarities in the aims and preoccupations of the two administrations evidenced through evaluative choices. For example we see how the Chinese Foreign Ministry would like things to be done by the key items which are qualifying evaluative adverbs jointly, mutually, properly, smoothly, resolutely, sternly, firmly, earnestly. Such qualifying adverbs are missing in the US data where most adverbials are stance adverbials, such as unfortunately, frankly, hopefully, obviously, clearly, basically, certainly and thus do not give an idea of how the State department likes things to be done, but rather signal dialogistic positioning.

Categorising keywords in discourse – a case study of texts on bacterial resistance

Natalie Dykes, Joachim Peters & Stefan Evert
FAU Erlangen-Nürnberg

Eligible for nomination to a student researcher award

Keywords are a standard method of corpus-based discourse analysis. However, there is no ‘best practice’ for generalising from words to discursive strategies (cf. Baker 2004) and the need for prior knowledge about the topic is unclear. We compare two keyword analyses which either do or do not assume prior familiarity.

We explore a German web corpus of 14 million tokens on multidrug-resistant organisms (MDRO). We compare mass media articles and sources on alternative medicine. The first analysis was based on a manual pre-study of similar texts, yielding a categorisation scheme of hermeneutically identified categories (metaphors, actors and topoi). The second analysis focuses on formal linguistic rather than topic-based criteria (word formation patterns, jargon, proper nouns and semantic tags (cf. Rayson et al. 2004)).

We find that both studies lead to similar conclusions. For instance, the first study found that newspapers often attribute MDRO to a lack of hygiene, while alternative sources tend to promote unconventional ailments. In the second study, the newspaper keywords commonly feature the tag B4 (Cleaning and personal care) and a higher frequency of compounds with constituents like hygiene, while alternative sources more commonly mention substances and plants.

We propose that an analysis guided by general semantic and grammatical criteria is a promising entry point to discourse. Moreover, the use of content-independent criteria for categorisation may help to reduce researcher bias, enhance comparability across corpora and give more weight to the role of linguistic form.

Linguistic othering of Arabs and Muslims through adjectival modifiers in the press

Irene Elmerot
Stockholm University

irene.elmerot@slav.su.se
<https://www.su.se/english/profiles/irel5167-1.364672>

Eligible for nomination to a student researcher award

This research paper examines the polarization between a national majority and a group of minorities through adjectives that modify nouns over 30 years' time. Charlotte Taylor (2019) and Zlinská (2019) have inter alia shown, in different ways, how such polarization through the naming of others may cause vivid discussions and even conflicts.

The aim here is to analyse whether the nouns for Arabs and Muslims are mainly modified by negative or positive adjectives in the material, compared to the reference group. And are there fluctuations over time, and if so, do they reflect the conflicts?

The study uses a partly corpus-based, partly corpus-driven (Baker and McEnery 2015: 47), critical discourse analysis, with the help of proximity measures (Cvrček 2014). The corpus searches are performed with a user-specific journalistic subcorpus from the Czech National Corpus (Hnátková, Křen, Procházka and Skoumalová 2014). The searches combine 105 nouns for Arab and Muslim nations and persons, with 773 adjectives from the Subjectivity Lexicon (Veselovská, Hajič and Šindlerová 2014) for Czech. The objective is to get a more comprehensive view of the polarization between these people that is reflected in the news press discourse, and these are therefore compared with a reference group comprising 2 619 nouns for many different kinds of humans and nations, that also co-occur with the same adjectives.

The results show how polarization is created in the studied news language, and how international events shape the trends for these denominations, but also how a group that get a consistent negative reporting are easily made into a typical out-group. The cumulative effect of this and similar othering is here shown for a Central European language that we need to know more about to be able to understand how its society works, and as a result better understand its people."

Personalising stock responses: A corpus-based study of 'mixed' replies to patient feedback on the NHS Choices website

Craig Evans
Lancaster University

Eligible for nomination to a student researcher award

As a form of patient-engagement, the NHS encourages patients to share their healthcare experiences through websites like NHS Choices. However, such engagement is undermined by NHS staff using stock replies to respond to feedback as a time-saving measure (Baker, Brookes and Evans, 2019). Lack of resources may preclude staff producing individually written replies, so 'mixed' replies that combine efficient stock text with more personal elements indicate a possible compromise.

In this research paper, I investigate the extent to which mixed replies represent such a compromise by considering 22,907 mixed replies (1,580,455 words) produced by healthcare staff. First, a keyword analysis reveals evidence of non-specific individualised elements, i.e. a balance between general and personal. Other keywords represent verbatim recap of the wording in patient feedback (e.g. ease as in pleased to hear you felt at ease), which indicates an efficient personalisation strategy. Keywords reveal another reason for mixed replies: the addition of individualised details to stock text for promotional purposes (e.g. five as in five-star to draw attention to comments' good score ratings).

Secondly, a variation analysis of 50 pairs of mixed replies with reused text in common (a qualitative sample identified using Duplicate Contents in WordSmith 7) produces findings that corroborate the personalisation and promotional tendencies highlighted by keywords. This analysis also reveals how mixed replies can involve subtle relational work, such as withheld personalisation in reply to highly critical feedback. In this respect, factors in the local interactional context can influence language choices as much as the desire to balance efficiency and individualisation.

A Corpus-Driven Analysis of Responsibility in Irish Oral History Documents

Christopher Fitzgerald
University of Limerick

Eligible for nomination to a student researcher award

This paper outlines the practicalities of constructing a representative corpus from a large archive of oral history documents and presents an analysis of how responsibility is linguistically conveyed within these statements. The Irish Bureau of Military History was established by the Irish Minister for Defense in 1947 to gather as much information as possible from those involved in the Irish independence movement. This resulted in one of the largest oral history collections of its kind ever undertaken, comprised of over 36,000 pages of statements. Since becoming publicly available in 2003, these oral histories have been used as valuable sources of historical data relating to the 1916 Rising and War of Independence, but have hitherto not been investigated linguistically. By constructing a representative corpus composed of these statements, a corpus-driven approach (Tognini-Bonelli 2001) was undertaken that identified linguistic devices that signal agency. These devices, namely; hedging (Hyland 1996), pronominal choice (Wilson 1990), passive voice (Bohner 2001) and reported speech are analysed for the extent to which these are employed to distribute responsibility (Hill and Irvine 1993) for actions depicted in the statements. It is seen that these devices, especially hedging through expressions of memory (Achugar 2016), are both explicitly and covertly used by participants to attribute responsibility for historical events to themselves and others. This presentation will also outline the potential contribution of oral histories to the field of applied linguistics while establishing the generic linguistic characteristics of oral histories. Finally, a case is made for the potential value of interdisciplinary collaboration between oral historians and applied linguists and points out the areas in which both fields overlap.

“There's only one pot of money it can come from”: A corpus based analysis of the International Baccalaureate in Canada’s provinces

Saira Fitzgerald
Lancaster University

s.fitzgerald@lancaster.ac.uk

This study explores the influence of the International Baccalaureate (IB), one of the actors in the “global education industry” (Verger et al., 2016), on Canada’s public education system. The study has a dual purpose: (i) to show how computer assisted discourse analysis and a 1.5 million word specialized corpus can provide insights into the beliefs and values of a particular social group, making possible a “new way of looking at old puzzles” (Stubbs, 2010); and (ii) to examine how the IB is represented in the Canadian provincial context and make visible the way values and attitudes associated with the IB have real world consequences for local provincial education systems.

In Canada, education systems are the mandate of provincial and territorial governments, but there is surprisingly little research comparing the IB across provinces. Nor is there much focus on costs and funding. The present study seeks to address this gap by comparing representations of the IB in all 10 provinces to gain insight into the variation in IB uptake, and shed new light on local decision-making about how and why the IB is adopted.

Keyword analysis was conducted using AntConc (Version 3.4.3) (Anthony, 2014), comparing different provincial subcorpora to the corpus as a whole, and helped to identify the different preoccupations or discourses surrounding the IB in each province, and the different roles the IB plays in furthering particular agendas. Although keywords across provinces revealed lexical differences in how the IB is represented, in-depth contextual analysis indicated a similarity of concerns, particularly regarding funding and cost.

Ideological Discourses in a 21st Century, Multi-register U.S. Migrant Education Corpus

Shannon Fitzsimmons-Doolan

Texas A&M University-Corpus Christi

Because more than 25% of the US K-12 population are children of immigrants (Migration Policy Institute, 2018), immigration discourses affect educational policy and practice. The 21st century has witnessed the sharp rise of both pejorative migration discourses and restrictive migration policies internationally (Wodak, 2015; Pierce et al. 2018), portending negative educational outcomes for migrant students and children of migrants around the world. Yet, while there has been a recent increase in studies of migration discourse generally, there are very few studies that focus on the intersection of migration and education or on cross-register discourses. This study addresses these gaps and asks, Which ideological discourses are present in a 21st century, multi-register corpus of texts about immigrant education?

The study identifies ideological discourses from the U.S. Immigrant Education Corpus—a more than nine million word corpus composed of five register-based subcorpora on the topic of immigrant education (newspaper comments, national newspapers, regional newspapers, federal Department of Education (DOE) webpages, selected state DOE webpages) spanning 2003 to 2018. To identify ideological discourses —“conventional ways of talking that both create and are created by conventional ways of thinking” (Johnstone, 2009, p. 3)—the author identified collocates of the node **migr** which met frequency, syntactic, and text dispersion criteria. Normed frequency counts of these collocates per text were used in a multidimensional (MD) analysis resulting in 10 factors. Use of highly loading collocates were analyzed qualitatively within and across texts, identifying ideological discourses such as government programs serve children in need and U.S. immigration policies are problematic, but there is no consensus for solutions.

This paper will present the findings and implications of this analysis but will focus most on the insights and considerations resulting from the methodological approach of using lexical variables in MD analyses to identify ideological discourses in a multi-register corpus.

Strategic presupposition in the U.S. Congressional reauthorization of the Higher Education Act

Kristen Fleckenstein

Coastal Carolina University

kfleckens@coastal.edu

@keflecks

This paper explores how presupposition has been used as an ideological marker in U.S. Congressional debate surrounding the reauthorization of the Higher Education Act (HEA) to indicate an affective evaluative position on the speaker's part (Degano 2007; Cepollaro & Stojanovic 2016). It takes a diachronic approach, comparing corpus data from the last four reauthorization hearings in 1998, 2008, 2015, and 2019 to mark how shifting attitudes toward campus sexual assault are reflected in the language used by members of Congress.

The analysis presented in this paper uses the integrated methodology of corpus linguistics and discourse analysis (Baker 2006; Partington et al. 2013). A first-step keyword analysis and concordance line analysis conducted using the AntConc software package (Anthony 2019) revealed that in the most recent hearing, there is a shift toward a discursive focus on the rights of the accused. In addition, campus sexual assault is framed with regard to legal policy implications in the 2019 reauthorization hearing.

The presupposition analysis presented in this paper draws on the set of presupposition triggers outlined by Levinson (1983) and focuses on stance expression in interpersonal meaning, especially as representative of the attitudes and values (Thompson and Hunston 2001; Degano 2007) expressed by members of Congress in the reauthorization hearings. Within the corpora gathered for this study, I examine verbs of judging, iteratives, factive verbs, implicative verbs, and change of state verbs as well as the specific types of evaluative presuppositions that they trigger; of particular interest is the use of *we know* (rather than *I know* or *I believe*) across all four corpora to invoke the presupposition that the culture of sexual assault on campuses is an immutable fact, and further that this is a widely accepted belief rather than a controversial statement that would seem to undermine the purpose of the HEA.

Evaluation and ideology in the representation of anarchists in two comparable corpora of Greek and English journalistic texts

Georgia Fragaki

University of Peloponnese

g.fragkaki@uop.gr

<http://phil.uop.gr/index.php/en/people>

This paper studies the representation of the social group of anarchists in two comparable corpora of journalistic texts, published on the Greek and English websites of the right-wing newspaper “I Kathimerini”. The corpora cover a period of 19 years (2001–2019) and were compiled by extracting all texts in which a number of relevant search terms occur (5,175,774 words in Greek and 1,506,861 words in English).

My study focuses on the lemmas *anarçikos* and *anarchist*, including both noun and adjective uses, occurring 4,503 and 2,350 times respectively. Their frequency of occurrence progressively increases in the period studied (generally following a similar pattern in the two corpora), possibly in response to important social events (namely the “December 2008” revolt, the onset of the Greek crisis and the rise of SYRIZA to power).

The representation of anarchists is largely consistent in both corpora; they are constantly realized as agents, mainly involved in violent and illegal acts, as well as groups of people rather than individuals. Most significantly, anarchists in both corpora are modified by evaluative adjectives such as *self-styled*, *self-proclaimed*, *self-described*, etc., which function as comment adjectives, in Fragaki’s (2010) terms. This suggests that, despite their self-identification as anarchists, their identity is constantly put into question by the newspaper. Moreover, although the portrayal of anarchists is negative, focusing on violence, the adjective is also found with a positive evaluation in the Greek corpus when ascribed to artists, writers etc., who do not use the term for themselves. Thus, the newspaper perceives the meaning of the words in a way that excludes those who explicitly identify themselves as anarchist(s).

The study has several implications for corpus-assisted discourse analysis (cf. Partington et al. 2013), regarding the analysis of conflicting discourses in the Greek context, the relation between evaluation and ideology and the significance of comparable corpora.

Synergies between Corpus Linguistics and Ludolinguistics: Using Corpora to Explore the Representation of Gender in Videogames.

Frazer Heritage
Lancaster University

Eligible for nomination to a student researcher award

This paper argues for the application of corpus techniques in the examination of the representation of gender videogames. Within the field of ludolinguistics (the study of language and videogames), a number of studies have utilised corpus linguistics to examine the representation of gender and sexuality. However, these have been conducted in the past, this is also typically confined to the study of language within videogame paratext (for examines Ensslin, 2012; Potts, 2015; though, see Author [2019; forthcoming] for an exception).

One of the few exceptions to this (Author [2019]) has used corpora to examine the representation of gender in a single videogame. This paper diverges from this line of inquiry and instead examines how gender is represented across ten of the best-selling fantasy videogames published between 2012 and 2016.

I demonstrate that gendered pronouns were within the top 50 statistically key words, and therefore analyse the pronouns he and she in more detail. The results show that male and female social actors are represented in different ways, with male characters collocating with words which denote physical violence. Concordance line analysis of these collocates suggest that it is these male characters who enact this violence, thus linking to the idea of hegemonic masculinity (see Connell, 2005). By contrast, the only time she collocates with a word denoting physical violence is hit. The concordance line analysis of this shows that it is the female social actor who is patient of this this verb, rather than the agent. This therefore leads to an analysis of transitive clauses containing he and she, in which I argue that female social actors are more likely to be positioned as having actions done upon them.

Army Deserters or Fat Kids?: A Corpus-assisted Critical Metaphor Analysis of Obesity News in Chinese Official Media

Xiang Huang

Universitat Pompeu Fabra, Barcelona, Spain

xiang.huang01@estudiant.upf.edu

Eligible for nomination to a student researcher award

Obesity has become a pressing public health concern in China. The media are a central site for socio-political construction of obesity to negotiate body-related values, ideologies and beliefs. Despite being a frequent topic in Chinese media, little is known about how obesity has been discussed in Chinese news. Even less is known about how obesity is conceptualised by Chinese authorities particularly via metaphor in mobilising national efforts in fighting obesity. Taking critical metaphor analysis (Charteris-Black, 2004) as the theoretical framework, this study analyses 109 obesity-related editorials from the past decade (2010-2019), comprising approximately 136,000 Chinese characters, in three major official newspapers from China's Mainland, i.e., People's Daily, Xinhua News Agency, and Guangming Daily. Two preferred source domains of metaphor can be identified — War and Family. The War metaphor, typically shown in 身体是革命的本钱 (The body is the capital for revolution), envisages obesity as an alien enemy to mobilize a public fight against it. The War metaphor positions the Chinese authority as the supreme commander among the public and implicitly labels people with obesity as “army deserters”; on the other hand, the Family metaphor, frequently embedded in 少年强则国强 (strong youths build a strong country), constructs the authority as a giant parent who aims to nurture strong kids, while being concerned with her “fat kids”, the obese Chinese people who need more health education. Both metaphors achieve the goal of building a bigger notion of “We” around the obesity issue in China. This study underscores the importance of critically interrogating the representation of obesity in Chinese media, especially the manipulative force of metaphor, which has discursively extended obesity from a biomedical issue to a political one in post-reform China.

“You can tell he’s a baddy just by looking at him!” Appearance and identity in children’s literature

Sally Hunt

University of Reading

Children's literature has been investigated for what it can tell us about "popular culture, children's worlds, stratification, and socialization" (McCabe et al 2011:198). Part of an ongoing project looking at fiction for children in English, this paper shows what corpus methods can reveal about discourses of identity, specifically appearance, embedded in this popular genre. The appearance of a person is one of the first aspects of a person that we perceive, in life and in literature, so it is not surprising that immediately visible features are amongst the most frequent descriptors of characters in children's literature in English. Indeed, many characters never move beyond the categorical membership labels they are given as their initial identifiers e.g. the girl with the red hair or the old wizard. In this paper I present research into appearance in a corpus of commercially successful series in English children's literature, with an emphasis on gender. Using collocation as the main tool, I show which descriptors are most closely linked to gendered terms like girl, boy, woman and man. Age and hair colour are frequent across all series as ways of identifying and describing characters, and there are significant patterns in terms of the meanings attached to certain appearance clusters. For example, in the Harry Potter series, families are unified by sharing physical features such as complexion and hair colour e.g. the Weasleys or the Malfoy males. Furthermore, whether a character is identified as good or bad is often indicated via their appearance. Descriptions of appearance also suggest the depiction of race and the extent to which Caucasian Anglo-American people are foregrounded and normalised, and less powerful demographic groups are marginalised.

“Russians are very sweet and nice”: a corpus-assisted multimodal discourse analysis of the representation of people in online travel reviews about Moscow

Ekaterina Ignatova
Lancaster University

e.ignatova@lancaster.ac.uk

@IgnatovaKat

<https://www.lancaster.ac.uk/linguistics/about/people/Ekaterina-Ignatova>

Eligible for nomination to a student researcher award

The paper explores how guests and hosts are represented in online travel reviews about Moscow.

Tourism provides an opportunity to get acquainted with the sociocultural background of other nations and potentially to improve international relations. Moscow, the capital of Russia, is sometimes viewed as an unfriendly or unsafe destination and the Russian Government aims to increase the popularity of the city. However, there are concerns that modern tourism discourse contributes to the maintenance of asymmetrical guest-host power relations. Guests are often accused of consumerism while hosts are frequently backgrounded or represented as servants or cultural markers. Such representation can lead to client-servant attitude and even cause discrimination against hosts.

While online travel reviews are considered an important genre of tourism discourse, most studies analyse the representation of people in promotional or media discourse. Considering that multimodality is an integral feature of tourism discourse and that the analysis of discourse patterns allows exploring the meanings widely shared by the society, the study utilizes a corpus-assisted multimodal approach by analysing the representation of people in headlines, texts, images and image captions of a corpus of online travel reviews.

The analysis corroborates previous conclusions that guests tend to be represented as consumers enjoying themselves while hosts are perceived as friendly servants. However, the study provides evidence that tourists can background not only hosts but also themselves or other tourists. Moreover, the results reveal that in contrast to promotional and media discourse, guests can also portray themselves as active, solving problems while sometimes representing guests as rude or unwelcoming. The results also show that the representation of people can vary across the modes of the same document.

The study concludes that user-generated tourism discourse reveals a complex picture and can express resistance to the dominant institutional imagery.

Different Roles of Role as a Signalling Noun in the Introduction Sections of Experimental Medical Research Articles: A Concordance-Based Study Using a Move-Specific Corpus

Tatsuya Ishii & Takeshi Kawamoto
Kobe City College of Technology / Hiroshima University

t-ishii@g.kobe-kosen.ac.jp

Signalling Nouns (SNs) are abstract nouns that are important in academic discourse (Flowerdew & Forest 2014). However, few studies have investigated the relationship between the behaviour of SNs and the functions of moves and steps. Thus, to describe the phraseologies of SNs in moves, we created a move-specific corpus, a corpus according to move analysis (1,612,735 words in total). In the introduction sections of experimental medical research articles, three moves were identified, following Swales (1990): (IM1) establishing a territory (51,014 words), (IM2) establishing a niche (74,714 words), and (IM3) occupying the niche (34,954 words). Using Antconc (Version 3.5.8) and comparing each move with the whole corpus as a reference corpus, we identified the keywords of each move. Although the SN role was found in the three moves, the concordance-based analysis revealed that it seemed to exhibit quite different behaviours, which were strongly associated with the function of steps. In (IM1), the collocation role in formed part of the phraseology play a (central/important) role in, indicating the step for introducing the topic of articles. In (IM2), the collocation role of was included in the phraseology the (functional) role of [material] remains to be address or the (functional) role of [material] remains unknown, which functioned as a step for indicating a gap. In (IM3), the collocation role of led to the phraseology in this study we sought to investigate the role of, as the step for outlining purposes. Moreover, the collocation role for was used for the phraseology here, we describe a role for as the step for announcing the present research. Collectively, this phraseological study demonstrates the possibility that move-specific corpora will reveal how the behaviour of SNs is strongly related to the function of steps. In our presentation, more phraseologies of the SN role will be shown

National Identity in the Press: Britain vs. Europe Dichotomy

Anna Islentyeva

Free University of Berlin

anna.islentyeva@gmail.com

[https://www.geisteswissenschaften.fu-](https://www.geisteswissenschaften.fu-berlin.de/en/we06/institut/mitarbeiter/Islentyeva/index.html)

[berlin.de/en/we06/institut/mitarbeiter/Islentyeva/index.html](https://www.geisteswissenschaften.fu-berlin.de/en/we06/institut/mitarbeiter/Islentyeva/index.html)

There is perhaps no other issue that currently enjoys more British media attention than the UK's relationship with the EU and the ongoing Brexit negotiations. Britain's decision to leave the EU not only marks a crucial point in its history, it also evidences the rise of right-wing populism in both Britain and Europe. It is particularly revealing to trace the media representation of Europe and Britain in the context of the EU membership referendum.

This corpus-based study focuses on the metaphorical patterns and discursive strategies employed by the British liberal and conservative press in 2016–2018 to construct concepts of Europe and Britain. The data represents a specialised corpus containing 500 editorials, opinion pieces and news reports from six national British newspapers. The Guardian, The Observer and The Mirror represent left-wing ideology, while The Telegraph, The Mail and The Sun represent conservative ideology.

The analysis identifies the linguistic mechanisms that ensure that the opposing ideologies promoted by these newspapers remain adaptable in the changing socio-political environment of Britain. In terms of methodology, the research applies a mixed-method approach that combines CDA and corpus linguistics (Stefanowitsch 2006) with a focus on the metaphorical patterns (Charteris-Black 2004) employed in the construction of the idea of Europe in the context of Brexit (Charteris-Black in press).

The analysis reveals a range of metaphors applied in reference to Europe, with EUROPE AS A CONTAINER, EUROPE AS A UNION and EUROPE AS A HUMAN occurring most frequently and DEATH OF EUROPE, EUROPE AS SUICIDAL and RELATIONSHIP WITH EUROPE AS A (BROKEN) MARRIAGE as the most creative, as well as the most negative. The analysis also shows the importance of the question of national identity (Bayley & Williams 2012) and the prevalence of discourse surrounding British exceptionalism, a concept that stands in opposition to the notion of a united Europe.

The Polish Catholic Church on life of Poles. Corpus Assisted Critical Discourse Analysis

Victoria Kamasa

Adam Mickiewicz University in Poznań

The Catholic Church plays an important role in Polish public life, strengthened only by the political change, which brought to power the right-wing conservative Law and Justice party. This role could be seen as a reflection of Poland's religious landscape (92% declared Catholics (CBOS, 2018) and 38% attending a mass every Sunday (ISKK, 2020)). Yet, it is accepted only in restricted domains of life: while 61% of people accept the Church taking the stance in matters of morality, 55% disapprove of the Church voicing opinions about the Parliament's bills (CBOS, 2013) and as many as 83% oppose the priests telling them how to vote (CBOS, 2015).

This study aims to analyse which of these domains are represented in the official statements of the Polish Catholic Church and how the concept of life is constructed within them. To achieve this aim, the study, situated within the corpus-supported Critical Discourse Analysis paradigm (van Dijk, 1993, Baker, 2006), combines methods of collocation and concordance analysis. Collocation based analysis of quantitative tendencies in semantic prosody led to formation of restrained hypotheses concerning the discursive construction of life. These hypotheses were further elaborated through qualitative (in terms of patterns and re-occurring categories) and quantitative (in terms of relations between these categories and semantic preference) analysis of concordances.

The results showed the presence of both religious and secular dimensions of life, although in different manners. The former has a rather individual character, often related to particular religious practices, whereas the latter concerns public life, especially the critique and demands aimed at politicians. Overall, the results support the view of the Polish Catholic Church as an administrator of values (Graff, 2010) and point to the great self-involvement of the Church. Finally, the role of these discourses in legitimizing and reinforcing the power of Catholic Church in Poland is shown.

Personal and impersonal authorial references in research articles: Variation across four disciplines

Mohsen Khedri
Sohar University, Oman

mkhedri@su.edu.om

This study took an interdisciplinary approach to investigate the use of personal (exclusive first-person plural pronouns) and impersonal (abstract rhetors, periphrastic passives, and it-clauses) authorial references in a corpus of 160 empirical research articles in Applied Linguistics, Psychology, Environmental Engineering, and Chemistry. To be specific, the study aimed to i) compare research articles in the four selected disciplines to determine to what extent authors make their presence in the text explicit using personal authorial references; ii) scrutinize the utilization of impersonal authorial references to see whether they are prone to variation across the disciplines; and iii) investigate the context within which academics belonging to different disciplinary communities choose subjective over objective constructions and vice versa, in order to ascertain whether any cross-disciplinary differences in the occurrence of personal and impersonal authorial references is related to the adoption of different strategies of interpersonality. The results indicated marked discrepancies across the disciplines in both the frequency and use of personal and impersonal authorial references, especially when the disciplines were compared for contrasting effects. While in the corpus at hand the applied linguistics and psychology writers favoured the use of exclusive first-person plural pronouns as an interpersonal strategy to construct their authorial persona, the environmental engineering and chemistry writers preferred a more detached interpersonal style by opting predominantly for periphrastic passives. Also, the results showed differences in the incidence of use of personal and impersonal authorial references across discourse functions, which could be attributed to the adoption of different interpersonal strategies, objectivity or subjectivity, within the four academic discourse communities. The present results are expected to extend our understanding of disciplinary variations towards the use of personal and impersonal authorial references in tandem with discourse functions in research articles in the four selected disciplines, particularly in the relatively unexplored disciplines of chemistry and environmental engineering.

Towards a Dictionary of Polish Political Discourse 2015-2019

Sonia Kropiowska

European Master in Lexicography

sonia.kropiowska@gmail.com

Although Corpus Linguistics has been used extensively in combination with a variety of academic disciplines, its methods have been applied to Discourse Lexicography, a subdiscipline of Lexicography developed by Heidrun Kämper, only to a limited degree, i.e. mainly for easy access and viewing of data. Drawing on Corpus Linguistics methods, specifically collocational analysis between keywords (Scott and Tribble 2006), I would like to introduce new considerations for Discourse Lexicography in terms of its methodology, namely in establishing the dictionary's main semantic structure (Kämper 2013), in a way consistent with the discipline's theoretical assumptions.

To demonstrate my approach I analyse the topic of LGBT in the context of a controversial period in modern Polish history, namely part of the eighth term of the more powerful upper chamber of the Polish Parliament (Sejm) spanning between July 1st 2018 and June 30th 2019. In recent years Poland has been marked by numerous political changes, which affected the state of media landscape as reported by Freedom House and Reporters Without Borders. Through the means of lexicographic representation I show that Polish media outlets, being far from the ideal of objectivity, manifest radically different ways of framing the phenomenon of LGBT.

'Social workers dismissed concerns': A corpus-assisted discourse study of the portrayal of a profession in UK newspapers

Maria Leedham
Open University

maria.leedham@open.ac.uk

@marialeedham

@ouwisp

<http://www.open.ac.uk/people/mel245>

<http://www.writinginsocialwork.com/>

Previous research indicates that social workers are portrayed negatively in the UK press, particularly in child protection cases (e.g. Reid & Misener, 2001). But what is the nature of this negativity? And are social workers also mentioned in more positive contexts? To explore this, I compiled 'SWP2019', a corpus of 3 months of articles from Nexis-UK seeded around social worker(s) from the UK national press. Almost 1000 occurrences of the search term were located within 736 articles (5.1m words). Each instance was categorised within WordSmith Tools (Scott, 2019) as 'positive' (5% of instances), 'negative' (25%), or 'neutral' (71%) (91% intra-rater reliability). Occurrences of social worker(s) are predominantly 'neutral', simply referring to professionals doing their job. In a significant minority of occurrences, however, the story concerns an individual who happens to be a social worker and who has behaved badly in their personal life; this suggests that, in contrast to previous studies, the press tend to portray social workers as 'good' people held to a higher moral code.

Further categorising of concordance lines through iterative rereading and sorting, employing substantial co-text, indicates that negative mentions are dominated by children's social work and by the subcategory 'failure to notice/act' (42% within 'negative') in contrast to 'over-zealous' behaviour (17%). The observation that social workers are criticized more for inaction is supported by the keyness of the semantic category failure, failing, losing in SWP2019 when semantically tagged (Rayson, 2009) and compared to the BNC written sampler.

The presentation also draws on findings from 81 interviews from the XXX project (Author). These point to social workers' fear of vilification in the national press and the potential emphasis on 'defensive' recording (Garcia-Maza, Lillis, & Rai 2010), indicating the importance of journalists and social workers developing a shared understanding of social workers' role in society.

Representation of China's Reform in English News Media of China and the West: A Corpus-assisted Discourse Studies

Longxing Li & Vincent X. Wang
University of Macau

Eligible for nomination to a student researcher award

Reforms have played an important role in the development of China in its long history. The year 2018 witnessed the 40th anniversary of China's Reform and Opening-up, which has not only profoundly reshaped China but also impacted the world. As a powerful driving force for social progress, reform is an everlasting and heated research topic in the broad field of social sciences, especially in economics, politics, administration, law, and history. By far, linguistic studies focusing on the discourse of reform are rare, especially from the Chinese perspective. Therefore, this study targets at the discourse of China's reform and intends to investigate how the reform is represented in the English media of China and of the west. Specifically, this study will address the following research questions:

How is the reform represented linguistically and conceptualized metaphorically in the English news media discourse in China and the west?

How are the relevant elements of the reform, such as participants, events, countries and regions represented or conceptualized in the English news media discourse in China and the west?

What are the underlying factors leading to the similar and different representation of the reform in English news media in China and the west?

To address the above research questions, a corpus of English media news reports of China's reform in 2018 will be compiled using the data from major news agencies of China (Xinhua News Agency) and the west (AFP, AP, BBC, DJ, FT, PRN, RT, etc.). The data will be analysed with the help of Sketch Engine, an online corpus tool with functions like word sketch, wordlist, keywords, n-grams. The analysis is expected to reveal the similarities and differences between China's self-representation and the western media's representation of reform in China, and to find out the possible underlying reasons.

Testing Agentless Passives' Ability to 'Mystify' Social Actors in the Minds of Readers

Will Lingle

University of Aizu, Japan

Passive verbs which lack an agent, or 'doer' of an action, have long been considered potentially ideologically significant in Critical Discourse Analysis due to their ability to omit or 'background' social actors in ways which reduce readers' comprehension of their role in events—a process known as 'mystification'. This 'mystification' of social actors is central to the potential ideological significance of agentless passives in CDA. Yet information absent from texts may be recovered by readers via inference generation. Drawing on research in cognitive linguistics and psycholinguistics, the Idealized Reader (IR) framework (O'Halloran 2003) identifies types of inferences which are generated easily by readers when reading for gist, and others which are unlikely. In the latter case 'mystification' may be said to occur. Two corpora of US newspaper editorials on the Greek financial crisis developing in 2013-2015 were compiled and searched for passive verbs. When these passives were identified and analyzed along with their co-text according to the IR framework, only 7% were predicted to mystify agency for casual readers. In order to test this method of inference prediction, a reader response study was carried out in which agents named in published editorial texts were deleted via the addition of agentless passives and nominalization. Although the altered texts reduced the likelihood of correct inference generation, deleted agents were correctly inferred by readers in an average of 60%-70% of cases. These results support the predictions of the IR framework and suggest that agentless passives in news texts may not exert a 'mystifying' effect on casual readers in most cases.

Transgressive Pleasures: Identity and desire in the discursive construction of gender-nonconforming characters in porn

Alon Lishinsky & Kat Gupta

Oxford Brookes University / University of Roehampton

alon@lishinsky.net

@alischinsky

kat@mixosaurus.co.uk

@mixosaurus

While popular claims about the negative effects of pornography (such as inducing sexual violence and emotional detachment) have been entirely discredited, empirical explorations of users' engagement with porn have shown its potential influence on beliefs, attitudes and behaviours (Billard, 2019): audiences use pornographic materials to inform themselves about sexual anatomy, attraction and pleasure, especially when other sources of information are lacking or perceived as judgemental (Albury, 2014:174). In the social construction of sexual desires and identities, porn provides one of the key spaces for articulating peripheral forms that fit poorly within the mainstream (Escoffier, 2011: 279; Rubin, 1992).

Porn has been especially important in constructing public images of trans and other gender nonconforming (GNC) identities. While trans representations in mainstream media are scarce and predominantly derisive (Richardson 2010:131), GNC characters and bodies have been conspicuous in pornography both visual and textual (Escoffier 2011:271-2); for many audiences it represents their first encounter with GNC, and certainly the first that celebrates it. At the same time, porn representations can be problematically loaded with normative assumptions (such as that all GNC people experience dysphoria or aim for surgical/hormonal transition; Jones, 2019; Lovelock, 2016). Steinbock (2017:29–30) argues that trans pornographic films reproduce misogynistic tropes of objectification, hypersexualisation and femininity as submissiveness. However, these claims have not been explored in written pornography.

In this paper, we build on prior work on the linguistic construction of desirability in porn (Baker, 2005; Bolton, 1995; Koller, 2015; Marko, 2008; Morrish & Sauntson, 2007; Motschenbacher, 2010) by exploring patterns of labelling and description of GNC characters. Using a 10-million-word corpus of erotic stories, we identify the range of terms denoting GNC identities and performances and explore their discourse prosodies and textual functions, seeking to identify distinctive ways in which GNC can be construed and appraised. Results suggest that characteristic lexical signals reflect different subcultural manifestations of gender variance. This diversity should be taken into account in order to avoid eliding intersections between gender and other cultural categories.

Time and place in news discourse: a linguistic approach

Azad Mammadov

Azerbaijan University of Languages

azad19622003@yahoo.com

The paper aims to explore the role of time and place in news discourse. It focuses on the linguistic representations of time and place from the perspective of their multiple functions in this genre. The analysis of the recent literature in the field (Neiger and Tenenboim-Weinblatt 2016, Tenenboim-Weinblatt and Neiger 2018, Sanders and Krieken 2019, etc.) suggests that the study of time in news is more common than the study of place. The main focus in these studies is on time perception in news as well as on the role of time in the construction of news narratives. On the other hand, place is not considered as the central issue playing mostly supportive role in time perception or in the construction of narrative in news discourse. Meanwhile, this paper prioritizes neither time, nor place equally focusing on the study of time and place in news discourse especially on their correlation from the linguistic perspective. This correlation builds temporal and spatial harmony, which is crucial in the process of perception by the participants in the news discourse.

Thus the central issue in the paper is the study of the functional-pragmatic and cognitive aspects of the tense forms of the verb, adverbs of time and place, linguistic metaphors and metonyms in order to reveal their role in the representation of time and place in the news discourse. For this purpose, various news texts (political, economic, social, daily life, cultural, sport, etc.) produced by the leading news agencies (Reuters, AP. etc.) newspapers (Financial Times, USA Today, TV-s (CNN, BBC, etc.) and internet sites (Buzz Feed News, etc.) have been collected for analysis based on both qualitative and quantitative methods.

The statistical analysis of the data suggests that the headlines contain more elements from figurative language especially metonymy whereas metaphor is common in the main body of the news text. The reason why the producers of news tend to use linguistic metaphor and metonymy so frequently is that they have strong pragmatic and cognitive effect on the audience. They help the audience to understand the news text in a more specific and sometimes culturally marked (White House, Scotland Yard, Pentagon, Christmas, etc.) way as all parameters of time and place matter in this genre.

A lone wolf or a red herring? Understanding construals of the so-called lone-wolf terrorist

Daniel Malone
Edge Hill University

Eligible for nomination to a student researcher award

Although lone-wolf terrorism is not a modern concept, its re-emergence in modern media discourses stems from American right-wing terrorist circles of the 1990s (Hamm & Spaaij 2017). Over the past decade, the metaphoric lone wolf has been used increasingly in connection with terrorism in the British press and yet it is an area that has received little academic attention. Its increased usage suggests both a change in meaning and possible conceptual realignment.

Employing a corpus-assisted discourse approach, this study identifies the representations of the actors and events which serve as a 'touch-stone' for newspaper discourses of lone-wolf terrorism; that is, the instances which have become discursive artefacts and represent archetypes of the phenomenon. The culture of a discourse community is composed of shared systems of meaning and understanding communicated via shared language (Holland & Cole 1995). Therefore, how events and connected actors are construed by the media and their underpinning ideologies contribute to schema building which can be both directive and evocative for the discourse community.

The corpus under analysis comprises articles collected via ProQuest Newsstream from seven UK national newspapers spanning a ten-year period between 2010 and 2019. To identify salient actors and events which have resonated intertextually, periods of high article frequency are initially focussed upon as they are likely to represent intense phases of media focus on 'big news' terrorist events. These actors and events identified are then qualitatively and quantitatively evaluated to elucidate the recurring themes and assigned values foregrounded in their representations.

“The truth is, as you know, I’m definitely not lying, believe me.” The linguistics of Donald Trump’s epistemic management strategies. A corpus-based study

Georg Marko

Karl-Franzens-University Graz

There is a strong epistemic dimension to Donald Trump’s presidency. This dimension is mostly negative, i.e. it is primarily concerned with the absence of truth and credibility, as becomes obvious from some expressions characteristic of the discourses by and about the President and his supporters, e.g. “fake”, “dishonest”, “hoax”, “con”, “lie”. Given the prominence of the absence of truth, it is interesting to look at whether President Trump feels the need to promote the truthfulness of his statements and, by implication, the credibility of him as a political persona and whether this can be studied by looking at his language use.

The proposed paper starts from the concept of epistemic moves. These are elements that fulfil a function in the management of truth and credibility of an utterance without substantially contributing to its propositional meaning or illocutionary force. Epistemic moves might be realized by added (“believe me”), subordinated (“as you know”), or superordinated (“the truth is”) clauses, by modal adverbs (“definitely”), and by many other means. Realizations might be ad hoc and creative (e.g. the above quote as a whole if added to a statement) or formulaic (e.g. discourse markers such as “believe me” or “as you know”). And there are differences in the epistemic function with respect to truthfulness, commitment, evidentiality (availability and status of evidence), and direction (enhancing or relativizing).

I will look at how Donald Trump uses these different categories of epistemic moves, taking into account quantitative (how often does a category feature in his language) and qualitative (which meanings are presented with epistemic moves) aspects.

Assuming that personal and spontaneous language is more interesting for my research purpose, I will use two self-compiled and self-annotated corpora of Trump’s non-scripted speech (news conferences, interviews, remarks to the press, etc.; 800,000 words) and his tweets (600,000 words) as my data.

A methodological proposal to realize a Systemic Functional Linguistics exam through Corpus Linguistics: comparing the textual strategies of the political discourse in English and Spanish

Virginia Mattioli

Pontificia Universidad Católica de Valparaíso

virginia.mattioli@pucv.cl

<https://pucv.academia.edu/VirginiaMattioli>

Although Corpus Linguistics (CL) is extremely effective and exploited in both Linguistics and Translation Studies, its synergy with the Systemic Functional Linguistics approach is still rare (among other previous attempts, see Hansen-Schirra, Neumann, & Steiner, 2007). As for the tools available for the studies based on both areas, they are very limited and the ones freely available are still implemented only for English.

To respond to such challenge, in this paper CL is used to compare the textual theme strategies used in English and Spanish political discourse, in order to confront the impact that they have on the audience. The research pursues two main objectives: (a) designing a methodology with the currently available tools that permits to analyse electronically the thematic structure of a Spanish corpus; (b) obtaining comparable data about the textual strategies used in the two languages.

With this aim, two comparable corpora are compiled including a selection of Obama's weekly addresses, in English and Spanish, respectively. The exam consists of four steps. Firstly, the English corpus is automatically analysed with the UAM Corpus Tool (O'Connell, 2008). Secondly, an original methodology originated from the combination of Treetagger (Schmid, 1995) and AntConc (Anthony, 2018) is designed to analyse the Spanish textual structure through a POS tagging and a context specific search of the connectors in the concordance list to identify the textual themes. Thirdly, the methodology is applied to the Spanish corpus and, finally, the results are compared.

The outcomes corroborate the effectiveness of the proposed methodology. In fact, it offers comparable results, particularly regarding the most frequent textual themes of the two corpora, providing information about the number, type, function and context of use of the textual themes identified that can be fruitfully used to draw conclusions about the impact of the political discourse in both languages.

Revisiting key-key-words: proposing a method for identifying unique keywords in a collection of corpora

Mark McGlashan & Alexandra Krendel
Birmingham City University / Lancaster University

Mark.McGlashan@bcu.ac.uk
@Mark_McGlashan
www.MarkMcGlashan.org

Eligible for nomination to a student researcher award

This paper proposes extending Scott's (1997) key-key-word method to consider keywords unique to a (sub)corpus; something we call complement keywords. Whereas key-key-words – keywords found to be “key in a large number of texts of a given type” (Scott, 1997, pp. 236-7) – consider shared patterns of distinctive word usage across multiple (sub)corpora when compared against the same reference corpus, complement keywords focus on keywords that are exclusive/idiosyncratic to a particular (sub)corpus when compared against the same reference corpus. The proposal of complement keywords draws on set theory by understanding key-key-words as the product of intersecting multiple keyword list sets, whereas complement keywords are the product of set difference (keywords found in only one keyword list set and in no other sets). The suggestion here is that, by drawing on notions from set theory when comparing keyword lists, both key-key-words and complement keywords can be identified (see Figure 1). The proposed method is explored through a case study of ~11m tokens used across the comment sections of 1,000 of the top most upvoted threads of five different forums (subreddits) on the website <https://www.reddit.com> related to the ‘manosphere’ – a loose online network of sites and forums dedicated to discussion (often anti-feminist and misogynistic in character) of issues relating to men and masculinity (see Table 1). Keyword lists for each subcorpus' comments section (see Table 1) are compared to find their intersecting key-key-words (Figure 1, set 1) as well as complement keywords (Figure 1, sets 2-6). Findings suggest that the method identifies both content-generalisable (Egbert & Biber, 2019) key-key-words – including a salient focus on social actors in Reddit manosphere communities, especially concerning women (girls, female, bitch, slut, chicks) and relational roles (girlfriends, boyfriends) – and content-distinctive (ibid.) complement keywords for each subreddit which reveal community-specific language usage.

Do denominations of people on the move have the same meaning in Belgian French- and Dutch-speaking newspapers?

Valériane Mistiaen

DESIRE-Vrije Universiteit Brussel

vmistiae@ulb.ac.be

<https://www.researchcenterdesire.eu/valeacuteriane-mistiaen.html>

Eligible for nomination to a student researcher award

According to Siblöt (2001) naming social phenomena depends on the point of view taken by the speaker and is always a choice. Moreover, language can be conceptualized as a form of “social practice” (Fairclough & Wodak, 1997). As a consequence, relations determine discourses but are also reflected in discourses (Fairclough, 2001).

This research focuses on media discourse because the media play an important role in mental representations as it constitutes an important knowledge among others about many issues, contributing to the “common sense” understanding of the world (d’Haenens & Mattelart, 2011: 237). The so-called “refugee crisis” has received a huge media attention in Belgium but contrary to other European country, there is no Belgian coverage of the situation. Indeed, the Belgian media landscape is divided between a Dutch-, French- and a much smaller German-speaking community, which harbour different journalistic traditions (Jacobs & Tobback, 2013: 408). For these reasons, Belgium seems an excellent case study to tackle the research question, namely, how media denominations participate in the construction of a social issue such as immigration? To explore it, an exhaustive corpus composed by press articles was collected between 2015 and 2017, in both French- and Dutch-speaking communities of Belgium (8557 articles; 4,136,075 words). The analysis combines (Critical) Discourse Analysis ((C)DA) and Corpus linguistics (CL), as they complement one another. If DA provides the conceptual and theoretical frameworks to analyse discourses, CL is used to “provide a ‘map’ of the corpus, pinpointing areas of interest for a subsequent close analysis” (Baker et al., 2008: 284-285).

At first glance, the repertoire of common nouns in each sub-corpus seems similar but differences lie in the frequency of denominations used to qualify people on the move but also in the collocations that construct their meaning. One important finding lies in the difference in frequencies of the word migrant, as well as the usages associated with it.

'YOT Talk': Analysing discourse in youth justice assessment interviews.

Ralph Morton
Loughborough University

ralph.morton@bcu.ac.uk
[@ralphmortonlang](https://www.bcu.ac.uk/english/staff/ralph-morton)
<https://www.bcu.ac.uk/english/staff/ralph-morton>

In this paper I will present findings from the Leverhulme Trust-funded project 'Exploring the language barrier to engagement in youth justice assessment interview practice' which ran from March 2017 to February 2019 at Loughborough University and Swansea University. The project set out to examine the role of communication in the assessment interviews that young people undergo when they come into contact with the Youth Justice System.

Previous studies had discussed communication challenges in this context in terms of learning deficits and attitudinal problem on the part of the young people. We wanted to look more broadly at the range of contextual and linguistic factors that enable and hinder communication between young people and professionals. We did this in two ways: firstly by using questionnaires and focus groups to ask young people and professionals about their experiences of communication in assessment interviews; secondly by observing, audio recording, and transcribing assessment interviews, in order to carry out linguistic analyses of the data. The results of these two strands of analysis were then brought together to inform a 'toolkit' for effective communication.

This presentation will report primarily on the second strand of the project during which a combination of Conversation Analysis and Corpus Linguistics was used to analyse the language of assessment interviews. By drawing on these compatible and complimentary methods, I was able to provide a detailed account of how speakers co-construct the exchanges, and investigate quantitative patterns across the dataset (examining frequent n-grams in young person and practitioner speech, and comparing these sub-corpora through keyword analyses.)

This research demonstrates the value of applying discourse analysis to data from challenging communicative contexts, and how the insights provided by this kind of analysis can help support best practice in a professional setting.

Pragmatic annotation for digital discourse analysis – value, quality criteria, further development

Marcus Müller & Michael Bender
TU Darmstadt, Germany

marcus.mueller@tu-darmstadt.de

@marcus_dislab

@MimoBender

<https://www.linglit.tu-darmstadt.de/institutlinglit/mitarbeitende/marcusmueller/index.en.jsp>

Corpus approaches to discourse analysis rely mostly on measuring surface phenomena such as patterns of lexicogrammar or keywords. Quantitative data are then interpreted and contextualized: “What we do when we analyse discourse using corpora ‘is a qualitative analysis of quantifiable patterns’” (Marchi & Taylor 2018: 6). While those measurements are often performed on data, which are (automatically) annotated with part-of-speech or semantic categories, pragmatic annotation (Aijmer and Rühlemann 2015) plays a minor role in corpus approaches to discourse. In this paper, we will argue that pragmatic annotation is highly valuable not only for the study of spoken interaction or speech acts (Weisser 2018) but also for (critical) discourse studies: Annotations operate as interfaces between segments of text and sense. They externalise and clarify interpretation, thus enabling collaborative hermeneutics. However, manual annotation is still best suited for detecting implicit phenomena in discourse. This is time consuming and hardly applicable to large corpora, which makes it desirable to achieve automation of such procedures. We will discuss quality criteria of pragmatic annotation for discourse studies (segmentation standards, inter-annotator agreement) and argue possibilities and limitations of automation by machine learning techniques (recommender systems and deep learning frameworks). An annotation study on heuristic textual practices in academic discourse will serve as an example. It is based on the introductions of 65 German dissertations out of 13 faculties. Heuristic textual practices are linguistic implementations of decision routines in research processes. We studied the distribution and combinatorics of such practices across disciplines, applying collaborative pragmatic annotation.

“We used to go there but now we fear it”. A corpus-assisted study of the development of market as a concept over time

Claes Ohlsson
Linnæus University

claes.ohlsson@lnu.se

This research paper aims for the conference theme of comparative studies of discourses over different periods of time. The theme of the presentation is how the concept of market has developed over time. Today, the concept of market has immersed a vast amount of everyday life in most of world, e.g. as in “market speech” or by the frameworks of new public management and financialization (Martin 2002; Leary 2018). The “marketization” of many types of societal activities further indicate that “market” is used in new and expanded ways (see Fairclough 1993). A conclusion is that the origin of market as a concrete place in time now also encompasses market as an abstract entity, capable of actions and even emotions.

In this presentation, the development of market as a concept or “Begriff” is addressed in a study of the usage of market-related words and of market as a keyword with importance for several discourses over time (see Schröter 2008). The methodological and theoretical framework combines discourse analysis and corpus methodology with a large sample of Swedish text material. The project aims to cover a time period from the early 17th century to present day. A major goal is to map and discuss the gradual development of how and when the meaning and usage of “market” has changed over time in line with the aforementioned trajectory from concrete and passive to abstract and active. The description of the timeline of the concept of market is further discussed in in-depth cases of “turning points” in time (Hollander et al 2008) where changes in usage and meaning become visible in certain discourses or areas of contemporary society. Examples of such discourses are politics, education and the general public debate in print media.

Same-sex marriage, gay marriage, or equal marriage? Exclusion or inclusion by premodification in a corpus of 21st century UK press reports

Laura L. Paterson & Mark McGlashan
The Open University / Birmingham City University

Mark.McGlashan@bcu.ac.uk
@Mark_McGlashan
www.MarkMcGlashan.org

Changes to marriage legislation in countries across the globe have received much academic and public attention (Bachmann 2011, Kania forthcoming, Ku forthcoming, etc.). However, one under-researched area, highlighted by Paterson & Coffey-Glover (2018) in their analysis of UK newspapers, is that the choice of labels used to refer to the expansion of marriage legislation (particularly same-sex/gay marriage) means some groups (i.e. people who are bisexual, transgender, etc.) are linguistically excluded from and/or backgrounded in wider debates. The exclusionary nature of such labels was also implied (albeit in a different sense) by MacCulloch (2019) in a British Academy blog post on the fifth anniversary of UK same-sex marriage legislation. He proposes that '[L]inguistic scholars will be able to map for us the shifts in language that have resulted from this momentous change: when 'same-sex marriage' becomes 'equal marriage', and finally, just marriage'.

This study takes up MacCulloch's call. We analyse the premodification of marriage using a 179 million-word corpus of UK newspaper articles (2000-2018) which contained the terms marriage(s) and/or civil partnership(s). Using corpus-based discourse analysis we ask:

What are the most frequent premodifiers of marriage (same-sex, gay, equal, etc.) and do trends in usage change over time?

What discourses are associated with each premodifier + marriage?

Is the exclusionary nature of any premodifier + marriage directly challenged?

Our findings will contribute to the increasing body of knowledge on the systematic erasure of particular social groups from wider public discourse.

The representation of immigrants in Spanish judicial decisions: using data to refute hate speech.

María José Marín Pérez & Ángela Almela Sánchez-Lafuente
University of Murcia, Spain

mariajose.marin1@um.es

@MJ_MP

<https://webs.um.es/mariajose.marin1/miwiki/doku.php>

angelalm@um.es

The phenomenon of immigration and its depiction in media texts have been examined profusely within the field of corpus-based discourse analysis (Gabrielatos and Baker (2008); Baker et al. (2013); Taylor (2014); Blinder and Allen (2016)). This research seeks to provide a different perspective on the phenomenon of immigration as presented in a corpus of 600 judicial decisions issued by Spanish courts in the years 2016 and 2017. This analysis was motivated by the rise of extreme right-wing parties in Europe in the recent years, which dehumanise immigrants and portray them as a threat to the welfare state. On a first approach, the results appear to dissociate immigration and crime since, apart from general legal terms (c. 80%), a considerable percentage of the keywords obtained (c. 20%) revolves around three major topoi, namely, family, territory/access, and legal punishment, not showing evidence of any major offences or crimes amongst the top ranking lexicon. The analysis of the collocate networks of the KWs within the category legal punishment confirms our initial perception, in fact, out of a list of 21 items, only the word delito (crime) itself, as was to be expected, collocates with terms referring to typified crimes such as violencia (violence), lesiones (personal injury) or trata de personas (human trafficking). In parallel, the data were triangulated using the text-classification software UMLTextStats (García-Díaz et al., 2018). Its structure is similar to the well-known text analysis tool Language Inquiry and Word Count –LIWC (Pennebaker & Francis, 1999)–, yet it includes a linguistic basis of European Spanish and several categories that are not word-based. We were also able to add an ad-hoc category, namely crime, whose results confirm our initial observations: only 0.03% of the words out of the whole corpus fall into this category, which stands in stark contrast to other meaningful ratios, like that of family (0.32%) and job (0.58%).

Gypsies and Travellers and the discourse of dirt and cleanliness

Roberta Piazza
University of Sussex

It is well known how the media represent mobile Gypsies, Romani and Travellers (GRT) as a negative group (Piazza 2017) that continuously undermines the values and beliefs of majoritarian settled society. Generally such othering of these groups is predicated on the grounds of a logos of 'dirt' (Newell & Green 2018) that has a strong emotive traction on mainstream people. This fear and horror of dirt, filth and mess often leads to racial hatred for individuals and groups that are 'filthy' and produce 'debris' or 'waste' like GRT. The paradox, however, is that Gypsies and Travellers are obsessed with cleanliness; women spend most of their time washing and polishing their possessions and train their girls to follow this custom from an early age. This study explores the narrative of dirt that settled society has created around GRT mobile communities through a small corpus of national quality and tabloid newspapers; the results are then compared with the response of a group of female Irish Travellers speaking in interviews collected since 2012 to the present. The aim is to investigate the relation between the mainstream society's discourse of dirt and the counter-discourse of cleanliness associated with that of moral rectitude that is at the basis of the GRT's identity construction (De Fina 2003).

Hausa discourse markers in computer-mediated communication

Tristan Purvis

American University of Nigeria

Drawing on a corpus of Whatsapp chats between Hausa speakers from Nigeria, this study surveys the distribution of various discourse markers found in Hausa chat exchanges. Special focus is given to so-called pragmatic markers (Fraser, 1990)—including fillers, expletives, and other interactive lexical items—to assess the degree to which computer-mediated communication emulates such distinctive cues of interaction found in spoken conversation.

Hausa is among the more well-documented African languages with hundreds of linguistic resources of all sorts. However, there is hardly any documentation on usage of discourse markers. Perhaps the only published study that focuses in part on pragmatic markers in Hausa is hidden in a volume on spoken Arabic: “Conversation Markers in Arabic-Hausa Codeswitching” (Owens & Hassan, 2010). Owens and Hassan evaluate seven so-called “conversation markers” in exchanges between multilingual Arabic-Hausa-English speakers in Maiduguri: *to*, *ee*, and *wei* from Hausa; *ahaa* and *inzeen* from Arabic, and *okkee*. (English-derived borrowing).

For this study, a range of lexical items are tracked with each occurrence being coded for certain contextual factors: (1) sociolinguistic: speaker’s sex and regional affiliation; (2) textual: location in ‘utterance’, preceding DM, co-occurring DM(s), and language context (where code-mixing applies); and (3) apparent pragmatic/semantic function. The targeted items include expletives such as *oho* (expression of doubt) and *wallahi* (‘I swear’), interactive conversation markers as found in Owens and Hassan’s study (*ee*, *to*, *wai*), conjunctions such as *amma* (‘but’) and *domin* (‘because’), so-called tags such as *ka ga(ni)* (‘you see’) and *ka san* (‘you know’), and markers of degree and emphasis such as *kawai* (‘only’) and *fa* (‘indeed’), as well as equivalent forms of English and Arabic where code-mixing occurs. The effects of the aforementioned factors on choice of DM is evaluated, with special focus on a subset of the most frequently used conversation markers.

Firing Away: Risk Communication and Climate Change in Australian Media Discourse

Katherine Russo

Università degli Studi di Napoli "L'Orientale"

During the last decade, news media have intensified their role as a channel for the communication of environmental risk. Risks are “threats to outcomes that we value” (Fischhoff and Kadvani, 2011: 22), yet while some outcomes, such as car mortality, are defined as risks, other outcomes such as climate change are contested and their measurement often leads to debate (Russo, 2017, 2018). Debates about environmental risks are in turn mediated by news media, which call lay people to think about them either as active participants or interested observers. In order to make risk decisions, citizens seek information, which is circulated through broadsheet news media, and later arises in social media such as facebook and twitter, or face-to-face conversations. This spreading of news across ‘genre chains’, is central to the present study on the recontextualisation of studies on bushfires and climate change in news discourse since it may affect the citizens’ evaluation of risk decisions (Fairclough 2003; Blommaert 2005). News chains and networks are particularly important to this study since they contribute to the systematic transformation and recontextualisation of information from genre to genre, and possibly affect its evaluation.

The paper provides an analysis of the recontextualisation and appraisal of climate change discourse in a media genre chain regarding the 2019-2020 Australian bushfire crisis. The analysis was carried out by analysing a corpus (2014-2020), specifically compiled to represent different interrelated news media discourse genres. The data is analysed according to an approach which draws upon findings in Critical Discourse Analysis, Appraisal Linguistics, and Corpus Linguistics (Baker 2006; Martin and White 2005; Thomson and White 2008). Accordingly, the paper situates quantitative analysis and qualitative analysis within a wider analytical framework which includes extra-linguistic social variables.

Investigating professional diplomatic discourse: a qualitative, corpus-based analysis of pragmatic function in the diplomatic cable genre

Jessica Stark

Aix-Marseille University

Eligible for nomination to a student researcher award

Corpus-based and discourse analytical methods are commonly used in languages for specific purposes research to highlight the prototypical characteristics of professional discourse. This paper proposes a genre analysis of the diplomatic cable – a routine, institutional genre specific to the diplomatic community (Loriol 2018). In order to characterise a genre which is usually inaccessible to the public, a small corpus (335,980 words) of cables selected from the Wikileaks organisation website, published in 2010, and a more recent corpus (58,002 words) of leaked European Union cables released in 2018 were selected for comparative analysis using the concordancer LancsBox (Brezina et al. 2015).

The study is based on a functional understanding of specialised genre as a set of communicative events corresponding to an overarching goal (Askehave & Swales 2004; Bhatia 2017). Qualitative methods of corpus analysis are used to lay out a preliminary generic prototype for the diplomatic cable drawing on regularly occurring lexico-grammatical features and structural and pragmatic characteristics. Detailed qualitative investigation of the corpus data further aims to demonstrate how stylistic devices such as evaluative markers revealing author stance and engagement (Hyland 2005) and metadiscursive structuring devices (asides, bracketing) are used to achieve a dual specialised purpose encompassing an informative, reporting function and an analytical, advisory orientation. The perceptions diplomatic specialists themselves have of the cable are also factored into the study to provide additional insight on the genre through a series of interviews with professional diplomats conducted between 2015 and 2018.

Styling a leader: Tracing pronoun clusters in the construction of persona

Esther Surenthiraraj

University of Lausanne / University of Colombo

Eligible for nomination to a student researcher award

Discourse and corpora combine in multiple ways to manifest a diverse set of analyses. This paper explores how the application of corpora studies practices complements a critical discourse analysis-driven approach. This paper is part of a larger study of Independence Day addresses delivered by two Sri Lankan Heads of State during the decade following the end of a civil war. The larger study takes a discursive approach to unpacking the articulation of the post-war nation, and argues that this articulation shapes and is shaped by the persona portrayed by Sri Lanka's post-war presidents. This paper concentrates on how one of these presidents – Mahinda Rajapaksa – styles himself. Methodologically, the paper applies selected corpus linguistics techniques and discourse analysis to a small set of speech data. It uses corpus linguistic techniques to identify pronoun usage and recurring patterns of lexical items that cluster with pronouns in six Independence Day speeches delivered by President Mahinda Rajapaksa (2009-2014). Thereafter discourse analysis is employed to trace the relevance of patterns identified in the corpus onto the larger speech event and post-war discourse centred on the nation. The paper finds that when used in conjunction, discourse analysis adds contextual depth to items recurrent in the corpus, thereby making findings richer. A close analysis of pronoun clusters used to construct the self reveals that these clusters function as rhetorical devices. They construct a Rajapaksa persona centred on a particular set of beliefs, a valorising of the present, a drawing on shared knowledge between himself and his audience, and acts of remembering. The paper finds that when pronoun clusters are read as part of Rajapaksa's persona construction, it provides a commentary on how the president's styling of himself as leader in the current conjuncture feeds into and sustains his articulation of a post-war Sri Lanka.

Combining Collocational Analysis and Semantic Prosodies in a Large-Scale Corpus Study of Metaphor

Stefanie Ullmann

University of Cambridge

In this paper, I present the triangulatory methodological framework I developed to perform a corpus-assisted critical discourse study of metaphors in a multi-million-word data set. While corpus approaches have become more and more frequent in critical metaphor studies, we still lack large-scale investigations of authentic natural language data. This talk details the incorporation of contextual information, in the form of collocation and semantic prosody, to enable a more explicit method for metaphor identification and analysis in a large corpus. The data set comprises a total of 11 million words of international media and political language from on the Arab Spring collected between December 2010 and September 2013. I identified a total of 74,176 linguistic metaphors, which could be separated into 14 conceptual metaphors. My central research question was as follows: Which representations of reality have the media and politics constructed in their discourse in relation to the AS protests? In order to enable a detailed and representative analysis of the different media and political representatives in the corpus and their uses of metaphors, I set up a framework – specifically for this corpus – distinguishing between positive, negative, mixed and neutral semantic prosodies of the metaphorical expressions used in context. This way, I was able to draw conclusions such as: In the media, negative representations were most frequent in UK liberal news reports and most commonly realised through the use of the natural force metaphor. In contrast, positive representations were most frequent in Arab state-owned news sources and most commonly realised using the journey metaphor. In this talk, I seek to demonstrate how collocation and semantic prosody can help us conduct more reliable corpus-assisted studies in critical discourse studies.

An ecological, cross-linguistic study of the *2030 Agenda for Sustainable Development*

Elena Valvason

University of Pavia / University of Bergamo

Eligible for nomination to a student researcher award

On 25th September 2015 the United Nations released *The 2030 Agenda for Sustainable Development*, a resolution meant to commit the world's countries to cooperate towards sustainable practices of developing and living (United Nations 2015). These sustainable practices were described in relation to seventeen Sustainable Development Goals aimed at improving the conditions of “people”, “planet”, “prosperity”, “peace”, and “partnership” (United Nations 2015). The preliminary reflections on these practices encouraged the involvement of world citizens, who were asked to contribute to this paramount endeavour (Fox and Stoett 2016). Nevertheless, the relationship that should tie citizens to the natural world in order to advance towards real sustainability was sometimes neglected in these reflections (Naeem et al. 2016).

The present paper explores thus the discursive construction of the relationship between the world citizens and the natural world in the English, Hungarian, and Italian versions of the 2030 Agenda. The goal of the paper is two-fold. First, it identifies and classifies the actors involved in (un)sustainable practices (i.e. people, animals, plants, or other natural elements) by hinting at the salient and the erased ones (Stibbe 2015). Second, it investigates the discursive construction of these actors in terms of agency and transitivity (Halliday 2004).

The study is carried out within the framework of the ecological analysis of discourse (Alexander and Stibbe 2014) and with a corpus-assisted approach (e.g. Partington et al. 2013; Marchi and Taylor 2018). The dataset consists of the English, Hungarian, and Italian versions of the *2030 Agenda for Sustainable Development*. The corpus is divided into three small, parallel subcorpora totalling roughly 15,000 tokens each. The corpus is explored with AntConc (Anthony 2019) and with #LancsBox (Brezina et al. 2018) in search for the most frequent or infrequent types of the collection (for the identification of the actors of (un)sustainable practices) and for their collocational networks. The collocational networks are expanded on to trace recurrent patterns of agency and transitivity for the aforementioned actors.

Developing CL-CDA methodologies for investigating language ideology in a general news corpus

Babette Verhoeven

University of Huddersfield

There is much qualitative research on language ideologies (LIs), but relatively little on LIs using quantitative methods, especially those from corpus linguistics (CL) (Edwards, 2018; Fitzsimmons-Doolan, 2011, 2014; Freake, 2011; Vessey 2015a, 2015b, 2016). Meanwhile, combining CL and Critical Discourse Analysis (CDA) methodologies to investigate ideologies generally suggests a possible manner of identifying LIs along similar lines (e.g. Baker et al., 2012; Baker & McEnery, 2015; Bednarek & Caple, 2014; Gabrielatos & Baker, 2008; Partington, 2015; Salama, 2011). Yet, existing CL-[C]DA LI research tends to focus on very specific discourses, usually with smaller corpora rather than the wide-ranging discourses of general corpora (Edwards, 2018; Fitzsimmons-Doolan, 2011; 2014; Vessey 2015a, 2015b, 2016). Possible reasons for the relative paucity of LI research with general corpora are that LIs are mostly implicit (Vessey, 2016) and omni-present (Cameron, 1995). This makes it hard for researchers to apply the standard CL-CDA tools:

1. Selecting nodes for investigation - there are not many obvious choices apart from language name(s) (Vessey, 2015b, 2016);
2. extracting keywords - any LIs in the focus corpus are also likely to be present in the reference corpus, resulting in keywords that are indicative of general ideological concerns, but not LIs.

This paper will explore approaches to overcome these challenges by investigating LIs in the general English news discourses of the Time-stamped JSI Web Corpus English 2014-2019 (Kilgarriff et al. 2004, 2014; Sketch Engine, n.d.). Existing qualitative sociolinguistic LI research is used to suggest nodes. This is alongside cross-language comparisons from several JSI corpora, while the quantitative corpus tool of word frequency lists is employed to locate other potential nodes. Finally, to illustrate how LIs can be identified, a selection of nodes' collocations, colligations, and discourse prosodies (Bartsch & Evert, 2014; Firth, 1957; Hauser & Schwartz, 2016, 2018; Stewart, 2010) will be explored.

Theorising corpus absence in a language policy context: identifying benign neglect in the United Nations (1970-2016)

Rachelle Vessey & Lisa McEntee-Atalianis
Birkbeck, University of London

Rachelle.vessey@carleton.ca

@rachellesvessey

http://www.bbk.ac.uk/linguistics/our-staff/academic-staff/rachelle_vessey

Although the United Nations officially promotes multilingualism, the de facto reality has seen an imbalance between the official and working languages and a growing preponderance of English (Piron 1980; McEntee-Atalianis 2016). Trying to address the challenge raises the question of agency in the organisation's language policymaking. In this paper, we propose that agency can be considered as having a 'voice' and the active use of that 'voice' to lobby for specific language causes (i.e. through metalinguistic discourse). At the same time, we are also conscious of the powerful role of silence and the related issues of absence and inaction (Duguid and Partington, 2018; Schroeter and Taylor, 2018). We contend that persistent silence on language issues relates to wider issues of 'benign neglect', i.e. the belief that the state, or in this case, the organisation, should not intervene to encourage or discourage particular linguistic choices by its citizens/membership (Kymlicka and Patten, 2003: 32). In this paper, we focus on benign neglect as it manifests through silence (and/or absence of discussion) and its potential effect on the language policy of the United Nations.

We use corpus linguistics to compare and contrast metalinguistic discourse from two levels of UN organisational membership over the course of 46 years (1970-2016): Secretary General annual statements (2,212,297 words) and member states' contribution to the annual General Debates (22,070,872 words). Both are charted against organisational language policy changes and developments between 1970-2016. Findings indicate that few language issues are raised by both groups over similar time periods and member state metalinguistic discourse rarely aligns with organisational language policy shifts, suggesting that "bottom-up" discourse is less influential than that emanating from "above" (the Secretariat/ Secretary General). Nonetheless, we conclude that "bottom-up" persistent silence is influential in that the silent majority serves to maintain the monolingualising status quo. Thus, silence serves as a means of exerting agency in terms of "benign neglect" language policy.

'Black or gay or Jewish or whatever': A diachronic corpus-based discourse analysis of how the UK's LGBTQI population came to be represented as secular, cisgender, gay, white and male

Mark Wilkinson
Lancaster University

m.wilkinson7@lancaster.ac.uk
@MarkJoseph_82

Eligible for nomination to a student researcher award

In spite of the inherent diversity within queer populations, critics have argued that media representations of LGBTQI people are 'embedded in whiteness' (Johnson 2014:280). Through an analysis based on corpus data from *The Times* between 1957-2017, this paper provides evidence for the ways in which the LGBTQI population in the UK has been discursively constructed as largely secular, cisgender, gay, white, and male. By rooting the analysis in post-structuralist discourse theory (PDT) (Laclau and Mouffe 1985), it is argued that such gendered and (de)racialised subject positions are the result of a 'sedimentation' of discourse that "'forget(s)" and "conceal(s)" the ultimately contingent nature' (Phelan and Dahlberg 2011:23) of all hegemonic formations, including identity.

In the current study, PDT's concept of discursive 'sedimentation' is complemented by considering both similarity (Taylor 2013) and absence (Schröter and Taylor 2018) in the corpus data. After dividing the corpus into 3 time periods, the consistent multiword keywords that occur across the sub-corpora are highlighted so as to ascertain which discourses have persisted over the past 60 years. Their c-collocates (Gabrielatos and Baker 2008) are then compared, revealing that the most consistent collocational patterns pertain to race and religion. Terms that denote whiteness or secularity are conspicuously absent in the data, suggesting that a term like black is marked while a term like white is assumed. Trechter and Bucholtz (2001:5) contend that 'as a cultural sign, whiteness works much like a linguistic sign, taking its meaning from those surrounding categories to which it is structurally opposed'. I argue that the absence of whiteness is in fact a sign in and of itself and, that the tacit assumption of whiteness, is therefore constitutive of the erasure of diverse queer subjectivities – a process which has occurred through consistent choices in representation occurring over a sustained of time.

Examining and comparing English textbooks and high-stakes exams in Turkey: A Corpus-Based Study

Xiaoli Yu

Middle East Technical University

xiaoli@metu.edu.tr

<https://fle.metu.edu.tr/en/xiaoli-yu>

Textbooks, as the de facto syllabi in elementary and secondary education, largely influence the pedagogical practice in classrooms. Hence, textbook research is of considerable importance in shaping educational constructions. Meanwhile, research in assessment, in particular standardized examinations, has revealed substantial washback effects on both teaching and learning. However, in the field of English language education in Turkey, few empirical studies have been conducted to investigate the relationship between textbooks and high-stakes standardized exams. Furthermore, no studies have been performed based on corpus data.

This ongoing research study aims to examine whether the English textbooks used in Turkish K-12 public schools align with the high-stakes high school and university entrance exams for the past decade. In other words, the study aims to investigate the connection between what is supposed to be taught and what is tested. Corpus-based quantitative analysis will be conducted to compare the textbooks and exams from vocabulary, sentence, and discourse levels. Moreover, sociolinguistics aspects will be considered in the comparison as well. Pedagogical implications as well as insights for curriculum revising, textbook writing, and exam designing will be provided based on the findings of the study.

The preliminary results of this study have shown limited overlap between the high school textbooks and the university entrance exams regarding most frequent vocabulary items, including lexical bundles. With respect to lexical sophistication and lexical diversity, significant differences between the textbooks and the university entrance exams have also been detected. The lexical sophistication and diversity level of the textbooks were significantly lower than the exams. The final research findings are expected to be concluded by May, 2020.

Czech broadsheets discourse about female politicians in comparison with male politicians

Adrian Zasina
Charles University

adrian.zasina@ff.cuni.cz

<https://ucnk.ff.cuni.cz/en/institute/people/adrian-jan-zasina-2/>

Nowadays it is obvious to incorporate women into politics, however, there are still some dissimilarities in the representation of both genders – not only in a government but in public discourse as well. Hence my study (Zasina 2018) aims at female politicians' (in contrast to male politicians') representation in the Czech daily broadsheets. It also explores a possible stereotypical image of women in politics. The research uses corpus methods to analyse adjectives that premodify two lexemes: politik 'male politician' and politička 'female politician'. The analysis is based on a subcorpus of the SYN corpus version 4 (<https://wiki.korpus.cz/doku.php/en:cnk:syn:verze4>) containing five different Czech daily broadsheets from the years 1991–2014. There are two case studies analysing adjectival collocations of both examined lexemes. The first one examines the positive and negative meanings of these adjectives to reveal the similarities and differences between male and female politicians. The second one analyses the top 20 collocates to capture the prevailing discourse. In comparison with similar studies on the English language (Pearce 2008, Caldas-Coulthard and Moon 2010) that emphasise gender differences, the results of this study provide a more complex view on the gender issue in politics. It shows that in this material both male and female politicians are portrayed mostly in a positive light and female politicians are not always presented in a stereotypical way. Moreover, the properties of adjectival collocates are more subtle than obvious stereotypes and are not dichotomous in nature. They also show that women, in some areas, are seen as having a potential, albeit they are portrayed as holding power as well. The present research brings a new insight because it is one of few attempts (cf. Elmerot 2017) at corpus linguistic analysis of gender in politics in the Czech language.

Investigating activism on Twitter

Angela Zottola & Andressa Rodrigues Gomide
Aston University / Lancaster University

angela.zottola@unito.it
@angela_zottola

a.rodriquesgomide@lancaster.ac.uk
@gomidear

As the right-wing agenda of politicians worldwide keeps growing stronger examples of bigot and discriminating behaviours are increasing. Some of the political campaign, which have been flooding the media, explicitly violate basic human rights, and more often than not the lives of people belonging to specific minority groups are being put at risk. More specifically, gender and sexual identities have been the target of many politicians, from Trump in the USA to Salvini in Italy, to Bolsonaro in Brazil.

This study aims at analysing the way in which activism is contrasting this wave of violence and discrimination through the use of Twitter and more specifically of hashtags. We will focus on a case study related to the transgender community in the USA which has been repeatedly attacked since president Donald Trump was elected in November 2016. In October 2018 the Trump administration announced that they were considering narrowing down the definition of gender as a biological and immutable condition determined by the sex assigned at birth.

As a response to this explicit and fast-growing hostile behaviour, the National Center for Transgender Equality organized a protest launching the hashtag #WontbeErased, a catchphrase which represents the way transgender people were feeling following this announcement. The hashtag immediately became very popular on social media and a symbol of the fight against gender identity discrimination. This study analyses a corpus of tweets containing the hashtag collected from October 21st 2018 onward. Through keyword analysis, the corpus will be compared against a large corpus of general tweets to identify the discourse, linguistic and semantic patterns used by transgender people and allies to articulate counter-discourses to hate and hostility and construct agency. In this talk we will also address the challenges and issues related to working with online data.

Short paper presentations

A Cross-linguistic Study of Metadiscourse Markers in English Academic Writing of Saudi EFL Students and UK Native Speakers of English

Nasser Alqahtani
Cardiff University

Eligible for nomination to a student researcher award

The use of metadiscourse (MD) in academic writing is a concern for students due to its explanatory and rhetorical nature (Nash, 1992). MD markers are words (e.g. first and surprisingly) that guide readers in a text and show the stance of the writer. MD has two main functions: textual and interpersonal, which are based on Halliday's (1973) macro-functions of language. On the textual level, MD guides the readers in the text to organize and interpret the content, and on the interpersonal level, MD turns the readers to the writer's perceptions and attitudes making texts more reader-friendly (Crismore & Farnsworth 1989). Crucially, the use of MD markers varies from one group to another depending on the cultural and institutional contexts (Hyland, 2005; 2017). Therefore, specific practices are witnessed in Saudi students' writing as they come from a different society and culture (Alowayid, 2018). A few studies have looked at Saudi students' use of discourse markers, and concluded that students underuse and misuse markers which result in low quality of their writing (Daif-Allah & Albeshir, 2013; Alowayid, 2018). Most MD studies investigate MD based on a functional approach i.e. communicative function (Hyland, 2005) but this study combines the functional approach and the basic constituents of the clause (subject, predicate, etc.) to determine the functions of the markers syntactically and communicatively. Using a corpus of 30 MA dissertations in Applied Linguistics, this study explores the differences and similarities of MD usage between three groups of participants: Saudis in Saudi Arabia, Saudis in the UK, and UK native speakers of English. This research adds to the emerging L2 academic writing literature on the use of MD cross-culturally. The results can be used to develop materials for teaching in KSA and to see what affect students' use of MD more the discipline or the institutional context (Saudi universities VS UK universities).

Weaponising coronavirus online

Nicole Baumgarten
University of Sheffield

The presentation introduces a new project on the potential evolution of coronavirus-related online hatred on German Twitter. The project will trace if and how lexical items initially referring to the medical and geographical realities of a news item, such as 'corona', 'COVID- 19', and 'Wuhan', have acquired through their context-based use on Twitter a hateful discourse prosody (Louw 1993), in parallel with the spread, peak, and aftermath of the outbreak of the pandemic in Germany and worldwide. Using an existing corpus of the full German-language Twitter, the project adopts a corpus-assisted discourse analysis approach to provide empirical evidence and near real-time insight into a) the construction of rhetorical repertoires grounded in the health threat and b) if and how these are weaponised in stigmatization processes that target particular socio-ethnic and national groups, and are used to disseminate socially disruptive responses to broader socio- and geopolitical processes, such as globalisation, migration, Sino-Western relationships, and cultural difference. The aim of the project is to contribute to overcoming current limitations in the understanding of language and culture-specific determinants of the production of online hatred, which are the necessary preconditions for closing the gap between institutional conceptions of hate speech and its linguistic and textual forms within the constraints of particular social media (Twitter) and in particular lingua-cultural environments (here, the German-speaking public), which can serve as the basis for computational and legal modelling of online hatred.

Military frames for the pandemic in Central and Eastern Europe: A corpus-linguistic perspective on figurative battles, historical battles, and the rhetorical potential of WAR

Ksenija Bogetic

University of Belgrade, Serbia / Research Centre of the Slovenian Academy of Sciences and Arts, Slovenia

ksenija.bogetic@zrc-sazu.si

This work in progress draws on corpus linguistics and Critical Discourse Analysis to explore the militarisation of public discourse in Central in Eastern Europe in the face of the covid-19 pandemic. It draws on specialised corpora of official political leaders' address in Serbia and Slovenia, and citizens' online comments and Twitter posts. Preliminary findings point to several WARFARE keywords in ambiguous political frames of metaphorical-cum-historical battles reinforcing nationalist imaginary, in contrast with what seem to be much more diverse framings of the crisis in the citizen data. The analysis aims to (a) dissect the rhetorical meanings of the WAR metaphor for the pandemic in the post-Yugoslav, post-conflict locales, where memory of non-figurative wars is still very much alive; and (b) highlight the nature of discursive frames of the pandemic as locally meaningful, dynamic, and socio-historically contingent, even when they appear globally resonant.

The good and the bad of Italian pandemic pedagogy: a corpus-driven analysis of the representation of Covid-19 emergency online teaching in the newly developing discourse of the Italian media

Giovanna Carloni

L'Università di Urbino Carlo Bo

This presentation aims to investigate the main features of the discourse the Italian media have developed to represent the emergency online teaching in which instructors and students – first in Europe – have engaged at secondary and tertiary level to handle the COVID-19 mass emergency education. A corpus of digital texts produced since the Italian national lockdown in early March has been compiled. Keyness, collocations, lexical bundles, and semantic prosody have been investigated to examine the main characteristics of the discourse the Italian media have been rapidly shaping to portray how the emerging pandemic pedagogy has been experienced by its participants (namely instructors, students, and families) engaged in the new highly challenging digital and emotional teaching/learning processes. A thought-provoking profile of the discourse patterns identified seems to emerge, which deserves a thorough investigation.

Stancetaking in Vietnamese and English texts in the light of appraisal theory

Tieu Thuy Chung

University of Queensland

tieuthuy.chung@uqconnect.edu.au

<https://www.researchgate.net/project/Writing-with-Attitude-a-Learner-Corpus-Study-of-Appraisal-Resources-of-Vietnamese-and-Khmers-L2-English-Writing>

Eligible for nomination to a student researcher award

Whether L2 English student writers lack essential English knowledge in expressing their stance is an issue educators may be aware of. Developed by Martin and White (2005), appraisal 'as a discourse semantic resource for meaning' has been employed to understand the interpersonal stance in L1 and L2 written texts. In Vietnam, appraisal theory has been seldom adopted for the analysis of stancetaking in L1 Khmer / Vietnamese language varieties and the L2 English writing of speakers with Khmer and Vietnamese backgrounds. The present study employs this framework to understand how tertiary student writers from these language backgrounds convey their stance, focusing on essays written in the same tasks under the same conditions. A Vietnamese corpus of around 4,200 words² and an English corpus of around 5,400 words³ built from twenty-four essays were analysed in my preliminary study, prior to the main annotation of my PhD project data. The preliminary study reveals that L2 English writing of Khmer-Vietnamese and their official Vietnamese had similar distribution of affect, judgement and appreciation. This may suggest there may be cross-linguistic similarity across the uses of attitude in these two languages. Additionally, more expansion was used in L1 English and L1 Vietnamese while in the three L2 groups contraction was preferred. The implications of these findings are that the dominant Vietnamese may have certain influence on appraisal employment of Khmer-Vietnamese's L2 English and L2 English learners at lower levels lack required vocabulary to adopt L1 English norms, overusing / underusing stance categories / lexis common to L1 English.

Covid-19 and social media: a diachronic discourse analysis for the modeling of linguistic patterns during crises

Louise-Amélie Cougnon & Louis de Viron
Université catholique de Louvain / Datatext

louise-amelie.cougnon@uclouvain.be
@Amelaaay
<https://uclouvain.be/en/research-institutes/ilc/miil>

louis@datatext.eu
@devironl
<https://www.datatext.eu/>

Our project aims at measuring the impact of the Covid-19 crisis on the French-speaking Belgium population through the analysis of discourses on social media. In particular, we qualitatively and quantitatively process a continuously growing corpus of (today) 4.000 tweets from politicians, 30.000 tweets from the media and 100.000 from the population. All messages include a Covid-19 related expression and were sent since 1 January 2020.

In this presentation, we will focus on the diachronic evolution of linguistic framings of the pandemic from the emergence of the disease in China to reopening plans, through lockdown measures in Belgium. We will notably try to identify formula of disincentives to the acceptance of political measures by citizens. Through the analysis of messages over time, our aim is to find out possible discourse/reaction curves that could model the linguistic pattern of the incentives-disincentives variation that happens during the different stages of the pandemic crisis.

Social media discourse and conspiracy theories during the COVID-19 pandemic

Natalie Dykes & Stefan Evert
FAU Erlangen-Nürnberg

Eligible for nomination to a student researcher award

With nationalist populist movements on the rise and the instrumentalisation of Twitter by parties like the German AfD and politicians like Trump, social media are considered central in spreading opinions contradicting the societal status quo. In this context, misinformation and conspiracy theories have gained traction, and COVID-19 is no exception. Such content is spread by a diverse range of fringe groups including esoterics, anti-vaccinators and nationalists – with increasing success in mobilising the general public to participate in (illegal) demonstrations.

We explore voices opposing mainstream scientific and political stances towards COVID-19 in Germany using two large corpora of Twitter and Reddit posts. We contrast “alternative” and “mainstream” language use regarding lexical, argumentational and phraseological style, exploring the link between conspiracies/misinformation and linguistic strategies. The tweets can also be used to track the spread of misinformation over time and across the social media network.

Wake Up It's a #Plandemic: Covid-19 and Conspiracy Clique Communities on Twitter

*Dr. Joseph Flanagan
University of Helsinki*

This presentation will explore how systemic functional linguistics, network analysis, and natural language processing can give us insight into the ways in which social media is used to create what Michelle Zappavigna describes as "ambient affiliation" among promoters of a broad spectrum of conspiracy theories. Using a dataset of approximately half a million tweets mentioning "plandemic," a conspiracy theory video about covid-19 released on social media in May 2020 that quickly became viral across many social media platforms, I will show how a hashtag co-occurrence network can reveal the thematic links among QAnon, anti-vaccine activists, and other conspiracy-minded tweeters. A further examination of cliques and related structures among this network will reveal how alliances get formed within and across these groups.

What's love got to do with it? Gender accommodation versus flirting in teenagers' online conversations

*Lisa Hilte, Reinhild Vandekerckhove & Walter Daelemans
CLiPS research center / University of Antwerp*

Lisa.Hilte@uantwerpen.be
<https://www.uantwerpen.be/nl/personeel/lisa-hilte>

Reinhild.Vandekerckhove@uantwerpen.be
<https://www.uantwerpen.be/nl/personeel/reinhild-vandekerckhove>

The present study analyzes the phenomenon of linguistic accommodation, i.e. adapting one's language use to one's conversation partner (see e.g. Dragojevic et al. 2015; Giles & Ogay 2007). Accommodative adjustments are generally driven by a desire to facilitate interaction and to regulate (i.e. increase/decrease) social 'distance' (see e.g. Dragojevic et al. 2015; Giles & Ogay 2007). While accommodation has been widely investigated in face-to-face interactions, it has largely been neglected with respect to written social media discourse, which is the focus of this paper.

In a large corpus of private and spontaneous instant messages, we compare Flemish teenagers' language use in two conversational settings: same-gender (including only boys or only girls) and mixed-gender online conversations (including at least one girl and one boy). We examine whether boys adopt a more 'female' and girls a more 'male' writing style in mixed-gender talks, i.e. whether teenagers converge towards their conversation partner, specifically with respect to expressive/emotional writing. The analysis covers two distinct repertoires that the teenagers have at their disposal to express emotionality: a 'digital' repertoire including several typographic chatspeak markers (e.g. emoticons) and a 'traditional' repertoire consisting of verbal expressions of sentiment (e.g. emotion words). We compare both types of expressive/emotional writing in the teenagers' online interactions and look for potential patterns of gender convergence. Finally, we tackle a theoretical challenge and try to distinguish between 'general' linguistic gender accommodation and flirting strategies from a quantitative as well as a qualitative perspective.

A Corpus-based Comparative Research on Covid19 News Discourse Representation in Voice of America and China Daily

Yuying Hu

Guangdong University of Foreign Studies

200511444@oamail.gdufs.edu.cn

<https://www.gdufs.edu.cn/>

This is a comparative corpus-based study on pandemic news discourse underpinned by critical discourse analysis (CDA) framework, with a focus on the salient lexical profile features of the top 20 positive keywords and their discourse functions in covid19 news reporting from Voice of America (VOA) and China Daily. The research aims at exploring the focus and attitude of how the pandemic is reported in two media, and interpreting the potential ideologies behind discourse representations in two media. There are 1,421,461 tokens in VOA and 1,048,149 tokens in China Daily respectively. The sub-corpus of news section in Corpus of Contemporary American English (COCA) is used as a reference corpus for keywords extraction. The research method is both quantitative and qualitative due to the combination of the corpus technique and the triangulation framework in CDA proves to be a “useful methodology synergy” (Baker et al. 2008).

Formulaic sequences in Early Modern English: A corpus-assisted historical pragmatic study

Ding Huang
Heidelberg University

ding.huang@stud.uni-heidelberg.de
@elaine_d_huang

Eligible for nomination to a student researcher award

This PhD project asks two main questions: 1) What are the form and functions of formulaic sequences (FSs) in Early Modern English (EModE)? 2) How do they characterise different text-types and types of communication? With a corpus-assisted approach, the project compares the use of FSs in dialogues and letters in early modern England. It aims at making a further contribution to research on FSs in EModE with an attempt of 1) defining FSs in a way that can be applied in research from various perspectives and 2) enhancing the methodology of identifying FSs semi-automatically in historical texts.

This particular poster presentation will report on the current progress of this project, focusing on the following issues: 1) defining and identifying FSs in a historical linguistic context; 2) the relationship between FSs and other similar theories such as construction grammar (e.g. Goldberg, 2003) and pattern grammar (e.g. Hunston and Su, 2017); and 3) examples from a pilot study. More specifically, for 1), Conrad and Biber (2005) summarised six characteristics of FSs: fixedness, idiomaticity, frequency, length of sequence, completeness in syntax, semantics, or pragmatics, intuitive recognition by native speakers of a language community. Among them, does “fixedness” of FSs mean the same thing in PDE and EModE? Is “intuitive recognition” reliable in historical texts? What kind of digital or computational approaches can be used to identify FSs and what are the criteria? For 2), this project agrees with Buerki (2016) that FSs are constructions and argues that FSs are a conventional mapping of form meaning and function. The project also argues that despite a useful way to investigate FSs in a massive amount of historical texts, lexical bundles are, to some extent, different from FSs, hence could not fully account for the form-meaning-function mapping of FSs.

Yellow or Blue? Ideological divergence on Anti-extradition Bill Protest in Hong Kong English News Media

Zeping Michelle Huang
Hong Kong Baptist University

huangzeping@hkbu.edu.hk

@DoraZeping

<http://www.coms.hkbu.edu.hk/faculty-staff/faculty/huangzeping/>

This research investigates the extent to which two major English-language news media in Hong Kong framed the biggest social turmoil in five decades triggered by the then-withdrawn extradition bill. The two news outlets under investigation are South China Morning Post (SCMP) and Hong Kong Free Press (HKFP). We collected two corpora of news articles on the anti-extradition bill protests published from January 2019 to 23 October 2019, the date when the Hong Kong government officially withdrew the extradition ordinance. Combining with corpus linguistics (CL) and critical discourse analysis (CDA) approach, this study focuses on the discursive constructions of the anti-extradition bill protest between the two news outlets. Keywords and collocation patterns are compared both quantitatively and qualitatively, followed by an in-depth concordance analysis of the keywords in context. The concordance lines are analyzed based on Discourse Historical Analysis framework (Reisigl & Wodak, 2009), i.e., 1) normalization 2) predication, 3) perspectivation, and 4) argumentation. Political stances and ideologies in SCMP and HKFP will be investigated and compared, followed by a discussion on how the two news media framed the unprecedented political crisis.

Exploring the relationship between COVID-19-related sinophobia and representations of Chinese food culture in the UK press

Ursula Kania
University of Liverpool

Police have recorded a sharp increase in hate crimes against Chinese people since January 2020. This rise has been partially driven by the claim that the origins of COVID-19 lie in Wuhan's Huanan Seafood and Wildlife Market, which is potentially connected to the stigmatisation of Chinese food culture on a more general level.

This project uses corpus-assisted critical discourse analysis (Baker & McEnery, 2015) to explore harmful stereotypes around Chinese food culture in UK press coverage. Preliminary results obtained through searching COVID-19-related data on Nexis (from 01/2020 onwards) indicate that lexical choices (e.g., 'exotic meat(s)') contribute to a construal of Chinese food culture as 'other' and hence potentially dangerous.

Further analyses (keywords/collocates/discourses, using AntConc/ProtAnt) will draw on a newly-compiled newspaper corpus (c300,000 tokens), with a historical corpus (for comparative purposes) also in preparation.

'What's So Special About The Circus?'

Katharine Kavanagh
Cardiff University

kavanaghk@cardiff.ac.uk
[@bustingfree](#)
<http://TheCircusDiaries.com>

Eligible for nomination to a student researcher award

This research is a corpus-assisted critical discourse study into the values attributed to the experience of attending circus performances. Within linguistic studies, underlying concepts of value are seen to materialise through cognitive stances that are realised in texts as evaluation (Thompson and Alba-Juez, 2014:10; Bednarek and Caple, 2017:78). Evaluation is also a key registerial feature of promotional genres, both those explicitly aimed at marketing a product and those aimed at reviewing a product—otherwise known as 'interested' and 'disinterested' genres (Shaw, 2009). My previous research has shown that, in the circus field, reviews most readily available to the public have only represented a limited variety of the values that can be attributed to the experience of attending circus productions (and, moreover, do so in a non-dialogic, authoritative fashion that occludes the possibility of alternatives). My PhD project extrapolates from these findings that there are values held by members of circus audiences that are not publicly represented. To discover the extent to which this is the case, and the way any such discrepancies manifest, this research combines Audience Research and Corpus Linguistics methods to compare the evaluation realised in three subcorpora: a subcorpus of audience interview transcripts; a subcorpus of circus reviews; and a subcorpus of publicity blurbs corresponding to the shows represented in the review subcorpus. Analysis will be conducted by triangulating results of Keyness (Gabrielatos, 2018) and APPRAISAL (Martin and White, 2005) analyses. Key semantic domains for each subcorpus will be identified through application of the UCREL Semantic Analysis System (Rayson et al., 2004), with the full circus evaluation corpus as reference. Each subcorpus will be proportionately sampled and the samples coded for APPRAISAL features using UAM CorpusTool (O'Donnell, 2008). This poster presents the rationale for this research and the proposed methodology.

Metaphor Use and Efficacy of Science Communication [planned multidisciplinary project]

Saskia Kersten, Alana Jelinek & Naomi Forrester-Soto
University of Hertfordshire

In an artist-led, multi-disciplinary project, we plan to explore the use of metaphors by a range of populations, including virologists, and how this shapes the discourses around viruses. An artist, a virologist and a linguist will work together on this project which is currently in its planning stage.

One of the methods for obtaining data on metaphors will be visual and crowd-sourced in addition to language data. The linguistic aspect of the project will focus on which metaphors virologists use to talk about viruses when teaching students, how and when these metaphors are used and which impact this may have on lay-expert communication, particularly at times of crisis such as the current Covid-19 pandemic. We would like discuss the role corpus linguistics may play in addition to methods from socially engaged art practice in the context of this pilot project.

Corpus-Based Critical Discourse Analysis of the Covid-19 Pandemic Discourse in the World Health Organization Director-General's Speeches

Baramee Kheovichai
Silpakorn University

kiao_ra@yahoo.com

<http://www.arts.su.ac.th/คณาจารย์ภาควิชาภาษาอังกฤษ/บารมี-เขี้ยววิชัย>

This on-going research aims to investigate the discourse surrounding the Covid-19 pandemic produced by the Director-General of the World Health Organization. It seeks to identify: 1) what key concepts are discussed in the discourse of Covid-19, 2) how Covid-19 is discursively constructed, and 3) how the World Health Organization is represented in discourse. To this end, the speeches delivered by the Director-General were collected from January to June 2020 to form a corpus. The corpus will then be compared against the BE06 corpus to identify keywords, which represent salient concepts in discourse. Words referring to Covid-19 are then investigated, using collocation and concordance functions. Moreover, words related to the World Health Organization are then analyzed to determine the self-presentation of the World Health Organization. Preliminary findings of keyword analysis based on the pilot corpus will be presented.

Textures of John Clare's sonnets: A corpus-based structural comparison between three master sonneteers

Kazutake Kita

Tokyo University of Science

John Clare (1793-1864), a 19th-century British peasant poet whose prolific literary output has gained increasing scholarly attention for the recent decades, could be counted as one of the major sonnet writers of his era. Some of these 19th-century sonneteers, among whom are included Clare as well as Wordsworth and Keats, are known to have made innovative attempts in the traditional 14-line poetic form, working on their own principles and ideas about sonnet and thus complexifying its generic identity. Literary researchers have not been slow to give descriptions and explanations for the innovative contribution of the poets, whereas their discussions have often focused on the poets' choice of rhyme scheme (patterns found in the vertical direction), leaving unaddressed the issue of rhythmical variations in each line (patterns found in the lateral direction). This poster presentation reports a piece of research designed to capture both of these two types of structural pattern observed in the sonnets composed by the above-mentioned three poets, with Clare, presumably the least known of them, as the main focus. Using mini-corpora compiled from all their sonnets included in the Major Works editions of the Oxford World Classics series (117 of Clare's, 57 of Wordsworth's, and 48 of Keats's), and drawing on the vector space model analysis and the keyword analysis as two principal methods, the research has found that: (1) Clare's rhyme scheme patterns, with his all-couplet sonnet as the signature form, evince outstanding uniqueness among the three; (2) Clare shows a stronger inclination than the others towards regular alternating rhythm; and (3) Clare's regular rhythm may have a close connection with his characteristically frequent use of "the" and "and". The findings suggest a possibility of more variety inside the genre than previously supposed and thus point towards a new direction for future sonnet studies.

The Development of COVID-19 Word List from the Perspective of Emergency Language Services

Longxing Li
University of Macau

Eligible for nomination to a student researcher award

The COVID-19 pandemic around the globe made us acutely aware of the importance of emergency language services in public health emergencies. Terminology support provided in the form of word list development and terminology management is an important way of emergency language services, which ensures accurate and smooth communication. In order to help prevent and control the pandemic from the perspective of emergency language services, 381 single-word terms and 176 multi-word terms of COVID-19 are extracted from the COVID-19 corpus using the corpus tool of Sketch Engine. From the practice of producing this word list, we put forward a seven-step procedure of academic word list development and point out future directions for academic word list making and research in the hope of providing reference for terminology support in emergency language services.

Body Parts as Means of Enacting Agency in Transgender Fiction

Eloisa Lillywhite
University of Sussex

E.Lillywhite@sussex.ac.uk

Eligible for nomination to a student researcher award

Contemporary society's understanding of gender as a spectrum has led to a significant rise in the inclusion of non-binary characters within fiction. This work in progress contributes to awareness of gender representation by considering the portrayal of transgender characters in young adult fiction.

The physical body is a vehicle of gendered representation in popular media such as children's books (Jackson and Gee, 2006) and magazines (Ricciardelli, Clow and White, 2010). In fiction, characters' bodies define their interactions with others, and their perceived status; the body, and its parts, is a means of demonstrating physical agency, or a lack thereof.

Using corpus analysis and critical discourse analysis methods, a pilot study was conducted with a corpus of four award-winning, and therefore widely accessible, books featuring transgender protagonists. The key findings so far are that cisgender females are nurturing to all but assertive over transgender characters, while transgender females are overtly feminine and physically passive. Cisgender males are aggressive and physically dominant, while transgender males are underrepresented and internally focused.

The current stage of the project extends findings with a significantly expanded corpus. To explore differences in levels of agentic behaviour, keyness scores are compared between gender groups, and against a reference corpus. Then, keywords are coded with attention to agent-subject relationships. Using an inductive approach, the concordance lines are analysed for differences and similarities between the gendered groups in regards to their physical agency.

A new corpus Rusidiolect as a tool to study a variety of discourses produced by the same person

Tatiana Litvinova

Voronezh State Pedagogical University

centr_rus_yaz@mail.ru

en.rusprofilinglab.ru

In this paper we describe the development of a corpus Rusidiolect designed to investigate a problem of distinctiveness and uniqueness of an idiolect of a typical Russian speaker as well as a level of stability of idiolectal features under different conditions (change of mode, genre, addressee, etc.). Thus it is crucial to obtain as many texts from the same person as possible. The corpus includes both natural written and oral samples (e.g., Instagram posts and vlogs transcripts by the same author) and texts produced under instruction.

In this paper, apart from describing of corpus statistics, we focus on comparing written and oral texts by the same person in order to reveal idiolectal features which are stable to mode shift. Additionally, we report the results of our experiments on distinctiveness and uniqueness of the authors' use of cohesion devices as idiolectal feature in different types of discourses (both oral and written). We performed a manual annotation of the texts throughout 45 local cohesion indices. As our experiments suggest, cohesion level is instrumental for authorship attribution although rarely used [1] in this field due to lack of appropriate corpora and a relatively low accuracy of automatic tools for this type of annotation.

The corpus is freely available at <https://rusidiolect.rusprofilinglab.ru/>

Interpersonal positioning in written academic ELF – different degrees of self-assurance

Monika Manakova
University of Ostrava

The language of academia has long been predominantly English and indeed if we want to elicit environment where English is used as lingua franca (ELF), the academic environment would be the one which undisputedly springs to mind among the first ones. The study of the use of ELF in various environments has been gaining on importance in the recent years (Jenkins, 2006; Mauranen, 2012, 2018; Seidlhofer, 2011; Swam, 2012, 2017) with the academic one in the forefront.

This paper investigates how authors assert themselves in the text expressing different degrees of self-assurance and certainty – interpersonal positioning and falls back on definition of voice as author's self-representation by Ivanič and Camp (2001). The research question focuses on distribution and frequency of use of verbal choices (tenses and modality) mainly in connection with personal pronouns as a mean of communicating authors' opinions.

By the use of a corpus-based method, I investigate the frequency of the use of the author-reference pronouns and possessive adjectives and the verbal choices in humanities and social science articles written by academics from different cultural backgrounds for the international audience of the global environment as presented in the corpus SciELF of Helsinki University. In order to compare and contrast the results, a reference corpus of articles published in high impact journals of respective fields was compiled and possible dependencies determined by a statistical analysis.

The preliminary results into author's interpersonal positioning based on single-author articles show that there are clear dependencies between the use of various modal verbs and the type of corpora with the ELF authors favouring the use of modals as hedges as well as boosters (they comprise 25.4% of the uses out of all verbs used) over the high impact journals authors who used modals in 5.4% of the cases.

Multimodal Corpus Analysis of Tourism Promotional Communication Online

Elena Mattei
University of Verona

elena.mattei@univr.it
[@ElenaMattei10](#)
<http://www.dlls.univr.it/?ent=persona&id=22467>

Eligible for nomination to a student researcher award

This project in Digital Humanities focuses on carrying out a cross-disciplinary and corpus-based investigation into three travel agencies' multimodal communicational strategies through mixed methodological approaches. The objective is to demonstrate how the multimodal, promotional orchestration of different semiotic resources transfers positive attitudes towards destinations to the target audience and influences its behaviour. At this stage of the project multimodal data has been collected from the official websites and Instagram and Facebook Social Media channels, in order to compile three different sub-corpora pertaining to three different travel organizations, which are “Western Australia”, “Explore Canada” and “Tourism Ireland”. These sub-corpora, featuring both images and texts, are being analyzed separately with Sketch Engine and Multimodal Analysis Image software technologies. On the one hand, image analysis is being carried out after developing a new tag system, based on Kress and Van Leeuwen’s Visual Design theory. On the other hand, Corpus Analysis focuses on the investigation of the three written sub-corpora, each one with a three-part internal division per channel, through concordancing, frequency, collocation and keyword analysis tools. The findings, after undergoing a detailed cross reference for agencies and channels, will be compared with one main corpus and its varietal subdivision, i.e. the Ten Ten Corpus, and with small tourism corpora such as Manca’s (2017) and Bianchi’s (2017). The aim consists in finding commonalities and discrepancies between the collected dataset and informal written tourism language, which comprehends trope, ego-enhancing, syntactical, pragmatic and lexical strategies. The retrieved data is also undergoing qualitative visual and Systemic Functional Linguistics analysis. The results, which have been achieved so far, show the presence of designed and interconnected multimodal techniques, such as the employment of persuasive lexicon and illocutionary patterns together with images that prompt imagination and the consumption of intangible services.

COVID-19 memorials

Mark McGlashan
Birmingham City University

Mark.McGlashan@bcu.ac.uk

@Mark_McGlashan

www.MarkMcGlashan.org

The COVID-19 pandemic is difficult to make sense of. Lockdowns and social distancing, panic buying, stretched public services, working from home, juggling childcare, accessing healthcare... all complex issues without any simple solutions. Guidelines are constantly evolving (“following the science”) and routes forward seem unclear. There is an entire new language to keep up with (R numbers and key workers; <https://www.bbc.com/worklife/article/20200522-why-weve-created-new-language-for-coronavirus>).

With over 30,000 deaths, the UK has one of the worst COVID-19 per capita death rates in the world. As a number, this is difficult to comprehend. But behind this number are people. This talk examines data collected from an online public memorial site (<https://www.rememberme2020.uk/>) to explore how the dead are remembered publicly and to understand the language of remembrance in the context of COVID-19. Some preliminary work, which explores frequent linguistic features (tokens and ngrams) in these memorials can be found here: <https://sites.google.com/view/mark-mcglashan/projects/covid-19>

Sentiment Analysis of hatred speech towards female politicians on Twitter during the UK elections

Carla Fernández Melendres & Aroa Orrequia Barea
King's College London / Universidad de Jaén

carla.fernandez_melendres@kcl.ac.uk

@carlafdz

orrequia@ujaen.es

@OrrequiaAroa

<https://www.ujaen.es/departamentos/filing/contactos/orrequia-barea-aroa>

Eligible for nomination to a student researcher award

Social media platforms such as Twitter play an essential role in politics and social movements nowadays. The shift from public squares to online conversations broke geographical barriers and made communication easier. However, social media maintains the ethnicity and social class struggles that were found on street debates before, and has led to manipulation of information (fake news), racism, and sexist hatred speech.

The aim of this paper is to compare and contrast the language usage of male and female politicians on Twitter during the last UK elections of December 2019 in order to raise awareness of gender inequality in politics. Our sample is formed by tweets mentioning Boris Johnson (Conservative Party), Jeremy Corbyn (Labour Party), Jo Swinson (Liberal Democrats), and Nicola Sturgeon (SNP). Methodologically, we start by obtaining the tweets that mention the politicians during the day of the elections and compiling our corpus with Sketch Engine (Kilgarriff et al., 2014). We then look at certain linguistic metrics commonly employed to characterise the sophistication of linguistic expression, and at the axiology of the messages of said candidates following Sentiment Analysis techniques. For this study we have used two lexicon classifications: Bing (Hu and Liu, 2004) and NRC (Mohammad and Turney 2010, 2013), which classifies words into eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive). We then employ programming language R and different libraries, such as tidyverse, tidytext, udpipe, and ggplot to visually represent the results. We finally supplement this analysis by a qualitative in-depth study of verbal abuse with concrete examples employing traditional discourse analysis methods. Results unequivocally show differences in the language directed to male and female politicians on Twitter. The tweets mentioning Jo Swinson and Nicola Sturgeon in our corpus are misogynist, using abusive language compared to the tweets mentioning Boris Johnson and Jeremy Corbyn. These results should not be overlooked, and solutions like Twitter bots which detect abusive language and send supporting tweets to influential women (Cuthbertson et al., 2019), must be implemented in a larger scale.

Intermarium goes viral: a corpus-driven approach to Twitter-based political communication on COVID-19 pandemic of Austrian, Croatian, Polish, and Serbian political leaders

Piotr Mirocha

Jagiellonian University, Cracow

Eligible for nomination to a student researcher award

This preliminary research is focused on political communication of Central European leaders (Austrian Chancellor S. Kurz, Croatian PM A. Plenković, Polish PM M. Morawiecki, Serbian President A. Vučić) with the use of keyword analysis and part-of-speech frequency analysis, performed on a corpus of tweets published between 2020-02-20 and 2020-04-15. References to representative concordances are used to clarify meanings of the most prominent keywords. These simple tools generate relatively comparable results, suitable for a multilingual analysis. The outcomes are then interpreted in the context of discursive representations of state's functions in the time of a crisis, having in mind often cited more or less pronounced authoritative tendencies of the governments in the respective countries. Finally, the research poses questions about new approaches to corpus-driven analysis of multilingual discourses.

Using topic modelling and collocation analysis for studying political and economical aspects of Croatian and Serbian media representations of Europe after 2007

Piotr Mirocha

Jagiellonian University, Cracow

Eligible for nomination to a student researcher award

The aim of the presentation is to summarise results obtained in a corpus-driven study of modelling of political and economical aspects in media representation of Europe in Croatian and Serbian newspapers after 2007. These aspects are understood as connotations distinct from merely locative denotation of the notion 'Europe'. In order to study this case, a large corpus of ca. 20,000 articles from ideologically diverse daily newspapers was employed.

One of the most important research questions concerns recurrent discursive structures constructing political aspects of Europe – whether topoi of Europeanisation, Eurocracy, or Europe of values are present in the corpus and what the differences between the respective countries and ideological profiles of the analysed media are. A similar question is posed with regard at economical aspects of the analysed notion. Here it is discussed whether the contrast between Europe of values and Europe of interests is relevant for some of the subcorpora and if any marked divergences emerge between countries or ideologies present in the corpus.

Methodologically, the study starts with an analysis of collocation of the lexeme 'Europe' and categorisation of the most common collocates with respect to their semantic domains, determined via a concordance analysis. This preliminary analysis is followed by an LDA topic modelling, enabling a refinement of the categorisation, drawing most typical sample texts, and observation of temporal trends.

Last but not least, contextualised interpretations are proposed, making use of existing theories of discourses on Europe and the Balkans, including the discourse on modernisation and Europeanisation, and a nexus of conservative topoi – of own country as a depository of true European values or of the 'rotten West'. Presence of pan-European discourses of Euro-enthusiasm and Euro-scepticism is also briefly discussed.

Coronavirus in Translated Media Discourse

Dr. Fahad Otaif & Dr. Waleed Alsubhi
King Khalid University / Jeddah University

Otaiff@kku.edu.sa
@Otaiff

woalsubhi@uj.edu.sa
@WaleedALSubhi

Media discourse has always been a source for misrepresentation of social actors and issues, e.g. (Fowler, 1991, Meyers, 1997, Koller, 2004; Charteris-Black, 2006; Semino, 2008, Otaif, 2015), in the current paper, we discuss the spread of the WAR metaphors in media discourse about the pandemic of Covid-19. Based on a corpus-based analysis of more than 105 Arabic news articles, we found that the spread of this metaphor is pervasive in the Arabic discourse given by a number of major international news agencies i.e. BBC, CCN and RT, yet, this use was found to adhere to certain socio-political norms which were found to be either related to a translation or discourse practice or both in some cases. The study concludes with a number of recommendations for translators and journalists as well as give a number of implications for cross-cultural discourse.

'Plotting', 'robbing', 'revolting': 'Remainers' and national identity in the pro-Brexit press.

Tamsin Parnell
University of Nottingham

tamsin.parnell@nottingham.ac.uk

@tamsinparnell

https://www.midlands4cities.ac.uk/student_profile/tamsin-parnell/

Eligible for nomination to a student researcher award

Since David Cameron's announcement in 2013 that there would be a referendum on the UK's membership of the European Union, talk about 'Brexit' has dominated British political and media discourse. Questions about what it means to be British have been at the core of these discussions, such that even regional news outlets like ITV News East Midlands have broadcast segments on public opinions of Britishness (ITV, 2020). While linguists have begun to explore national identity construction in Brexit-related political discourse (Charteris-Black, 2019; Wenzl, 2019), there is yet to be a corpus-based critical discourse analytical study of how emerging political identities (such as 'Remainers') are related to constructions of Britishness.

Using Wmatrix's semantic domain analysis, this research poster will examine representations of British national identity in a pilot corpus of 50 pro-Brexit newspaper articles (totalling 98,360 tokens) from The Sun, The Telegraph, The Daily Mail, The Express and The Daily Star. It will demonstrate the relationship between pro-EU identities and the construction of Britain as a divided nation, threatened by national "traitors" and 'rotten politics,' through the statistically key semantic fields of 'Crime', 'Damaging and Destroying' and 'Formal/Unfriendly'. Through the poster, I address the role of the "traitor" identity in disparaging a version of Britishness that involves EU membership in order to sustain public resistance to a second referendum and weaken support for the Labour party. Finally, I reflect on how my corpus approach has helped to identify discourses of national identity that otherwise might have remained hidden or overlooked, and outline the next steps of the project.

Challenges in historical corpus analyses of discourse relations: a case study in Old Church Slavonic (OCS)

Nilo Pedrazzini
University of Oxford

nilo.pedrazzini@st-hughs.ox.ac.uk
@samovar_samovar

Eligible for nomination to a student researcher award

In this paper I attempt to establish whether the discourse relation Background is able to singlehandedly define the function of absolute constructions, a type of non-finite adverbial clause attested in several historical Indo-European languages. The phenomenon is inherently indeterminate in its semantics, while its syntactic status is often considered unclear due to the occasional violation of alleged cross-linguistic constraints (e.g. subject co-reference with the matrix and overt subordinating/coordinating conjunctions).

Segmented Discourse Representation Theory (SDRT) (Asher et al. 2007) conceives Background both as a stage-setter (i.e. the context with respect to which another eventuality is to be interpreted) and as an expedient to accommodate presuppositions in the discourse structure. SDRT thus incorporates what is treated as multiple functions in other frameworks (notably, Background/Circumstance in Rhetorical Structure Theory) into a single discourse relation. This allows to easily operationalise a Background annotation in corpus research. As a case study, I analyse the dative absolute (DA) in Old Church Slavonic, exploiting the TOROT Treebank and devising a simple scheme that attributes a Background +/- tag to all DAs and to two seemingly competing constructions: conjunct participles and finite temporal subordinates. These tags are added to a subcorpus which already contains detailed annotation for several levels of linguistic analysis, allowing to check for correlations between Background and other factors.

I consider previous informal attempts at describing the DA in terms of discourse functions and test these on the annotated subcorpus. The results confirm that, while syntactic criteria fail to define it across the board, the consistent attribution of Background to the DA counteracts its semantic indeterminacy. This shows that formal analyses of discourse structure can provide fresh perspectives to the study of historical syntax and suggests that discourse annotation schemes are worth developing for historical corpora.

What counts as trolling? – A corpus-based analysis of the discourse around the communicative actions associated with perceived trolling in the comment threads of British political blogs

Marton Petyko
Aston University

m.petyko@aston.ac.uk
<https://research.aston.ac.uk/en/persons/marton-petyko>

This research paper investigates the metapragmatic discourse around trolling, a negatively marked online behaviour (Hardaker, 2013), on British political blogs. It aims to identify the communicative actions that users attribute to those they call trolls in 1,713 comment threads. These threads include 740,841 comments. The threads were published on 27 British political blogs, such as Guardian Politics blog, LabourList, and Guido Fawkes, in 2015. The paper is also concerned with how the actions associated with trolling affect the ways in which trolling is depicted and trolls are portrayed in the users' comments.

The analysis focuses on 2,144 action-related metapragmatic comments taken from these 1,713 threads. In these comments, participants call other users trolls or identify comments as trolling and also discuss the specific actions in which the alleged trolls engage. Consequently, the paper approaches 'trolling' and 'troll' as metapragmatic labels that participants use to describe, conceptualise, and evaluate others' communicative behaviour (Haugh, 2018). Using the concordance lines of the search term *troll*, the study first presents a taxonomy of the linguistically marked communicative actions in these troll action comments and then it applies this taxonomy to annotate the comments.

The paper identifies four complex communicative activities ascribed to trolls. These include spamming, ignoring or withholding information, flaming, and dishonesty, which in total cover sixteen specific communicative actions. The paper also points out that users employ action attribution as a behaviour- and identity-building device to construct trolling and trolls in various ways in their comments (Fichman & Sanfilippo, 2015). A common feature of these different constructions is that users generally depict trolling as a non-normative and manipulative behaviour while trolls are portrayed as bad debaters and uncooperative troublemakers. This suggests that users attribute actions to the trolls not only to conceptualise their behaviour but also to belittle and discredit them.

A Study of the Concept of 'Menopause' in Thai society

*Prapaipan Phingchim
Chulalongkorn University*

Research in the social sciences has seen the notion menopause as subject to a complex and variable set of representations, all ideologically loaded politicized and subject to cultural variation (Harding, 1997; Hunt, 1984; Hunter and O' Dea, 1997; Kaufert, 1982; Lupton, 1996; Wei Leng, 1996; Worcester and Whately, 1992). In Thai society, women have different ways of interpreting menopause experiences while Thai people construct meaning relating to menopause variously. This paper uses a close pragmatic analysis to examine how the term "Wai Thong" in Thai (menopause) represent the concept of 'Menopause' in Thai society. For this purpose, a quantitative analysis of a corpus that represents the concept was conduct. Since Twitter remains a salient platform of social practice, the data under scrutiny of Twitter corpora was collected between 2009 and 2019, resulting 1195 transcripts.

Preliminary result shows that concepts of 'Menopause' mostly correlate to a perspective on mental health or exacerbate a preexisting mental illness. There are several linguistic strategies representing the notions namely metaphor, implication, presupposition manipulation, presupposition denial, irony, rhetorical question and teasing. Additionally, the relationship between linguistic devices and representations of "Menopause" in Thai Twitter platform can be elaborated more by using critical discourse analysis.

The balance between reassuring and scaring: the discourse of government health advisers and politicians during the pandemic in the UK

Roberta Piazza & Ivor Gaber
University of Sussex

Covid-19 has been defined metaphorically a 'physical assailant' (Johnson) and the biggest threat to the UK since WWII. The defeat of this 'invisible mugger' (Johnson) has been presented as depending on the effort, sacrifice, 'forbearance and altruism' of the British people who have had to withstand an unprecedented reduction to their individual freedoms. We plan to investigate how, through the initial and later stages of the pandemic the most authoritative of TV networks, BBC 1, presented the politicians and experts' messages to the people of Britain balancing off prescriptions and horrendous stories of failure with an invocation to solidarity and reassurance that the NHS will cope and the country will survive. On a macro level, we will focus on human interest stories constructing a discourse of heroism and scientific information, while on the micro level we will identify the discursive strategies encoding the governmental reassuring prescriptions (e.g. metaphors, pronouns and implicatures)

Dementia in the media: a corpus and multimodal analysis of how dementia is represented by non-profits and newspapers in Britain.

Emma Putland

University of Nottingham

emma.putland@nottingham.ac.uk

@Emma_Putland

https://www.midlands4cities.ac.uk/student_profile/emma-putland/

Eligible for nomination to a student researcher award

Currently incurable, and likely to remain so (Kenisberg et al 2016), dementia is an increasingly significant health issue and one that is heavily stigmatised (Alzheimer's Disease International 2019). Attending to how dementia is discursively positioned and exploring how people respond to these discourses thereby offers a valuable means of addressing stigma to help improve life for people with dementia.

My work builds upon a growing field in sociolinguistics that examines social representations of dementia. I combine quantitative and qualitative approaches (of corpus-based discourse analysis and multimodal critical discourse analysis (MCDA) respectively) to chart how dementia is depicted in British news articles and in national non-profit organisations' blogs and news outputs between January 1 st 2017 and December 31 st 2019. I will present some preliminary findings from my two specialised corpora in this research poster, exploring how the two institutions discursively construct dementia and how these constructions relate to one another. Alongside this, I will illustrate some relevant visual and linguistic examples from a qualitative MCDA perspective and consider these in relation to the corpus findings.

I will briefly situate this analysis within the larger context of my project, which aims to bring my analysis of dementia discourses into dialogue with the perspectives of people with a direct experience of dementia (whether lived or as a carer/supporter). I do this through conducting focus groups and interviews that are grounded in visual and linguistic stimuli from non-profits and newspapers. I explore attitudes towards the two social institutions' discourses, as well as how people with direct experience position themselves in relation to current discourses for dementia. I propose that this variety in perspectives and methodology creates a project that can better speak to the complex interrelationships of discourses, social institutions and individuals, contributing further insights into how dementia is discursively positioned in British society.

Stability and change of concepts in legal speech: Corpus-based sociolinguistic analysis of the Old Bailey Voices Corpus

Justyna Robinson, Julie Weeds, Fraser Dallachy & Tim Hitchcock
University of Sussex / University of Glasgow

In this presentation we explore concepts characterising courtroom speech of different demographics in 19th-century London. We particularly focus on the analysis of conceptual variation and change across the century and on discussing methodological innovations needed to complete this task. The key challenge facing corpus-based sociolinguistic research concerns scarcity of textual metadata in corpora including socio-demographic information on authors of the corpus content. The second challenge concerns ways of operationalising concepts. Corpus analyses of concepts have primarily been based on lexical co-occurrence and lexico-grammatical (collostructional) patterns. However, these techniques are not always reliable in the light of polysemy.

In this presentation we address these challenges by carrying out a historical-sociolinguistic distribution of concepts in the Old Bailey Proceedings as represented in the Old Bailey Voices Corpus (OBVC). The OBVC is a unique database containing a wealth of sociolinguistic information on speaker demographics. Additionally, the OBVC represents a historically real, yet linguistically-controlled dataset restricted to one genre. Since the OBVC takes a consistent generic form (the trial); and since judicial speech aims at maximal transparency by minimising ambiguity, the OBVC is well suited for testing methods of concept identification, variation, and change.

In order to carry out automatic concept tagging we use the SAMUELS semantic tagger (Piao et al. 2014), which is sensitive to historical word use. Second, the distributions of concepts are modelled in terms of the co-occurring concepts and grammatical relations between concepts. We use SpaCy (Honnibal and Johnson, 2015) to annotate the text in terms of grammatical dependency relations and then extract triples which represent co-occurrences of concepts and grammatical relations. Third, we use localised positive pointwise mutual information (PPMI) to rank co-occurrences (Evert, 2005). These rankings provide characterisations of concepts (and grammatical relations) according to the corpus or sub-corpus from which they are derived.

We exemplify the use of the method by analysing variation and change of gendered use of concepts in the courtroom across a century. [For example, do men and women use different concepts in the context of a trial? How do these concepts function in relation to each other? Do these relations change in time and what the meaning of the change is in the context of developments in legal process in the 19th century?].

Health, Diaries, and Digital Technologies

Justyna Robinson & Rhys Sandow
University of Sussex

Eligible for nomination to a student researcher award

In this presentation, we showcase the potential for digital humanities to enhance corpus and discourse studies through innovative methodological and analytical techniques. Using the archive of the Mass Observation Project (MOP), hosted opposite the University of Sussex by The Keep, we built a socio-demographically tagged corpus of >1 million words pertaining to the topics of health and wellbeing. The autobiographical, diary-type, narratives of the panel of the MOP's 'observers' provide a large repository of longitudinal qualitative data. Observers periodically respond to a series of thematically related open-ended questions, or 'directives'. These responses provide semi-controlled narrative content. We use the responses to consider the cognitive and conceptual representations of a range of concepts by analysing affect and the use of metaphor.

Through an analysis of the corpus we identified areas of conceptual interest such as mental health, cancer, and death, including the theme of 'good death'. We used recent geo-mapping techniques to identify variation across region and, in addition, this can be filtered by time or by socio-demographic criteria, including age, gender, socioeconomic class, and marital status. This serves to explore the way in which wellbeing is conceptualised, cognitively, conceptually, and socially.

The meta-data enables us to study synchronic social variation as well change in the way that health and wellbeing are conceptualised. For example, we mined the data for attitudes towards euthanasia, and code the response for positive or negative affect. We were then able to divide the responses by socio-demographic profile and empirically analyse the data which reveal that young working-class panel members hold the most positive views towards euthanasia and older middle-class members, the most negative. Such findings enable insight into the thoughts, feelings, and experiences of the general British public.

Discourses of ‘suddenness’ in 55,000 early English books

Justyna Robinson, Susan Fitzmaurice & Seth Mehl
University of Sussex

This paper presents the key findings of the Linguistic DNA project (UK AHRC AH/M00614X/1) which maps ‘discursive concepts’ in Early Modern English texts, by employing computational methods alongside close reading. In this talk, we define the notion of a discursive concept, discuss methods and issues surrounding the corpus use of digitised, and share findings on discursive concepts related to the notion of suddenness.

The project investigates concepts in Early Modern English texts not as word senses, but as discursive constructs, and context-dependent meanings. In this case, meaning is linked to discourse context, beyond the level of the utterance or sentence, as well as to real-world knowledge (cf. Evans 2009). We begin with a bottom-up, data driven computational analysis of every word in every text of Early English Books Online (specifically, EEBO-TCP), which contains 55,000 printed Early English texts, and over one billion words, hand-keyed by the Text Creation Partnership. The project pre-processes this data using MorphAdorner (Burns 2013), which tokenises texts, regularises spelling, and tags words with lemma and Part of Speech information. We calculate Pointwise Mutual Information (PMI) for co-occurring lemmas across relatively large proximity windows, i.e. 100 tokens, reflecting the project’s interest in meaning constructed in discursive contexts well beyond the level of the utterance or sentence. We also calculate PMI not just between pairs of co-occurring lemmas, but also trios and quartets.

This approach allows to identify sets of strongly related lemmas. With careful interpretation, these sets of lemmas in turn suggest the emergence, shift, and decline of particular culturally important concepts within particular contexts and discourses.

We present the example of discursive concepts that relate to the lemmas sudden and suddenly, and their strong co-occurrences, a category of meaning whose vocabulary expanded dramatically in Early Modern English according to the Historical Thesaurus of English. We discuss the relationship between sudden and the discursive contexts it occurs in, identifying multiple discourses related to suddenness, and providing textual examples to connect these discourses to historical and social contexts in meaningful ways.

Risk communication and covid-19: affect and nationalist/xenophobic discourse

Katherine Russo

Università degli Studi di Napoli "L'Orientale"

During the last decade, news media have intensified their role as a channel for the communication of risk. In order to make risk decisions, citizens seek scientific information, which is circulated through online news media, and later arises in social media or face-to-face conversations. The paper aims to investigate the role of affect and ideology in the recontextualisation of scientific discourse by providing a preliminary analysis of the appraisal of covid-19 in the News on the Web Corpus (NOW). In order to communicate related risks, the news media coverage of the outbreak arguably conveyed feelings of anxiety, vulnerability and alarm, yet this study would like to investigate whether the perceived threat was related to specific affected populations. More specifically, whether it fuelled the re-irruption of nationalist and xenophobic discourses in the aftermath of the outbreak. The data will be analysed according to an approach which draws upon findings in Critical Discourse Analysis, Appraisal Linguistics, and Corpus Linguistics.

The Russian adjectives antiasadovskij, antibuševskij, antilukašenkovskij, antiputinskij, antisaddamovskij and antitrampovskij in Russian media in the beginning of the 21st Century

Thomas Samuelsson
Stockholm University

Eligible for nomination to a student researcher award

The aim of this talk is to present an investigation of Russian adjectives that are composed with the Russian prefix anti- ‘anti-’ and have roots that refer to presidents: antiasadovskij ‘anti-Assad’, antibuševskij ‘anti-Bush’, antilukašenkovskij ‘anti-Lukashenko’, antiputinskij ‘anti-Putin’, antisaddamovskij ‘anti-Saddam’ and antitrampovskij ‘anti-Trump’.

Corpus-assisted methods (CADS) are applied in order to study changes of usages of these adjectives in Russian on-line media. The corpus is compiled of Russian news articles from the beginning of the 21st century. The news outlets represent different ideologies including nationalism, communism, liberalism and state-controlled media from the Putin era 2000–2018.

The analysis reveals changes in the discourse and sheds light over the construction of opposition to the presidents in the corpus.

COVID-19 pandemic in Polish social media: a mixed methods approach to analysis of collective emotions in digital communication

Maja Sawicka & Agnieszka Karlińska
University of Warsaw

Polish society, along almost the entire world, has recently gone through an acute social crisis caused by the COVID-19 pandemic. The recent pandemic is seen as an exemplary case of a critical, threatening situation, in which individuals communicate largely through social media. In this presentation we will demonstrate a design and preliminary findings of a study oriented at capturing the emergence of collective emotions in digital communication in such circumstances, and evolution of these emotions. The study is based on a multilayer analysis of data from the Polish social media and uses a mixed methods approach. In quantitative analysis, we employ text and opinion mining, including measurements of keyword trends, time series analysis, sentiment analysis, and topic modelling. The subsequent qualitative step will cover an in-depth thematic content analysis. In the process of data analysis, we will create a tool for analysis and recognition of text sentiment.

Political discourse in the era of digital communication: a corpus-assisted discourse analysis of Italian and American populism on Twitter

Ester di Silvestro
University of Catania

ester.disilvestro@phd.unict.it
<http://www.disum.unict.it/dottorandi/ester.disilvestro>

Eligible for nomination to a student researcher award

Populism is a heterogeneous political phenomenon that is currently growing all over the world. It arose in Russia during the 19th century as a political movement that aimed to establish a rural socialism. Later the word populism has been used in the 21st century to indicate left-wing presidents in Latin America, right-wing challenger parties in Europe, and both left-wing and right-wing presidential candidates in the United States (Mudde & Kaltwasser, 2017:1). Nowadays the increase of populism is due to several reasons such as the socio-economic situation, the migration phenomenon and all the terrorist attacks that followed 9/11. Moreover, populism is strictly connected to social media such as Twitter, since they allow politicians to address the electorate directly and without the help of traditional media (Enli, 2017; Van Kessel & Castelein, 2016), and to disseminate very easily their ideologies (Kreis, 2017: 1-2). This work is a critical discourse analysis of the Italian and the American populist discourse on Twitter. Specifically, I will analyse qualitatively (Machin & Mayr, 2012) and quantitatively (Partington A., Duguid A., Taylor C., 2013) the corpora to highlight possible differences and similarities between the two political scenarios. More precisely, the analysis focuses on the linguistic strategies used by Donald J. Trump, Matteo Salvini and Luigi Di Maio in order to gain consensus during the last months of electoral campaign and the first months of government. This work investigates also the evolution of populist discourse; indeed, some traditional speeches will be analysed as well.

HUM19UK Fiction Corpus: enhancing the methodological reliability of corpus stylistic studies

*Fransina Stradling, Brian Walker, Dan McIntyre, Michael Burke, Elliott Land & Hazel Price
University of Huddersfield*

fransina.stradling2@hud.ac.uk

@f_stradling

linguisticsathuddersfield.com/hum19uk-corpus

Eligible for nomination to a student researcher award

One recurring problem for corpus stylistic research is locating a suitable norm against which to compare the language of the text or corpus under investigation. Corpus stylisticians can create their own reference corpora to fit their particular research questions, but for practical reasons often make use of ready-made, publicly available corpora.

This poster describes the construction, organisation and source material of the HUM19UK corpus, which was created to function as a ready-made corpus representative of 19th-century British prose fiction. While there are ready-made, publicly available reference corpora of 19th-century British fiction already in existence, these appear to have been created with a specific purpose in mind, and in our experience are not always ideal for projects in stylistics. Some existing corpora, such as the Corpus of English Novels (CEN), are representative of a much longer period of time (late 19th and 20th-century novels), while others, for example the CLIC 19th century reference corpus, are representative of a limited number of 19th-century authors from the Western Canon. Other ready-made reference corpora that have been used for the study of 19th-century fiction, for example the British National Corpus (BNC) and the Freiburg, Lancaster, Oslo, Bergen (FLOB) corpus, do not contain 19th-century fictional texts at all. Because of the current limited choice faced by stylisticians, it is pertinent that a reference corpus that is more representative of 19th-century fiction is available. Crucially, the HUM19UK corpus was constructed for use in a variety of projects by ensuring easy division into smaller, more specialised corpora (e.g. a corpus of 19th century fiction written by female authors). In this way, the HUM19UK corpus aims to enhance the methodological reliability of corpus stylistic studies.

A Corpus-based Analysis of Frames and Metaphors in the Media Reports of COVID-19

Shuyi Amelia Sun

The University of Queensland, Australia

shuyi.amelia.sun@uq.net.au

https://www.researchgate.net/profile/Shuyi_Sun4

Eligible for nomination to a student researcher award

Coronavirus disease 2019 (COVID-19) has now spread globally and resulted in an ongoing pandemic. This study employs corpus-based methods to analyze frames and metaphors (Lakoff, 2006), as well as their corresponding linguistic features and effects in the cross-cultural (i.e. China vs. Australia) media reports of COVID-19. By constructing and analyzing a specialized corpus which incorporates 60 reports collected from China Daily and Australian Broadcasting Corporation (ABC) News, this study aims to find out:

- (1) Based upon frames and metaphors, what linguistic features are displayed in the two countries' media reports? Are there any cross-cultural similarities and/or differences?
- (2) What roles and effects do frames and metaphors play in the online coverage of COVID-19?

Address term and of stance-taking in the discussion of Wuhan Diary in Weibo

Yingnian Tao
Lancaster University

y.tao4@lancaster.ac.uk
[@Daphne68947756](https://www.researchgate.net/profile/Yingnian_Tao2)
https://www.researchgate.net/profile/Yingnian_Tao2

Eligible for nomination to a student researcher award

On 13 May, a Chinese writer Fang Fang asked Zhang Boli, the academician of the Academy of Engineering, to apologise to her for he reprimands Fang's motive in disseminating false information in her Wuhan Diary about the national campaign against the Covid-19 pandemic during the Wuhan lockdown. The hashtag #Fangfang asked Zhang Boli to apologise went viral on Weibo (Chinese twitter) with netizens taking sides distinctively. The current project takes interest in how netizens address the two characters whilst expressing their stance (i.e. anti-Fang, pro-Fang, neutral). It explores three aspects of address terms: 1) types address terms e.g., honorifics, derogatives, parody; 2) position of address terms, viz., sentence-initial, sentence-middle and sentence-final; 3) collocations of address terms, e.g., address terms in a noun phrase. It also explores the statistical relationship between three aspects of address terms/independent variables and the netizens' stance/dependent variable by using R functions such as decision tree and multinomial logistic regression. This study aims to enrich the research of stance-taking in online forums by introducing address terms as well as statistical analysis.

Don't Interrupt Me When I am Speaking: A Multidisciplinary Approach to Interruption Force in Chinese Everyday Conversations

Yingnian Tao
Lancaster University

y.tao4@lancaster.ac.uk
@Daphne68947756
https://www.researchgate.net/profile/Yingnian_Tao2

Eligible for nomination to a student researcher award

Interruption is a common feature in everyday conversations where the hearer may “cut in” at various stages of the ongoing speech and insert his/her affiliation or disagreement with the current speaker. This paper aims to statistically derive an interruption force model from theoretically linguistic dimensions and social dimensions (e.g. gender differences, social roles). The linguistic dimensions will be extracted in pragmatics (interruption marker, speech act), cognitive linguistics (resonance) and conversation analysis (interruption timing, turn-constructive unit, interlocutor’s reaction). It uses 50 telephone conversations by native Chinese speakers in the CALLHOME Mandarin corpus. R functions will indicate the value that each dimension contributes to the final making of interruption force as well as correlations across these dimensions. Based on a 10-collection data, interruption force is evaluated in three aspects: (1) neutral interruption, out of immediate communicative needs, the interrupter asks for clarification of understanding or hearing in the current or prior utterance; (2) collaborative interruption, the interrupter helps the prior speaker/interruptee (e.g. prompting utterances which the interruptee is searching for); (3) intrusive interruption, the interrupter hurts the needs of the interruptee in various ways (e.g. disagreeing with the interruptee). Each linguistic dimension is further classified. For instance, interruption timing is divided into utterance-initial, utterance-middle and utterance-final; interruption marker is divided into backchannels like en, en, hmm, dui, right, pragmatics markers indicating contrast (bushi, no), inference (suoyi, so), elaboration (jiushishuo, that is to say), and multimodal cues like laughter. The interruption force model is expected to instruct polite utterances in social interaction where interruption is unavoidable (e.g. doctor-patient interviews). That is, the interrupter can choose among many ways to deliver interruptions less intrusive and more acceptable – using markers huh, sorry can I interrupt, cutting in at the turn completion place with less directive mood in phrases rather than a complete sentence. (words: 298)

Pandemics and metaphor: A historical perspective

Charlotte Taylor
University of Sussex

@_ctaylor_

<https://profiles.sussex.ac.uk/p329327-charlotte-taylor/grants>

The use of metaphor in framing covid-19 and reactions to it has already attracted considerable attention and both academic and public debate. We have also seen extensive discussion of how this pandemic might be compared to past events such as the so-called Spanish flu of the 1910s, Asian flu of the 1950s, Hong Kong flu of the 1960s, and, most recently SARS, Swine flu and MERS. In this short paper, I attempt to draw these two strands together by identifying the metaphorical framings of these public health threats in media and political discourse in the UK. The project will address to what extent we have continuity in the use of metaphor across these events and whether the metaphors are shared between the two domains of press and politics. This will help cast light on the ways in which current framings draw on existing discourses or show a new departure.

British and American English – A Reciprocal Influence

Franco Tondi
University of Catania

franco.tondi@unict.it

The main goal of this study is to emphasize the reciprocal influence of British English on American English, and also to delineate the differences between both varieties. Great Britain and the United States ought to be considered two nations separated by a common language influencing each other.

This survey, mostly consisting of statistical analyses, can be regarded as an in-depth study of my previous work “The Americanization of English” (Tondi 2018). The British English influence on American English is a recent trend that has been denominated the Bricitisation of English, which has also been defined as Britspeak or Anglocreep.

I intended to demonstrate the lexicographic influence of British English on American English vocabulary by undertaking lexicographic investigations and corpus analyses, as well. In order to conduct corpus analyses, I used “The Corpus of Contemporary American English”, “The British National Corpus”, “The Corpus of Historical American English”, and “The Corpus of Web- Based Global English”.

The on-line dictionaries I consulted were “Merriam Webster’s Dictionary” and “The Cambridge Dictionary”.

I also wanted to pay tribute to Ben Yagoda, an illustrious American writer and educator, professor of journalism and English at the University of Delaware. Thus, I included in the final part of this work an excerpt of his famous blog on the internet “Not one-off Britishisms” (<https://notoneoffbritishisms.com/>) in order to provide a concrete example of the Bricitization of English.

Framing advice in the Islamic Sermons Online corpus

Cipto Wardoyo
Coventry University

wardoyoc@uni.coventry.ac.uk / ciptowardoyo@uinsgd.ac.id
<https://ppsuisngd.academia.edu/ciptowardoyo>

Eligible for nomination to a student researcher award

The internet is playing an increasingly important role in expanding religious teaching, and it is used for participating in religious rituals and strengthening spiritual life (El-Nawawy and Khamis 2009). Kesvani (2019: 18) claims that YouTube has a central role in exploring and understanding the Muslim faith.

Van Noppen (2015) points out that oral ritual activities such as praying, and preaching are primary linguistic behaviours which can be analysed using speech act theory. One of the key functions of sermons is to advise the audience on appropriate behaviour; as Hashem (2010: 55) states, sermons are full of messages for purifying one's heart and calls for doing good deeds.

In order to investigate the linguistic means which global Islamic preachers on YouTube use to give advice, my first step was to collect a corpus, which I have named the Islamic Sermons Online (ISO) corpus. I chose to transcribe YouTube sermons on the topic of family delivered by 10 male and 10 female global Islamic preachers, in English. These sermons were delivered to a live audience and then recorded and uploaded. The majority of the preachers have undertaken formal education in Islamic studies.

This poster will report on the initial stage of my study, the creation of a taxonomy of realisations of advice speech acts in the sermons. I will provide some initial findings on how preachers typically frame their advice, whether it is direct or indirect, which forms are most associated with the act and whether any variation can be seen by gender at this stage.

'The more we restrict contact, the more we slow the spread of infection': Ambiguity, responsibility and political action in the UK coronavirus briefings

Jamie Williams & David Wright
Nottingham Trent University

This paper reports a work-in-progress analysis of the U.K. government's televised coronavirus briefings. Taking a corpus-assisted approach to discourse analysis, the focus here is on the strategic use of pronominal reference by government ministers delivering the updates. As it stands, a corpus of 22 briefings (totalling just over 24,000 words) has been analysed dating from 9 March to 5 April 2020. A combination of cohesion and transitivity analyses (Halliday and Hasan, 1976; Halliday and Matthiessen, 2014) have been applied to identify (possible) referents and process types attributed to them. Early results indicate that ministers use these briefings to: (i) exploit the inherent ambiguity of pronouns (c.f. Petersoo 2007) when allocating responsibility for controlling the spread of the virus, and (ii) obscure the precise nature of the interventions they have made as a means to create a sense of leadership in the face of a public health crisis.

Insight into the Opening Ceremony of the Olympic Games: A corpus-driven analysis of national image

Jincao Yu

University of Electronic Science and Technology of China

Eligible for nomination to a student researcher award

Internationally, national image of a country is understood as the representation of its positive and negative standing in media coverage of other countries. While lots of previous studies conclude that China has been, to some extent, negatively represented by western media, whether such a negative representation is targeted at China only or it happens to other less developed countries remains unclear. To access a more complete picture of national image, this corpus-driven study tries to make a comparison and contrast of the media coverage of two BRICS countries, namely P. R. China and Brazil. Two specified corpora are built based on the UK media coverage about the opening ceremonies of 2008 Beijing Olympic Games and 2016 Rio de Janeiro Olympic Games throughout the game days. Within the framework of critical discourse analysis, the study will apply the approach of Sinclair's (2004) descriptive model of lexical items and make a full analysis of the most frequent words and the keywords generated by the software Lancsbox (Brezina, V., Timperley, M., & McEnery, T., 2018) as well as and their concordances. A linguistic framework of transitivity based on SFL (Halliday, 1994) will also be applied to discover the recurrent patterns of "Beijing/Rio De Janeiro does what to whom". Hopefully, this study can identify the patterns by UK media when reporting China and Brazil respectively and explore the possible reason behind such representations.

Looking or not looking for? A corpus-assisted analysis on discriminative and homonormative language used by British and Italian men having sex with men (MSM) in personal ads in “PlanetRomeo” dating app

Federico Zaupa

University of Modena and Reggio Emilia

federico.zaupa@unimore.it

Eligible for nomination to a student researcher award

The increase in the number of location-based real-time dating (LBRTD) apps' users has raised interest within many interdisciplinary fields. Within computer-mediated communication (CMC) research area, scholars have mostly investigated on dating apps used by men having sex with men (MSM), with particular attention to the discursive construction and expression of their identity and sexual desire (Birnholtz et al. 2014; Blackwell et al. 2014). With reference to language use in MSM personal ads, attention from a diachronic perspective has been given by Baker (2005) and successively by other scholars: Callender et al. (2012) focused on racialized language reproduced in profiles of an Australian sex-seeking website designed for gay and bisexual men; Borrelli (2018) conducted a corpus-driven research investigating how masculinities are linguistically performed in personal ads within a number of straight and gay dating apps.

With the aim of providing a further contribution to this specific research topic, this corpus-assisted study examines how hegemonic masculinity and other social phenomena such as ethnical racism and the stigmatisation of older people and sex-workers are reproduced in British and Italian MSM profiles. In order to fulfil the task, profiles in “PlanetRomeo”, currently one of the most popular gay dating apps, have been randomly chosen from both small towns and big ones to build an ad hoc corpus. Using a corpus-driven approach to this data, the purpose is also to observe whether or not such discourse practices previously mentioned vary from both an intercultural perspective and within two different intracultural urban and social environments.

The Discourse of Human Rights in China's News Media

Zihuan Zhong

Queen Mary University of London

z.zhong@qmul.ac.uk

@ZhongZihuan

Eligible for nomination to a student researcher award

My research project investigates the discourse of human rights in mainland China's news media: The People's Daily (Chinese language) and China Daily (English language). I focus on how the discourse is represented in the Chinese context, how it has changed over a period of 40 years -- from December 1978 (i.e., since the policy of Reform and Opening-up) to the present day -- and how the discourse differs between the two languages in the two publications.

I am building a specialized corpus of Chinese human rights reporting from the news media, and analyzing the occurrences of words of interest, including human rights, Chinese human rights and human rights in China, using corpus linguistic tools. The data will be examined critically, looking at the lexical choice, topic focus, stance, ordering of information, and representation of voices in the news accounts. For example, the English-language stories always mention X, while the Chinese-language versions never mention Y. In addition, interviews with both Chinese and non-Chinese people who live in China of different ages, genders, educational backgrounds and nationalities will be included, aiming to explore their perceptions of and attitudes towards human rights in China and to the discourse of Chinese human rights in the media. The sociocultural issues that accompany translation (House, 2015) is also explored.