RiCL Research in Corpus Linguistics

Review of Fuster-Márquez, Miguel, José Santaemilia, Carmen Gregori-Signes and Paula Rodríguez-Abruñeiras eds. 2021. *Exploring Discourse and Ideology through Corpora*. Bern: Peter Lang. ISBN: 978-3-0343-3969-8. https://doi.org/10.3726/b17868

Carmen Maíz-Arévalo
Complutense University of Madrid / Spain

Given its undeniable complexity, discourse (and language in general) can be analysed from very different perspectives ranging from the employment of invented examples to million-word corpora. Undoubtedly, the increasing development of technology has helped in the latter direction and corpus linguistics has gained its place (and reputation) as a more reliable way to tackle this complexity. Hence, the development of *Corpus-Assisted Discourse Studies* (CADS henceforth) was a welcome and natural step, joining both corpus linguistics and critical (and non-critical) discourse analysis and promoting the synergy between automatised analyses and the fine-grained, manual work of analysts; between the 'armchair' and the 'machine' (Partington 2008). CADS is thus defined as "the set of studies into the form and/or function of language as communicative discourse which incorporate the use of computerised corpora in their analyses" (Partington *et al.* 2013: 10). Furthermore, the 'beauty' of CADS also lies in the fact that it finds a neat balance between quantification and qualitative approaches, providing a real equilibrium between the two perspectives that is far from purely cosmetic (Bryman 2017).

The current volume bears witness to the rapid expansion of the discipline, which as rightly pointed out by the editors themselves in the introductory chapter, offers "a powerful instrument of social inquiry" (p. 7). Besides the introduction, the book encompasses ten chapters, which provide readers with a welcome variety of examples,

ranging from political to gender-based studies, among others. Furthermore, this variety extends to the kind of software tools employed —e.g. *AntConc* (Anthony 2019), *UAM Corpus Tool* (O'Donnell 2012), *Lingmotif*,[1] etc.— and to the size of the corpora under scrutiny. Nonetheless, all the chapters are consistently and coherently linked together by the fact that they all focus on 'burning' social topics such as violence against women, child sexual grooming or the increasing popularity of extreme right-wing parties such as the Spanish party Vox.

In the first chapter, Alan Partington presents a fascinating account of the notion of delegitimisation and some of the most common strategies and structures employed to that purpose (e.g. the use of the prefix *post-* in terms like *post-truth* or *post-democracy*). By relating this notion of delegitimisation to the classic Aristotelian model of 'logos', 'pathos' and 'ethos', as well as to facework and positive and negative face, the author illustrates with a myriad of examples obtained from *Lexis Nexis*[2] and *The Corpus of Contemporary American English* (COCA; Davies 2008), and assisted by *WordSmith version 5* (Scott 2008), different strategies used to delegitimise individuals and groups such as undermining their 'ethos', especially in the case of female politicians. The chapter not only provides readers with a useful overview of CADS but shows how such an approach can indeed shed light not only at a micro but also at a macro level.

In a rather radical topic shift, the second chapter employs CADS to uncover the narrative of drought in the nineteenth century British media, thus also helping us understand present public attitudes to drought. To this purpose, Tony McEnery, Helen Baker and Carmen Dayrell exploit an over five-billion-word corpus retrieved from eight historical newspapers from the *British Library Newspaper*[3] collection by means of the corpus analysis system *CQPweb* (Hardie 2012), a powerful and flexible tool. The authors combine this tool with a technique known as geo-parsing, which allows to exclusively focus on texts dealing with droughts in Britain instead of somewhere else. This detailed study combines a quantitative and qualitative approach and shows that CADS can interdisciplinarily shed light onto other —even apparently unrelated— fields such as environmental science.

The next two chapters take the reader back to the field of politics. In chapter 3, Salvador Enguix-Oliver and Beatriz Gallardo-Paúls focus on journalistic discourse, more specifically on the notion of 'media populism' developed by Mazzoleni (2008), and apply it to the increasing popularity of the Spanish ultra-right party Vox. This party moved from not having any parliamentary representation to occupying a central position in Spanish politics, with 52 MPs after the November 2019 Spanish elections. To this purpose, the authors resorted to *Factiva*[4] and *Nexis*[5] to gather a corpus of 1,186 news from the most relevant Spanish written press, which was first analysed by means of the sentiment analysis software *Lingmotif*. Interestingly enough, the authors' fine-grained analysis revealed that a great deal of the (negative) evaluation present in the texts was implicitly conveyed (e.g. by means of presuppositional triggers such as factive and change-of-state verbs, but also flouting the Gricean maxims of manner and quality (Grice 1975). The results prove that disproportionate media coverage (albeit negative) has indeed helped to boost Vox' success, a tendency that seems to prevail throughout Europe (see also Ellinas 2018 and Mondon and Winter 2020, among others).

In Chapter 4, Ana Belén Cabrejas-Peñuelas and Rosana Dolón zero in on the expression of evaluation in Theresa May's three seminal Brexit speeches. Following Huston's (2000) classification of types of averral and attribution, they support their analysis with the freeware programme *UAM Corpus Tool* (O'Donnell 2012). Interestingly, a progressive tendency towards a more factual and directive speech attitude is observed from May's first to third speech. Despite the limited size of their corpus, especially in contrast to the prior chapters, the authors demonstrate that the issue of how large a corpus should be is still debatable and even a smaller corpus like theirs can render statistically significant results. It is odd, however, that the authors do not mention the myriad of articles that these same three speeches have given rise to (see Atkins and Gaffney 2020 and Marlow-Stevens and Hayton 2021, among others). This absence might be derived from their focus on evaluation rather than on other aspects like the populist ring of May's speeches (Stefanowitsch 2019) or because these papers appeared afterwards. What is clear, nonetheless, is that Brexit still draws scholarly attention from different disciplines, linguistics being one of them.

---

[4] https://www.dowjones.com/professional/factiva/
[5] https://www.nexis.com

Chapters 5 to 7 focus on gender issues, with a special emphasis on burning issues like sexual abuse, Violence Against Women (VAW henceforth) or online child sexual grooming. In Chapter 5, Leanne Victoria Bartley analyses the case of British footballers Ched Evans and Clayton McDonald, both at the forefront of one of the most controversial rape cases ever, as the former was found guilty whilst the latter was not. Interestingly, this is yet another example of how the same 'reality' can be linguistically represented in such different ways resulting in controversial accusations with life-changing consequences —i.e. Ched Evans had to serve two years. Using a CADS approach, Martin and White's (2005) Appraisal Theory and Bednarek's adjustments (2008), the author analyses the British press representations of Ched Evans and the alleged victim at three different stages for a four-year period, with a special emphasis on the system of attitude. Not surprisingly, there is a change in Evans' representation from a more negative to a more positive attitude after his retrial in 2016, with a proportional shift from a positive to a more negative light towards the alleged victim. Despite the relative predictability of her results, the study shows how the combination of CADS and a detailed qualitative analysis can solidly and reliably demonstrate expectations. As in the previous chapter, however, the reader seems to miss reference to other authors that have also greatly contributed to the study of emotion and evaluation (e.g. Alba-Juez 2018).

Chapter 6 depicts the way media outlets discursively represent female victims of VAW. By employing an impressive corpus of circa 20,000 articles gathered over a ten-year period (2005–2015) and consisting in 14.5 million words, Sergio Maruenda-Ballester contrasts these discursive representations in English and Spanish within the comprehensive and relatively recent (but blooming) framework of Discursive News Values Analysis (DNVA henceforth) developed by Bednarek and Caple in 2014 and refined in 2017. His study hence fills an under-researched gap, as VAW had not been approached from the DNVA and CADS perspectives (more specifically, Maruenda-Ballester makes use of *AntConc*). His analysis renders extremely interesting results, out of which there are two aspects that particularly draw the readers' attention. On the one hand, while the Spanish press tends to stress the victims' inner emotional suffering, the British press emphasises their emotional endurance. On the other, the author interestingly finds a preference by the Spanish dailies to stress impact by employing phrases describing signs of extreme violence. This contrasts with the UK corpus, where deceased victims are often referred to by their geographical location and/or identification. These differences

may be pointing out to cultural differences worth further research. To my view, the inclusion of Spanish is another major asset of this chapter, as most of them focus exclusively on English.

Within the related theme of Intimate Partner Violence (IPV), Alfonso Sánchez-Moya takes a different stance in Chapter 7 by analysing the online discourse of women having suffered this kind of traumatic experience and contrasting it with that of other female online users that have never experienced it. His analysis is supported by the text analysis software tool *Linguistic Inquiry and Word Count* (LIWC henceforth) developed by Pennebaker *et al.* (2007). The author convincingly shows how LIWC can help obtain reliable quantitative results in such a 'slippery' field as the analysis of emotion (reflected through the use of language). In fact, a fascinating finding is that IPV survivors tend to vary in the 'analytical' and 'tone' variables, with these users displaying a more narrative-oriented and personal discourse, in contrast to a more logical and formal hierarchical thinking patterns by the non-IPV users. Nonetheless, this is the only chapter in the whole volume where the approach is purely quantitative, and the reader misses a more qualitative perspective. The author is aware of this shortcoming himself and specifies that a qualitative analysis is envisaged as the next methodological step. However, another major asset of the chapter is that the author provides Internet scholars with valuable references and a set of guidelines on how to comply with good research practice in ethical terms.

Although all the chapters in this volume show how the study of discourse (especially from a CADS approach) can indeed provide a deeper understanding of burning social issues, this is particularly more evident in the case of Chapter 8, where Nuria Lorenzo-Dus and Anina Kinzel apply Lorenzo-Dus *et al.*'s (2016) clear model of Online Child Sexual Grooming (OCSG henceforth), and combine it with CADS to help identify these sexual abusers and thus protect such a vulnerable community as children are. The authors show a solid trajectory in the research of OSGD and, in the present chapter, zero in on the importance of implicitness (and vague language) in the communication of sexual intent within OCSG, by means of which sexual groomers may be trying to avoid being caught. By analysing an impressive corpus of circa 3.3 million words scraped from the *Perverted Justice*[6] website by means of *Python*, the authors further employed *CQPweb* to analyse their data, showing that vague expressions were

[6] http://www.perverted-justice.com/

often employed next to sexually loaded terms (e.g. *foreplay and other stuff*), hence mitigating the illegal act of engaging in sexual activity, whilst indirectly referring to it as a new category that the authors name 'Explicit-Vague'. Another major asset of this chapter is the clear definition and theoretical underpinnings of the complex concepts of 'implicitness' and 'vagueness'. Following Zhang (2013), the authors also display a comprehensive and fully operational taxonomy of linguistic realisations and pragmatic functions of vague language, all of them clearly illustrated by examples. This taxonomy, together with the new categories they identify, may indeed help inform further research not only in OSGD but also in other issues such as VAW, political discourse, and so on.

The last two chapters in the volume focus on *Twitter*. In Chapter 9, Stefania Maci analyses the narrative of the anti-vaccination campaign on *Twitter* while Alotaibi and Mulderrig focus on the *Twitter* campaign against the 'Male Guardianship' system in Saudi Arabia. Given its focus, it is relatively unclear to the reader why Chapter 10 has not been included within the group of chapters dealing with gender. A strong editorial reason might be that both chapters study *Twitter* anti- campaigns. Having said that, Maci touches upon the burning issue of conspiratorial theories against the validity of vaccines and the role played by social media (specially *Twitter*) in easily and rapidly spreading distorted information and 'fake news'. By means of semantic annotation supported by *WMatrix* (Rayson 2009) and qualitatively supplemented, the author finds out that, besides the expected semantic fields (e.g. disease, medical treatments, physiology, etc.), there were other semantic fields completely unrelated to the semantics of vaccination, such as 'families or parents', 'power' or 'dead' (and related terms like *death* or *died*). All of them contribute to spread negative ideas —often fake— about vaccines.

The volume closes with Chapter 10, where Nouf Alotaibi and Jane Mulderrig focus on the key role played by social media (especially *Twitter*) in voicing Saudi women's rights activists against the 'Male Guardianship' system, according to which Saudi women are forced to be provided written consent by a male close relative if they wish to participate in different activities ranging from enrolling in education to accessing bank services. Their work shows how Saudi women (both in favour and against this Male Guardian System) are textually (and discursively) depicted. Using *AntConc*, van Leeuwen's (2009) socio-semantic model, and Halliday and Matthiessen's (2014) well-known transitivity model, their results confirm our expectations as female users fighting against the Male Guardianship system tend to represent other Saudi women as passive

and beneficiaries, hence depicting them as weak participants. In contrast, female users in favour of Male Guardianship tend to represent their fellow women as active and capable agents. An interesting finding, however, is the common 'metaphorical' representation of Saudi women by supporters of the Male Guardianship system as 'queens' or 'pearls', following the traditional Islamic discourse in an attempt to justify the fact that women are precious and hence should be protected.

As this review has tried to show, the present volume encompasses a collection of well-written, reader-friendly papers that provide readers with an impressive collection of platforms and software tools to carry out CADS (e.g. *AntConc*, *WordSmith*, *CQP Web*, *Lingmotif*, *LIWC* or *WMatrix*), which any reader interested in discourse analysis from a mixed-method approach will indeed find extremely useful. However, I believe that the current collection of chapters will be relevant not only to those readers interested in CADS, but also in burning social issues ranging from sexual violence and sexism to climatic phenomena like droughts.

REFERENCES

Alba-Juez, Laura. 2018. Emotion and appraisal processes in language: How are they related? In María de los Ángeles Gómez González and J. Lachlan Mackenzie eds. *The Construction of Discourse as Verbal Interaction*. Amsterdam: John Benjamins, 227–250.

Anthony, Laurence. 2019. *AntConc* (version 3.5.8). Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software/antconc/

Atkins, Judi and John Gaffney. 2020. Narrative, persona and performance: The case of Theresa May 2016–2017. *The British Journal of Politics and International Relations* 22/2: 293–308.

Bednarek, Monica. 2008. *Emotion Talk across Corpora*. New York: Palgrave Macmillan.

Bednarek, Monika and Helen Caple. 2014. Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society* 25/2: 135–158.

Bednarek, Monika and Helen Caple. 2017. *The Discourse of News Values: How News Organizations Create Newsworthiness*. Oxford: Oxford University Press.

Bryman, Alan. 2017. Quantitative and qualitative research: Further reflections on their integration. In Julia Brannen ed. *Mixing Methods: Qualitative and Quantitative Research*. London: Routledge, 57–78.

Davis, Mark. 2008. *The Corpus of Contemporary American English* (COCA). https://corpus.byu.edu/coca/

Ellinas, Antonis A. 2018. *Media and the Radical Right*. Oxford: Oxford University Press.

Grice, Herbert Paul. 1975. Logic and conversation. In Peter Cole and Jerry L. Morgan eds. *Syntax and Semantics: Speech Acts*. New York: Academic Press, 41–58.

Halliday, Michael and Christian Matthiessen. 2014. *Halliday's Introduction to Functional Grammar*. London: Routledge.

Hardie, Andrew. 2012. CQPweb— Combining power, flexibility and usability in a corpus analysis tool. *International journal of Corpus Linguistics* 17/3: 380–409.

Hunston, Susan. 2000. Evaluation and the planes of discourse: Status and value in persuasive texts. In Susan Hunston and Geoff Thompson eds. *Evaluation in Text: Authorial Status and the Construction of Discourse*. Oxford: Oxford University Press, 176–207.

Lorenzo-Dus, Nuria, Cristina Izura, and Rocío Pérez-Tattam. 2016. Understanding grooming discourse in computer-mediated environments. *Discourse*, *Context & Media* 12: 40–50.

Marlow-Stevens, Samuel and Richard Hayton. 2021. A rhetorical political analysis of Theresa May's statecraft on Brexit. *Parliamentary Affairs* 74/4: 871–889.

Martin, James R. and Peter R. White. 2005. *The Language of Evaluation: Appraisal in English*. Basignstoke: Palgrave MacMillan.

Mazzoleni, Gianpietro. 2008. Populism and the media. In Daniele Albertazzi and Duncan McDonnell eds. *Twenty-first Century Populism*. London: Palgrave Macmillan, 49–64.

Mondon, Aurelien and Aaron Winter. 2020. *Reactionary Democracy: How Racism and the Populist Far Right Became Mainstream*. London: Verso Books.

O'Donnell, Michael. 2012. *UAM Corpus Tool*. http://www.corpustool.com/. (25 October 2021.)

Partington, Alan. 2008. The armchair and the machine: Corpus-Assisted Discourse Research. In Carol Taylor Torsello, Katherine Ackerley and Erik Castello eds. *Corpora for University Language Teachers*. Bern: Peter Lang, 74–95.

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

Pennebaker, James W., Roger J. Booth and Martha E. Francis. 2007. *Linguistic Inquiry and Word Count (LIWC): LIWC2007*. http://liwc.wpengine.com/(25 October 2021.)

Rayson, Paul. 2009. *Wmatrix*: *A Web-based Corpus Processing Environment*. Computing Department. Lancaster University. http://ucrel.lancs.ac.uk/wmatrix/

Scott, Michael. 2008. *WordSmith Tools version 5*. Liverpool: Lexical Analysis Software.

Stefanowitsch, Anatol. 2019. Delivering a Brexit deal to the British People: Theresa May as a reluctant populist. *Zeitschrift für Anglistik und Amerikanistik* 67/3: 231–263.

Van Leeuwen, Theo. 2009. Critical Discourse Analysis. In Jan Renkema ed. *Discourse*, *of Course: An Overview of Research in Discourse Studies*. Amsterdam: John Benjamins, 277–292.

Zhang, Grace. 2013. The impact of touchy topics on vague language use. *Journal of Asian Pacific Communication* 23/1: 87–118.

Reviewed by
Carmen Maíz-Arévalo
Complutense University of Madrid
Department of English Studies: Linguistics and Literature
Ciudad Universitaria s/n
28040. Madrid
Spain
E-mail: cmaizare@filol.ucm.es