CORPUS-ASSISTED DISCOURSE STUDIES (CADS) AND CORPUS-BASED GENRE ANALYSIS (CBGA): EXPLORING CONNECTIONS

LYNNE FLOWERDEW
BIRKBECK, UNIVERSITY OF LONDON

ABSTRACT

Corpus-assisted discourse studies (CADS) and Swalesian corpus-based genre analysis (CBGA) are two distinct discourse fields. The focus of CADS tends to be on political and media discourses and press briefings, and more recently on environmental concerns such as climate change. In contrast, the Swalesian notion of genre is generally concerned with academic and professional written genres. Over the past decade, Swales's framework of genre-based move structures has been applied to corpus research and pedagogy in an English for specific purposes (ESP) context. This paper aims to show, with reference to key work in CADS and CBGA, that these two distinct discourse fields do, in fact, share commonalities with respect to their linguistic enquiries. These are illustrated using three main linguistic insights gleaned from corpus investigations in CADS work: co-selection/co-occurrence, lexical priming and evaluative cohesion. Future work in both areas is discussed, which centres on very similar issues.

CONTACT

Lynne Flowerdew, School of Creative Arts, Culture and Communication, Birkbeck, University of London, WC1E 7HX, UK. flowerdewlynne@gmail.com

ORCID

0000-0002-1771-4048

Corpus-assisted discourse studies (CADS) and corpus-based genre analysis (CBGA): Exploring connections

Lynne Flowerdew

Birkbeck, University of London

1. Introduction

McEnery and Brezina (2023) propose a set of 48 fundamental principles of corpus linguistics, one of which is stated as follows: 'Principle 20: Corpus linguistics proceeds by convention to use methods, data sets, ontologies etc. accepted by a community of scholars as fit for the investigation of language' (p. 64). In common with the field of corpus linguistics, corpus-assisted discourse studies (CADS), a term first coined and formulated by Partington (2004, 2008) and genre analysis (Swales, 1990, 2004), of which many investigations are corpus-based, also have their respective 'community of scholars'. CADS researchers are concerned with political and media discourses such as press briefings, and more recently environmental issues such as climate change. In contrast, researchers working in corpus-based genre analysis (CBGA) focus on academic and professional genres such as lab reports and research articles with much of their work being pedagogically motivated. Given that these distinctive fields of discourse studies and genre analysis are triangulated with corpus linguistics, it is not surprising that they make use of the same methodological toolkit. For example, the main corpus linguistic tools for CADS described in Gillings et al. (2023), i.e., frequency, concordance analysis, collocation analysis, and keyword analysis, are also those that feature in CBGA. In this paper, corpus linguistics is viewed as a methodology (but see McEnery & Hardie, 2012, for a discussion on corpus-as-theory versus corpus-as-method).

However, to date, as far as I am aware, there is no study which has examined how closely these two fields are aligned in terms of the linguistic areas examined. The main aim of this paper is thus to illustrate that, despite the different types of corpora used, there is much alignment between the two fields in terms of their linguistic enquiries and methodologies employed. The following section gives a brief overview of each field and the types of studies chosen to exemplify their commonalities in terms of linguistic enquiries. In the main part of the paper, three areas are chosen for illustration, co-selection/co-occurrence, lexical priming and evaluative cohesion, which, as Partington and Marchi (2015) have underscored, are the three main linguistic insights gleaned from corpus investigations in CADS. The final section of the paper discusses how both CADS and genre analysis researchers express similar views on future developments in their respective fields.

2.   Overview of genre analysis and CADS

2.1.   *Genre analysis*

In the field of applied linguistics, Swales (1990) notes that the term 'genre' refers to 'a distinctive category of discourse of any type, spoken or written, with or without literary aspirations' (p. 33). He elaborates that genres are conceived as a goal-driven communicative event, typically employed by particular discourse communities, characterised by typical structures, stages of development and lexico-grammatical features. Swales's use of the term 'discourse' also encompasses contextual features and, in this respect, it has some affinity with two other traditions of genre, namely rhetorical genre studies (RGS) and the Sydney School systemic-functional linguistic approach. RGS is associated with genres in the fields of rhetoric and composition and privileges a situational contextual approach over a linguistic orientation (see Devitt, 2015). The Sydney School is concerned with relationships between language forms and their meanings in social contexts. Described as a 'staged, goal-oriented social process' used by cultures to carry out various functions (Martin, 2009, p. 13), it shares features with the Swalesian tradition in which genre is conceived as a goal-driven communicative event.

Swales's (1990, 2004) theory of genre has been widely applied in the context of English for specific purposes (ESP) research and pedagogy, a field which spans a wide range of academic and professional genres (e.g., thesis writing, lab reports, research articles). His well-known CARS (create-a-research-space) model for writing article introductions aptly illustrates his concept of genre and framework of analysis. According to Swales's (1990) model, article introductions progress through a series of rhetorical moves which lay the groundwork for the research to be presented. Three main moves are identified:

1.   Establishing a territory,
2.   Establishing a niche, and
3.   Occupying the niche.

These are then broken down into 'steps'. For example, Move 1, Establishing a territory, has the following three steps:

Step 1, Claiming centrality (e.g., *Recently, there has been wide interest in…*) 'are appeals to the discourse community whereby members are asked to accept that the research about to be reported is part of a lively, significant or well-established research area' (p. 144).

Step 2, Making a topic generalisation (e.g., *The aetiology and pathology of … is well known*) 'represents statements about *knowledge* or *practice*, or statements about *phenomena*' (p. 146).

Step 3, Reviewing items of previous research (e.g., *A review of the recent literature yielded only two reports of…*) is 'where the RA [research article] author needs to relate *what has been found* (or claimed) with *who has found it* (or claimed it)' (p. 148).

Over the last decade Swales's seminal work on genre analysis has inspired much corpus-based research on academic genres for pedagogic applications. One large-scale CBGA project is reported in Cotos et al. (2015) based on a corpus of 900 research articles (30

articles published in high impact journals in 30 disciplines). The researchers devised a move/step framework for all sections of the research article genre and took a rhetorical top-down perspective at the outset with manual annotation of the moves/steps. This was followed by more bottom-up lexico-grammatical analyses. For example, in Cotos et al. (2016, p. 45), prototypical linguistic realisations for the move 'Explicating results in the Discussion section' are given as follows:

(1) *The results indicate that* incumbents do indeed react pre-emptively to Southwest's entry threat. [ECON]

(2) *The validation results presented suggest that* STEMS-Air *can be applied to* both short-term and long-term modelling of PM10. [ENVE]

However, many CBGA initiatives on academic writing are much smaller in scale and make use of sub-components of either freely available academic corpora or small, specialised corpora compiled by the ESP practitioner. Depending on the search engine used, these ESP corpora can be searched by genre, discipline, or sub-section. To note is that CBGA practitioners tend to use untagged corpora for classroom exploitation for a variety of data-driven learning (DDL) activities, thereby adhering to the Sinclairian principle of letting the discourse speak for itself, as it were (Sinclair, 2004; see Flowerdew, 2023a for a review of key CBGA studies). The types of linguistic enquiries carried out using small disciplinary corpora are discussed in Section 3 of this paper.

2.2.   CADS

The terms CADS and a related discourse area, Critical Discourse Analysis (CDA), are not uncomplicated and have sometimes been conflated. Gillings et al. (2023) state that 'What all CADS projects have in common is that they have a social question at their centre rather than a purely linguistic one. That question may involve an issue such as inequality, poverty, racism, or other social ills' (p. 1). However, as Partington (2023) notes, this definition reflects rather the key tenets of Critical Discourse Analysis (CDA) 'in which "discourse" is a near synonym of social or political "narrative", and which focuses largely on studying power differentials in society' (p. 54). Other leading corpus linguists make a similar distinction. With reference to Partington et al. (2013), Hunston (2022) comments thus: 'In contrast to CDA, proponents of corpus-assisted discourse studies (CADS) explicitly distance themselves from a specific political agenda' (pp. 210–211; see Flowerdew, 2023b for a brief overview of key research in CDA and CADS).

CADS, as defined by Partington et al. (2013), is the focus of attention in this paper. In brief, the CADS approach, first formulated by Partington (2004, 2008), is conceived as building on traditional, qualitative linguistic analysis by combining statistical overview analysis involving, for example, frequency listing of words and clusters, with close reading to bring added value. Such analyses can uncover non-obvious meaning, unavailable to conscious awareness and enable the analyst to gauge how work is typically realised in a particular discourse (see Partington et al. 2013). Here, Hunston (2022) sees commonalit-

ies between CDA and CADS as both 'share the practice of identifying consistencies in how events and people are represented in news discourse, with particular attention paid to attitudes that are apparent only when a large amount of evidence is considered' (p. 111).

In Section 3, specific reference is made to the work of Partington and various CADS projects on political media discourses associated with a team of researchers at the universities of Siena and Bologna. Much of their corpus research makes use of the SiBol Corpus (see Partington, 2010), an English language newspaper corpus containing around 850 million words from 18 newspapers, spanning the years 1993–2021. While the initial version contained UK broadsheets, it was extended in 2017 and 2021 to include UK tabloids as well as newspapers from other countries such as the USA and India. Importantly, searches can be made according to a specific year, newspaper, author, or date, thus allowing for comparable investigations.

3.   Linguistic enquiries in CADS and CBGA

As previously stated, this section discusses how three types of linguistic enquiries, namely co-selection/co-occurrence, lexical priming and evaluative cohesion, are treated in CADS and CBGA to illustrate the close alignment between the two fields. In addition to the types of linguistic enquiries, another common thread linking the two discourse approaches is that, as in political and media discourses, attitudinal meanings and evaluation have great import in academic writing and have been the subject of much corpus-based work (see Bondi, 2008; Hyland, 2000, 2005, 2016). Both fields view language as socially-situated and are sensitive to contextual information (Halliday, 1994), as will be seen. While the CADS examples are from research-related work, the CBGA examples are taken from DDL classroom-based tasks as CBGA is essentially a pedagogically motivated field.

3.1.   Co-selection/co-occurrence

Partington and Marchi (2015) define the principle of co-selection/co-occurrence thus: 'The principle of co-selection or co-occurrence states that a far greater proportion of the language of most discourse types is made up, not of the accretion of individual items chosen from the mental lexicon, but of prefabricated or semi-prefabricated collections of items; "chunks" if we prefer' (p. 217).

By way of example, Partington and Marchi (2015) provide an analysis of concordances of the item *job* in a 1.3-million-word corpus of press briefings of the Obama administration from the year December 2010 to the end of November 2011, a period towards the end of the Iraq war. They note that the spokesperson, i.e., White House Press Secretary, uses *job* to praise a member of the government or service personnel. They found that of the 250 occurrences, 23 were of the form *do** an [intensifier] *job*, for example, 'Our troops are *doing an outstanding/superb/terrific job*' (p. 227). Other instances of *job* were also found in messages of praise positively evaluating a person or people (*we greatly appre-*

ciate the job they are doing; we will make sure our troops have all the tools/resources they need to get the job done/do their job; p. 227).

In common with political and media discourses, academic writing also has a significant evaluative component consisting of repeated sequences fulfilling a particular function. For instance, Diani (2012) devised corpus-based tasks to develop her linguistics students' awareness of pragmatic functions of a research paper to sensitise them to how lexical and rhetorical choices are used to express particular values. To this end, she made use of a sub-corpus of 10 linguistic papers consisting of 45,000 words from the freely available MICUSP (Michigan Corpus of Upper-Student Papers, 2009), a corpus consisting of around 2.6 million words of A-grade university student papers spanning 16 disciplines within four academic divisions (Humanities and Arts, Social Sciences, Biological and Health Sciences and Physical Sciences) and various genres (e.g., research report, proposal). One of the tasks asked students to expand the concordance lines for *argue** and to extrapolate from this 'extended unit of meaning' (Sinclair, 1996) that this verb tends to co-occur with a negative evaluation of the reported claim, for example, 'Cortazzi *argues that narratives are introduced into turn-by-turn talk …While very insightful, Cortazzi's model seems somewhat narrow*' (p. 57). This two-part structure would correspond to the 'creating a research gap' move, setting up the following rhetorical move for the researcher in which to position their own work. The positive evaluation (*While very insightful*) is overturned by a hedged negative evaluation. Hyland (2005) states that this kind of hedged evaluation acts as a face-saving device out of deference to the author's peers or humility in the case of postgraduate students, thereby emphasising the interactional nature of academic writing.

A somewhat similar kind of evaluative patterning is reported in Duguid and Partington (2017). Their analysis of the lexical item *flagship* on the gov.uk website giving 645 occurrences from 2010 to 2017 was found to be used metaphorically indicating schemes initiated by the government as inherently successful, for example, 'Academy has "new sense of purpose" after receiving new building through government's *flagship £4.4 billion rebuilding programme*' (p. 86). This repetition of key lexis and phrases 'to drive a message home' is referred to as 'forced priming' (see Section 3.2). However, a *flagship policy* in the Sibol newspaper corpus from the year 2013 is invariably negatively evaluated, for example, 'David Cameron's *flagship policy is not the panacea* for the deep problems in our education system [*Guardian* 2013]' (p. 87). The following example: 'The figures represented an improvement on results last November but the **slow rates of progress** will lead **to fresh criticism of the flagship government flagship policy** [Telegraph 2013]' (p. 87), mirrors the dialogic evaluative patterning in Diani's examples. For instance, in the example given previously (*While very insightful, Cortazzi's model seems somewhat narrow*) the negative evaluation is preceded by a positive one. However, the pragmatic force of the two examples is quite different, carefully calibrated to take into account differences in discourse types, their communicative purposes and intended audiences.

In contrast to the above example, some two-part rhetorical moves in research articles display a very tight lexico-grammatical patterning, as exemplified in the account in

Flowerdew (2015). This pedagogic initiative illustrates how a sub-corpus was used to aid postgraduate science and engineering students in writing the 'Discussion' section of their PhD theses. The corpus used was the freely-available Hong Kong Poly U *Corpus of Research Articles* (CRA) consisting of about 5.6 million words. This corpus contains 780 empirical research articles from 39 different disciplines from two broad fields: Engineering and Applied Sciences and Humanities and Social Sciences. The corpus can be searched by field, discipline or by section, i.e., Introduction, Literature review, etc. (see Lin & Evans, 2012, for a cross-disciplinary research study). The sub-section of 'Discussions', consisting of 2.3 million words, was used for consultation by students. Whilst acknowledging theses differ from research articles in terms of their purposes, scale, audience and requirements to be met, it is argued that there are significant areas of overlap in lexico-grammar and rhetorical functions and that judicious pedagogic use can be made of this corpus which is a part-related genre. To illustrate, one inductive hypothesis-testing task required students to write the beginning of a sentence showing how they might use the item *surprising* for commenting on their data. Without fail, students wrote a phrase along the lines of *This finding is surprising…; It is surprising that….* However, a follow-up search revealed that *surprising* had the following collocates, as shown in Table 1.

Left collocates for *surprising* (126 occurrences)							
not	47	is	13	hardly	9	be	7	somewhat	6
most	4	a	4	more	4	the	4	with	4
been	2	are	2	too	2	also	2	all	1
although	1	especially	1	particularly	1	passage	1	perhaps	1
quite	1	seem	1	seems	1	so	1	some	1

Table 1. *Left collocates for the 126 occurrences of surprising in the Discussions section of the Corpus of Research Articles (Flowerdew, 2015, p. 63)*

The software accompanying the CRA allows for advanced searches where various parameters can be set. The words immediately to the left of *surprising* (126 occurrences) were sorted in alphabetical order, which were then listed by the number of instances, as shown in Table 1. This tabular data provides an ideal starting point for taking a closer look at the collocational patterning, for which the most common collocate was *not*, for example, *This is not surprising*. It then suddenly dawned on students that this was entirely appropriate for the context as data commented on as 'surprising' might indicate a flaw in a project's research design, thus echoing Louw's (2007) observation that collocates are not always obvious through introspection alone but become so with '20:20 hindsight' (p. 333). This example also serves to illustrate a key tenet of the CADS approach—the uncovering, in the discourse type under investigation, of 'non-obvious meaning' (Partington et al., 2013, p. 13). Students were then asked to predict what might follow and suggested the item *because*, signalling logical reasoning. While they identified the concept, they were unable to

suggest 'given that' in this two-part move structure: *This is not surprising **given that**...*, revealing that lexico-grammar as well as meaning may be 'non-obvious'. However, students commented that they had encountered this phrase from their reading, suggesting that it is part of their receptive priming rather than productive priming repertoire (see Section 3.2). This tight lexico-grammatical patterning in this two-part move structure lends weight to Sinclair's (1991) dictum that there are preferred ways of saying things, especially in highly conventionalised genres.

One type of co-occurrence is lexical bundles, defined by Biber et al. (1999) as fixed sequences, usually of three or more words (e.g., *at the end of*), which commonly go together in natural discourse (p. 990). However, they are not a major feature of CADS research, for which the reason lies largely in Partington and Marchi's (2015) observation that 'the kind of repeated sequences we see here are very often longer and syntactically more complex than Biber et al.'s examples of prefabs (or "lexical bundles" as they term them) from conversation such as, *Can I have a...? Do you know what...?*' (p. 227).

Neither do lexical bundles feature much in CBGA initiatives despite the not inconsiderable amount of insightful research on this phenomenon in English for academic purposes and English for specific purposes contexts (see Cortes, 2023, for a review). This may well be due to the fact that most research of lexical bundles has focused on their use as building blocks at the sentence level rather than at the level of discourse. The investigation by Cortes (2013) on their occurrence in move structures of journal articles is an exception. For example, for the move structure 'Indicating a gap in the field', Cortes found there to be a strong connection with this move and lexical bundles such as *it is difficult to, it is necessary to, little is known about the, there is no* (Cortes, 2023, p. 230). Another reason for their lack of uptake, as put forward by Byrd and Coxhead (2010), could be the lack of information on their use in context. They investigated the bundle *on the basis of*, a key bundle found in many corpus-derived academic wordlists in the literature. Their finer analysis of this bundle in the 3.5 million-word corpus from which Coxhead's AWL (Academic Word List) was derived reveals it to have three possible uses (adapted from Byrd & Coxhead, 2010, pp. 53–54):

a. Used at the beginning of a sentence to provide a transition and specify methods used to carry out a process, for example, ***On the basis of** his analyses, Clyne also identifies a number of...*

b. Used as adverbial of reason in a passive sentence or clause to explain the way that a decision was made or data handled, for example, *Meanwhile, unskilled and unassisted migrants, most notably from Southern Europe were accepted **on the basis of** nomination by relatives in Australia...*

c. Used with an adverbial for strengthening or diminishing the force of an utterance, for example, *Only for L. notosaurus was the decision on its specific distinction made solely **on the basis** of allotropic data.*

Byrd and Coxhead's analysis of *on the basis of* illustrates the importance of taking into account the positioning of this lexical bundle in the sentence and its co-text for determin-

ing its meaning. Here, the authors have touched on Hoey's (2005) theory of lexical priming, the focus of the following sub-section.

The linguistic examples discussed in this sub-section show that, while different terminologies have been used, co-occurrence/co-selection in the case of CADS and phraseologies or lexico-grammar in the case of CBGA, the linguistic enquiries are remarkably similar across the two discourse types.

3.2. Hoey's theory of lexical priming

Collocations have been briefly discussed from a text-based perspective in the previous section. Hoey's (2005) theory of lexical priming also encapsulates a psychological perspective with respect to this linguistic phenomenon:

> We can only account for collocation if we assume that every word is mentally **primed** for collocational use. As a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context. (Hoey, 2005, p. 8)

Hoey also makes a distinction between productive and receptive primings:

> Productive primings occur when a word or word sequence is repeatedly encountered in discourses and genres in which we are ourselves expected (or aspire) to participate and when the speakers or writers are those whom we like or wish to emulate. Receptive primings occur when a word or word sequence is encountered in contexts in which there is no probability, or even possibility, of our ever being an active participant. (Hoey, 2005, pp. 11–12)

Whereas CADS research is concerned with receptive primings, CBGA is, in the main, concerned with productive primings. To test Hoey's theory of collocational priming. Durrant and Doherty (2010) carried out an experiment with a group of native speakers of English to investigate the extent to which corpus-identified collocations exhibit mental 'priming'. Their finding that high frequency collocations found in a large corpus are likely to have psychological reality have important implications for language teaching.

Hoey's (2005) observation that 'priming occurs, in principle, within specific domains and/or genres' (p. 19) is of relevance to CBGA pedagogy. For instance, take the example from Flowerdew (2015), *This finding is not surprising, given that....* When 'surprising' co-occurs with 'not' and is followed by 'given that' signalling logical reasoning, it is associated with the move structure 'commenting on the data' in the discussion sections of research articles and theses.

Hoey continues thus:

> The same applies to word sequences built out of these words; these too become loaded with the contexts and co-texts in which they occur. I refer to this property as **nesting**, where the product of a priming becomes itself primed in ways that do not apply to the individual words making up the combination. (Hoey, 2005, p. 8)

To illustrate that 'primings nest and combine' (p. 10), Hoey takes the word *word* which 'collocates with *say, say a word* in turn collocates with *against*, and *say a word against* collocates with *won't.*... In this way, lexical items (Sinclair, 1999, 2004) and bundles (Biber et al., 1999) are created' (p. 11). Bundles can also be involved in textual priming, as illustrated by Byrd and Coxhead's (2010) analysis of the bundle *on the basis of*, which, when used to specify methods to carry out a process tended to occur in sentence-initial position. An example of textual priming in a CBGA task is given below.

Charles (2007) used her two self-compiled corpora of successful theses (eight MPhil theses in politics/international relations and eight doctoral theses in materials science) to sensitise students to the rhetorical function of 'Defending your work against criticism'. Tasks involved students in performing controlled, context sensitive searches. In one task they were asked to retrieve *while* in sentence-initial position with the context terms *appear*/seem*/may* and identify the two-part structure – the writer first concedes the possibility of criticism and then moves to neutralise its potentially negative effect. For example, (1) Anticipated criticism: **While** *I acknowledge in some cases that the distinction* **may** *seem rather arbitrary*; (2) Writer's defence: *such political actor subjects are not the focus of interest in this thesis.*

A follow-up task asked students to take a close look at the lexico-grammar and notice the most typical patterning. In these two corpora there were no occurrences of *appear*, with *may* the preferred verb, used together with *seem* in a few instances. Students were then asked to perform a similar task to the first one, but this time instructed to look for *while* within a sentence rather than sentence-initial position. This search revealed there to be far fewer concordance lines, only two of which were used to construct a concession. Finally, students were asked to explain why the *while* clause tends to occur at the beginning of the sentence and formulate, in their own words, the following explanation: 'By ordering the clauses in this way the writer downgrades the importance of the conceded proposition and puts emphasis instead on the information in the main clause which counters the concession and provides a justification for the writer's view'(p. 294).

CBGA initiatives based on Swales's concept of move structure patterning would thus seem to chime with Hoey's theory of lexical priming, although none make reference to this explicitly. In contrast, in the CADS approach associated with the SiBol group, this theory has been taken up and extended to encompass the concept of 'forced primings', as briefly referred to previously in research by Duguid and Partington (2017).

To return to the CADS examples given in Section 3.1 on co-selection/co-occurrence, it was noted that the White House Secretary used *job* to praise a member of the government or service personnel (e.g., *Our troops are doing an outstanding/superb/terrific/job*). That *job* is primed to collocate repeatedly with a set of very positive evaluative adjectives is seen as a type of 'forced priming', a concept first introduced by Duguid (2009). Partington and Marchi (2015, p. 225) comment thus on the phenomenon of forced priming:

> … the strategic flooding by linguistic means of messages favorable to speakers or their
> clients into an ongoing discourse—and also the related phenomenon of competition
> amongst speakers to have their messages, their reading of events, accepted by either

interlocuters or an audience of beneficiaries (the party for whom the language event is taking place: Halliday 1994: 144; Partington 2003: 57-58)

Another example of forced priming in the 1.3-million-word corpus of press briefings from the Obama administration is the repeated use of *we + WORK* occurring with positive evaluating intensifiers, for example, *we are working avidly, we have worked assiduously / diligently / aggressively / very hard / every day*. However, this kind of 'semi-subliminal persuasion' (Partington & Marchi, 2015, p. 225) to 'portray and evaluate the White House and their political affiliates generally as active to the point of workaholism' may not wash with journalists and prompt a put-down rebuttal: 'Q: Why does the Congress and the President and Washington generally act like a college kid and wait until the last minute to get everything done?' (Partington & Marchi, 2015, p. 227).

A contrastive study by Marakhovskaiia and Partington (2019), discussed in Partington and Duguid (2021), of Chinese and US foreign affairs press briefings has also uncovered the use of this semi-subliminal persuasive strategy in a 495,635-word corpus of 223 press briefings (translated into English from Chinese) in 2018 of the Chinese Foreign Affairs Ministry (CMFA). Their keyword list revealed positive items signalling 'a value system built up by repeated evaluative choices' (p. 129). Values were expressed using lexis such as *stability, benefit, openness* and qualities such as *positive, constructive, fruitful, beneficial*. Qualifying evaluative adverbs found to be key included *mutually, properly, smoothly* with *manner* found to collocate with items such as *timely, responsible, orderly* and *fair*. A weaving together of these repeated positive terms into set phrases, i.e., 'forced lexical priming', projects 'China's positive national face as cooperative and constructive'. (p. 129). In contrast, their keyword analysis of a comparable corpus of press briefings from the US State Department uncovered the use of stance adverbials such as *hopefully* and *certainly* rather than qualifying evaluative adverbs.

A somewhat parallel concept to 'the strategic flooding by linguistic means of messages favorable to speakers' (Partington & Marchi, 2015, p. 225) has been expressed by Lee and Swales (2006), but in a much attenuated form. They comment that with respect to the academic writing of doctoral students, 'what may be mostly missing is fine-tuning of lexical and syntactic subtleties, particularly in terms of their strategic and rhetorical effect' (p. 57). An example from Flowerdew (2015) on the use of conditionals for signalling the move 'Limitations of the study' is a case in point. She found that the concgram (a type of concordance listing ocurrences of two words or more) 'if/would' revealed that in the Discussion sections of the CRA corpus, second conditionals instead of third conditionals tended to be used for signalling limitations. The second conditional, she suggests, is used as a means of strategically downplaying the limitations, for example, *Also our design would be more objective if the exact final outcome analyses were fully specifed, but this would be a very difficult task in the context of* (p. 64). As Zhou and Liu (2021, p. 44) point out, based on their corpus findings, this kind of rhetorical conditional quite often serves as a response to possible criticisms from the readers, as signalled by *but this would be a very difficult task in the context of* in the above example. However, follow-up consultation with students' supervisors would be helpful to gauge their acceptance of this use in thesis writing as re-

peated use of this structure may not have the desired effect and might be construed as a 'cover-up' of the limitations in the student thesis.

A more clear-cut example of 'forced priming' can be found in research by Hyland and Jiang (2021) using a corpus of 360 articles in leading journals in four disciplines at three periods over the past 50 years. In order to determine whether the intense pressure and competitiveness in research publishing has led authors to rhetorically 'sell' or 'hype' their studies, Hyland and Jiang traced the use of 400 'hyping' words 'which seek to promote, embellish or exaggerate aspects of research papers' (p. 189), such as those in bold in the following examples: 'Our **striking** results **demonstrate** that longitudinal scaling of all long bones is **clearly** isometric throughout elongation (Biology); Drawing on an **exceptionally high-quality** longitudinal data source, this article provides **important contributions** to understanding variation in family behaviour (Sociology)' (p. 190). Their results showed that this kind of hyperbolic language had increased by 19% over the years and was especially marked in the hard sciences in certain broad functional categories such as author's prior research and research outcome. While not using Swales's framework of move structures, their categories have some overlap; for example, 'author's prior research' would correspond to Step 3 'Reviewing items of previous research' in Move 1 'Establishing a territory' in the CARS model (see Section 2.1). Like Partington and Marchi (2015), they caution that such hypes may backfire and have an undesired effect of undermining arguments.

The concept of lexical priming and especially the extended notion of 'forced priming' has usefully informed much work in CADS. While the term 'priming' is not used in CBGA studies, the above discussion has shown that it is very much present in the DDL work of Charles (2007) and Flowerdew (2015), although not explicitly referred to. The corpus study by Hyland and Jiang on research articles (2021) shows that 'forced priming' is also of relevance for corpus analysis of academic discourse, although, again, this term is not used in the study.

3.3. *Evaluative cohesion*

Alessi and Partington (2020) describe two interrelated systems of cohesion ('sticking together'), namely entity/propositional cohesion, i.e., standard and evaluative cohesion. They point out that standard cohesion, which has been much more extensively analysed than evaluative cohesion, 'can tell us a great deal about *how* a text is rendered coherent' (p. 1). On the other hand, the study of evaluative cohesion, sometimes referred to as evaluative harmony (Partington, 2017), sheds light on the communicative intents of the speaker or writer, i.e., '*why* it is they wish to communicate what they do' (p. 1). They point out that evaluative meaning often spreads across grammatical boundaries, suggesting that 'texts cohere evaluatively as well as propositionally' (p. 9). In the example below from the SiBol corpus (Alessi & Partington, 2020, p. 9), a negative impression of George Bush is built up through a chain of negatively evaluated items, indicated by square brackets: '*George Bush is talking again and [I don't have a clue what he's saying]. It's not that [he's mangling his syntax]. That's [par for the course]* (SiBol 05)'.

Evaluative cohesion has also been studied in short complete texts, as illustrated by Partington (2017, p. 197) of the publisher's description of the celebrated work, *The Bottom Billion* (Collier, 2007).

> (Global poverty is falling rapidly); *but* [in fifty or so failing states the world's poorest people—the 'bottom billion'—face a tragedy that is growing inexorably worse. Why do these states defy all attempts to help them? Why does current aid seem unable to make a difference?]

> (In his award-winning bestseller, Paul Collier pinpoints the issues of corruption, political instability, and resource management that lie at the root of the problem, and offers hard-nosed solutions and real hope for a way of solving one of the great crises facing the world today.)

Partington (2017, pp. 196–197) deconstructs the above text to illustrate how evaluative cohesion is realised, using round brackets for positive entities and square ones to indicate negative ones. He argues that 'writers and speakers employ—and readers and listeners in turn recognize and employ—the two processes of firstly, *evaluative embedding* and secondly *contagion*, in order to maintain what we have called consistency or 'harmony' of evaluation over long segments of text' (p. 196). In brief, contagion refers to items which acquire evaluation in context and evaluative embedding is illustrated with reference to text segments. Partington annotates the first sentence as follows: ([Global poverty] is falling rapidly), maintaining that it is the outer level determining the overall polarity. This positive evaluation is then reversed by the adversative *but* signalling a chain of negative elements. The second paragraph contains highly negatively evaluated items, but these are all subsumed within positively evaluated outer levels: (pinpoints [the issues of corruption, political instability]) and (solving [one of the great crises]). The SiBol corpus was consulted for the item *pinpoint*, indicating that it 'has a positive semantic prosody, strongly associated with uncovering remedies and solutions' (p. 197), and would most probably be interpreted as such by experienced readers. Importantly, Partington makes the point that competent readers would not read this text word by word, but evaluative block by block, processing the first paragraph as two evaluative units, i.e., a short opening positive block followed by a single negative one, signalling *problem*, and the second paragraph as a positively evaluated block, i.e., *solution* (Hoey, 1983).

Likewise, Hoey's Problem-Solution pattern, in an elaborated form, can be mapped onto the rhetorical move structures in CBGA tasks, as in the following grant proposal abstract analysed in Flowerdew (2016, p. 7): Situation ('research' and 'real world' territory), Problem (niche), Partial Solution (part filling of niche), Solution (proposed research) and Evaluation (promising outcome). See Table 2 below.

Title: The development of a *new* plasmid partition system to effect *high-efficiency* stabilisation of recombinant plasmids in various bacterial strains used as hosts in biotechnological fermentations.

Text	Move structure
Some plasmids express special functions which contribute to their stability in bacterial cells. These functions are of *scientific interest.*	'Research' territory
Also, they are of *practical value*, as many *commercial* recombinant products are expressed from plasmids, borne in bacteria growing in fermentation culture.	'Real world' territory + Benefits
Problems with plasmid stability *may be encountered* during the many generations of growth required for attainment of satisfactory cell densities.	Niche
The best methods of improving plasmid stability involve the use of the natural plasmid stability systems (par functions) to increase plasmid stability. *Existing par systems increase plasmid stability to some extent*, but *the best system has been described by us.*	Part filling of niche (Toning down of negative evaluation + foregrounding of proposers' previous research)
In the proposed study, this system will be exploited to provide **a routine and simple means** of stabilisation of recombinant plasmids….	Proposed research + Promising outcome

Table 2. *Move structure analysis of a grant proposal abstract (Flowerdew, 2016, p. 7)*

From a pedagogic perspective, students, who had already been introduced to the Swalesian concept of move structure analysis in a previous course on thesis writing, were given free rein to carry out a 'quick and dirty' move structure analysis (Swales, 1990). Some students took the initiative to map Hoey's Problem-Solution pattern onto the text. Students were thus encouraged to decipher the text in evaluative blocks, in line with Partington's (2017) observation of the processes competent readers use to read text. Follow-up language work focused on propositional cohesion such as the use of adversatives (see Alessi & Partington 2020), patterns of lexis across the title and text involving the items *stability* and *system* (Hoey, 1991) and evaluation. For example, students were asked to explain why the negative evaluation was couched indirectly (*Existing par systems increase plasmid stability **to some extent***), most likely as a face-saving device and out of respect to the professor's peers. The MICUSP (see Section 3.1) was used as a source for answering students' self-initiated queries such as more information on the use of the bundle *to some extent*, here an example of contagion as the bundle is not inherently negative but acquires a negative prosody in this context. Parallels can thus be drawn between the move structure analysis of a grant proposal abstract and the construction of evaluative cohesion in a publisher's book description, also informed by Hoey's Problem-Solution pattern.

Both analyses use the corpus as a standby resource for verification of specific queries, thereby emphasising the corpus-*assisted* nature of the CADS approach.

This section has illustrated that the two distinct discourse approaches of CADS and CBGA share commonalities in terms of the linguistic insights gleaned from their corpus investigations. The main linguistic insights from CADS research involve co-occurrence/co-selection, lexical priming and evaluative cohesion. It has been shown that these three areas are also prominent in CBGA initiatives, although the analyses are couched using different terminologies.

4.   What are the futures of genre analyses and CADS?

Proponents of both CADS and genre analysis, (and by extension CBGA), express similar views on future work, putting greater attention on contextual features. In the case of journalistic discourse Marchi (2022) advocates going beyond the message to focus on institutional and journalistic practices by examining the rules that journalists and editors abide by in the production of the message. In a similar vein, Swales (2019) has suggested that genre studies tend to be 'too textual' and advocates taking as much of an insider 'emic' approach as possible by interviewing scholars to gain an insight into their disciplinary and discoursal practices. Both, not surprisingly, consider that future studies will also encompass more exploration of how digital and multi-modal genres are utilised. Interdisciplinarity is another aspect singled out for future investigations. Swales suggests more attention be paid to emerging interdisciplinary fields (e.g., biostatistics and museum studies). Marchi also views interdisciplinarity as a burgeoning area but here it is used in a somewhat different sense (see Ancarno, 2018, for a case study on collaboration between an anthropologist and a CADS researcher grounded in anthropological work and methods.) On a final note, in his personal essay, Swales (2019) expresses the following: 'I suggest that more attention could usefully be given to a) syntactic and phraseological patterns and uses, and b) to local cohesive elements that will increase the "flow" of student texts' (p. 81), aspects taking centre stage in CADS work associated with the SiBol group. Contrastive studies investigating cross-cultural/cross-linguistic aspects of discourse is an under-researched area at present. The CADS research by Marakhovskaiia and Partington (2019) of Chinese and US foreign affairs press briefings briefly discussed in Section 3.2 is an exception. In their methodology paper, Moreno and Swales (2018) have commented on the lack of multilingual corpora for ESP genre-based research and pedagogy. Bondi and Yu's (2019) cross-cultural study on textual voices in corporate reporting and Yu's (2022) CBGA study investigating Chinese, Italian and English CSR (corporate social responsibility) reports, together with a suite of pedagogic applications, are exceptions in the field of ESP.

Both discourse analysis researchers and DDL practitioners have considered the impact of ChatGPT on future corpus work, which clearly has implications for CADS and CBGA. In brief, ChatGPT, launched in November 2022, is a deep learning model that is trained on a vast corpus of text data and generates responses through prediction. With

respect to corpus approaches to discourse analysis, Curry et al. (2024) present three replication studies of previously published work designed to investigate the affordances of using ChatGPT for conducting automated qualitative analysis. In two of their cases studies, an analysis of concordance lines and identification and analysis of direct and indirect questions, which rely on context, ChatGPT was found to perform poorly. Meanwhile, Crosthwaite and Baisa (2023) explore the potential advantages and disadvantages of GenAI (generative artificial intelligence) vis-à-vis the use of corpora for DDL. They note that 'One of the main advantages of corpora is that we know exactly the domain of texts from which the corpus is derived, something that we cannot track from current large language models underlying applications like ChatGPT' (p. 1). They also note that collocations can be difficult for GenAI to handle and question the authenticity of the language generated, which may not be contextually or register appropriate (see also Lin, 2023). These would seem to be significant drawbacks for the type of CADS research and CBGA activities examined in this article. Both approaches are context-sensitive and rely on carefully compiled specialised corpora which can be searched according to a specific year, newspaper, author or date in the case of CADS research using the SiBol corpus for comparative studies, or, in the case of genre-based enquiries, by discipline, genre, sub-section or move structure to investigate genre and disciplinary differences.

To conclude, in an early paper Partington (2004) draws attention to the aims of the initial call for papers for the first CamConf 2002, held at the University of Camerino, Italy: "[…] to examine how it is possible to use concordance technology and the detailed linguistic evidence available in corpora to enhance the study of the discourse features of particular genres of the language and of the communicative strategies used by speakers and writers to pursue their designs" (p. 12).

Chapters in the ensuing conference volume covered discourse organization, evaluation, pragmatics, rhetoric and the 'real world', and lexical priming, topics still of great relevance today. This synthesis of corpus linguistics and discourse studies was formulated as CADS, 'a most congruous beast', in Partington's words. Two decades later this 'most congruous beast' is alive and kicking. Partington (2023) comments thus: 'In reality CADS is not confined by topic or the stance or school of the researcher/s, whether it be ontological and descriptive, or normative and political. CADS is an *omnivore*' (p. 55). I hope this piece has shown, through the discussion of three key linguistic insights gleaned from CADS research and investigated in CBGA initiatives, that CADS is indeed an *omnivore* and can bring insights and 'added value' to another discourse field.

Acknowledgments

Competing interests

The author has no competing interests to declare.

References

Alessi, G. M., & Partington, A. (2020). Two ways of sticking together and getting along: Propositional cohesion and evaluative cohesion in discourse. Italy: Mattioli 1885. Retrieved from https://iris.unimore.it/handle/11380/1203509.

Ancarno, C. (2018). Interdisciplinary approaches in corpus linguistics and CADS. In C. Taylor & A. Marchi (Eds.), *Corpus approaches to discourse: A critical review* (pp. 130–56). London: Routledge. https://doi.org/10.4324/9781315179346-7

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Essex, UK: Pearson Education.

Bondi, M. (2008). Emphatics in academic discourse: Integrating corpus and discourse tools in the study of cross-disciplinary variation. In A. Ädel & R. Reppen (Eds.), *Corpora and discourse: The challenges of different settings* (pp. 31–55). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.31.04bon

Bondi, M., & Yu, D. (2019). Textual voices in corporate reporting: A cross-cultural analysis of Chinese, Italian, and American CSR reports. *International Journal of Business Communication, 56*(2), 173–197. https://doi.org/10.1177/2329488418784690

Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney papers in TESOL, 5*, 31–64. Retrieved from https://www.sydney.edu.au/content/dam/corporate/documents/faculty-of-arts-and-social-sciences/research/research-centres-institutes-groups/uos-papers-in-tesol/volume-5/article02.pdf

Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes, 6*(4), 289–302. https://doi.org/10.1016/j.jeap.2007.09.009

Collier, P. (2007). *The bottom billion*. Oxford: Oxford University Press.

Corpus of Research Articles (CRA) 2007. Research Centre for Professional Communication in English, Hong Kong Polytechnic University. Retrieved from http://rcpce.engl.polyu.edu.hk/RACorpus/default.htm

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes, 12*(1), 33–43. https://doi.org/10.1016/j.jeap.2012.11.002

Cortes, V. (2023). Lexical bundles in EAP. In R. R. Jablonkai & E. Csomay (Eds.), *The Routledge handbook of corpora and English language teaching and learning* (pp. 220–233). London: Routledge. https://doi.org/10.4324/9781003002901

Cotos, E., Huffman, S., & Link, S. (2015). Furthering and applying move/step constructs: Technology-driven marshalling of Swalesian genre theory for EAP pedagogy. *Journal of English for Academic Purposes, 19*, 52–72. https://doi.org/10.1016/j.jeap.2015.05.004

Cotos, E., Link, S., & Huffman, S. (2016). Studying disciplinary corpora to teach the craft of Discussion. *Writing & Pedagogy, 8*(1), 33–64. https://doi.org/10.1558/wap.v8i1.27661

Crosthwaite, P., & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning? Not so fast! *Applied Corpus Linguistics, 3*(3), 1–4. https://doi.org/10.1016/j.acorp.2023.100066

Curry, N., Baker, P., & Brookes, G. (2024). Generative AI for corpus approaches to discourse studies: A critical evaluation of ChatGPT. *Applied Corpus Linguistics, 4*(1), 1–9. https://doi.org/10.1016/j.acorp.2023.100082

Devitt, A. J. (2015). Genre performances: John Swales' Genre Analysis and rhetorical-linguistic genre studies. *Journal of English for Academic Purposes, 19*, 44–51. https://doi.org/10.1016/j.jeap.2015.05.008

Diani, G. (2012). Text and corpus work, EAP writing and language learners. In R. Tang (Ed.), *Academic writing in a second or foreign language: Issues and challenges facing ESL/EFL academic writers in higher education contexts* (pp. 45–66). London: Continuum.

Duguid, A. (2009). Insistent voices: Government messages. In J. Morley & P. Bayley (Eds.), *Corpus-assisted discourse studies on the Iraq conflict: Wording the war* (pp. 234–260). London: Routledge.

Duguid, A., & Partington, A. (2017). Forced lexical primings in transdiscoursive political messaging: How they are produced and how they are received. In M. Pace-Sigge & K. J. Patterson (Eds.), *Lexical priming: Applications and advances* (pp. 67–92). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.79.03dug

Durrant, P., & Doherty, A. (2010). Are high frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory, 6*(2), 125–155. https://doi.org/10.1515/cllt.2010.006

Flowerdew, L. (2015). Using corpus-based research and online academic corpora to inform writing of the discussion section of a thesis. *Journal of English for Academic Purposes, 20*, 58–68. https://doi.org/10.1016/j.jeap.2015.06.001

Flowerdew, L. (2016). A genre-inspired and lexico-grammatical approach for helping postgraduate students craft research grant proposals. *English for Specific Purposes, 42*, 1–12. https://doi.org/10.1016/j.esp.2015.10.001

Flowerdew, L. (2023a). Corpora for EAP writing. In R. R. Jablonkai & E. Csomay (Eds.), *The Routledge handbook of corpora and English language teaching and learning* (pp. 234–247). London: Routledge. https://doi.org/10.4324/9781003002901-19

Flowerdew, L. (2023b). Corpus-based discourse analysis. In M. Handford & J. P. Gee (Eds.), *The Routledge handbook of discourse analysis* (2nd ed., pp. 126–138). London: Routledge. https://doi.org/10.4324/9781003035244-11

Gillings, M., Mautner, G., & Baker, P. (2023). *Corpus-assisted discourse studies*. Cambridge Elements. Cambridge University Press. https://doi.org/10.1017/9781009168144

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.

Hoey, M. (1983). *On the surface of discourse*. London: Allen and Unwin.

Hoey, M. (1991). *Patterns of lexis in text*. Oxford: Oxford University Press.

Hoey, M. (2005). *Lexical priming. A new theory of words and language*. London: Routledge. https://doi.org/10.4324/9780203327630

Hunston, S. (2022). *Corpora in applied linguistics* (2nd ed.). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108616218

Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press. https://doi.org/10.3998/mpub.6719

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173–192. https://doi.org/10.1177/1461445605050365

Hyland, K. (2016). Writing with attitude: Conveying a stance in academic texts. In E. Hinkel (Ed.), *Teaching English grammar to speakers of other languages* (pp. 246–265). London: Routledge.

Hyland, K., & Jiang, F. K. (2021). 'Our striking results demonstrate…': Persuasion and the growth of academic hype. *Journal of Pragmatics, 182*, 189–202. https://doi.org/10.1016/j.pragma.2021.06.018

Lee, D., & Swales, J. (2006). A corpus-based ESP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes, 25*(1), 56–75. https://doi.org/10.1016/j.esp.2005.02.010

Lin, L., & Evans, S. (2012). Structural patterns in empirical research articles: A cross-disciplinary study. *English for Specific Purposes, 31*(3), 150–160. https://doi.org/10.1016/j.esp.2011.10.002

Lin, P. (2023). ChatGPT: Friend or foe (to corpus linguists)? *Applied Corpus Linguistics, 3*(3), 1–5. https://doi.org/10.1016/j.acorp.2023.100065

Louw, B. (2007). Truth, literary worlds and devices as collocation. In E. Hildalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 329–362). Amsterdam: Rodopi. https://doi.org/10.1163/9789401203906_021

Marakhovskaiia, M., & Partington, A. (2019). National face and facework in China's foreign policy: A corpus-assisted case study of Chinese foreign affairs press conferences. *Bandung: Journal of the Global South, 6*(1), 105–131. https://doi.org/10.1163/21983534-00601004

Marchi, A. (2022). Corpus linguistics in the study of news media. In A. O'Keeffe & M. J. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (2nd ed., pp. 576–588). London: Routledge. https://doi.org/10.4324/9780367076399-40

Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education, 20*(1), 10–21. https://doi.org/10.1016/j.linged.2009.01.003

McEnery, T., & Brezina, V. (2023). *Fundamental principles of corpus linguistics.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781107110625

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511981395

Michigan Corpus of Upper-level Student Papers. (2009). Ann Arbor, MI: The Regents of the University of Michigan. Retrieved from http://micusp.elicorpora.info/

Moreno, A. I., & Swales, J. M. (2018). Strengthening move analysis methodology towards bridging the function-form gap. *English for Specific Purposes, 50*, 40–63. https://doi.org/10.1016/j.esp.2017.11.006

Partington, A. (2003). *The linguistics of political argument: The spin-doctor and the wolf-pack at the White House.* London: Routledge. https://doi.org/10.4324/9780203218259

Partington, A. (2004). Corpora and discourse: A most congruous beast. In A. Partington, J. Morley, & L. Haarman (Eds.), *Corpora and discourse* (pp. 11–20). Bern: Peter Lang.

Partington, A. (2008). The armchair and the machine: Corpus-assisted discourse research. In C. Taylor Torsello, K. Ackerley, & E. Castello (Eds.), *Corpora for university language teachers* (pp. 189–213). Bern: Peter Lang.

Partington, A. (2010). Modern diachronic corpus-assisted discourse studies (MD-CADS) on UK newspapers: An overview of the project. *Corpora, 5*(2), 83–108. https://doi.org/10.3366/cor.2010.0101

Partington, A. (2017). Evaluative clash, evaluative cohesion and how we actually read evaluation in texts. *Journal of Pragmatics, 117*, 190–203. https://doi.org/10.1016/j.pragma.2017.06.008

Partington, A. (2023). Book review: Gillings, M., Mautner, G. & Baker, P. (2023) *Corpus-Assisted Discourse Studies*. Cambridge Elements. *Journal of Corpora and Discourse Studies, 6*(1), 53–60. https://doi.org/10.18573/jcads.109

Partington, A., & Duguid, A. (2021). Political media discourses. In E. Friginal & J. A. Hardy (Eds.), *The Routledge handbook of corpus approaches to discourse analysis* (pp. 116–135). London: Routledge. https://doi.org/10.4324/9780429259982-8

Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and meanings in discourse: Theory and practice in corpus-assisted discourse studies (CADS)*. Amsterdam: John Benjamins.

Partington, A., & Marchi, A. (2015). Using corpora in discourse analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 216–234). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.013

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. M. (1996). The search for units of meaning. *Textus 9*(1), 75–106.

Sinclair, J. M. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge. https://doi.org/10.4324/9780203594070

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524827

Swales, J. M. (2019). The futures of EAP genre studies: A personal viewpoint. *Journal of English for Academic Purposes, 38*, 75–82. https://doi.org/10.1016/j.jeap.2019.01.003

Yu, D. (2022). *Cross-cultural genre analysis: Investigating Chinese, Italian and English CSR Reports*. London: Routledge. https://doi.org/10.4324/9780429353154

Zhou, H., & Liu, Y. (2021). If it had been conducted with a larger database…: A comparison of If-constructions in Chinese L2 learners' theses and published research articles. *English for Specific Purposes, 64*, 37–48. https://doi.org/10.1016/j.esp.2021.07.005