Application of Swalesian Genre Analysis to Academic Writing Pedagogy: A Corpus Perspective

Lynne Flowerdew[a*]

[a] flowerdewlynne@gmail.com, Department of Applied Linguistics and Communication, Birkbeck, University of London, UK
[*] Corresponding author, flowerdewlynne@gmail.com

Abstract

This short article reviews key corpus-based pedagogic initiatives in the spirit of the Swalesian tradition of genre analysis. Pedagogic genres covered include report writing, thesis writing, writing a grant proposal and legal essay writing. More recently, attention has been paid to the writing of research articles by postgraduate students for whom published articles are becoming a prerequisite for a PhD to be awarded. These pedagogic endeavours invariably commence from a top-down perspective. Genre moves are first identified followed by more bottom-up corpus searches for the identification of prototypical lexico-grammatical patterning for specific move structures. In a few cases, relevant sub-sections of freely available corpora, such as the Michigan Corpus of Upper-level Student Papers (MICUSP), are used. However, most pedagogic initiatives make use of self-compiled specialised corpora to meet the needs of students from a specific discipline. Genre-based research findings can also usefully inform classroom-based corpus activities.

Introduction

It goes without saying that the Swalesian approach to genre has had a great impact on EAP/ESP research and pedagogy over the past thirty years (see Swales, 1990). Also, there is now a substantial literature on applying genre-based, rhetorical move structure analysis to corpora of written texts, most notably the research article (RA) genre (see Samraj, 2016 for a review). In this respect, Kanoksilapatham's (2005, 2007) corpus research of rhetorical moves in 60 biochemistry RAs is of note as it was one of the first studies to undertake coding and analysis of the moves in all the four sections of the RA (introduction, methods, results, discussion) using a discipline-specific corpus. As pointed out by Kanoksilapatham (2007), the main advantage of using corpus linguistic techniques is that they can uncover the prototypical lexico-grammatical patterning associated with specific moves, an aspect missing from early genre studies where the focus tended to be on identification of rhetorical moves rather than linguistic devices realising the moves.

Research has also been carried out on pedagogic genres (e.g., essay writing) using sub-components of freely available corpora of highly-graded student writing such as the British Academic Written English (BAWE) corpus and the Michigan Corpus of Upper-level Student Papers (MICUSP). For example, Durrant and Mathews-Aydınlı (2011) examined formulaic language in various move structures in 96 academic essays using the essay sub-component of the BAWE corpus. They first annotated the essays for communicative functions and then identified recurrent patterns associated with each function. Three main generic functions were identified in the essay introductions: *Background information*, *justifying research*, and *essay focus*. Each move was then sub-divided into several steps. For example, the *essay focus* move describing the content of the paper was sub-divided into the following five steps: *Stating focus*, *stating limitations*, *indicating structure*, *stating approaches*, and *stating the argument*. Another sub-corpus of the BAWE corpus, namely 60 laboratory reports, features in research by Parkinson (2017) who identified key moves and their typical lexical features.

While research studies combining a Swalesian approach with corpus analysis have tended to dominate the field of academic writing, it is only in recent years there have been more reports in the literature on the pedagogic applications of such an approach. The main aim of this

article is to take stock of such initiatives, which have employed either freely available corpora or self-compiled corpora (While freely available corpora come with their own search engines requiring minimal preparation, the compilation of small, specialised corpora is not an insurmountable task for the class tutor; see Anthony, 2019). To note is that the starting point for such pedagogic applications tends to be with top-down, initial pen-and-paper rhetorical awareness-raising activities followed by more bottom-up, direct 'hands-on' corpus consultation of language features. The most interesting and innovative aspects of the reports will be highlighted in the following two sections.

Using Freely Available Corpora

One account on using a freely available corpus of research articles (CRA) together with MICUSP is given in Flowerdew (2015). Flowerdew made use of the discussion sections of the CRA for a workshop for postgraduate science and engineering students on writing up the Discussion section of a thesis. To note is that she did not use an 'exemplar' corpus of theses due to unavailability but an 'analogue' corpus of research articles, a corpus as close as possible in terms of genre, discipline etc. to the student disciplinary writing (Tribble, 2002). The focus was on meta-disciplinary writing rather than subject-specific instruction, mainly because the students were not a homogeneous group from the same discipline (see also Flowerdew, 2016). The corpus tasks were aimed at honing advanced students' mastery of lexico-grammatical patterns for core move structures. Many of the tasks were hypothesis testing ones as these were found to be the most successful and generated the most discussion. For instance, students were asked to write a phrase for the move structure 'showing comparison with previous work' beginning with *This finding*…'. Some students were unsure about tense use and which preposition to use after 'consistent'. Students then compared their suggestions with the concordance output in Table 1 below. On scrutiny of the concordance lines one student asked whether it was possible to say: 'This finding is compatible with…' (which did not appear in the corpus data). Students were reminded to consult other corpora such as MICUSP and found that this phrase, while seemingly synonymous with 'This finding is consistent with', was used in a different context.

Table 1

Selected Concordance Output for 'This finding' from the CRA

1	This finding	is consistent with Meyer, Becker, and Vandehbergen
2	This finding	is consistent with observations in the theoretical
3	This finding	is consistent with previous work concerned
4	This finding	is in agreement with those of previous studies
5	This finding	is in keeping with numerous other studies
6	This finding	is similar to prior findings that federal relief
7	This finding	is supported by research conducted by
8	This finding	has been supported by
9	This finding	has not been previously established.

Note. From L. Flowerdew (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*, 62.

An initiative using MICUSP for students majoring in Linguistics is reported in Diani (2012). Diani used a sub-corpus of linguistic papers from MICUSP to help develop her students' awareness and understanding of the pragmatic functions of a research paper. This approach entailed reading beyond the truncated concordance lines. In order to sensitise students to how rhetorical move structures and lexical choices are used to express particular values, students were encouraged to expand the concordance lines for the verb *argue* and extrapolate from this 'extended unit of meaning' that the verb *argue* tends to be used for creating a negative evaluation of the reported claim, e.g., *Cortazzi argues that* narratives are introduced into turn-by-turn talk… *While very insightful, Cortazzi's model seems somewhat narrow*… (Diani, 2012, p.57). While Diani used a ready-made corpus of linguistic papers in her disciplinary writing course, most reports on subject-specific writing make use of corpora compiled either by the class tutor or in some cases the students, the focus of the following section.

Using Self-Compiled Corpora

An early corpus initiative, specifically legal essay writing by undergraduates, is reported in Weber (2001). First, Weber's students were inducted into the genre of legal essays by reading through whole essays

taken from the University of London LLB Examinations written by native speakers, and identifying some of the prototypical rhetorical features, e.g., identifying and/or delimiting the legal principle involved in the case. They were then asked to identify any lexical expressions which seemed to correlate with the genre features. This was followed up by consulting a corpus of legal essays compiled by the class tutor to verify and pinpoint regularities in lexico-grammatical expressions. For example, items such as *assume, consider*, and *regard* in various constructions, were all found to act as signals in an opening-type move, delimiting the case under consideration before the principle involved in it was defined, as exemplified by the extract below in Table 2.

Table 2

Concordance Lines for "Delimiting the Case Under Consideration"

received Brian's letter.	Assuming	the offer does remain open, Brian's Thursday
proceeding on the latter to discuss. I now have to	assumption consider	In order to discuss the law related to whether B's message, left on the answerphone
third party. Bata v. Bata. With second part of the story.	Considering regard to Regarding	first the story about The BCDs, can the ba contracts *ex facie* illegal it is necessary to the potential claim of Evangeline, it is subm

Note. Adapted from Weber (2001). A concordance- and genre-informed approach to ESP essay writing. *ELT Journal*, *55*(1), 17.

In an interesting departure from the usual type of ESP work, Weber's students were also exposed to corpora of different, non-legal genres in order to sensitise them to the highly specific use and patterning of certain lexical items, such as *held* and *submit*, in legal genres.

Another initiative in the area of legal genres is that reported in Bhatia et al. (2004), who propose various move-specific concordancing activities for one genre of legal English, the problem-question genre, written by students within an academic setting. However, Bhatia et al. (2004, p. 224) caution that as legal genres are highly complex and 'dynamically embedded in one another', one has to go beyond the immediate textual concordance lines and look at discursive and

institutional concerns and constraints to fully interpret and by extension become a skilled writer.

A few pedagogic endeavours focus on the writing of research articles for postgraduate students. In view of the fact that PhD students are facing increasing pressure to publish and in some parts of the world published articles are a prerequisite for a PhD to be awarded, several accounts address the burgeoning area of English for Research Publication Purposes (ERPP). An early account can be found in Cortes (2007). While the class was a mixed-discipline one, Cortes accommodated a discipline-specific perspective in her teaching by compiling ten sub-corpora of research articles reflecting the disciplines of her students for genre-based corpus activities.

Turning to another ESP area, that of psychology, Bianchi and Pazzaglia (2007) adopted a genre perspective in a cycle of activities for helping Italian students write psychology research articles in English. Acknowledging the continuing debate on English as an International Language, they state that their 500,000-word corpus drawn from the areas of language acquisition and developmental psychology 'should be representative of the language of the psychology community, which includes authors from different nationalities using English as a *lingua franca*' (p. 265). An innovative feature of this writing instruction cycle is that students were asked to subdivide their choice of written article into moves and annotate it themselves using a functional and meta-communicative coding system devised by the authors. This was followed by data-driven guided writing tasks, which focused on the lexico-grammatical patterning of key words related to the concept of research and verb tense usage in different moves.

Dong and Lu's (2020) corpus-based course was designed for electrical engineering and mechanical engineering master's students. Of interest is that the instructor and students collaboratively compiled a corpus of research article introductions which was then annotated for rhetorical moves using AntMover (see Anthony, 2019). Their report also covers an in-depth evaluation of the course using triangulated data, addressing two key questions: (1) Do guided corpus-based genre analysis activities using a self-compiled specialised corpus contribute to improvement of EFL learners' genre knowledge and genre-based writing skills, and (2) How do EFL learners perceive their experience with the corpus-based genre analysis activities? (p. 140).

While all the pedagogic applications above make use of small, specialised corpora compiled by the individual class tutor, Cotos et al. (2015) report on a large-scale study in which a move analysis was carried out of all RA sections in a corpus of 900 articles spanning 30 disciplines. Pedagogic exploitation of this cross-disciplinary corpus using online tools can be found in Cotos et al. (2016). This large-scale project is thus notable for bridging the gap between research and pedagogy. Another report of a sophisticated online writing tool based on a corpus of research articles annotated for moves and steps can be found in Mizumoto et al. (2017).

Conclusion

The focus of this short article has been on the use of corpora for academic writing pedagogy carried out within a Swalesian genre-based framework (see Flowerdew, in press 2022 for applications using other approaches). It has been shown that the pedagogic initiatives are quite wide-ranging, encompassing a variety of academic genres and disciplines. A few make use of off-the-shelf corpora while others use self-compiled corpora to meet the needs of students from a particular discipline. Some small, specialised corpora have been compiled by the class teacher for their own individual use, whereas other larger-scale corpora form the basis of sophisticated online writing programmes. Importantly, research findings such as those by Kanoksilapatham (2005, 2007), Durrant and Mathews-Aydınlı (2011) and Parkinson (2017) mentioned in the introduction of this article offer valuable insights on lexico-grammatical pattering of moves in specific genres and disciplines and can be used for the preparation of hands-on concordancing tasks. On a final note, it is hoped that these accounts illustrate the affordances of using corpora for academic writing pedagogy.

About the Author

Lynne Flowerdew: An honorary visiting research fellow in the Department of Applied Linguistics and Communication, Birkbeck, University of London.

References

Anthony, L. (2019). Tools and strategies for data-driven learning in the EAP writing classroom. In K. Hyland & L. Wong (Eds.), *Specialised*

English. New directions in ESP and EAP research and practice (pp. 179–194). Routledge.

Bhatia, V.K., Langton, N., & Lung, J. (2004). Legal discourse: Opportunities and threats for corpus linguistics. In U. Connor & T. Upton (Eds.), *Discourse in the professions. Perspectives from corpus linguistics* (pp. 203–231). John Benjamins.

Bianchi, F. & Pazzaglia, R. (2007). Student writing of research articles in a foreign language: metacognition and corpora. In R. Facchinetti (Ed.), *Corpus linguistics 25 years on* (pp. 259–287). Rodopi.

Cortes, V. (2007). Exploring genre and corpora in the English for academic writing class. *The ORTESOL Journal, 25,* 8–14.

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.

Cotos, E., Link, S. & Huffman, S. (2016). Studying disciplinary corpora to teach the craft of Discussion. *Writing & Pedagogy*, 8, 33–64.

Diani, G. (2012). Text and corpus work, EAP writing and language learners. In R. Tang (Ed.), *Academic writing in a second or foreign language* (pp. 45–66). Continuum.

Dong, J. & Lu, X. (2020). Promoting discipline-specific genre competence with corpus-based genre activities. *English for Specific Purposes*, *58,* 138–154*.*

Durrant, P. & Mathews-Aydınlı, J. (2011). A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes, 30,* 58–72.

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.

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.

Flowerdew, L. (in press, 2022). Corpora for EAP writing. In E. Csomay & R. Jablonkai (Eds.), *The Routledge handbook of corpora in English language teaching and learning.* Routledge.

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes, 24,* 269–292.

Kanoksilapatham, B. (2007). Rhetorical moves in biochemistry research articles. In D. Biber, U. Connor & T. Upton (Eds.), *Discourse on the*

Move. Using corpus analysis to describe discourse structure (pp. 73–119). John Benjamins.

Mizumoto, A., Hamatami, S. & Imao, Y. (2017). Applying the bundle-move connection approach to the development of an online writing support tool for research articles. *Language Learning, 67,* 885–921.

Parkinson, J. (2017). The student laboratory report genre: A genre analysis. *English for Specific Purposes*, 45, 1–13. https://doi.org/10.1016/j.esp.2016.08.001

Samraj, B. (2016). Research articles. In K. Hyland & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes* (pp. 403–415). Routledge.

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

Tribble, C. (2002). Corpora and corpus analysis: New windows on academic writing. In J. Flowerdew (Ed.), *Academic discourse* (pp. 131–149). Longman.

Weber, J. J. (2001). A concordance- and genre-informed approach to ESP essay writing. *ELT Journal*, *55*(1), 14–21.