

Use of Corpus-Based Classroom Activities in Developing Academic Awareness in Doctoral Students

DOI: 10.15804/tner.2017.48.2.02

Abstract

The corpus-based approach in language education has driven an array of studies over the past few decades, including its application in developing academic awareness in second-language learners of English for academic purposes. Most publications that show ways of incorporating corpus-based classroom activities (C-BCA) concern concordance lines and lexico-grammatical studies with reference to teacher feedback. The present study examines postgraduates' perception of C-BCA in English for academic purposes classes. A class of 10 postgraduate students from Samara University, Russia, participated in the research. Questionnaire results showed a positive perception of using C-BCA as a tool for writing assignments. The study also concerns an analysis of the students' results in the writing assignments as a part of their academic awareness. Specifically, this sample study shows interdependence between the positive dynamics of postgraduates' perceptions of C-BCA and good results of a writing examination.

Keywords: *corpus, corpus-based classroom activities, writing, academic awareness, language awareness*

Introduction

Scientific communication is increasingly mediated by the English language, thus inevitably requiring that all participants understand specific aspects of English academic register and discourse. Academic awareness is also becoming

increasingly crucial today because academics whose mother tongue is other than English are less interested in publishing in their own languages (Benesch, 2001; Goatly, 2000; Luke, 2004; Wallace, 2003) due to declining citation of non-English scientific literature. For instance, references to English language publications now constitute 85% of all citations in French academic journals; in the Science Citation Index, English now accounts for 95% of all publications (Hyland, 2007). Therefore, irrespective of location, a sufficient level of academic awareness becomes a prerequisite for a distinguished scholarly career and should be pursued as an essential goal in all academic English teaching programmes, particularly at PhD level. This idea has been reflected in a number of studies undertaken by researchers in the fields of education, language teaching and linguistics (Mcgowan, 2014; Bucheister, 2016; Ayazgok & Aslan, 2014; Munoz-Luna & Taillefer, 2014; Gordon, 2015). However, to the best of our knowledge, investigations addressing practical approaches to promoting and developing academic awareness have so far been rather limited.

Research Problem

In the last three decades, a significant body of research has developed that supports the use of corpus-based classroom activities (C-BCA) with second language learners. Some work has been carried out to relate corpus-based approaches to the principles of second-language acquisition (SLA) and locate these practices within the communicative language-teaching (CLT) paradigm (Saeed & Waly, 2009). It has been shown that uses of corpora within the classroom setting can range from teaching individual linguistic features to exploring the specific characteristics of texts from various discourses, including business memos, biology lab reports and campaign speeches. The implementation of a corpus and a concordance in a language classroom has shown to be useful for SLA purposes (Floweredew, 1994; Tribble & Johns, 1997). Individual cases of successful adaptation and implementation of corpus-based computer-aided linguistic learning programmes have been presented (McEnery, Baker & Wilson, 1995). In contrast, relatively little attention has been paid to learners' attitudes towards actual corpora use in EAP classes for developing writing skills. Johns (2000) argues that using corpus analysis methods with second language learners has the additional consequence of reassigning traditional classroom roles, whereby students become linguistic 'researchers', and teachers become directors or coordinators of research.

In order to ensure that doctoral students are able to communicate effectively in international contexts, a means should be found to encourage students to use aca-

demically appropriate vocabulary in their papers. Corpus technology has demonstrated great potential for L2 writing instruction through integrating vocabulary, grammar and discourse patterns of given types of writing into the teaching of academic writing (Gledhill, 2000). As will be mentioned in the literature review section, corpus linguistics has expanded rapidly in the areas of second language research and teaching, with its largest contribution to-date being in the area of vocabulary, e.g., the teaching of collocations, semantic prosody and the Academic Word List (AWL) (Coxhead, 2000).

Research Focus

The main purpose of this study was to increase the academic awareness of doctoral students by using a corpus-based teaching approach. In using existing corpora to create their own repertoires of linguistic “chunks”, students may gain an understanding of the idea of corpus-based classroom activities. Due to their adaptability to individual research fields, these corpus-derived linguistic constructions are applicable for academic writing purposes.

In Russia, one of the most challenging aspects of EAP programmes is deciding which linguistic material to focus on in written papers. This is primarily because students tend to be more familiar with technical vocabulary specific to their own fields than with the general academic English lexicon, since the latter occurs with lower frequency than general-service vocabulary items (Chernykova, 2011; Koptyaeva & Oveshkova, 2016; Worthington & Nation, 1996; Xue & Nation, 1984).

In this study, the Academic Word List (AWL) (Coxhead, 2000) was used as a methodological resource. The AWL, which was developed in 2000 by A. Coxhead, contains 570 word families that account for approximately 10.0% of the total words (tokens) in academic texts, but only 1.4% of the total words in a fiction-based corpus of the same size. By highlighting the words that university students encounter across a wide range of academic texts, the AWL shows EAP learners which words should be focused on and studied. The list also provides a useful basis for further research into the nature of academic vocabulary.

Research Methodology

Research General Background

There are many examples in the second-language acquisition (SLA) literature that promote the use of corpus-based activities (C-BA). In conforming C-BA with SLA principles, various approaches have been taken that place it within the com-

municative language-teaching (CLT) paradigm. Corpora were first used in language teaching at Aston University in 1969. In the early 1980s, COBUILD concordances were used in language classrooms; these tools gave learners access to corpora so they could find multiple examples of language (McEnery & Wilson 1997, p 12). Tim Johns (1994) proposed the term “data-driven learning” (DDL) to describe software that can learn from new data. These days, data gathering is much easier as computers and machine-readable texts are readily available. DDL is used within computer-supported classrooms to help students explore patterns in the target language, with activities and exercises created on the basis of existing corpora.

Over the last 15 years, some researchers (Conrad, 1999; Cortes, 2007; Saeed & Waly, 2009) have described how the use of corpora in English Language (EL) classrooms can be an effective way of showing all the peculiarities of “actual language usage” (Bennett & Bricker, 2006). The provision by authors of a sample of authentic text in the appropriate field of communication can be a valuable resource for novice scholars as well as academics. Johns (2000) additionally argues that corpus analysis methods used with second language learners has the added effect of reassigning traditional classroom roles, whereby students become linguistic “researchers”, and teachers become directors or coordinators of research. The idea of assembling one’s “own corpus” provides a fertile opportunity for students to discuss the best ways of compiling and encoding corpora for their own purposes (“Teaching and Learning by Doing Corpus Analysis”, 2000). Over the course of the present study, the benefits of integrating linguistic corpora in student classroom activities were explored.

A compilation of research exploring different manners of implementing corpus-based learning activities of L2 shows that this teaching tool may be used directly in the classroom and indicates how corpus research may be applied to inform syllabi and classroom materials. Conrad (1999) aims to make the ideas of C-BCA accessible to second language teachers, and teacher-trainers working with language instructors. To understand and apply C-BCA in language teaching, it is essential to comprehend what a corpus is. A corpus is defined as the set of written or spoken samples of actual language (Nadja Nesselhauf, 2011). Although ‘corpus’ can refer to any systematic text collection, it is commonly used in a narrower sense today to primarily refer to text collections that have been systematically computer-processed. Nevertheless, it has been shown that it is possible to use a ‘data-driven’ approach (Johns, 1994) to language learning and teaching even if one does not have access to computer-based corpus resources. The present study investigated the benefit of using corpora as a resource for language learners’ foreign language skills development using ‘self-made’ corpora for their specific area.

Research Sample

In this paper, the scope of studies that present the concept and application of using corpus in EL education will be extended into the micro-context of Russian universities. Currently, the changing nature of higher education in Russia, characterised by a growing emphasis on the importance of publishing in peer-reviewed high impact factor journals, increasingly implies the need to develop and support academic writing. In Russia, there has not been a strong tradition of developing academic writing skills. The first attempts to get acquainted with contemporary academic writing approaches were made when the Unified State Exam (“EGE”) was introduced in Russia in 1996. The task was to write two types of essays (a “for and against essay” and an “opinion essay”). However, partly due to a lack of established approaches or methods used for developing academic writing skills in a second language (L2) in Russia, the task was challenging. Today, university students, postgraduate students and novice scholars are expected to write to communicate ideas, thoughts about their research and study with their colleagues and supervisors. One of the most challenging aspects of designing an EAP course is how to improve students’ writing proficiency in such a way that deliberate and relevant usage of academic vocabulary becomes second nature. Hyland (2003) argues that there are very many difficulties in L2 writing caused by misunderstanding of a text’s register, grammar and lexical diversity. Another reason why academic writing is hard to develop is because it is used to express acquired knowledge in a specific field. This means that one needs to understand specific vocabulary of a given research area. When participating in global scientific communication, a researcher must be able to use appropriate language, the meaning of which is clear to colleagues around the globe. Since it tends to be serious in nature and often used to advance particular theories or arguments in relation to a specified discourse, academic vocabulary plays a crucial role in developing postgraduate writing skills in EAP courses. Researchers (Cohen, Ferrara, & Fine, 1988) assert that incorrect vocabulary usage of a lexis-specific or etymological nature can cause misunderstanding in academic communications where technical vocabulary is required. On these grounds, it is necessary to understand the terminology of a certain scientific or research area.

This study sets out to address the issues outlined in the above review, to explore postgraduate students’ perspectives on C-BCA and to examine whether, by using a corpus-based teaching approach, the awareness of academic language on the part of doctoral students can be enriched.

Instrument and Procedures

The main research question underlying this study was whether corpus-based activities enhance doctoral students' academic language awareness, and if so, in what way they impact on the development of writing competencies.

The study was conducted at Samara University, Russia, 10 participants in this study were postgraduate students attending an EAP course. Classes were held twice weekly for two hours over a 15-week period. The postgraduate students who took part in the study were from different academic backgrounds (History, Sociology, Psychology, Pedagogy, Management and Economics). The students were studying English in order to pass their postgraduate (EAP) exam and improve their English skills for sharing their research through papers, presentations and other forms of scientific communication. There were four female and six male student participants, with ages ranging from approximately 23 to 39. According to the results of their placement test at the beginning of the course, their proficiency level was: B1–8 students; and B2–2 students. These students are Russian, therefore Russian is their first language. All the students are literate in their first language.

The primary target of the course is to prepare students to be part of the global scientific community through academic communications such as analysis and discussion of scientific papers in their field. At the end of the course they are required to write a literature review consisting of a text written to consider the critical points of current knowledge including substantive findings, theoretical and methodological contributions to a topic in the field of the particular postgraduate student' research. Since, as Cooper (1998) points out, a literature review usually precedes both a research proposal and results section and is associated with academic-oriented literature such as a thesis or peer-reviewed article, the task of writing literature reviews is a staple for research in every academic field. The activities of the EAP classes were selected to address the course objectives (from the course syllabus). Additionally, a variety of original instructional activities were devised, in which the postgraduate students focused on improving their academic English skills, many of which involved authentic reading and interpreting materials, such as articles from journals and monographs. Most student output in the classroom consists in answers given to show their work and research through the medium of the academic text format. Student performance was monitored and assessed formatively through assignments, performance during classroom tasks and summatively through unit exams, a midterm and a final exam. The students had to achieve an overall mark of 70% to pass this 3-credit course.

There were four main tasks (each 25 % of the total sum). Four academic sources were selected to work with as no single text would involve all elements of repeated

language characteristic of a larger corpus. (According to Byrd & Coxhead (2007), studying the stylistic choices of language in a sample of writing by one author is insufficient to discern the whole range of formulaic phraseology characteristic of a subject-specific lexicon: in order to reveal lexical bundles and other formulaic linguistic patterns, it is necessary to study large corpora). During the course of study – during the 3rd, 7th, 12th and 14th weeks – the students were required to write four critical reviews of an article from their field of study. The evaluation of the articles consisted of two peer assessment tasks (3rd and 12th weeks) using IELTS band descriptors. The students were asked to take part in a survey (Appendix) three times (at the beginning of the study, in week 5 and at the end) about the efficacy of applying C-BCA in the execution of their critical review writing assignments.

Data Analysis

The primary purpose of this research was to investigate the students' evaluations of the efficacy of C-BCA as used in their writing assignments in terms of whether it had increased the level of their academic awareness. Due to the lack of an appropriate empirical unit of learning, the students' perceptions of C-BCA efficacy was selected as the best method to gauge achievement. In so doing, it was assumed that the implementation of C-BCA is dependent on the actual educational conditions, which are in turn dependent on students' perceptions of the educational usefulness of their writing assignments. The outcomes of this survey were used to contribute to our understanding of the overall effectiveness of this learning method.

As a regular classroom activity, the students were asked to write an opinion essay about the difficulties of using academic vocabulary in their writing. The objective was to learn about the students' understanding of the term "academic vocabulary". Out of ten students, only two showed a competent understanding of the notions of "special words for writing articles" and "the words which occur frequently in a range of academic subjects" (including commerce, management, and marketing). Then it was explained to the postgraduate students participating in the activities what a corpus is, the concept of corpus-based learning and how corpora can be valuable for understanding linguistic variations due to subject-specific and register factors.

In our case we would use a corpus of informal conversation for academic writing instruction. One of the main tools in C-BCA are activities which address students' response to corpus-based classes (for feedback).

For the purposes of this study, it was assumed that an important aspect of effective delivery of EAP would rely on an understanding of students' preferred

learning styles. Attempts were made to diversify the format of different tasks: some students worked online, while others used printed copies of corpora. The idea was to present the students with a methodological approach to vocabulary AWL using (Coxhead, 2000), followed by getting them to produce their own examples from their analysed papers.

In order to determine the students' responses to the tasks, a final critical review was conducted at the end of the course, at which point the students' surveys about completed activities were collected. By means of these opinions, it was possible to ascertain whether or not the activity had enabled them to use written language to describe their thoughts and ideas in the most appropriate manner.

Research Results

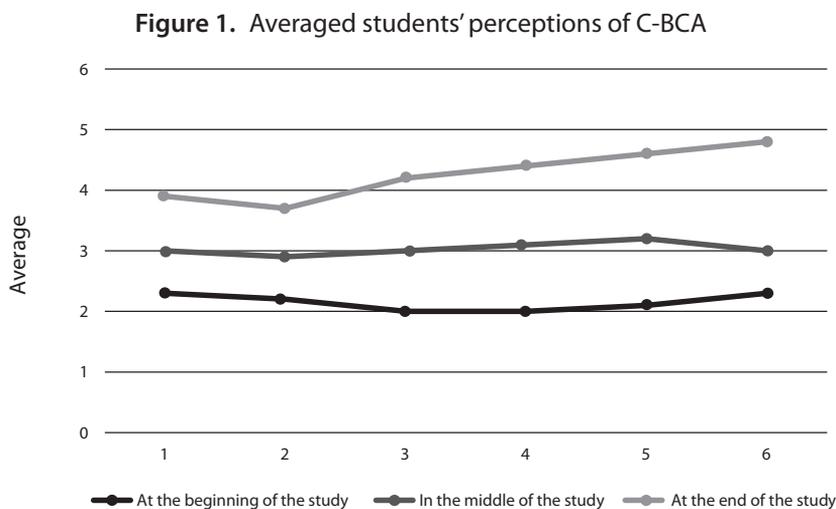
This study used both qualitative and quantitative methods of analysis. The students' critical reviews and reflective surveys (N=10) were triangulated and analysed.

The surveys of the postgraduate students, which were conducted three times during the study: at the beginning before the intervention, in the middle when the idea of C-BCA was introduced and put into practice and at the end after the final writing exam, were collected and analysed (Appendix). This discussion mentions a chronology of events (Table 1).

Table 1. Analysis of the students' perceptions of C-BCA

At the beginning of the study						
Number of questions/results	1	2	3	4	5	6
Average mean	2.22	2.08	1.8	1.7	1.65	1.7
Variation	0.49	0.54	0.31	0.46	0.31	0.28
In the middle of the study						
	1	2	3	4	5	6
Average mean	3	2.84	3.05	3.3	3.35	3.1
Variation	0.34	0	0.22	0.34	0.32	0.32
At the end of the study						
	1	2	3	4	5	6
Average mean	3.87	3.68	4.35	4.63	4.75	4.8
Variation	0.10	0	0.16	0.23	0.10	0.08

Table 1 shows that the students' mean perception of the necessity of C-BCA grew over the course of the study. When analysing the statistics, it was found that > 0.35 of the students demonstrated an awareness of C-BL and C-BCA as evidenced by their answers to questions 1, 2 and 4. Following the implementation and practice of C-BCA, the students' perceptions of C-BL and C-BCA appeared to grow (Figure 1).



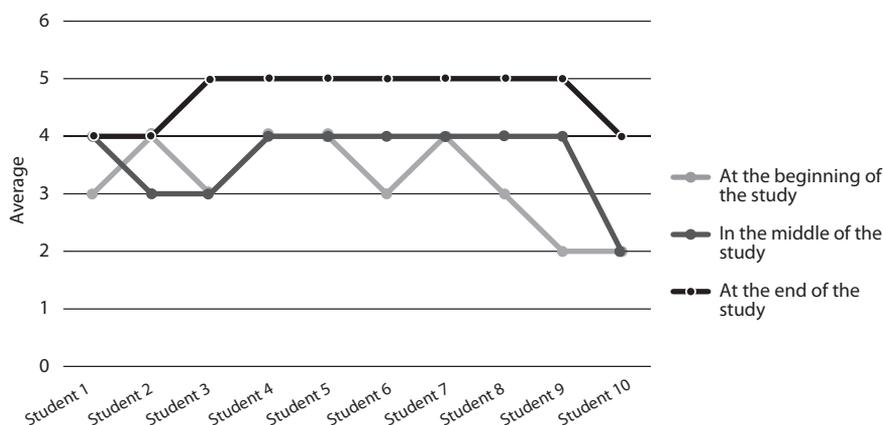
All sources of qualitative data are represented in a narrative description; therefore, the data obtained through tests (writings) can be analysed by means of descriptive statistics (Burns, 2010) and the IELTS scale. The critical review scores were tabulated from different time periods and the results compared and analysed (Table 2).

Table 2. Students' critical review assignment results

At the beginning of the study	
Average mean of the students writing papers	3.2
Variation	0.25
In the middle of the study	
Average mean of the students' writing papers	3.6
Variation	0.19
At the end of the study	
Average mean of the students' writing papers	4.7
Variation	0.1

Table 2 shows the dynamic of the writing assignment results of 3 critical reviews, which were conducted three times during the study. Positive results can be seen with a reduction in variation and an increase in mean value (Figure 2).

Figure 2. Mean values of the level of the students' writing assignments



All the data was taken from independent sources concerning one category, then it was triangulated, thus the research findings have an established validity. In other words, the data from the surveys validated the data obtained from the critical review papers. The students indicated that C-BCA had a positive impact on their academic writing skills and would be useful for their future participation in scientific communication abroad.

Discussion

Specifically, this study examined the students' perceptions of C-BCA and results of their writing assignments. The study set out to discover the strengths and weaknesses of corpora as an EAP writing tool. The study's qualitative and quantitative data indicate that the students perceived the corpus approach as beneficial to the development of EAP skills, specifically in terms of increased confidence towards L2 academic writing tasks. The findings of the study support research on the positive perception of C-BCA and C-BL as effective teaching tools in EAP classes for postgraduate students. The following quotations from reflection papers

of the participants in this study provide evidence that C-BCA had enhanced their awareness in academic writing and academic language as a whole:

- “C-BCA helped me in improving my critical review writing assignments. Firstly, it became easier to understand academic articles. Now I see that the language of science is a specific one. Secondly, now I can formulate my thoughts better in, which is important, an appropriate style.”
- “Before this course I did not even think that scientific papers are written using specific words. In my earlier writing activities I used usual, everyday phrases. Now I understand why I never had my papers published”.
- “It is important to pay special attention to AWL because it can help develop and vary our vocabulary, also AWL helps while writing or reading academic articles.”
- “My first critical review was brief and did not include high-frequency academic words. The last critical review was larger and it was easier for me to include necessary vocabulary to express ideas. I hope that now my future papers will look much more scientific.”

Conclusions

The participants were asked to take part in three feedback surveys throughout the study. Had the survey only been conducted at the end of the course, the dynamic and change of perception would not have been observable.

The study, conducted over half a year, investigated changes in the students' perceptions about C-BCA practices and growth in the level of their academic awareness, allowed us to report its findings. The primary purpose of this study was to examine the impact of the students' exposure to corpus technology on their perceptions concerning the development of their academic writing skills and how this awareness, in turn, influences their use of academic language itself. The findings reveal that C-BCA not only had an immediate effect of helping students solve writing/language problems but additionally boosted their understanding of the relationship between lexis and grammar and increased their general linguistic awareness. Therefore, these teaching practices clearly had a positive effect on the development of doctoral students' academic awareness. Our findings support the assumption that C-BCA and EAP significantly contribute to L2 writing pedagogy inasmuch as the acquisition of vocabulary is a key L2 writing competence objective.

Appendix

Dear friends! Evaluate the statements according to your perception. 1 – is the lowest level and 5 – is the highest one. Thank you!

1	It would be advantageous for me if there were more C-BCA	1.....2.....3.....4.....5
2	C-BCA advantages for me	1.....2.....3.....4.....5
3	C-BCA requires significant student-led changes	1.....2.....3.....4.....5
4	C-BL will help me in my writing assignments	1.....2.....3.....4.....5
5	It is necessary to use AWL for academic writing	1.....2.....3.....4.....5
6	I would like to have more C-BCA in my EAP classes	1.....2.....3.....4.....5

References

- Benesch, S. (2001). *Critical English for academic purposes: Theory, politics, and practice*. Mahwah, NJ: Lawrence Erlbaum.
- Bennett, G., Bricker, M. (2006). Book review: Corpus-based Classroom Activities. *HEIS News July 2006: 25, Issue 2*.
- Bucheister, J. Toward a theory of academic self-awareness. *Teaching and Learning together in Higher Education*. US: Scholarship, Research, and Creative Work at Bryn Mawr College, 2016.
- Burns, A. (2010). *Doing action research in English language teaching: A guide for practitioners*. New York: Routledge.
- Cohen, A., Glasman, H., Rosenbaum-Cohen, P.R., Ferrara, J., & Fine, J. (1988). Reading English for specialized purposes: Discourse analysis and the use of standard informants. In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading*, 152–167.
- Conrad, S. (1999). The importance of corpus-based research for language teachers. *System*, 27(1), 1–18.
- Cooper, H. (1998). Synthesizing Research: A Guide for Literature Reviews. *Applied social research methods series, Vol. 2*
- Coxhead, A. (2000). A new AWL. *TESOL Quarterly* 34(2), 213–238.
- Coxhead, A., Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 16, 3, 129–147.
- Crystal, D. (2003). *English as a global language*. Cambridge: Cambridge University Press.

- Gledhill, C. (2000). The discourse function of collocation in research article introductions. *English for Specific Purposes*, 19(2), 115–135.
- Creswell, J.W. (2003). Research design: Qualitative, quantitative, and mixed methods approaches (2nd ed.). Thousand Oaks, CA: Sage.
- Goatly, A. (2000). Critical reading and writing: An introductory coursebook. London, England: Routledge.
- Hyland, K. (2003). Second language writing. Cambridge: Cambridge university press.
- Hyland, K. (2007). Is there an “Academic vocabulary”? *TESOL Quarterly*, 41,2, 235–253.
- Johns, T. (1994). From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin, ed., *Perspective on pedagogical grammar*, 293–313. Cambridge: Cambridge university press.
- Koptyaeva, N., Oveshkova, A. (2016). The use of web technologies in academic writing. 5th Information school for young researchers (Eds.). Ekaterinburg: UB RAS Central Scientific Library. P. 55–60.
- Luke, A. (2004). Two takes on being critical. In B. Norton & K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 21–29). Cambridge, England: Cambridge University Press.
- Reppen, R. (2010). Using corpora in the language classroom. New York, NY: Cambridge University Press.
- Saeed, N.A. & Waly, S. (2009). Corpus for Classrooms: Ideas for Material Design, *Proceedings of the 10th METU ELT Convention*. Retrieved from: <http://dbe.metu.edu.tr/convention/>
- Teaching and Learning by Doing Corpus Analysis (2000). Proceedings of the Fourth International Conference on Teaching and Language Corpora, Graz 19–24 July, Vol. 1 Retrieved from: