

Uptake of corpus tools in the Spanish Higher Education context: a mixed-methods study

Pascual Pérez-Paredes · Purificación Sánchez-Hernández
University of Cambridge / UK · University of Murcia / Spain

Abstract – This paper examines the introduction and use of corpus consultation in the course of a training initiative sponsored by the Professional Training Unit of a medium-sized University in Spain. ‘Introducing Research Articles (RA) Writing’ was a 12-hour module that offered researchers the opportunity to gain insight into the nature of the research articles (RA) across different disciplines. The researchers (n=25) and the instructors met three times in two-hour sessions during a two-month period. All participants completed two post-task questionnaires and a delayed questionnaire. An interview was completed two years after the end of the course. After task 2, 64 percent of the participants found corpus tools to be of great help when writing their research articles. No significant differences between B1 and B2–C1 groups were found in their assessment of the writing tools provided. Increased familiarity with the corpus tools did not result in a better appraisal of these resources and all participants seemed to favour the use of the curated list of vocabulary provided. The delayed questionnaire and subsequent delayed interviews (n=5) revealed that the use of corpora had had limited or no impact on the writing practices of these researchers. We argue that the use of corpora in professional writing contexts requires careful planning as well as continued institutional support.

Keywords – DDL, corpus linguistics, language education, academic writing, Higher Education

1. INTRODUCTION

English is the lingua franca of the academia and is widely referred to as the language of science. As 95 percent of the publications in the Science Citation Index are written in English (Hyland 2013), non-native speakers of English scientists very frequently face difficulties (Hanauer and Englander 2013) to produce quality academic writing (Connor 1999). In particular, academic writing in English is challenging for Spanish students and professionals for several reasons (Basanta 2005; Pérez-Paredes and Sánchez-Hernández 2005; Pérez-Paredes 2010). As reported by Gea-Valor et al. (2014), it seems difficult for many Spanish researchers and scholars to publish their research in international journals in part due to the need of training in academic writing and also because of the low levels of proficiency in English (Pérez-Paredes 2010; Ferguson et al. 2011).

This study originated when the authors were commissioned the design and teaching of a 12-hour course to University researchers across different fields of knowledge. The main purpose of this study is to examine the University researchers’ evaluation of corpus-consultation techniques and the uptake of such practice immediately after the training sessions, and both two months and two years after the completion of the training initiative. This research, together with Charles’ (2014), is one of the few studies that makes use of mixed methods to explore a delayed evaluation of corpus resources. We set out to contribute to the debate about how to implement Data-Driven Learning (DDL) pedagogy in the context of Higher Education (HE)



and, specifically, how University researchers may benefit from DDL in training initiatives across universities in Spain and beyond.

2. CORPORA AND SCIENTIFIC WRITING IN ACADEMIC CONTEXTS

2.1. *Corpora, register description and L1/L2 writing in academic contexts*

Different research efforts have pioneered the use of corpus linguistics methods in the description of general English, scientific language and academic writing (Altenberg 1987; Bathia et al. 2011; Boulton et al. 2012). The Create a Research Space (CARS) model (Swales 1990) has been applied, for example, to the research in different moves and steps in PhD literature reviews (Flowerdew and Forest 2010), the use of research articles corpus consultation (Charles 2014) or the analysis of the use of corpus-based research (Flowerdew 2016). Phraseology has been widely investigated in English for Academic Purposes (EAP) (Staples et al. 2013) and corpus studies have confirmed the strong link between academic (sub)registers and lexical choice (Staples et al. 2016). Cortes (2004) and Hyland (2008) have discussed the links between lexical bundles in writing in different fields and pedagogical implications. This significant body of research seems to suggest that corpus linguistics is not only of relevance for the description of combinatory grammatical-lexical nature, but also of great relevance for the learning of language features that are specific to scientific writing.

Although the use of corpora is not a mainstream practice in the language classroom (Boulton and Pérez-Paredes 2014), DDL (Johns 1991) has successfully illustrated language patterns by prompting learners to analyse concordances, frequency lists, distribution patterns, etc. These inquiry methods have been extensively used in HE settings (Boulton 2010a; Chang 2014) to improve the learning of EAP (Tribble 2002; Lee and Swales 2006; Charles 2007), and to teach grammar to HE students (Chambers and O'Sullivan 2004; Yoon 2008; Boulton 2009, 2010b; Pérez-Paredes et al. 2011; Pérez-Paredes et al. 2012). Many of these studies have focused on classroom experiences and found benefits in grammar instruction with DDL. In this mood, Huang (2014), Smart (2014) and Karras (2016) argue that using DDL is beneficial to improve learners' grammar and vocabulary knowledge.

Researchers have profusely looked at the students' attitudes towards DDL. Boulton (2009) and Geluso and Yamaguchi (2014) discuss attitudes toward DDL, positive in most of the cases, and Yoon and Hirvela (2004) and Kennedy and Miceli (2010) report positive feedback from students using corpora autonomously as reference sources for in-class writing tasks. Some studies (Hill et al. 2000; Coxhead 2008) found some resistance on the part of students to search for vocabulary and phraseological patterns in collections of texts. However, not many studies report on the use of corpora some time after the learning/teaching experience. Charles (2014: 39) reports the use of personal corpora by students one year after the course and states that most respondents did "get the corpus habit".

2.2. *Training programs in L2 writing for researchers*

In Spain, the acquisition of English academic writing skills remains an important issue. Gil Ortega (2002) analysed the difficulties Spanish students have to face when studying in higher education institutions in Britain and reports that communicating in English is one of the most important problems Spanish students tackle when studying in British Universities. This is so because a large number of students has never been taught academic writing skills before (Gil Ortega 2002). Gea-Valor et al. (2014) confirm that in Spain specialized courses on academic writing are not available in most universities and research institutions. This results in deficient writing practices during and after HE. Kaufhold (2015) investigated the conventions in postgraduate academic writing demonstrating "the increasing alignment to the epistemological approaches in their research fields" (2015: 132). Her conclusions are in agreement with Belcher and Hirvela (2005), who argue that postgraduate writers would benefit from explicit guidelines for orientation when writing academic texts.

Further investigations are still needed to assess the use of corpus consultation to improve writing as well as the attitudes of students and researchers towards the use of corpora in the context of HE. Our research questions are the following:

- RQ1: Do researchers favour corpus consultation over other resources during the course of a training course on academic and scientific writing? Does English communicative competence play a role in their assessment of the usefulness of corpus consultation?
- RQ2: What is the uptake of corpus consultation after the end of a training course on academic and scientific writing?

3. METHODOLOGY

The following sub-sections deal with the research setting of this experience (3.1), the participants (3.2), the training activity completed by those researchers taking part in this study (3.3), as well as the instruments used (3.4).

3.1 Research setting

Our research was carried out in a Spanish medium-sized public University with around 2,000 teachers and 31,500 undergraduate and graduate students in areas as diverse as medicine, law, physics or nutrition. Spanish is the L1 of most of the students and teachers of this University. Only 4.2 percent of the students during the 2014-2015 academic year were born in a country other than Spain.

Given the growing pressure to publish in high-impact international journals, the Unit in charge of teacher training commissioned the authors the design and development of a 12-hour module ('Introducing Research Articles (RA) Writing') that could offer volunteering HE researchers the opportunity to gain further insight into the nature of RAs across disciplines. The module was devised as a blended-learning initiative. The trainees and the instructors met three times in two-hour sessions during a two-month period. The first two sessions were face-to-face. Information on the structure of research papers, moves within each RA section, main features of scientific/academic writing and some points of grammar relevant for successful scientific writing were discussed. The next two distance sessions were organized so that students could complete exercises to practice the points raised in the first two sessions. They were also asked to write an abstract and a conclusion section in their own field of specialty. For this purpose, the participants were given access to a corpus of abstracts and another one of research articles¹ as well as precise instructions on how to access and use them. Session number 5 was devoted to checking the accuracy of the answers provided by the students to the exercises completed in previous sessions. The last session was entirely dedicated to feedback on exercises and to an analysis of a model abstract section and discussion sections. The training activity was carried out during the second semester of the 2012-2013 academic year and was open to any faculty staff irrespective of their rank or discipline. The authors did not have any sort of responsibility or influence on the selection of the participants.

3.2 Participants

Our sampling was a convenience sample where the "members of the target population are selected for the purpose of the study if they meet certain practical criteria, such as (...) the willingness to volunteer" (Dornyei 2007: 98–99). In this research, the criterion was that all participants should be enrolled in the training module described in Section 3.1. Ultimately, 28 researchers were selected by the unit in charge of teacher training. All the participants were native speakers of Spanish and none of them was bilingual. As three of the original participants did not attend all the sessions and/or did not complete all the post-task questionnaires, we decided to exclude them from our analysis. Of the remaining participants (n=25), 12 were either Professors or Associate Professors, while 13 were non-tenure academic staff. In terms of stated language proficiency, 60 percent said that they had obtained or were studying to obtain a B1 level of The Common European Framework of Reference for Languages (CEFR), while 40 percent said that they had achieved a B2 or C1 or were studying to achieve a B2 level. Figure 1 shows the participants' proficiency across all 5 disciplines involved.

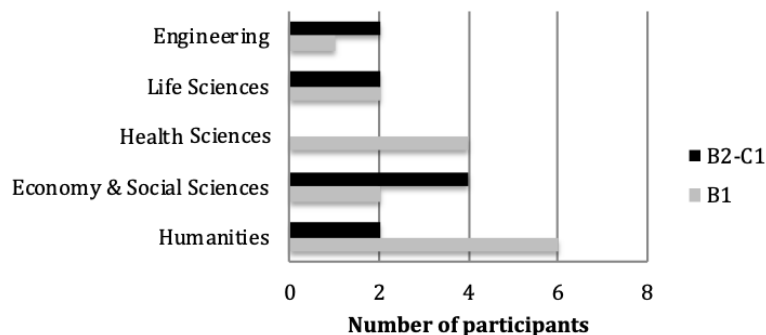


Figure 1: Participants' self-perceived EFL proficiency across all 5 disciplines involved

¹ <http://www.lexutor.ca/conc/eng>

3.3. Training module

The aim of the training module was to offer trainees an account of how RAs are structured across different disciplines and how to approach their writing, with special emphasis on the abstract and discussion sections. In the face-to-face sessions the instructors dealt with the topics outlined in Section 3.1 and set two writing tasks to be completed individually at home by each of the participants. The first task consisted of writing a 150-word abstract in the disciplinary field of the researcher. The second task consisted of writing around 150 words of discussion in the disciplinary field of the researcher. For each of these two tasks, the researchers were given (i) a list of section-specific vocabulary from Glasman-Deal (2010) (Appendix 1); (ii) a comprehensive description of the nature and characteristics of the sections involved; and (iii) guidelines to query and interpret a corpus of academic and scientific English (Appendix 2). During the face-to-face sessions, the students were given precise indications on the affordances of corpus-consultation as found on the lextutor.ca website. Trainees were briefed on concepts such as keywords, concordance lines, left and right sorting and, in general, how to make sense of accumulated frequency and word patterning. For this purpose, we used some of the words in the word lists provided as a resource and transformed them into concordance lines discussing how to interpret them. Our trainees were then introduced to the corpus of academic abstracts and the Research Articles Corpus (RAC) on the lextutor.ca website (Cobb 2019) and were walked through a selection of queries as those previously discussed. This website is easy to access from any laptop or desktop and users are not required to download and install any new software. Liu and Jiang's (2009) suggestions to present modelling and warm-up activities that helped contextualize corpus work were followed by the instructors.

3.4. Instruments

Our research combines qualitative and quantitative methods in a single research project, both sequentially and at the data analysis stage. Table 1 illustrates the data collection points as well as the instruments used for data gathering.

Data collection time	Instrument	Type of data
After submission of writing task 1: writing an abstract	Questionnaire Appendix 3	QUAN
After submission of writing task 2: writing the conclusions section	Questionnaire Appendix 3	QUAN
Two months after the end of the training module	Questionnaire Appendix 4	QUAN + QUAL
Two years after the end of the training module	Structured interview Appendix 5	QUAL

Table 1: Data collection points and instruments used

Some of the benefits of this combination of research methods include the use of different research paradigms to gain better understanding of our specific research context and a multi-level analysis that looks at both the broader societal context and the individual and improves validity (Dornyei 2007: 45–46). Mertens (2005) maintains that, when an issue is embedded in a complex educational or social context, mixed methods are of particular value. For example, surveys and questionnaires are widely used to “gather data at a particular point in time with the intention of describing the nature of existing conditions (...) or determining the relationships that exist between specific events” (Cohen et al. 2011: 256), while interviews allow us to “engage, understand and interpret the key feature of the life worlds of the participants (...) and be able to reveal and explore nuanced descriptions” (Cohen et al. 2011: 414).

Our students completed two post-task questionnaires (Appendix 3), a delayed questionnaire two months after the course (Appendix 4) and a delayed interview (Appendix 5), which was completed two years after the completion of the training module. Figure 2 offers an overview of our data collection methods and their chronological distribution.

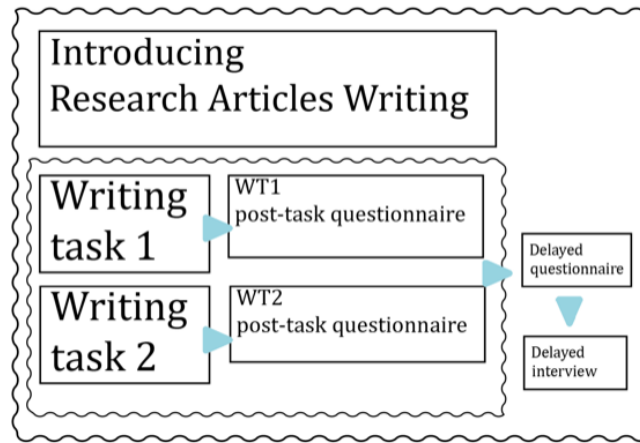


Figure 2: Overview of data collection methods and their chronological distribution

Five out of the 25 participants in the training module volunteered to take part in a structured interview devised to gain insight into the uses of corpus consultation during the 2-year period after the completion of the aforementioned module. The aim of the interview was to seek both description and interpretation of the uses (or failures to use) of corpora. Cohen et al. (2011) maintain that in this type of structured interviews all respondents answer the same questions, which increases the comparability of the responses. Similarly, this interview design reduces interviewer effects and bias. On the other hand, structured interviews offer less flexibility than informal conversations, where the salience and relevance of the questions is increased. The interview data was transcribed and analysed using grounded theory and grouping participants' opinions in themes.

4. RESULTS

In this section the results of the post writing 1 questionnaire (4.1), and the post-writing task 2 questionnaire (4.2) are presented, followed by the findings of the delayed questionnaire (4.3) and the delayed interview (4.4).

4.1. Post-Writing Task 1 questionnaire (WT1)

The participants completed an online questionnaire, where they were asked to assess the usefulness of both the list of words and online corpus consultation during the completion of the task. Table 2 shows the mean scores for WT1 questionnaire.

	Groups/ English competence	N	Mean	Std. Deviation
Wordlist WT1	B1	15	4.00	.535
	B2–C1	10	4.00	.471
Corpus WT1	B1	15	3.67	1.113
	B2–C1	10	3.70	1.160

Table 2: Usefulness of word list and corpus consultation during WT1 (Post Writing Task 1 Questionnaire)

An independent-samples t-test was conducted to compare researchers' assessment of the usefulness of the word lists in B1 and B2–C1 competence levels. There was not a significant difference in the scores for B1 researchers ($M=4$, $SD=0.535$) and B2–C1 researchers ($M=4$, $SD=0.471$); $t(23)=0.001$, $p=1.00$. The independent-samples t-test conducted to compare researchers' assessment of the usefulness of online corpus consultation yielded similar results. There was not a significant difference in the scores for B1 researchers ($M=3.67$, $SD=1.113$) and B2–C1 researchers ($M=3.70$, $SD=1.160$); $t(23)=-0.072$, $p=0.943$.

When asked about how helpful the word list had been, 12 percent decided not to agree or disagree, 76 percent agreed that the list of words had been of great help to write the abstract, while 12 percent strongly agreed that this resource had been of great help. As for the online corpus consultation, 24 percent disagreed

that the resource had been useful, 8 percent decided not to agree or disagree, 44 percent agreed with its usefulness, whereas 24 percent strongly agreed with the statement.

4.2. Post-Writing Task 2 questionnaire (WT2)

As with WT1, Table 3 shows the mean and standard deviation for the usefulness of the word list and online corpus consultation during WT2.

	Groups/ English competence	N	Mean	Std. Deviation
Wordlist WT1	B1	15	4.13	.640
	B2–C1	10	4.10	.568
Corpus WT1	B1	15	3.67	1.113
	B2–C1	10	3.50	1.080

Table 3: Usefulness of word list and corpus consultation during WT2 (Post Writing Task 2 Questionnaire)

An independent-samples t-test was conducted to compare researchers' assessment of the usefulness of the word lists in B1 and B2–C1 competence levels. There was no significant difference in the scores for B1 ($M=4.13$, $SD=0.640$) and B2–C1 researchers ($M=4.10$, $SD=0.568$); $t(23)=0.133$, $p=0.895$. An independent-samples t-test conducted to compare researchers' assessment of the usefulness of online corpus consultation yielded similar results. There was no significant difference in the scores for B1 researchers ($M=3.67$, $SD=1.113$) and B2–C1 researchers ($M=3.50$, $SD=1.080$); $t(23)=0.371$, $p=0.714$.

64 percent agreed that the list of words had been of great help to write the discussion section, 12 percent decided not to agree or disagree, while 24 percent strongly agreed with this statement. When asked about the usefulness of online corpus consultation, 24 percent disagreed with the statement, 12 percent decided not to agree or disagree, 44 percent agreed with it, whereas 20 percent strongly agreed with the statement.

4.3. Delayed questionnaire

The results of the delayed questionnaire submitted to the HE researchers are presented in this section.

4.3.1. Perception of the usefulness of different tools for academic/scientific writing

Two months after the completion of the course, the researchers completed a questionnaire that demanded an assessment of different tools for academic/scientific writing in English. Table 4 offers a summary of the percentage of perceived usefulness of different tools for academic/scientific writing across competence levels according to the delayed questionnaire.

	All students (n=25)			B1 students (n=15)			B2–C1 students (n=10)		
	Slightly useful %	Useful %	Very useful %	Slightly useful %	Useful %	Very useful %	Slightly useful %	Useful %	Very useful %
Dictionaries	8	60	32	6.7	66.7	26.7	10	50	40
Reading RAs	0	20	80	0	20	80	0	20	80
Studying EFL	4	60	36	6.7	53.3	40	0	70	30
Google	16	64	20	20	66.7	13.3	10	60	30
Corpus consultation	16	56	28	20	60	20	10	50	40

Table 4: Percentage of perceived usefulness of different tools for academic/scientific writing across competence levels according to the delayed questionnaire

Independent-samples t-tests were conducted to compare the students' assessment of the usefulness of the resources above in B1 and B2–C1 competence levels. No significant differences were found between both groups (Table 5):

	B1 researchers	B2–C1 researchers	t-test for Equality of means sig (2-tailed)
Dictionaries	M=2.20 SD=.561	M=2.30 SD=.675	t(23)=-0.403 p=0.691
Reading RAs	M=2.80 SD=.414	M=2.80 SD=.422	t(23)=0.000 p=1.000
Studying EFL	M=2.33 SD=.617	M=2.30 SD=.483	t(23)=0.144 p=0.887
Google	M=1.93 SD=.594	M=2.20 SD=.632	t(23)=-1.072 p=0.295
Corpus consultation	M=2.00 SD=.655	M=2.30 SD=.675	t(23)=-1.109 p=0.279

Table 5: Independent-samples t-tests to compare researchers' assessment of the usefulness of different resources two months after the end of the writing course

4.4. Delayed interview

Two years after the completion of the course, the participants in the study were contacted again to find out about their willingness to be interviewed on the corpus resources used in the course. Five researchers agreed to do so. The questions (Appendix 5) were related to whether or not they had used the corpus given to them, or other corpora for that matter, during their academic or scientific writing; how often they did so and if they would recommend other researchers the consultation of corpora to improve their academic writing. They were questioned about how they evaluated corpora as compared with dictionaries or vocabulary lists and the frequency of use. In case of a negative answer they were asked why this resource was not useful for them and which aspects of this resource were considered as the most positive for academic writing. Their answers are reported below.

Researcher A's (B2) area of expertise is Biochemistry and Molecular Biology and is an Assistant Researcher. After WT1, she stated that she had found the word list very useful to "start" and link paragraphs. She felt that she did not know how to make use of corpus consultation. After WT2, she offered almost identical answers. She continued to believe that she did not know how to make use of the online corpora provided and found it hard to integrate this resource into her writing. In the delayed interview, she confirmed that, after the training initiative, she never made use of corpora for her academic/scientific writing. She said that she did not remember the resource at all, but after reviewing her notes and the corpus tool, she said she would consider this resource in future writing. She went on to explain that she found corpus consultation useful to find the "right expression when you know what you want to say but do not know how to say it". Subject A highlighted that her command of the language prevented her from assessing whether her writing is "correct or not".

Researcher B's (B1) area of expertise is Psychology and is a tenured Lecturer. After WT1, he stated that he had found both the word list and online corpus consultation very useful. He explained that these tools "simulate how a native speaker approaches communication". He stated that when writing in English, "words come to mind in groups and that word association is very important". After WT2, he found that both the word list and online corpus consultation were useful to write good paragraphs and sentences. In the delayed interview, he confirmed that, after the training initiative, he never made use of corpora for his academic/scientific writing. He stated that he "probably never fully became aware of its potential uses". He said that during the course he preferred to focus on the guidelines to produce research articles in terms of "style and strategies", something that he could relate to easily as it suited his past learning experiences. Subject B said that he was now more curious about the "power" of the tool and that he might consider using it in the future.

Researcher C's (B2) area of expertise is Sociology and is a part-time Lecturer. After WT1, she stated that she had found both the word list and online corpus consultation useful to become more familiar with how English works. In particular, she stated that these tools gave her the opportunity to understand the "structure and the concordance lines". After WT2, she offered identical answers. In the delayed interview, she confirmed that, after the training initiative, she never made use of corpora for her academic/scientific writing. She said that she had not discussed the potential of the tool with anybody else, although she would

be “willing to do so should the opportunity arise”. She admitted that she did not remember the resource well and was “unable to tell the differences between corpora and dictionaries”. She said that she did not remember much from the course, but she had realized that corpora consultation offers “an academic perspective that is absent in other resources”. She also pointed that using corpora for academic/scientific writing might be interesting because they are “rigorous and trustworthy” sources of information.

Researcher D’s (B1) area of expertise is Psychology and is an Assistant Researcher. After WT1, she stated that she had found the word list very useful to “start” and to “have a baseline lexicon to rely on”. After WT2, she stated that she had “liked” both the word list and online corpus consultation, but she lacked the competence to benefit from their potential. Despite this, she explicitly said that these tools had helped her “a lot” in her writing of WT2. In the delayed interview, she confirmed that, after the training initiative, she never made use of corpora for her academic/scientific writing. In fact, she pointed that she did not recall using corpora consultation during the course. She stated that she used Google and Google Scholar on a regular basis and the “online journals that could be accessed from the University library” for her academic/scientific writing.

Researcher E’s (B1) area of expertise is Geography and is a non-tenured Lecturer. After WT1, he stated that he had found both the word list and online corpus consultation very useful to improve his “competence in English” and, more emphatically, very useful to “improve the translation of my research into English”. After WT2, he found that both the word list and online corpus consultation were useful to “understand how to express ideas in academic English”. He stated that writing academic/scientific English is “stylistically plain” and “more straightforward” than writing in Spanish in academic contexts. This researcher stated that using corpora is of use to know more about how to avoid direct and “rigid” statements in research articles. In the delayed interview, he confirmed that, after the training initiative, he made use of online corpus consultation for his academic/scientific writing “once or twice a year”. He maintains that online corpora are “more professional than Google Translator” and provide more “certainty” when translating into English. He would certainly recommend this resource to other researchers.

5. DISCUSSION

Our most significant finding is that, despite a general positive assessment of the role of corpora for academic writing in English in the post-task and two-month delayed questionnaires (RQ1), very few researchers continued to use this language resource once the training initiative had ended (RQ2). This is evidenced by the fact that four of the five participants interviewed had made no use of online corpus consultation during the two-year period after completing the course. Consequently, our results do not seem to confirm Yoon’s (2008) or Charles’ (2014) findings that contact with corpus consultation affects the writing habits of researchers, although this finding may be taken cautiously, as only 20 percent of the participants agreed to be interviewed. Similarly, our results confirm that the stated level of English communicative competence played no significant role in the researchers’ assessing of the usefulness of the different resources available (RQ1).

5.1. Do researchers favour corpus consultation over other resources during the course of a training course on academic and scientific writing? Does English communicative competence play a role in their assessment of the usefulness of corpus consultation?

The post-writing task questionnaires showed that the word list was perceived most useful after both WT1 and WT2. After WT2, the number of informants that strongly agreed with the usefulness of the word list rose from 12 to 24 percent. This was not the case of the online corpus consultation, where only 68 percent either agreed or strongly agreed that consulting corpora helped these researchers write their abstracts and discussions after WT1. The figure decreased after WT2, when only 64 percent either agreed or strongly agreed that online corpus consultation was useful for their academic/scientific writing. This shows that, despite the evidence that training and familiarity is essential when using corpus resources (Aşık et al. 2015), the same proportion of our informants still believed that online corpus consultation could help them to write in English in academic contexts after meeting the resource for a second time. This opinion is visibly in contrast with the informants’ more widespread acceptance of the usefulness of word lists for the same purpose. This finding corroborates the beneficial role of academic vocabulary lists for HE writing advocated by Durrant (2016). Also, our post-task questionnaires confirm overall positive attitudes towards corpus use (Yoon and Hirvela 2004) as long as the students remain engaged with classroom activities in the presence or under the guidance of a teacher.

Post-task questionnaires show that, after both WT1 and WT2, 25 percent of the informants actually disagreed that online corpus consultation was useful for their academic/scientific writing. This low endorsement was not found in the word lists, either after WT1 or WT2, which suggests that, at least for some of the HE researchers taking part in this experience, online corpus consultation was not as appealing as suggested by previous research. This finding suggests that an important number of L2 writers preferred support materials that do not demand their playing an overt active role, as DDL purportedly promote “the use of information technologies and active search strategies” (Pérez-Paredes et al. 2012: 483), and confirms previous reports on resistance to work with phraseological patterns (Hill et al. 2000; Coxhead 2008).

This apparent lack of appeal of corpus consultation resources happens to be so for 26.7 percent of the B1 researchers and for 20 percent of the B2–C1 researchers after both WT1 and WT2. However, no significant differences were found between the B1 and B2–C1 groups in this study. The fact that the percentage of rejection remains the same across the two levels of proficiency after completing the two writing tasks seems to point to the existence of a stable group of reluctant academic writers. Surely, hands-on consultation may not suit everyone. It seems reasonable to suggest that some of our subjects did not go past the initial corpus-query learning curve (Geluso and Yamaguchi 2014), which would have ensured a more autonomous use of corpora. Research experiences like Lee and Swales (2006) engaged students by means of activities that went on for a longer period of time, something which was not an option considering the specifications of the course provided by our HE institution. A longer exposure to corpus consultation and DDL may have had a positive impact on these researchers, as shown in Chang (2014), where students used corpora for 22 weeks. Similarly, more teacher support may have ensured more active engagement with the resources (Yoon and Hirvela 2004). However, this time span and allocation of human resources may not be feasible for most HE training initiatives and may call for a re-conceptualization of corpus work in the context of academic writing. As Pérez-Paredes and Alcaraz-Calero (2009) and Pérez-Paredes (2010) put it, the methodological transfer from corpus linguistics research to language learning has generally undergone no adaptation, and, as a result, learners are regularly presented with the same tools, corpora and analytical tasks as professional linguists. This approach may not simply work in most contexts, as it requires great investment in terms of time and understanding of how corpus linguistics methodology can be used. Leńko-Szymańska (2014: 260), working with would-be language teachers, concluded that “fourteen sessions (...) is not sufficient to encourage teacher trainees to use corpora in their future work if they have no contact with these resources and tools in other classes”. In other words, it takes a very long learning curve to become a proficient user of professional corpus linguistics tools and methods. Given the time frame of our course, it seems reasonable to suggest that this range of competences should be re-conceptualized in the context of pedagogically-friendlier DDL (Pérez-Paredes 2010).

5.2. What is the uptake of corpus consultation two years after the end of the training course on academic and scientific writing?

The lack of enthusiasm for online corpus consultation did not surface during the interviews in an explicit way, neither did the range of difficulties that are typically associated with corpus consultation. Subject A and C did not remember the resource. Once attention was drawn to it, A said she found corpus consultation useful to know more about how to convey ideas in academic contexts. C said that, after reconsidering its usefulness, corpus consultation was a very powerful academic resource. Subject D did not remember the resource either but, rather than offering reasons why she refused to make use of corpus consultation, she went on to discuss other tools that were actually of use in her English scientific writing. Subject B said that he most likely never became aware of the potential of the tool and admitted that he would consider using corpora for his professional writing from now on. This “never became aware” and lack of motivation to use corpus tools is unexpected in the light of research such as Charles (2014), which documents an extraordinary learners’ uptake of corpora a year after completing a course on corpus building and use. Charles (2014) reports that 70 percent of the 40 participants had used their corpus during that year. The fact that these users compiled their own corpus in their own realm of expertise and practice may have played a decisive role in their continuing to use corpus consultation for academic writing. This is shown by one of the case studies in Charles’ research, Ahmad, who is extremely keen on using his own compilation of RAs in his area of research. For him, it is a professional proofreading aid that helps him in some of the areas of writing that he has identified as problematic.

Like Ahmad in Charles (2014), researcher E explicitly endorsed the use of corpus consultation in his academic writing and described in positive, constructive ways how this resource contributed to enriching his writing experience, which corroborates Boulton’s (2011) claims that DDL may yield longer-term benefits for the language learner. Researcher E found a way to integrate corpus consultation into his everyday practice as he found it useful to “improve the translation of [his] research into English”. It must be stressed that translating into English (from Spanish) is common place in Spanish universities according

to Gea-Valor et al. (2014). In other words, researcher E did not significantly change the way he approached his academic writing, but he found that corpus consultation could actually improve the quality of his translations. Using Bordieu's terminology (Grenfell 2014), we may see this behaviour as *agency* examining *doxa* and acting in positive ways to improve everyday activities. This is exactly Ahmad's approach in Charles (2014). Ahmad is aware that articles in English is an area where he can still improve. A corpus of RAs is then of great help for him.

From a neo-Vygotskian perspective (Bax 2011), the attitudes of those that failed to use corpus consultation in the two-year period investigated can be explained in terms of engagement with the social group in which they manifest themselves. In other words, learning and development are socially motivated and happen in culturally formed settings. These ideas, initially in Mercer and Fisher (1997), may be instrumental in understanding how a given technology, corpus consultation in our case, may fail to make an impact on individuals that are very strongly tied to common, uniform practices across the board. Judging from the interview data, the rejection trend observed after WT1 and WT2 was somehow disguised as lack of interest. In this context, one of the issues that did not emerge from the interviews was the shortage of specialized corpora (Chang 2014) in the fields of academic and scientific practice of the interviewees, a fact which would partially explain why these subjects did not even consider online consultation as part of their writing routine. Similarly, although the issue was raised during the sessions, the subjects did not mention the possibility of building their own corpus and thus develop a sense of ownership through the control of its construction and consultation (Charles 2014). The second case study in Charles (2014), Piotr, the non-user, can be extremely useful when trying to understand this lack of interest in corpus consultation. For him, the main concern when writing is textual organization. Charles (2014: 38) put it this way: "[i]t should be recognised that take-up of the personal corpus may well depend on the student's perception of their main needs at any given time". Our results corroborate Charles' (2014) views on the incorporation of corpus consultation routines on existing writing behaviour.

The two-month delayed questionnaire, however, had painted a very different landscape, one where four of the interviewees had stated that corpus consultation was believed to be either very useful or useful. Subjects B and C said that they had found online corpus consultation very useful for their academic writing, D and E said that the resource was useful, while subject A found it slightly useful. These results are potentially consistent with, among others, Yoon and Hirvela's (2004) and Liu and Jiang's (2009) findings that learners improved their academic writing thanks to the use of corpora and, particularly, are consistent with general positive evaluations of corpus resources (Charles 2007; Smart 2014). But the delayed questionnaire also showed that our participants actually thought that there were other resources that were considered more useful for their English academic writing. In particular, reading RAs was very useful for 80 percent of the informants, by far the resource which received the highest score in the questionnaire. All in all, 84 percent considered online corpus consultation either useful or very useful, together with Google, the lowest score of all the resources evaluated. Particularly, reading RAs, a traditional recipe for academic writing (Connor 1999) across institutions where no specific training is available, was thought to be useful or very useful by every participant in the course. This is no surprising in the Spanish scientific context, where reading RAs is seen as a way to trace and learn from models. Our findings suggest that this prevailing socially motivated behaviour may prevent potential users of corpora from benefitting from corpus consultation as the culture of the institution seems to favour other types of learning (Mercer and Fisher 1997). This is corroborated by researcher D, who relied very heavily on the journal repository in the University to do her academic writing. Researcher A's claim that integration of corpus resources is difficult may be of interest when contrasted against researcher D's willingness to use corpora in her writing flow.

6. CONCLUSIONS

Corpus consultation was only used by one of our five informants in this study two years after the end of an academic writing course. Our results suggest that (i) language competence seems to play no role in assessing the usefulness of corpus consultation and (ii) we cannot totally understand corpus uptake if we do not fully appreciate the context – *doxas* – in which these users develop their academic writing in English. This approach has been particularly neglected in the past by researchers that have tried to focus on the affordances of corpus consultation instead of exploring how these uses were integrated in their learning or professional practice. More research is needed to understand how the uptake of corpus methods is affected by the writing culture of the institutions of the writers.

Despite the limited number of subjects, our interview data seems to imply that these researchers have not been properly trained to write in English (Gil Ortega 2002) and, what is more, they lack a proper methodology to do so. This is consistent with the findings in Gea-Valor et al. (2014) and points to an

important lack of awareness of conventions in the academic staff of Spanish universities. In particular, the interview data suggest that most of these researchers tend to rely on RA reading as writing model (Connor 1999). This model role did not emerge explicitly in most of the delayed interviews. Although researcher A brought up the issue of non-nativeness and found in corpora a way to come up with the “right expression”, researcher B referred to the more theoretical guidelines offered in the course as a more productive way to gain insight into academic writing. Similarly, D used online research databases to find published research that could, in turn, help her own writing, although more precise insights are missing in the interview data. Even when generally speaking models are discussed by the participants, there is no reference whatsoever to how language is accessed by these writers.

Our data need to be interpreted in a context where, despite the enormous pressure to publish in English, there is a lack of both institutional and individual appreciation of the complexities involved in writing in a foreign language (Gea-Valor et al. 2014). Accordingly, it may not be possible to generalize these results to other different HE contexts. The limited number of participants (n=25) as well as the length of the course (12 hours) may have undoubtedly conditioned our results. However, it is necessary to stress that the mixed-methods used in our research has allowed us to go beyond the often-used LIKERT questionnaire. Similarly, to the best of our knowledge, this may be the research that has examined the longest period of delayed uptake of corpus consultation habits, which offers new perspectives on the analysis of how these resources are employed by HE researchers. Other limitations of this study lie in the fact that we cannot be positively certain that all 25 researchers used the resources provided (word lists and corpora) during the training sessions, as previous tracking methods have shown that not all participants actually interact with the corpus resources provided by the instructors (Pérez-Paredes et al. 2011).

Our results may suggest that reading other RAs to learn how to write in English may effectively prevent these researchers from implementing more active ways to integrate corpus consultation in their everyday science making. Our post-task questionnaires and interview data informants’ data suggest that, for some of these potential users of corpus resources, attention to accumulated frequency and patterning (Flowerdew 2015) may not be of interest for their academic writing. This is supported by Pérez-Paredes et al. (2011) and particularly by Pérez-Paredes et al. (2012), who found that despite corpus guidance and training only a group of learners, 46 percent to be precise, showed a tendency towards language research-oriented behaviour. Further research should explore the relationships between the writing routines of these researchers and more effective integration of corpora and, particularly, how HE institutions may integrate corpus resources in the teaching of writing skills in feasible ways.

REFERENCES

- Altenberg, Bengt. 1987. Causal ordering strategies in English conversation. In James Monaghan ed. *Grammar in the construction of texts*. London: Frances Pinter, 50–64.
- Aşık, Asuman, Arzu S. Vural and Kadriye D Akpınar. 2015. Lexical awareness and development through Data Driven Learning: attitudes and beliefs of EFL Learners. *Journal of Education and Training Studies* 4/3: 87–96.
- Basanta, Carmen P. 2005. Assessing the receptive vocabulary of Spanish of English Philology: an empirical investigation. In José Luis Martínez-Dueñas Espejo coord. *Towards an understanding of the English language, past, present and future: studies in honour of Fernando Serrano*. Granada: Editorial Universidad de Granada, 545–564.
- Bathia, Vijai, K., Purificación Sánchez-Hernández and Pascual Pérez-Paredes eds. 2011. *Researching specialized languages*. Amsterdam: John Benjamins.
- Bax, Stephen. 2011. Normalisation revisited: the effective use of technology in language education. *International Journal of Computer-Assisted Language Learning and Teaching* 1/2: 1–15.
- Belcher, Diane and Adam Hirvela. 2005. Writing the qualitative dissertation: what motivates and sustains commitment to a fuzzy genre? *Journal of English for Academic Purposes* 4: 187–205.
- Boulton, Alex. 2009. Testing the limits of data-driven learning: language proficiency and training. *ReCALL* 21/1: 37–54.
- Boulton, Alex. 2010a. Data-Driven Learning: taking the computer out of the equation. *Language Learning* 60/3: 534–572.
- Boulton, Alex. 2010b. Leaning outcomes from corpus consultation. In María Moreno Jaén, Fernando Serrano Valverde and María Calzada Pérez eds. *Exploring new paths in language pedagogy: lexis and corpus-based language teaching*. London: Equinox, 129–144.
- Boulton, Alex. 2011. Data-Driven Learning: the perpetual enigma. In Stanislaw Gozdz-Roszkowski ed. *Explorations across languages and corpora*. Bern: Peter Lang, 563–580.

- Boulton, Alex and Pascual Pérez-Paredes. 2014. ReCALL special issue: researching uses of corpora for language teaching and learning. *ReCALL* 26/2: 121–127.
- Boulton, Alex, Shirley Carter-Thomas and Elizabeth Rowlye-Jolivet eds. 2012. *Corpus-informed research and learning in ESP. Issues and applications*. Amsterdam: John Benjamins.
- Chambers, Angela and Ide O’Sullivan. 2004. Concordancers in the EFL classroom. Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning* 16/1: 83–94.
- Chang, Ji-Yeon. 2014. The use of general and specialized corpora as reference sources for academic English writing: a case study. *ReCALL* 26/2: 243–259.
- Charles, Maggie. 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.
- Charles, Maggie. 2014. Getting the corpus habit: EAP students’ long-term use of personal corpora. *English for Specific Purposes* 35/1: 30–40.
- Cobb, Tom. 2019. Corpus Concordance English. <<https://www.lexutor.ca/conc/eng>> (accessed December 2018).
- Cohen, Louis, Lawrence Manion and Keith Morrison. 2011. *Research methods in education*. London: Routledge.
- Connor, Ulla. 1999. Learning to write academic prose in a second language: a literacy autobiography. In George Braine ed. *Non-native educators in English language teaching*. Mahwah, NJ: Erlbaum, 29–42.
- Cortes, Viviana. 2004. Lexical bundles in published and student disciplinary writing: examples from history and biology. *English for Specific Purposes* 23: 397–423.
- Coxhead, Averil. 2008. The best practice is practice. In Adel Jendli and Christine Coombe eds. *Best practice in English language teaching*. Dubai, UAE: TESOL Arabia, 3–14.
- Dornyei, Zoltan. 2007. *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Durrant, Philip. 2016. To what extent is the Academic Vocabulary List relevant to university student writing?. *English for Specific Purposes* 43: 49–61.
- Ferguson, Gibson, Carmen Pérez-Llantada and Ramón Pló. 2011. English as an international language of scientific publication: a study of attitudes. *World Englishes* 30/1: 41–59.
- Flowerdew, Lynne. 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: 56–68.
- Flowerdew, Lynne. 2016. A genre-inspired and lexico-grammatical approach for helping postgraduate students craft research grant proposal. *English for Specific Purposes* 42: 1–12.
- Flowerdew, John and Richard Forest. 2010. Schematic structure and lexico-grammatical realization in corpus-based genre analysis: the case of research in the PhD literature review. In Maggie Charles, Diane Pecorari and Susan Hunston eds. *Academic writing at the interface of corpus and discourse*. London: Continuum, 15–36.
- Gea-Valor, María L., Jesús Rey-Rocha and Ana Moreno. 2014. Publishing research in the international context: an analysis of Spanish scholars’ academic writing needs in the social sciences. *English for Specific Purposes* 36: 47–59.
- Geluso, Joe and Atsumi Yamaguchi. 2014. Discovering formulaic language through data-driven learning: student attitudes and efficacy. *ReCALL* 26/2: 225–242.
- Gil Ortega, Carmen. 2002. Spanish students in British Universities: developing a support resource to improve their academic writing skills. Paper presented at the ‘Setting the Agenda: Languages and Area Studies in Higher Education’ Conference. Manchester, 24–28 June.
- Glasman-Deal, Hilary. 2010. *Science research writing for non-native speakers of English*. London: Imperial College Press.
- Grenfell, Michael J. ed. 2014. *Pierre Bourdieu: key concepts*. London: Routledge.
- Hanauer, David I. and Karen Englander. 2013. *Scientific writing in a second language*. Anderson, South Carolina: Parlor Press.
- Hill, Jimmie, Morgan Lewis and Michael Lewis. 2000. Classroom strategies, activities and exercises. In Michael Lewis ed. *Teaching collocation: further developments in the lexical approach*, Boston, Mass.: Cengage Learning, 88–117.
- Hyland, Ken. 2008. Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics* 18: 41–62.
- Huang, Zeping. 2014. The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL* 26/2: 163–183.
- Hyland, Ken. 2013. Writing in the university: education, knowledge and reputation. *Language Teaching* 46/1: 53–70.
- Johns, Tim. 1991. Should you be persuaded: two examples of data-driven learning materials. *English Language Research Journal* 4: 1–16.

- Karras, Jacob N. 2016. The effects of Data-Driven Learning upon vocabulary acquisition for secondary international school students in Vietnam. *ReCALL* 28/2: 166–186.
- Kaufhold, Kathrin. 2015. Conventions in postgraduate academic writing: European students' negotiations of prior writing experience at an English speaking university. *Journal of English for Academic Purposes* 20: 125–134.
- Kennedy, Claire and Tiziana Miceli. 2010. Corpus-assisted creative writing: introducing intermediate Italian learners to a corpus as a reference resource. *Language Learning & Technology* 14/1: 28–44.
- Lee, David and John Swales. 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.
- Leńko-Szymańska, Agnieszka. 2014. Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. *ReCALL* 26/2: 260–278.
- Liu, Dilin and Ping Jiang. 2009. Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal* 93/1: 61–78.
- Mercer, Neil and Eunice Fisher. 1997. The importance of talk. In Rupert Wegerif and Peter Scrimshaw eds. *Computers and talk in the primary classroom*. Clevedon: Multilingual Matters, 13–21.
- Mertens, Donna M. 2005. *Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative, and mixed methods*. 2nd ed. Thousand Oaks, CA: Sage.
- Pérez-Paredes, Pascual. 2010. Corpus linguistics and language education in perspective: appropriation and the possibilities scenario. In Tony Harris and María Moreno Jaén eds. *Corpus linguistics in language teaching*. Bern: Peter Lang, 53–73.
- Pérez-Paredes, Pascual and José M. Alcaraz-Calero. 2009. Developing annotation solutions for online Data Driven Learning. *ReCALL* 21/1: 55–75.
- Pérez-Paredes, Pascual and Purificación Sánchez-Hernández. 2005. Examining English for academic purposes students' vocabulary output: corpus-aided analysis and learner corpora. *Revista Española de Lingüística Aplicada* Extra 1: 201–212.
- Pérez-Paredes, Pascual, María Sánchez-Tornel and José M. Alcaraz Calero. 2012. Learners' search patterns during corpus-based focus-on-form activities. *International Journal of Corpus Linguistics* 17: 483–516.
- Pérez-Paredes, Pascual, María Sánchez-Tornel, José M. Alcaraz Calero and Pilar Aguado Jiménez. 2011. Tracking learners' actual uses of corpora: guided vs non-guided corpus consultation. *Computer Assisted Language Learning* 34/3: 233–253.
- Smart, Jonathan. 2014. The role of guided induction in paper-based data-driven learning. *ReCALL* 26/2: 184–201.
- Staples, Shelley, Jesse Egbert, Douglas Biber and Bethany Gray. 2016. Academic writing development at the university level: phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication* 33/2: 149–183.
- Staples, Shelley, Jesse Egbert, Douglas Biber and Alyson McClair. 2013. Formulaic sequences and EAP writing development: lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes* 12/3: 214–225.
- Swales, John. 1990. *Genre analysis: English in academic and research setting*. Cambridge: Cambridge University Press.
- Tribble, Christopher. 2002. Corpora and corpus analysis: new windows on academic writing. In John Flowerdew ed. *Academic discourse*. London: Longman, 131–149.
- Yoon, Hyunsook. 2008. More than a linguistic reference: the influence of corpus technology on L2 academic writing. *Language Learning & Technology* 12/2: 31–48.
- Yoon, Hyunsook and Alan Hirvela. 2004. ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing* 13/4: 257–283.

Corresponding author

Pascual Pérez-Paredes
Faculty of Education · Cambridge Language Sciences
9 West Road
Cambridge CB3 9DP
e-mail: pfp23@cam.ac.uk

received: November 2018

accepted: December 2018

Appendix 1

[From Glasman-Deal, Hilary. 2010. *Science research writing for non-native speakers of English*. London: Imperial College Press.]

Writing abstracts

BACKGROUND INFORMATION

a number of studies exist(s)
frequently generally
is a common technique is/are assumed to is/are based on
is/are determined by is/are influenced by is/are related to
it has recently been shown that it is known that
it is widely accepted that occur(s)
often popular produce(s)
recent research recent studies recently
recently-developed

STATEMENT THESIS AIM

in order to our approach
the aim of this study to compare
to examine to investigate to study
with the aim of

PROBLEM

(an) alternative approach a need for
although complicated desirable difficulty disadvantage drawback essential expensive however impractical
inaccurate inconvenient
it should be possible to limited
not able to problem require risk
time-consuming
unsuccessful

WHAT THE PAPER DOES

In this study/paper/investigation, we address
analyse argue compare consider describe discuss emphasise examine extend introduce present propose
review show
This study/paper/investigation considers
describes examines extends includes presents reports reviews

SUMMARY OF METHODOLOGY/MATERIALS

was/were assembled was/were calculated was/were constructed was/were evaluated was/were formulated
was/were measured was/were modelled was/were performed was/were recorded was/were studied
was/were treated was/were used

RESULTS

caused
decreased had no effect increased
it was noted/observed that ... occurred
produced resulted in was identified
was/were achieved was/were found was/were identical was/were observed was/were obtained was/were
present was/were unaffected (by) yielded

ACHIEVEMENT/CONTRIBUTION

better consistent effective enhanced exact improved new novel
significant simple suitable superior achieve allow demonstrate ensure guarantee obtain validate
compare well with for the first time in good agreement

IMPLICATIONS

The evidence/These results ... indicate(s) that
mean(s) that suggest(s) that it is thought that
we conclude that
we suggest that can

APPLICATIONS

applicability can be applied can be used
make it possible to potential use relevant for/in

Appendix 2

Guidelines to query and interpret concordance lines

http://www.lextutor.ca/concordancers/concord_e.html

[Home](#) > [Concordancers](#) > Corpus search input (Eng)

Corpus Concordance English (v.6.5)

With precast link extractor/exporter

[French](#) [German](#) [Spanish](#) [English](#)

Keyword(s): In corpus: [Corpus descriptions](#)

CONTROLS :

Sort By word(s) to of keyword | Line Width Number of Lines Gapped?

OPTION : With associated within 4 words to side.

+ Scan for any recurring word (potential colloc.) within words presenting <= times

DEMOS : <<3+4 have new sort info >> |

Link Extractor [As discussed here](#)

1. Asegúrate que el corpus que estás consultando es el de “Academic Abstracts”.
2. Introduce la palabra o palabras que quieras consultar.
3. Obtén las concordancias.

Un ejemplo. En nuestro *abstract* queremos hacer referencia a los resultados obtenidos en nuestra investigación y, obviamente, vamos a usar la palabra *results* para ello. Pues bien, al buscar *results* en el corpus vamos a obtener unas líneas de concordancia donde *results* aparece como eje central o nodo. Esto nos va a permitir averiguar cómo se relaciona *results* con otras palabras y estructuras, esto es, estamos accediendo a información de índole no solo léxica sino también sintáctica, algo que, desafortunadamente no aprendemos cuando estudiamos el idioma a menos que hagamos un uso muy intensivo del mismo. Una vez generadas estas concordancias podríamos deducir lo siguiente:

- *results* va seguido de verbos como: *indicated, provide, support, indicate, obtained, showing, suggest*
- *results* va seguido de preposiciones como *from/of*
- si sigue *in*, es un verbo
- *results* va precedido de adjetivos como *previous* y *numerical*.

Todo esto nos da un perfil léxico-gramatical del nodo en cuestión, ya que podemos caracterizar la red de relaciones léxicas y gramaticales de una palabra o una cadena de palabras en concreto.

Appendix 3

Post-writing task 1 questionnaire

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I have found the list of words of great help when writing the abstract					
I have found the corpus tools of great help when writing the abstract					

Post-writing task 2 questionnaire

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I have found the list of words of great help when writing the abstract					
I have found the corpus tools of great help when writing the abstract					

Appendix 4

Delayed questionnaire

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Overall, I have found the resources provided in the course of great help when writing both the abstract and the conclusions					

	Slightly useful	Useful	Very useful
Overall, I have found dictionaries of great help when improving my academic writing skills in English			
Overall, I have found reading RAs of great help when improving my academic writing skills in English			
Overall, I have found studying EFL of great help when improving my academic writing skills in English			
Overall, I have found Google search of great help when improving my academic writing skills in English			
Overall, I have found online corpus consultation of great help when improving my academic writing skills in English			

Appendix 5

Structured interview questions

1) Have you used this corpus or other corpora to help you with your academic or scientific writing in English after the course?

How often?

Have you talked about this resource with other researchers?

Would you recommend this resource?

2) If you have used this resource, which aspects do you like best about corpus consultation as compared to other resources such as dictionaries or vocabulary lists?

3) If you have not used this resource, could you explain why you have not found it useful?

Despite not using it, which aspects of this resource do you consider as most helpful for academic writing?