Servicio de Publicaciones y Difusión Científica
Universidad de Las Palmas de Gran Canaria

Form-Focused Data-Driven Learning for Grammar Development in ESP Contexts

Alejandro Curado Fuentes[1]

Universidad de Extremadura

ABSTRACT

The use of corpus-informed language in English for Specific Purposes (ESP) / English for Academic Purposes (EAP) teaching has considerably grown over the past two decades, especially due to the proliferation of course book material based on corpora. However, in terms of learning approaches to grammatical items, much less has been published commercially to train teachers in the use of corpora and their exploitation. DDL (Data-Driven Learning) techniques are still something to foster and publicize for ESP scenarios, since DDL tends to work as a productive method among learners whose L2 courses usually present specific (academic or professional) purposes in language use, as previous studies have shown. This paper deals with DDL as approached via hands-on concordancing by university students in the double degree of Business Administration and Tourism (B/T) in our institution. The targeted linguistic items were accessed via form-focused activities (e.g., noticing usage differences with verb tenses and synonymous verbs). Two grammar tests were administered to compare their results with another group that focused on the same grammatical points as delivered by an upper-intermediate, corpus-informed, textbook. The statistical findings reveal that the DDL group performed slightly better in the acquisition of such grammatical points, whereas most students felt that lexico-grammatical knowledge had been gained by relying on DDL to discover and apply usage patterns. Grammatical study may thereby evolve as a different perception in their minds due to corpus data exploitation and teacher-mediated discovery.

Keywords: DDL, ESP, B/T, corpora, form-focused teaching, teacher mediation

RESUMEN

El uso de un lenguaje informado por corpora en clases de Inglés para fines específicos (IFE) / Inglés para fines académicos (IFA) ha aumentado considerablemente en las últimas dos décadas, sobre todo como consecuencia de la proliferación de libros de texto y material académico que se basan en corpora. Sin embargo, en cuanto a cómo se enfocaría el aprendizaje de esos elementos gramaticales, mucho menos se ha publicado de manera comercial de cara a la formación de los profesores en el uso de corpora y su explotación en

[1] Facultad de Empresa, Finanzas y Turismo. Avda. Universidad, S/N, Universidad de Extremadura, Cáceres 10071. Spain.

Email: acurado@unex.es

clase. DDL (Data-Driven Learning, o aprendizaje basado en datos, ABD) constituye aún un terreno donde hay que fomentar y publicitar más su aplicación para escenarios de IFE, ya que tiende a servir como técnica metodológica productiva para los discentes que buscan objetivos concretos en el uso de un idioma, como han mostrado diferentes estudios. Este trabajo trata sobre ABD y su enfoque concreto a través de concordancias electrónicas por parte de estudiantes universitarios procedentes del doble grado de ADE y Turismo en nuestra institución. Los elementos lingüísticos que se trataron fueron explotados en actividades centradas en la forma/estructura (en este caso, la diferenciación de uso entre tiempos verbales y entre verbos sinónimos). Se administraron dos tests gramaticales para comparar los resultados de este grupo de alumnos con otro grupo que se centró en los mismos elementos gramaticales según se presentaban y trabajaban en un libro de texto de nivel intermedio-avanzado e informado por datos de corpora. Los resultados estadísticos revelan que el grupo de ABD consiguió puntuaciones ligeramente superiores en la adquisición de esos puntos gramaticales, mientras que la mayoría de los estudiantes manifestaron que habían incrementado su conocimiento léxico-gramatical al haberse apoyado en ABD a la hora de descubrir y aplicar patrones de uso. El enfoque gramatical podría así evolucionar como una percepción diferente en sus mentes gracias al apoyo en datos de corpus y la inducción, colaborando con otros compañeros y el profesor.

1. Introduction

DDL (Data-driven Learning) has become an interesting approach to explore language and discourse in the classroom (e.g., Johns, 2002; Gavioli, 2006; O'Keefee et al., 2007; Charles, 2011; Boulton, 2012; Lenko-Szymanska & Boulton, 2015). Albeit most typically associated with on-line corpora use and application in learning situations, DDL may also involve hands-off concordances (e.g., printed excerpts from corpora), or even isolated corpus-driven examples, to help students with linguistic / grammatical points (Frankenberg-Garcia, 2014; Chujo et al., 2015).

Another important advantage analyzed in DDL is its tendency to foster a learner-centered focus, where autonomy, discovery, peer / teacher collaboration, inductive reasoning, and noticing are significant traits (Aston, 2001; Scott & Tribble, 2006; Boulton, 2009; Boulton, 2012). In language terms, this affordance generally allows for "extending or deepening knowledge of existing linguistic items, distinguishing close synonyms, detecting patterns of usage, collocations, colligations, morphology, and so on" (Boulton, 2009, p. 4). The higher the learners' language command, the handier the strategies seem to become for them, as previous studies have explored (Pérez-Paredes & Cantos-Gómez, 2004; Curado-Fuentes, 2007b; Charles, 2015). The noticing approaches to linguistic form (Robinson, 1995; Skehan, 2001) pointed that observation too, since, in them, an important prerequisite for language intake is noting, observing, and closely heeding linguistic items with a focus on form.

A form-focused approach may consist of different types of activities and tasks that may be of interest within a communicative methodology in the classroom, e.g., self- and peer- corrections in written form, oral face-to-face interactions with teachers and peers (Nassaji, 2000; Williams, 2001), and explicit concordance-based induction and discussion that may facilitate learner-centered focuses (Aston, 1997; Bernardini, 2001).

When the focus was / is made on grammatical forms (e.g., teaching verb tenses) according to the traditional form-focused approach, the teacher acts as a rule-knower and provider, and thus, he/she, acting this way, tends to be the opposite of what DDL strategies suggest for the L2 classroom. In Spain, too much concentration on the deduction of grammatical rules for acquisition (cf. Pérez-Basanta, 2005) has in fact often eschewed that type of learning fostered by DDL, and, among other things, the realization that vocabulary and grammar often go together (Pérez-Basanta, 2005, p. 547).

This paper describes the integration of DDL in a case study[2] with English for Specific Purposes (ESP) students in the double degree of Business Administration and Tourism studies (B/T) at University of Extremadura. The teaching scenario is the exploitation of three grammatical points, aiming to empower learners with a command for differentiating important language use-related grammatical features, such as the future (perfect and simple), past (past continuous in active and passive voices), and the synonyms *convince* and *persuade*. The three points were examined with a control group of 25 students by means of the upper-intermediate course book *Natural English*, a corpus-informed ESL resource. In contrast, the same three points were exploited with an experimental group of 25 other students at the computer lab. The lab students used COCA (Corpus of Contemporary American English; Davies, 2008-) for the extraction and analysis of concordances from specific texts and registers.

Before and after the grammar-targeted sessions, two tests presenting the three grammatical points were administered (the first one, one week before, and the other, three weeks after the sessions). The aim was to compare intra- and inter-group evolvement by evaluating scores and processing them statistically. The findings point to positive observations for both groups. However, the experimental group outperformed the control group with slightly higher marks. Some questionnaires were also handed out to the learners (a general one for all the students, and a more specific handout about DDL for the experimental students). Among other observations, most surveyed learners put forward that DDL appeals to them and can be profitable for grammatical development, and few in this group differed with negative impressions.

The following section will describe our Theoretical framework for the contextualization and justification of DDL in ESP. Then, the Case study section will describe the learning scenario, tools, and activities used, followed by the Results, where the test scores, group performance, and survey information are discussed. Finally, our Conclusions will identify key analytical aspects derived from this study, which may be extended and revised, possibly via other types of learners, and / or in other academic contexts.

2. Theoretical Framework

Starting in the late 1980s and early 1990s, important editorial projects (e.g., Baker et

---

[2] This study is an in-depth extension of the summary published by Curado-Fuentes (2016).

al., 1993) were based on large amounts of corpus data with the underlying goal of actually putting into practice the view that grammar and lexis could be approached simultaneously. The focus on collocations and colligations, and not on grammatical concepts for the analysis of structural and semantic levels of text and discourse (cf. Hoey, 1991; Hoey, 2005), also drew a line that separated from previous approaches based on explanations and dissections of "higher" grammar rules. As Hunston (2002; pp. 147-148) explained, the traditional approach was something that language analysis using corpora for teaching and learning could not target as an ultimate end.

Lexico-grammatical patterns in corpora have been often analyzed in connection with teaching / learning purposes (e.g., Hunston and Francis, 2000). The patterns of use that certain words feature across and / or within specific texts and registers seem to have opened the eyes of different researchers in classroom situations for the development of DDL as a possibly useful tool for language learning. In this scope, the validity of DDL as one more means of classroom instruction can be evaluated for language skills exploitation at different levels (Boulton, 2010; Huang, 2014; Smart, 2014).

Huang (2014) observed that academic writing could benefit from the learners' overt attention to how some abstract nouns typically behaved in those academic texts. He found that the students could gain usage knowledge from the corpus data in order to make lexico-grammatical corrections and improvements in their choice of word combinations when revising their compositions. Smart (2014) worked with the (re)construction of the passive voice after some ESL (English as a Second Language) students explored teacher-mediated hands-off concordances (in paper form), other students used corpus-informed textbook materials, and others non-corpus informed grammatical exercises. His findings revealed that greater improvement in the acquisition of the passive voice was made by the first group, compared to the rest of the learners.

DDL seems to work as a productive type of instruction in ESP. Learners majoring in disciplines other than English Philology may feel that they are being given "serious things to do" with the foreign language (Gavioli, 2009, p. 46). Using on-line tools and concordances constitutes a learning task in itself (Mauranen, 2004, p. 99), while the fact that the ultimate decisions on word meaning and use are left to the students may lead to nudging their sense of responsibility and cooperation for project-oriented work. This aspect of authenticity not only of language use but of the content and purposes of the tasks tends to work well, as Johns (1988) already observed before the use of computers. It was a good thing, in his view, that learning might result from "authenticity of script, authenticity [authenticity] of purpose, and an authenticity of activity" (Johns, 1988, p. 10).

Cheng et al. (2003) have appealed to academic competence enhancement as an advantage of using pattern induction and discovery in ESP. The language thus retrieved is examined as specific and relevant to the learners' own field (Cheng et al., 2003, p. 181). ESP students may also recognize advantages in corpus consultation not only for language use but for subject content exploitation. Undoubtedly, learning styles and preferences may vary among individuals, but a key nexus seems to have

been the weight posed by the discipline in the academic context (Flowerdew, 2001; Curado-Fuentes, 2007a). DDL would unlikely appeal, at least initially, in less interactive academic situations (e.g., lecture-led courses) or among those students who prefer to learn from rules instead of from data (Flowerdew, 2008, pp. 116-117).

Learning styles may vary according to types of disciplines and / or individuals. However, Pashler et al. (2008, p. 105) state that "the literature on learning styles is enormous" and that "very few studies have even used an experimental methodology capable of testing the validity of learning styles applied to Education". In ESP settings, e.g., English for Computer Science, some students tend to prefer to work individually with corpus-driven information as part of their final degree projects (Curado-Fuentes, 2007a), and yet, they can eventually cooperate and share information from concordances.

Previous local experiments (Curado-Fuentes, 2007b; Curado-Fuentes, 2015) also suggest the importance of specific tailor-made corpus affordances. i.e., *ad hoc* corpora, designed by the teacher or in cooperation with higher-level students to integrate content and language that is relevant to the academic context (Seidhofer, 2000; Charles, 2011). In addition, by using specific texts and registers, learners' sensitivity to collocation-related phenomena could be spurred via explicit frequency charts that may compare with the students' pre-existing linguistic knowledge (Huang, 2014).

We thus find that discipline-related corpus information via significant concordances, pre-arranged and selected in advance by the teacher, can lead to good results in fostering students' awareness of appropriate language use. Linguistic hypotheses can thus be tested and resolved via linguistic meaning exploration. Thomas (2015) finds this observation to apply in his case studies, where high-level students succeeded to produce effective lexical-grammatical associations within academic contexts. For instance, a word like "scholarship" could become primed in the association "a scholarship + be awarded [by an institution] [to a student] to study [a subject/skill] [somewhere]" (Thomas, 2015, p. 96).

Among lower-level language learners (e.g., B1, the level of most of the participants' in this study), an inadequate grammatical knowledge often hampers students' ability to construct coherent texts (both in written and oral form). University students thus tend to blame grammar as the reason for ineffective communication, or they may also view that traditional approaches made in the past to grammar (explanation of rules and mechanical, deductive work) were not productive enough (Pérez-Basanta, 2005).

Because our hypothesis is that DDL activities and tasks may trigger an effective focus on grammar, the following case study should examine real data. At lower-level stages of language learning, such procedures can be adapted, such as the one analyzed by Chujo et al. (2015), who chiefly relied on adapted corpus-driven sentences to offer students ease of use. The examples are then more familiar, leading to the realization of the connection between linguistic data and grammatical rules. The lower-level students in our experiment are mainly different from those of Chujo et al's (2015) in that they are operating in an ESP framework where formal and professional texts are

favored. In addition, on-line tools, ad-hoc corpora, and hands-on interfaces have become significant elements for the activation of digital interactivity in classroom situations. This blended scope of L2 learning across tertiary education scenarios thus seems increasingly suitable for the integration of DDL.

3. The Case Study

The 50 students that comprised this ESP course-based case study were third-year undergraduates in the double degree of Business Administration and Tourism (B / T) at University of Extremadura (Cáceres, Spain). 36 students were female and 14 male subjects. Only one student was not a Spanish native (Erasmus student from France, male). Their age ranged from 20 to 23, and their linguistic level was mostly B1, with 4 students at B2, and 5 others at A2. Their linguistic level was measured at the beginning of the course by having learners take on-line, freely available, standardized grammar tests at englishteststore.net, scaled to the Common European Framework of Reference for Languages (CEFRL) levels (Council of Europe, 2003).

The subject taken was Specific English III, the third compulsory ESP subject in their studies. The students carried out the grammar exercises during the first semester of the 2015 / 16 academic year. The experiment lasted four weeks, from early November to early December of 2015, and was described to them by the teacher as an obligatory task on which a substantial part of their participation grade (30 percent) would be based. This criterion (exam-oriented) has been found to tend to increase learners' engagement (cf. Milton, 1996) with the task. Also, this initial goal-oriented proposal in the course could have led to encouraging positive attitudes (Curado-Fuentes, 2015).

At this stage of B/T studies, the students are by and large familiar enough with specific texts in their discipline, but not as much with reading general English texts. Their reading scores tend to be higher than in previous years when business topics are discussed. However, in the case of writing tasks, they tend to make many linguistic mistakes, and some are quite basic, even corresponding to lower levels (e.g., lack of subject to verb correspondence, adjective order, and so forth). Needless to say, and as the description of skills in the CEFRL shows, each level may denote higher or lower fluency depending on the different skills exploited. Underdeveloped production (writing and speaking) point to, among other things, a lower grammatical competence, which is true in our setting.

Still, grammatical exercises are seldom carried out in our ESP courses, not even to revise and / or to clarify linguistic errors. Instead, self-study is recommended via an in-house Moodle platform, where remedial grammatical exercises are collected and displayed for learners to use. In addition, no specific course book is followed; content derives from multiple resources that directly or indirectly provide ESP tasks and activities in business / tourism contexts. In this type of curricular approach, an overt inclusion of grammatical exercises would deviate from the course aims and may cause organization problems within the scheduled semester.

Therefore, the integration of grammatical points in Specific English III was exceptional,

carried out in the Fall semester of 2015 with two groups of students: In one, 25 students (chosen in alphabetical order) were taken to the lab for grammatical activities via DDL (this was the EG, i.e. experimental group). The other 25 students (next people on the alphabetical list, henceforth the CG, control group) did some exercises with the same grammatical items as delivered by an upper-intermediate textbook, *Natural English Upper / Intermediate*, edited with corpus-informed language (Gairns and Redman, 2012). For two consecutive weeks (6 classroom hours), DDL was carried out in the computer lab for hands-on (on-line) concordancing with the Corpus of Contemporary American English (COCA). This group came to the lab with the instructor / author of this paper, while, for the same time period, in parallel sessions with a different instructor, the other group dealt with various course book grammar exercises in conventional classroom situations.

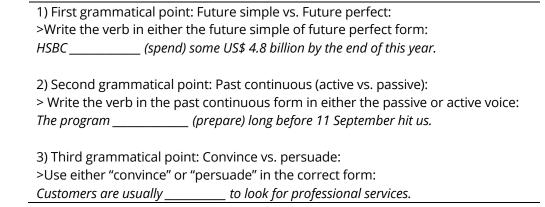| |
|---|
| 1) First grammatical point: Future simple vs. Future perfect:
 >Write the verb in either the future simple of future perfect form:
 HSBC _____ (spend) some US$ 4.8 billion by the end of this year.

 2) Second grammatical point: Past continuous (active vs. passive):
 > Write the verb in the past continuous form in either the passive or active voice:
 The program _____ (prepare) long before 11 September hit us.

 3) Third grammatical point: Convince vs. persuade:
 >Use either "convince" or "persuade" in the correct form:
 Customers are usually _____ to look for professional services. |

Table 1: Examples of sentences given in the tests.

The three grammatical points exploited and tested were: The future simple versus future perfect forms, the past continuous in the active versus passive voice, and the use of convince and persuade. Table 1 displays examples of sentences given in the two tests. Such examples were retrieved from on-line corpora related to business and economy, but not from the COCA corpus, so that the content would not be biased with a possible benefit for the EG students. Each test contained 10 sentences (5 for the first grammatical point, 3 for the second, and 2 for the third).

As mentioned, COCA was used with the EG students. In particular, the Academic and News sections (183 million words) were found suitable for the identification and selection of texts where the concordances could be run and explored. Such texts belonged to written economic and business sources. In addition, texts from the Social Sciences discipline were contrasted at some point in the activities to provide a wider range of examples within the academic register. Finally, the Spoken section of COCA was used in the last part of the DDL sessions.

Access to COCA was something new for all 25 students. They registered one week prior to the sessions, following explicit guidelines posted in the virtual section for the course. This way, on the first day at the computer lab, we could proceed to the main search features, followed by the identification of tagging procedures. The focus was placed on frequency charts and collocates in the concordances. Surprisingly (by comparison with other courses), most students caught on quite fast, understanding

the purposes and the features of the tool, especially when clear examples of key business words were displayed. A successful item was the combination of adjectives before nouns, such as adjective + *customer*(s), which confirmed some students' guesses (e.g., *regular customer* and *satisfied customer*, found at the top of the COCA frequency tables).

During this first (90-minute) session, the first type of form-focused activity was already handed out in paper form. It asked learners to scan frequency charts for the detection of verbs that occurred most frequently in either passive or active forms in the past continuous. The POS (part of speech) tags were clearly specified to them at all times. Table 2 displays the aims and skills targeted with this first type of DDL activity.

Activity type # 1

<u>Aim</u>: General (Newspapers and Academic registers):
<u>Skills</u>: *Identifying structures targeted / Comparing them across registers*
E.g. -Identify ONLY past continuous forms (first active & then passive).
    -Give 12 examples: 3 in the active and 3 in the passive in news, and 3 in the active and 3 in the
    passive in academic texts.
    -Use the tags:
    • [be] [v.ng]
    • [be] being [v.en]

Table 2: Activity type 1 (sessions 1 and 2)

Since there were many more verb forms in the present continuous, the past continuous verbs had to be first detected and then compared in the two registers (e.g., *were being fired* as the first passive form in newspapers versus were being studied in the academic section). The use of concordances with these forms was then practiced in order to induce structural and semantic senses.

Type 2 (Table 3) dealt with collocation and colligation searches in specified POS (e.g., seeking prepositions before and after future tense forms), pattern interpretation (making sense of structural / syntactic meaning), and reverse translation (Spanish to English).

Activity type # 2:

<u>Aim</u>: Specific (Money-related news).
<u>Skills</u>: *Identifying specific P.O.S. as colligate / Exploring structural / syntactic meaning / Translating*
E.g. -Find prepositions before future perfect (*will have*) verb forms: Which five are most frequent?
    -Then identify 3 examples with different prepositions (3 total): Do the prepositions indicate any
    reference to time within the clause/sentence? If so, what is the meaning?
    -Translate the following sentence into English: *Para la hora que llegue el pedido, ya habremos*
    terminado la reunión.
    -Tags: will have [vn*] / collocates (4 left & right): [jj]

Table 3: Activity type 2 (sessions 2 and 3).

Because the selection of the texts, in this case, was narrow, i.e., restricted to the "Money" topic, the most frequent preposition (of) occurred only 10 times, followed by *by* (7), *in*, *on*, and *over* (2 each). These few lines of concordance could favor a clearer vision of future tense references made with the use of *by*, although some confusion also arose when comparing in within time clauses (e.g., *in a given year*). In any case, the teacher did not interfere with the final answers, but mediated by asking specific questions in the process that sought the clarification of what the colligations might mean. Reverse translation was then used as a way of testing their knowledge in a production exercise.

Type 3 activities (Table 4) involved the use of longer co-text such as text excerpts, and even whole paragraphs for the use of thorough information. Some examples from specific text types (in the academic register) led to grammatical pattern identification. For instance, it could be observed that the past continuous occurred with the verb *build* more prominently in the passive than in the active voice within Social Sciences texts.

Since only nouns (tag: [nn*]) were sought as a subject in this case (within such a restrained section of Social Science texts), few occurrences popped up (e.g., *hotel*, 3 times; *condominiums*, 2; and *others* only once). Again, fewer examples led to a generally better realization of the use of the linguistic form in a specific context.

Activity type # 3:
Aim: Specific (Social sciences texts).
Skills: *Solving specific queries with given verb forms / Using context / Illustrating*

E.g. What was / were being built? What happened while it / they was/ were being built?
   -Give 5 examples.
   -Tags: [vb*] being built / collocates (4 left): [nn*]

Table 4: Activity type 3 (sessions 3 and 4).

Type 4 exercises (Table 5) seem to have strengthened students' ability to detect patterns, as lexical behavior was scrutinized well across different sections (broader sections with smaller sets of texts, leading to linguistic variation identification). For example, the verb *persuade* was found to occur within active clauses more frequently than in passive statements across News texts. However, when the exercise focused on the texts from *The New York Times* alone, the opposite was observed.

Activity Type # 4:
Aims: General and specific (News register, International news & New York Times texts):
Skills: *Finding X with Y Patterns /   Discovering lexico-grammatical use / Translating*

E.g. Find "persuade" with "to" in newspapers. Can you identify the more recurrent voice (passive or active) with those words together? Which one?
   -Observe its use in *International news* and then focus on *The New York Times* texts only:
    Identify 5 subjects for those sentences. What things were they persuaded to do?
   -Translate into English: *A los clientes se les persuadió para que dejaran de fumar, pero el conferenciante les persuadió a ellos para que hicieran lo contrario.*
   -Tags: [persuade] / collocates: to (2 left & right)

Table 5: Activity type 4 (sessions 4 and 5)

Some results for broad section questions in Type 4 were numerous (e.g., more than 2,000 lines in the News section), and this overwhelming amount discouraged or intimidated some students. At this stage, some clear directions by the teacher were crucial to call for discerning capabilities and pattern detection. After comparing persuade in its active versus passive forms (on separate screens), frequencies could be recorded, and a direct prompt such as "well, are there more in the active, or more in the passive?" seemed to work to re-direct amidst the concordance "noise". Thus, getting them to think about frequency and then by observing repetition turned out to be productive as long as the pattern counts could be made. In contrast, with smaller amounts of texts, the situation was eased, as, for example, the preposition *to* coming after *be + persuaded* was seen as a clear colligate. In this regard, the clearer the pattern showed on the screen, the more likely the activity became effective for grammatical development.

Type 5 questions (Table 6) engaged learners in the use of other, less specialized, registers. Pattern behavior was compared by relying on specific collocates. The key here was for students to realize that frequencies could change according to the kind of register selected. For example, News sources were compared with Spoken texts by observing the pattern *convince* + someone + that. Students could then spot differences in terms of what topics or subjects were typically associated. For example, students compared political references in News with the references to miscellaneous topics in spoken texts. In another example, they could identify more formal items (News) versus informal / personal references (Spoken).

Activity type # 5:
Aims: Register comparison (general use):
Skills: *Reflecting on frequency & collocations / Detecting variation / Illustrating*

E.g. Find "[*convince*] + *Noun* + that" in News Magazines, then in All News, and then in the Fiction + Spoken sections: Which noun is most frequent in all cases?
    -Now, find 2 different nouns / pronouns that occur most frequently in subject position for each
      SECTION (6 total):
     ? + CONVINCE + *the noun found* + THAT: Observe word use differences, and based on
      them decide what sentences are more or less formal / academic: Give 1 distinct example for
      each use.
    -Tags: [convince] [nn*] that

Table 6: Activity type 5 (sessions 5 and 6).

A common noun in Type 5 was people in all the sections. This highly frequent collocation (i.e., *convince people that...*) made it easier for students to be able to distinguish patterns. The students also felt more confident and motivated because the findings provided them with a distinct solution, which tends to work as a positive

feature for them in the ESP context.[3]

4. Results

Performance varied among learners in both classroom situations (EG and CG). The course book grammatical activities, expectedly, became easier and more recognizable for the higher-level students (B1-to-B2). The same thing happened in the EG sessions. Even if the type of approach was different, a better grammatical command usually made students feel more at ease with the exercises. Likewise, a heterogeneous grammatical level showed through the pre- and post-tests. For example, in the pre-tests, two students in CG and three in EG scored 9 or above (from 1 to 10), while three in CG and four in EG, 3 or below. In the post-tests, three students in CG and four in EG obtained perfect scores (10), but three students in CG and two in EG obtained 4 or below.

Given those differences in scores, a statistical examination was deemed as necessary, but such a view could not be based on a normal distribution of the data. In fact, as score data appeared as non-normally distributed, it was decided that non-parametric tests be applied to measure both intra-group (within group performance differences) and inter-group (between group comparison) values. First, a sign test (Table 7) was chosen to indicate whether there was a significant improvement from their performance in pre-tests to that of post-tests (for both groups). And, secondly, a Mann-Whitney U-test (Table 8) was applied to show whether one group did better than the other in the post-tests.

N	x	p
EG 20	2	0.016
CG 23	6	0.022

Table 7: Sign test for intra-group scores

Post-tests	Z-ratio	U-value	p
50	-1.469	227	0.049

Table 8: Mann-Whitney U- Test for inter-group comparison

The sign test compared the score differences (N) between test performances, adjusting progress according to p-value significance at <0.05. Both groups scored significantly better in the post-tests according to this measurement. For the inter-group examination with all 50 students, the non-parametric test applied (Table 8) is similar to a parametric t-score test. It calculates the number of score wins (sum of score ranks with U-value) to find whether that overall score difference is significant. The result is that a better performance was slightly significant for EG, with p-value significance at <0.05.

The score statistics thus spoke out in favor of the grammar sessions in terms of performance. As an additional measurement, the students' own reactions,

---

[3] This type of observation, i.e., that DDL can be a good instrument for the adoption of a problem / solution model of learning within the ESP framework, agrees with previous studies (e.g., Boulton, 2012).

impressions and comments were collected in class. The two anonymous questionnaires included for further analysis were handed out in class for both groups at the end of the semester. The EG questionnaires were given out one week after the lab sessions, and six weeks before the general (both groups) questionnaires. Both were open-answer surveys that sought their impressions and honest opinions. Table 9 displays the summarized answers for the EG questionnaires.

Questions	Summarized answers
1. Did you find this task useful?	Yes (88%) / No (12%)
2. If yes to 1, what for in language terms?	Vocabulary (50%) Grammar (36%) Not sure / don't know (9%) Translation (4%)
3. What did you like most about it?	Comparisons (44%) Discovering (24%) On-line work (16%) Miscellaneous (16%)
4. What did you like least about it?	Technical issues (64%) Lines and texts (12%) Miscellaneous (24%)
5. When compared to other search tools, What advantages can you see, if any?	Better for language (40%) Examples (28%) Context discovery (16%) Interface (16%)
6. Did you learn grammar with this?	Yes (72%) / Not sure (12%)

Table 9: Summarized EG questionnaire answers

A fair amount of positive answers could be expected from the onset, as most students had already orally voiced their favorable impressions within the lab sessions, especially because the experience was something new in the course. What the information from the questionnaires seemed to add was that the EG students had especially had a positive experience with the comparison of sections (Activity types 1 and 5), and also, that guess work had led students to pattern discovery and linguistic knowledge (point 3 in Table 9).

In contrast, the survey seems to indicate that the learners had found as less favourable the need for inserting POS tagging and collocation information (point 4 in Table 9). This had been something that, indeed, two students had already pointed out during the sessions, as they argued that things should be simplified in the interface, and that the exploration of concordances should improve. Such remarks may be interpreted as a demand that DDL activities be made more user-friendly and learner-autonomous. On the other hand, the advantage of having the teacher nearby was

conveyed, which seems to reveal that teacher mediation is acknowledged and appreciated.

As regards the miscellaneous answers given for questions 3 and 4 (Table 9), they varied; but overall, they seemed to indicate a preference for activity types 1 and 5 (comparing sections), and less for types 2 and 4, where a closer focus on concordance lines had to be made. Reverse translation based on the information searched was also found as a difficult exercise by some students.

Other answers valued the use of the tool as a language learning-priming resource (point 5 in Table 9), including its "possibilities", as some learners said in class, for written production activities. An example was put forward by the teacher about its possible benefit for final degree projects written in English, which was understood and considered by some learners. Some students also said that the DDL searches turned out to be better than dictionaries for clarification and vocabulary learning because of the wealth of examples. The students had thus noticed lexical advantages and recognized word choice enrichment with the tool.

Finally, the collected answers from both groups in a last survey on grammatical competence, six weeks later, attest to their overall satisfaction with grammar-focused activities. Table 10 shows the highest percentages for those answers.

	Questions	EG	CG
1.	Did you like this grammar section?	Yes (88%)	Yes (84%)
2.	If yes, why?	Command (68%)	Command (76%)
3.	Did you learn from it?	Yes (88%)	Yes (88%)
4.	If yes, what?	Tenses (60%)	The future (52%)
5.	Which activity did you like the most	Type 5 (44%)	Future tense (40%)
6.	Why?	Fun (52%)	Important (58%)

Table 10: Summarized answers from the general questionnaire

Other answers, not shown here, had lower percentages. In question 2, for example, the answers ranged from "vocabulary" and "work with language" (EG) to "sentence writing" and "translation" (CG). For question 4, other possibilities seen as advantages included "using words in context" (EG) and "verb use in sentences" (CG). Both groups liked the activities, but more answers pointed to concrete exercises (question 5 in Table 10). Most students enjoyed such activities because of different factors (e.g., ease of exercises, clear examples, familiar vocabulary, etc). Many EG students valued the collaborative part of the lab sessions, and most in CG considered the grammar sessions as valuable for language learning, among other reasons, because the activities were already familiar to them and / or met their expectations from other ESL courses.

5. Conclusions

Students seem to feel positive about the presentation of upper-intermediate grammatical points in ESP courses. As examined in the questionnaires, their answers

may imply a demand for the integration of grammatical approaches. However, it seems that for this type of specialized / semi-specialized context of language use (i.e., B / T), relying on a grammar textbook would deviate from course objectives. In contrast, the on-line DDL section has contributed to their sense of connection within the course. In other words, the lab sessions may have served a double purpose for them: To enhance not only their awareness of the need for a grammatical competence development, as it also did among the CG students, but also their realization that the hands-on concordance exercises could relate to their own field and discipline, where English tends to be used in characteristic ways.

The probable factor that motivated that conscientious thinking was that the learners could activate language discovery and inductive reasoning, in agreement with previous findings. The students did not seem to be intimidated by DDL, but quite the opposite, they found that the examples and patterns could clarify answers. Teacher mediation was undoubtedly crucial at different stages, as mentioned above, but also group collaboration and peer assistance came to work beneficially towards task accomplishment. The students' favorable impressions in the questionnaires reflect such perceptions.

It may be the case that the specific nature of the texts (academic and news sections / sub-sections) led them to integrate the task more effectively within the ESP course. This fostered their awareness of how language is used. Lexico-grammatical meaning could be induced after the solutions were reached for the exercises, and linguistic discovery was an aftermath in many cases. The students were able to learn by noticing on their own and by exchanging the information noticed. Another advantage seemed to be their access to already-known vocabulary in their studies, which contributed to a better identification of linguistic occurrences.

In terms of the teacher's guidance and supervision along the activities, this particular DDL group of students was quite open and communicative (not always the case in ESP courses). This combination of fluent peer-to-peer assistance and the teacher's mediation may have worked as a decisive factor that favored the results obtained. Some question type activities, especially those that encouraged more use of co-text and examples to demonstrate a concrete point, demanded teacher mediation and student discussions. The use of concordances, albeit sometimes regarded as hard and tedious, was also experienced as challenging and productive. The use of patterns succeeded to decode meaning mainly because the students could compare different word uses according to text sections. Mostly, they understood that text-to-word reciprocity, and the connection motivated a positive attitude.

The case study has thus led to positive results because most students liked the approach, even competing about who would first get the answers right. In this sense, when the clear examples surfaced, the activity was regarded as positive and "fun" (questionnaires). It was therefore found as crucial that DDL questions address pre-selected, specific content, easily identifiable. In summary, different general observations could be contemplated for ESP:

1)   The use of an approach that integrates lexis and grammar for the
            teaching of grammatical competence.

2)   The combination of general and specific text types and registers so that
            students can compare lexical-grammatical behaviour.

3)   The facilitation of pre-selected and pre-arranged concordances that
            lead to specific solutions, within specialized / semi-specialized contexts
            of language use that are familiar to the learners.

4)   The application of a task method that adjusts to a problem-solution
            model.

5)   The adaptation of content and language use to language levels by using
            authenticity for task development, mainly via exercises that encourage
            induction and collaboration.

We also consider that other scenarios should be further compared to corroborate, contradict, or extend our findings. DDL, in any of the forms considered, i.e., off-line corpus-driven examples, or large on-line collections, may, or may not, improve language learning skills. We think that this hypothesis should be further tested in longitudinal studies. More varied on-line corpora should be accessed for hands-on concordancing, e.g., in other disciplines, text types and genres, but printed material should also be compared so that the computer variable can be further evaluated as a motivating factor.

On the other hand, as more tailor-made corpus-based resources emerge for academic aims (e.g., FLAX, Flexible Language Acquisition, by Witten et al., 2015), their use could be integrated and compared. Phraseology learning, for instance, is already being exploited in mobile devices, which could benefit ESP learners within their field for self-study and discovery. Flexibility and ease of access are fundamental aspects in this sense. For ESP courses, a reinforcement of L2 grammatical knowledge may be made possible if, among other aspects, DDL can be integrated in the world of mobile technologies. An on-going challenge is to encourage and enable DDL flexibly and effectively beyond the classroom.

About the author

Alejandro Curado Fuentes works as a Senior Lecturer at University of Extremadura, Cáceres, where he teaches English for Specific Purposes at the Faculty of Business Administration, Finances and Tourism. He also teaches Corpus Linguistics and Research Methods courses in two Master programs at the Faculty of Humanities. His teaching aims to integrate past and present research experiences coming from an amalgamation of communicative and corpus-driven methods in the classroom. His research usually involves classroom applications and practice in the ESP context, ICT integration, task-centered and learner-centered approaches, among other aspects. He is currently involved in two European

projects dealing with intercultural issues in ESP and migrants, and is also the current president of the European Association of Languages for Specific Purposes (AELFE).

Bibliography

Aston, G. (1997). Small and large corpora in language learning. In B. Lewandowska-Tomaszczyk & P. Melia (Eds.), *Practical applications in language corpora* (*PALC 97*) (pp. 51-62). Lodz: Lodz University Press.

Aston, G. (Ed.) (2001). *Learning with corpora*. Houston: Athelstan.

Baker, M., Francis, G., & Tognini-Bonelli, E. (Eds.) (1993). *Text and technology*. Amsterdam: John Benjamins.

Bernardini, S. (2001). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15-36). Amsterdam: John Benjamins.

Boulton, A. (2009). Data-driven learning: Reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*, 35(1), 1-25.

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534–572.

Boulton, A. (2012). Corpus consultation in ESP: a review of empirical research. In A. Boulton, S. Carter-Thomas & E. Rowley-Jolivet (Eds.), *Corpus-informed research and teaching in ESP* (pp. 261-292). Amsterdam: John Benjamins.

Charles, M. (2011). Using hands-on concordancing to teach rhetorical functions: Evaluation and implications for EAP writing classes. In A. Frankenberg-Garcia, L. Flowerdew & G. Aston (Eds.), *New trends in corpora and language learning* (pp. 26-43). London: Continuum.

Charles, M. (2015). Same task, different corpus: The role of personal corpora in EAP classes. In A. Lenko-Szymanska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 131-154). Amsterdam: John Benjamins.

Cheng, W., Warren, M., & Xun-feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31(2), 173–186.

Chujo, K., Oghighian, K., & Akasegawa, S. (2015). A corpus and grammatical browsing system for remedial EFL learners. In A. Lenko-Szymanska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 109-130). Amsterdam: John Benjamins.

Council of Europe. (2003). *Common European framework of reference for languages*. Cambridge: Cambridge University Press.

Curado-Fuentes, A. (2007a). Digital reading strategies in computer science. *Revista Alicantina de Estudios Ingleses*, 20(1), 23-43.

Curado-Fuentes, A. (2007b). A corpus-based assessment of reading comprehension in English. In E. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 309–326). Amsterdam: Rodopi.

Curado-Fuentes, A. (2015). Exploiting keywords in a DDL approach to the comprehension of news texts by lower-level learners. In A. Lenko-Szymanska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 176-194). Amsterdam: John Benjamins.

Curado-Fuentes, A. (2016). Grammatical development via DDL at the upper-intermediate level in LSP contexts. In M.F. Litzer, J. García-Laborda & C. Tejedor-Martínez (Eds.), *Beyond the universe of languages for specific purposes: The 21st century perspective* (pp. 65-68). Alcalá de Henares: Universidad de Alcalá, Servicio de Publicaciones.

Davies, Mark. (2008-). The corpus of contemporary American English: 520 million words, 1990-present. <http://corpus.byu.edu/coca/> [20/2/2017]

Flowerdew, L. (2001). The exploitation of small learner corpora in EAP materials design. In M. Ghadessy, A. Henry & R. Roseberry (Eds.), *Small corpus studies and ELT*: *Theory and practice* (pp. 363-379). Amsterdam: John Benjamins.

Flowerdew, L. (2008). Pedagogic value of corpora: A critical evaluation. In A. Frankenberg-Garcia (Ed.), *Proceedings of the 8th teaching and language corpora conference* (pp. 115-119). Lisbon: AEICI.

Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL*, 26(2), 128-146.

Gairns, R. & S. Redman. (2012). *Natural English upper-intermediate*. Oxford: Oxford University Press.

Gavioli, L. (2006). *Exploring corpora for ESP learning*. Amsterdam: John Benjamins.

Gavioli, L. (2009). Corpus analysis and the achievement of learner autonomy in interaction. In L. Lombardo (Ed.), *Using corpora to learn about language and discourse* (pp. 39–72). Bern: Peter Lang.

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

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.

Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, 26 (2), 163-183.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hunston, S., & Francis, G. (2000). *Pattern grammar. A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.

Johns, T. (1988). Whence and whither classroom concordancing? In P. Bongaerts, P. de Haan, S. Lobbe & H. Wekker (Eds.), *Computer applications in language learning* (pp. 9-27). Dordrecht: Foris.

Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 107-118). Amsterdam: Rodopi.

Lenko-Szymanska, A., & Boulton, A. (Eds.) (2015). *Multiple affordances of language corpora for data-driven learning*. Amsterdam: John Benjamins.

Mauranen, A. (2004). Spoken corpus for an ordinary learner. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 89-105). Amsterdam: John Benjamins.

Milton, J. (1996). Exploiting L1 and L2 corpora for computer assisted language learning design: The role of an interactive hyptertext grammar. In S. Botley, J. Glass, A. McEnery & A. Wilson (Eds.), *Proceedings of teaching and language corpora* (pp. 233-243). Lancaster: Lancaster University.

Nassaji, H. (2000). Towards Integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *The Modern Language Journal*, 84(2), 241-250.

O'Keefee, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.

Pashler, H. McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, a Journal of the Association for Psychological Science*, 9(3), 103-119.

Pérez-Basanta, C. (2005). Assessing the receptive vocabulary of Spanish students of English philology: An empirical investigation. In E. Martínez-Dueñas (Ed.), *Towards an understanding of the English language, past, present and future: Studies in honour of Fernando Serrano* (pp. 545–564). Granada: Universidad de Granada.

Pérez-Paredes, P., & Cantos-Gómez, P. (2004). Some lessons students learn: Self-discovery and corpora. In G. Aston, S. Bernardini & D. Stewart (Eds.), *Corpora and language learners* (pp. 247-257). Amsterdam: John Benjamins.

Robinson, P. (1995). Attention, memory, and the noticing hypothesis. *Language Learning*, 45(1), 283-331.

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.

Seidhofer, B. (2000). Operationalizing intertextuality: Using learner corpora for learning. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 207-223). Frankfurt: Peter Lang.

Skehan, P. (2001). The role of a focus on form during task-based instruction. In C. Muñoz-Lahoz, M.L. Celaya-Villanueva, M. Fernández-Villanueva, T. Navés & O. Struck (Eds.), *Trabajos en lingüística aplicada* (pp. 11-24). Barcelona: Univerbook.

Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL*, 26 (2), 184-201.

Thomas, J. (2015). Stealing a march on collocation: Deriving extended collocations from full text for student analysis and synthesis. In A. Lenko-Szymanska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 85-108). Amsterdam: John Benjamins.

Williams, J. (2001). The effectiveness of spontaneous attention to form. *System*, 29(4), 325-340.

Witten, I.H., Wu, S., Franken, M., Brine, J., Brown, K., & Fitzgerald, A. (2015). FLAX: Flexible language acquisition. <http://flax.nzdl.org/greenstone3/flax?a=fp&sa=library> [8/31/2016].