E·S·P·Today
JOURNAL OF ENGLISH FOR SPECIFIC PURPOSES AT TERTIARY LEVEL

María José Marín-Pérez*

Universidad de Murcia, Spain
mariajose.marin1@um.es

María Pilar Aguado Jiménez

Universidad de Murcia, Spain
paguado@um.es

DDL LEARNING IN ESP TEACHING: AN EXPLORATORY STUDY ON MULTI-WORD UNIT USAGE IN BUSINESS LETTERS

Abstract

The present research explores the potential benefits that corpus-based instruction might have on the acquisition and usage of multi-word units (MWUs) in business English. To that end, a data-driven learning (DDL) methodology was designed and implemented to teach such units to a group of 55 informants divided into an experimental and a control group. The experimental group was taught about the usage of MWUs in business letter writing through the exposure and management of an *ad hoc* corpus of such expressions, while the control group went through the traditional sessions designed to that purpose. The comparison between the pre- and post-test results suggests that the experimental group excelled the control one in the proportion of correct MWUs used in their letters (48.19% of the units retrieved against 20.15%, respectively). On a qualitative level, the variety of MWUs used by the former was also greater, covering a larger number of functional categories than the control group, where we found almost 3/4 of the expressions falling under one single category. Finally, non-parametric Mann-Whitney U test was also run on the data related to individual achievement, pointing at the statistical significance ($p \leq 0.001$ for both groups) of the differences found on a global level and supporting our perception on the benefits of corpus-based instruction.

Key words

corpus linguistics, business English, DDL, multi-word units, business letters.

* Corresponding address: María José Marín-Pérez, Facultad de Letras, Campus de la Merced, C. Santo Cristo, Planta Baja, 30001 Murcia.

1. INTRODUCTION

In spite of the large amount of studies concerned with the design and implementation of data-driven learning (DDL) methodologies (Aston, 1997; Boulton, 2010; Chen & Flowerdew, 2018; Clerehan et al., 2003; Hadley, 2002; Pérez-Paredes et al., 2012; Todd, 2001; Argyroulis, 2022; Elkasović & Jelčić Čolakovac, 2023; Karpenko-Seccombe, 2018; Soto-Almela, 2024), focused both on the general and specialised fields of language, the number of these devoted to the area of business English is quite reduced (Curado Fuentes, 2002; Flowerdew, 2012; Hiltunen & Mäkinen, 2014; Parizoska & Rajh, 2017; Shi, 2013; Someya, 2000; Xia et al., 2023), which also applies to the amount of corpora available within this field. In addition, meta-analyses like Boulton and Cobb's (2017), Ueno and Takeuchi (2023) or Lee et al. (2019) also signal the scarceness of studies focused on the interaction between DDL and English for Specific Purposes (ESP) in general.

Multi-word units (MWUs), also known as *lexical bundles*, *formulaic expressions* or *n-grams* (amongst others), have been examined from different perspectives (Biber & Barbieri, 2007; Erman & Warren, 2000; Sympson-Vlach & Ellis, 2010; Wray, 2002; Wulff, 2018), being deemed essential in the way they inform on language acquisition processes (Forsberg & Bartning, 2010; Paquot, 2018, 2019) as well as performing a fundamental role in structuring spoken and written discourse (Biber & Barbieri, 2007).

The present research was thus conceived with the aim of trying to bridge the existing gap in the business field as regards the design and implementation of DDL methodologies which could improve the acquisition and usage of MWUs (given their linguistic and pedagogical relevance) by compiling an *ad hoc* corpus of such expressions which could be used as reference, due to the scarcity of such materials.

Therefore, the first part of the empirical study (subsection 3.2.), which is preceded by the literature review section on MWUs and an account of the corpora used in this research, presents the compilation process of a corpus of business MWUs. Such units were automatically retrieved from a collection of business letters taken from ESP textbooks (Ashley, 2003; Brieger, 2011; Lougheed, 2002; Sharma & Mohan, 2017; Shiyab & Halimi, 2015; Taylor, 2012). It is followed by the description of the experiment that was carried out with a group of 55 undergraduate students enrolled in an ESP module dealing, amongst others, with business letter writing (subsection 3.3.). The fourth section introduces the results of the pre- and post-tests given to the informants before and after the DDL methodology was implemented, as well as their discussion, all of which are summed up in section 5, which offers the final remarks.

2. DEFINING MULTI-WORD UNITS

Multi-word units (MWUs) have been examined from varied angles by scholars and have been labelled differently (*lexical bundles*, *restricted collocations*, *multi-word* *combinations*, *formulaic expressions*, *n-grams* or *p-frames*) on the basis of the perspective adopted for their identification and analysis. Within the field of Second Language Acquisition (SLA), Wray (2002, p. 9), who identifies such expressions as *formulaic sequences*, provides a comprehensive definition that goes as follows: "[they are] sequence[s], continuous or discontinuous, of words or other elements, which [are], or appear to be, prefabricated, that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar". Although it has been proven that native speakers learn and use formulaic language with greater ease than other similar more open patterns (Sympson-Vlach & Ellis, 2010, p. 492), in SLA, such units take longer to be acquired (Forsberg, 2021; Sinclair, 1991). As postulated by Forsberg and Bartning (2010) and Paquot (2018, 2019), the processing and usage of formulaic language reflect the learners' evolution in their L2 acquisition process, since the use of such expressions by L2 learners could be observed from the so-called upper-intermediate level upwards in L2 French and L2 English, respectively. This is even more so for those formulas displaying a higher degree of semantic transparency (Wulff, 2018, p. 30), given the fact that the more idiomatic (thus, less transparent) formulaic language tends to be, the harder it seems for L2 learners to grasp and acquire it, even at the highest competence levels, also causing an undesired ceiling effect in the learners' path towards native-like language proficiency (Erman et al., 2016; Granger, 2001; Pawley & Syder, 1983).

Amongst the research dedicated to study formulaic language from a phraseological perspective, as seen in Forsberg (2021), we find the early work by Cowie (1981), influenced by the Russian tradition of the 1940s, 50s and 60s, who relies on manual analysis and applies the 'idiomaticity' criterion as a point of departure to categorise such expressions. On a similar note, authors such as Erman and Warren (2000) resort to the concept of 'restricted exchangeability' whereby, for a string of words to be regarded as an MWU, at least one of its members cannot be substituted by a synonym, as it would alter the interpretation of the sequence. In any case, as Flowerdew (2012) already acknowledges, a theory of phraseology is already established and lexical items are often viewed by scholars like John Sinclair (1998) as units which deploy associations with other lexical elements through collocational and colligational processes, also displaying semantic preferences and prosodies (Partington, 2004).

Biber and Barbieri (2007) use the term *lexical bundles* to refer to "the most frequently recurring sequences of words (...). Lexical bundles are usually not structurally complete and not idiomatic in meaning, but they serve important discourse functions in both spoken and written texts" (Biber & Barbieri, 2007, p. 264). In sum, the basic trait that makes them stand out within a corpus is their high frequency while, concerning their form, they might sometimes seem truncated chunks of speech, although they play a pivotal role on a discursive level. As regards their meaning, Biber and Barbieri (2007) eliminate the idiomaticity principle that governs the phraseological approach to multi-word units, focusing merely on their statistical saliency. Depending on their discursive function, lexical bundles are classified by Biber and Barbieri (2007) into three major categories, namely, *stance expressions*, which are used to express the writer/speaker's position towards the propositional content of the text, *discourse organizers*, aimed at marking the relationship between the propositions that articulate discourse and *referential expressions,* whose major function is to "identify entities or specific parts of entities" (Biber & Barbieri, 2007, p. 265).

Concerning the structure of multi-word units and still within the frequency-based approach to MWUs, Glabasova et al. (2017) address a relevant issue when it comes to their definition, that of proximity. They employ the term *n-gram* to refer to word sequences formed by two or more elements whose adjacency leaves little room for variation, marking a difference with other methods to identify, for instance, collocational networks, where the network node, its collocates and particularly its co-collocates stand at a "mediated distance via an association with another word in the network" (Glabasova et al., 2017, p. 158), which allows to explore the context of usage of a word/term in much greater depth.

From a statistical perspective, the bonds generated amongst the constituents of MWUs can be quantified using different techniques. Software applications like *Wordsmith Tools* (Scott, 2022), *Antconc* (Anthony, 2022) or *Sketch Engine* (Kilgarriff et al., 2014) facilitate the extraction of word clusters from any corpus greatly, producing lists of MWUs. These units could be deemed representative of any corpus if their frequency or distribution can be compared with similar elements retrieved from other corpora, using tools like *key n-grams*, which is offered by the software package *Sketch Engine* (Kilgarriff et al., 2014). MWUs can be also ranked on the basis of factors such as keyness, which determines how statistically relevant an MWU may be in a corpus (the study corpus) when compared against other similar units found in a different corpus (the reference one). This technique is labelled as 'key n-gram extraction' in *Sketch Engine* and it was employed in the present research, whereby MWUs are understood as word chains made up of adjacent elements whose statistical bonds and saliency facilitate their automatic retrieval.

Thus, considering the above, our approach to the study of these word sequences leans towards the use of the label *multi-word units* for the following reasons. Firstly, they are regarded as units formed by adjacent constituents due to quantitative criteria, as the sequences extracted from the corpora introduced in subsection 3.1. were automatically identified as 'key n-grams' and then confirmed manually as such. Secondly, their unity can be understood in statistical terms as well as from a semantic perspective: they constitute units of meaning which also perform a specific pragmatic function in context, as will be described later in this article. Thirdly, once the lists of MWUs were filtered, truncated expressions such as *to hearing from you* or *writing to inform you* were completed for pedagogical purposes, therefore, they cannot be identified with the label *n-gram* or *lexical bundle*, as discussed above. Finally, the number of constituents in each unit retrieved from our corpora varies, as the sequences were completed if they were truncated, hence the term *multi-word*.

3. METHODOLOGY

3.1. The corpora

Four different corpora were employed in the present research. The first one, compiled *ad hoc*, consists in a collection of MWUs typically found in a set of business letters which, due to their written character, were compared against the British National Corpus (BNC), mostly composed of written texts from various fields, by using the tool 'key n-grams' included in the software package *Sketch Engine* (Kilgarriff et al., 2014).

The textbooks from which the corpus of MWUs was extracted (Ashley, 2003; Brieger, 2011; Lougheed, 2002; Sharma & Mohan, 2017; Shiyab & Halimi, 2015; Taylor, 2012) mostly contained specimen letter designed with a pedagogical purpose, except for Ashley's, which comprises authentic correspondence related to real business transactions carried out in the UK.

The third corpus contains a set of business letters written by 55 informants who were divided into an experimental and a control group as part of the experiment presented in subsection 3.3. As a learner corpus, it was contrasted with the British Academic Written English Corpus[1] (BAWE) also using *Sketch Engine*. The BAWE is a learner corpus of varied texts gathered by members of the University of Coventry and written by students from that university. The rationale behind the choice of this text collection and not the BNC basically stems from its similarity to the study corpus, as both were written by university undergraduate students from different academic areas. Therefore, due to these potential similarities, the degree of reliability implied in the automatic process of key n-gram extraction could be greater than if compared against a corpus of general language texts.

The main features of each text collection are displayed in Table 1 below.

CORPUS ROLE	CORPUS	TEXTS	TOKENS	TYPES	FIELD	GENRES
Study corpus	*Business MWUs*	1,261	177,632	12,113	business	letters
Reference corpus	*BNC*	4,000	100m	724,893	General: world affairs, leisure, arts, commerce and finance, social science, applied and	academic prose, news reports, editorials, drama texts,

[1] https://www.coventry.ac.uk/research/research-directories/current-projects/2015/british-academic-written-english-corpus-bawe/

					natural science (...)	narrative fiction (...)
Study corpus	*Learner corpus (experimental)*	19	5,126	935	business	letters
Study corpus	*Learner corpus (control)*	36	9,674	1,329	business	letters
Reference corpus	*BAWE (learner corpus)*	2,700	6.5m		Academic: Arts and Humanities, Social Sciences, Life and Physical Sciences (...)	essays, critiques, case studies, narrative recounts (...)

Table 1. The corpora

3.2. MWU extraction and classification

The automatisation of the process of extraction of MWUs presents many advantages, as the software employed for the process of identification, *Sketch Engine* (Kilgarriff et al., 2014), performs this task practically on the fly. Nevertheless, it may be argued that the algorithms implemented in MWU retrieval could ignore relevant units that might not be well distributed or display low frequency (particularly in small corpora like the ones used in this study). On the other hand, some irrelevant elements might also be brought to the foreground simply because they occur recurrently in the study corpus, in spite of them not having any specific value, hence the need to filter lists manually once they are produced by any software application.

In general, Automatic Term Recognition (ATR) methods have proved to be more efficient in retrieving single word terms/units in specialised environments,[2] particularly when resorting to corpus comparison. On the contrary, multi-word term identification has yielded large noise proportions in various experiments (many of the elements identified as MWUs could not be confirmed as such after human validation). In any case, handling large amounts of text manually in search of MWUs would be an unattainable task that could not even be performed on small corpora like the ones herein. This is why a decision was made to opt for the tool 'key n-grams', offered with *Sketch Engine* and not available in other similar applications. In order to obtain a list of key n-grams, a comparison between the study corpora (the corpus of business letters and the learner corpora above) and the reference ones (the *BNC* and the *BAWE*) was carried out. The automatic comparison of both sets of corpora offered results which were based on the statistical saliency of a number of elements in the study corpora. Such saliency might be interpreted as a sign of the relevance and representativeness of those units, that is, as their keyness.

Authors like Scott (2001, p. 110) relate the concept of keyness to that of *aboutness* (Hutchins, 1977) due to its capacity to inform about the major topics in a

[2] See Marín (2015) for a comparison amongst five different ATR methods used for MWU extraction.

text collection. Although scholars like Gabrielatos and Baker (2008) question this idea by referring (amongst others) to the inability of keyness measures to discriminate homonymous or polysemous terms, there is no doubt that, regardless of the method applied to extract them, keywords are useful to unveil relevant terminology/expressions that could hardly be spotted otherwise. When applied to the identification of MWUs, we could connect this concept of aboutness with the semantic unity MWUs display as well as to the pragmatic functions that they could perform in any text. By filtering[3] manually a list of MWUs, we could come up with repertoires of expressions that might be used in classrooms settings, like the one described below. This is precisely one of the objectives of the present research, creating an inventory of MWUs that could be deemed representative of business letters (as recorded in ESP textbooks) on the basis of its statistical saliency and not simply of prescriptive criteria. Some business English textbooks (Brieger, 2011) offer lists of expressions suitable for business letter writing yet, to the best of our knowledge, they were not generated applying corpus-based methods.

In *Sketch Engine* (Kilgarriff et al., 2014), the algorithm used for n-gram extraction is straightforward, as it relies on what the authors call 'simple maths'.[4] The system compares the relative frequency of a word sequence in the study corpus with the same parameter in the reference corpus. However, the method employed to obtain key n-grams cannot be adjusted, as the software only offers one option, the simple math model. The length of n-grams can vary, we adjusted it to 4 elements in this case, but it could not be determined whether implementing other measures/algorithms may have improved the efficiency of the system. Even so, after obtaining, revising and filtering the MWU output lists, the proportion of true positives out of the entire set of elements identified automatically was similar on average to that observed by Marín (2015, p. 12) as regards the precision attained by other ATR methods in MWU extraction (52.14%), if not slightly higher. *Sketch Engine* managed to retrieve 54.1% true MWUs from the business letter corpus, 54.9% from the experimental learner corpus and 55.23% from the texts written by the control group.

Having filtered the lists of MWUs extracted from the corpus of business letters and discarded those elements that could not be regarded as proper expressions, those that were truncated had to be completed for pedagogical purposes, i.e., word sequences like *thank you for your* (…), *I look forward to* (…) or (…) *for your letter dated* were searched as keywords in context (KWIC) and all the possible variants they might display in the corpus were grouped together. Then, based on their similarity in meaning and the pragmatic function they could perform, they were arranged and ranked using the average keyness value of each MWU and its variants as reference, e.g., the chunk *thank you for* and its variants *letter*, *enquiry*, *order* or

[3] The term *filtering* should be interpreted as the revision of the output lists produced by the software and the deletion of false positives.
[4] See the *Sketch Engine* website for more information on the 'simple maths' technique: https://www.sketchengine.eu/documentation/simple-maths/

email registered an average keyness value of 161.444 (the highest in the entire list), thus it stands at the top of the list.

A working taxonomy was created bearing in mind the practical character of the corpus at hand, aimed at helping students improve their business letter writing skills. Therefore, the categories presented in the taxonomy result from the classification of the MWUs retrieved from the corpus with regard to their meaning and function. This taxonomy could therefore be deemed corpus-driven, as it is based on the results obtained after extracting and filtering the lists of MWUs automatically identified by the software.

The categories proposed for our working taxonomy are the following: *thanking*, *closing formulas & expression of courtesy*, *introducing a topic & opening expressions*, *salutations*, *time reference*, *showing appraisal*, *apologising*, *regretting*, *making offers*, *requesting*, *complaining*, *warning action*, *informing* and *adding attachments/shipments.* The specific order in which they are enumerated corresponds with the average keyness value of each group of expressions and their variants. These values were calculated by the tool *key-ngrams* included in the software package *Sketch Engine* (Kilgarriff et al., 2014), as stated above.

Table 2 below illustrates the top two categories and all the elements comprised in them according to their meaning and function (column 2). The third column contains the average relative frequency assigned to each expression and its variants in the study corpus (the business letter collection), while the fourth column shows the same average value as found in the reference one (the BNC). Their average keyness score is displayed in column 5.

Rank	***Meaning/function***	***Relative frequency -study corpus- Average***	***Relative frequency -ref. corpus- Average***	***Keyness score Average***
	Thanking	1604.44067	0	161.444
1	Thank you for your:			
	- email dated .../email of ...			
	- letter dated/of ...			
	- order of ...			
	- enquiry/inquiry			
	- interest			
	- advice			
	- application			
	- assistance			
	- your attention (to this matter)			
	- concern			
	- consideration			
	- continued support			
	- time			
	Closing formulas/ courtesy			

2	I look forward to:	669.92432	0	67.992
	- hearing from you			
	- meeting you/with you			
	- receiving your reply/ comments ...			
	- seeing you			
	- working with you			
	- your prompt/early reply			
3	- If you have any (further) questions, please contact me at (email/phone number)	540.44318	0.47983	52.524
	- If you have any (further) questions, do not hesitate to contact me at (email/phone number)			
	- If you have any (further) queries, feel free to contact me at (email/phone number)			
	If you have any (further) questions, please let me know.			
4	Yours sincerely	95.70348	0	10.57
5	If there is anything (else):	84.44424	0	9.444
	- I can do, please let me know.			
	- do not hesitate to call.			
	- give me a call.			
6	If I can be of further assistance, please call.	45.03693	0	5.504
	- If we can be of further service, please feel free to call.			
7	Yours faithfully	39.4073	0	4.941

Table 2. Working taxonomy (sample)[5]

3.3. Experiment design

Concerning the design of the DDL experiment itself, a group of 3rd year undergraduate students of the Degree in English Studies taught at the University of

[5] The relative frequency of most elements in the reference corpus, the BNC, was recorded as 0, as the BNC does not include "commercial correspondence" as a text category. As regards the category "Thanking", some of the items comprised by it are used in the BNC, for instance, the relative frequency of the expression "thank you for your letter dated" in the BNC was 0.0000000071213 (it appears 8 times in a corpus of 112 million words). For practical reasons, we decided to approximate it to 0, as the statistical relevance of such elements in the BNC is almost none. The values displayed from columns 3 to 5 show the average score for the entire category, not individual values, otherwise, ranking the MWUs by categories using their keyness score as reference would not have been possible.

XXX was selected and divided into two subgroups: a control group (36 informants) and an experimental one (19).

The informants' L1 background was also taken into consideration with the aim of establishing a connection between such variable and the informants' achievement, yet, similarly to the students' productive vocabulary levels, such connection could not be established due to the lack of balance of both samples, as explained in greater detail below.

As Figure 1 illustrates, the groups were not homogeneous or compensated with respect to the distribution of the L1 backgrounds. The most numerous nationalities were Spanish (72.22% in the control group and 36.84% in the experimental one) and Italian (36.84% in the experimental group and 2.77% in the control group), followed by the Polish (11.11% in the control group) and the Lithuanians (10.52% in the experimental group), whereas other backgrounds were marginal, as we find, for instance, only one informant from Hungary or England in the control group. In fact, it was not possible to make the groups more balanced as regards their vocabulary level or nationality, as the students joined the experiment voluntarily and not all of them finished it; this is why the size of both groups is not similar either.

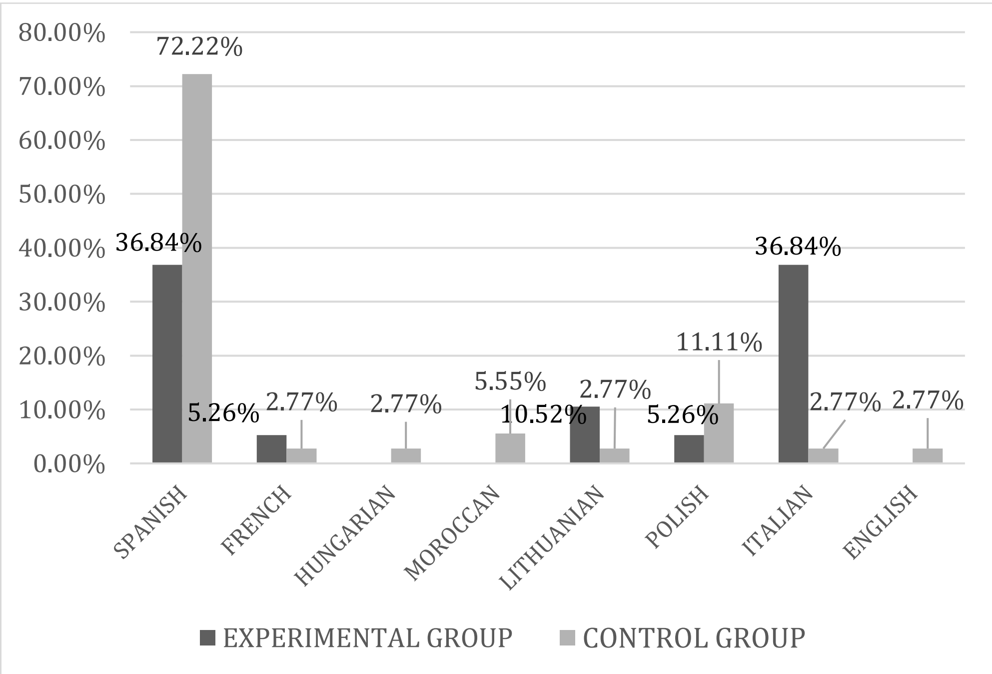

Figure 1. Students' language background (experimental and control group)

Even so, an attempt was made at trying to describe how such variables as the informants' L1 background (see Figure 1) or their productive vocabulary level (see Figure 2) might condition individual achievement in both groups. Although the lack of balance between the two datasets seemed quite evident in this respect, for instance, the distance existing between the proportion of informants assigned to the highest vocabulary level group and the lowest one was around 48 percentage points in the experimental group (it was even higher for the control one), the data were processed with SPSS and their descriptive analysis ratified our perception, thus preventing the inclusion of these variables in the statistical analysis discussed in subsection 4.2. Undoubtedly, future research along these lines should consider the influence of such variables on the informants' achievement by granting greater balance in the samples used.

The informants were all enrolled in an ESP module that lasts an entire semester (4 hours a week) and covers different ESP varieties and genres both from a theoretical and a practical point of view. They were lectured on the major traits of text types such as memoranda, business letters, emails, proposals, circulars or reports in approximately 10 hours. The major focus was set on business letters (4/5 hours were dedicated to them), which students worked on and had to produce in groups during the practice sessions assigned to those tasks. The control group would follow the traditional methodology, which consisted in following the sessions designed in previous years, where the theory on different genres was presented through genre characterisation, description and analysis. After that, they would take part in some practical sessions where they had to write business texts as part of their formative assessment process.

Meanwhile, the experimental group would attend the extra sessions described below. At the end of the semester, all the informants had to write a business letter as part of their final exam, as described in greater detail in subsection 3.1. The letters were digitalised and processed to determine whether the DDL experiment presented herein had benefitted the experimental group as opposed to the control one.

In two separate sessions, the members of the experimental group were trained to manage the corpus of business letters collected by the authors, which had been shared with them through *Sketch Engine* (Kilgarriff et al., 2014). In the first session (first half), they were instructed on the use of the KWIC tool provided with the software (which they had previously used during the second year of the degree) so that they could search and study the context of usage of the business MWUs included in the corpus sampled in Table 2 (they had a printed version of the entire table with all the categories and variants of each expression ranked according to their keyness). The corpus consultation and hands-on work developed afterwards followed the pattern defined by Thurstun and Candlin (1998, p. 272).

Phases 3 and 4 (Thurstun & Candlin, 1998) were completed at the same time at the end of the first session and also throughout the second session. In total, they had to complete 7 tasks, which required the use of different types of MWUs in several contexts. For instance, the students were asked to write a paragraph[6] where

[6] The training session consisted in writing paragraphs performing a specific function instead of full letters. Letter writing would be a final task used as part of the formative assessment process and

they had to start a letter in which they showed their interest in the products a company had manufactured and then, they had to request something ([1]); in addition, they were instructed to write a paragraph complaining about a particular situation related to their doing business with the addressee(s) ([2]), and to apologise for any problem they had been referred to in a previous letter ([3]). Here are some samples of student writing.

[1] Dear Mr. and Mrs Goodman,

> I am writing to you concerning the product your company offers. I am particularly interested in the VSync range of equipment your establishments have on sale. I would be delighted if you would like to meet me in order to arrange the selling conditions.

[2] On November 22, we placed an order with your company for 48 pieces of computer hardware. We have received a letter with the confirmation, copies of the invoice, and the information that the order will be shipped the next day; however, it has not arrived yet. In one of the previous letters, we emphasised that we need that equipment to be delivered by the end of December. Unfortunately, your firm has not fulfilled our request which resulted in postponing the opening of the new branch of our company, and consequently huge financial loss.

[3] Please receive our most sincere apologies. We encountered some unexpected issues with our packaging team, but have proceeded to resolve this issue by sending you the remaining 400 left-footed socks. Please allow for two business days for the package to arrive, we have sent them by first-class postage.

As regards the control group, they had had no exposure to the corpus materials, although they had followed the lessons on commercial correspondence planned by the lecturers and practised letter writing during the practice sessions.

In addition, in order to obtain a clearer language profile of the informants taking part in the study, both groups were administered the productive vocabulary test designed by Laufer and Nation (1999), – computerised by Tom Cobb[7] on *Lextutor*, his online corpus site – before actually starting working on the corpus of business MWUs, which allowed us to determine the breadth of their productive vocabulary practically on the fly. The reasons to use this test were mostly practical, as it took almost no time to complete it and obtain the results. We considered that, since we were trying to measure the students' writing abilities, getting to know about the breadth of their productive vocabulary would be a good indicator of their ability to express ideas properly and perform the tasks mentioned above.

would also be included as one of the practical questions in the final exam (the learner corpus used as the post-test was gathered using these letters).

[7] https://www.lextutor.ca/tests/levels/productive/

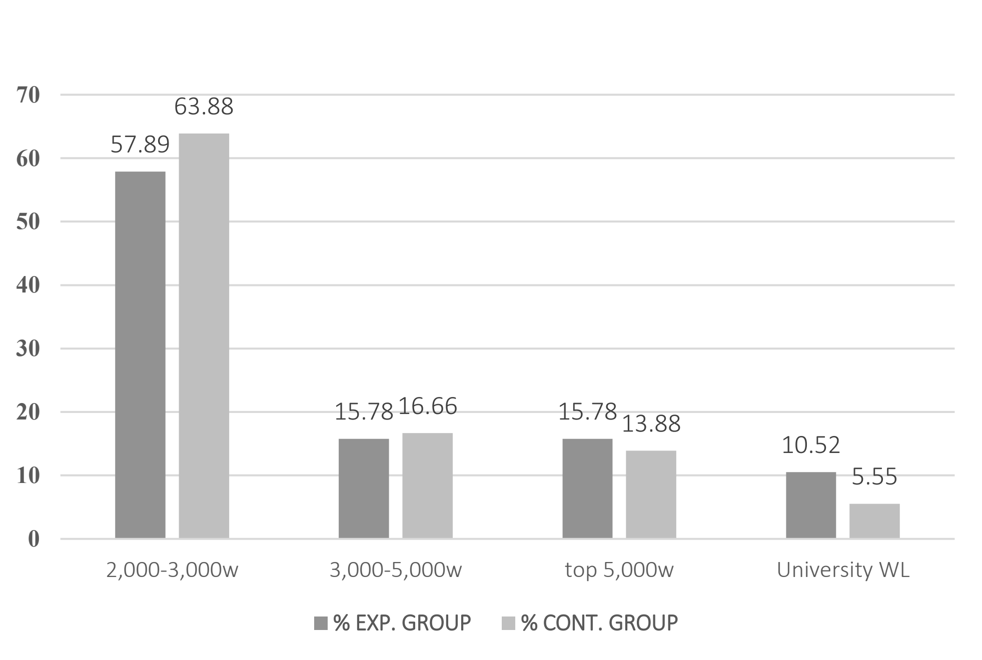

Figure 2. Productive vocabulary level test results (Laufer & Nation, 1999)

As shown in the graph above, the initial productive vocabulary level displayed by the informants both in the experimental and the control group was similar, yet, the experimental group contained a larger percentage of informants with a higher vocabulary level, that is, roughly 26% of the students in that group displayed a breadth of productive vocabulary knowledge between the most frequent 5,000 words and the University World List, whereas the control group stood slightly below, at around 19% within the same span. On the contrary, the control group presents a higher proportion of students whose breadth of vocabulary is smaller, almost 80% of them stood between the most frequent 2,000 to 5,000 words, as opposed to the experimental group, at 8 points below.

Lastly, the informants' L1 background was also taken into consideration with the aim of establishing a connection between such variable and the informants' achievement, yet, similarly to the students' productive vocabulary levels, such connection could not be established due to the lack of balance of both samples, as explained in greater detail in the final section.

4. RESULTS AND DISCUSSION

4.1. Pre-test results

As part of the practice sessions programmed for the subject, all the informants (both in the experimental and the control group) had to write a business letter that would be considered for their formative assessment. These letters were used as a pre-test to determine the proportion of business MWUs our students were capable of including in their texts prior to the implementation of the experiment above.

These letters were digitalised and processed using *key-ngram* extraction, a tool included in the *Sketch Engine* (Kilgarriff et al., 2014) software package, and the procedure for the automatic retrieval of MWUs, as depicted in subsection 3.2., was applied. In the first place, the results of the pre-test show a tendency for the experimental group to perform better, as the members of this group used 34.6% MWUs (out of the entire set extracted automatically by the software) which could be deemed typical of business correspondence, as they fit into the MWU categories defined in our corpus. On the other hand, the performance of the control group was considerably poorer, since they only employed 18.6% of these elements out of the whole set of MWUs detected by the software in their business letters.

In principle, such differences might be explained in connection with the results obtained by the students in the vocabulary placement test (see Figure 2), where the experimental group excelled the control one by 7 points as regards the percentage of informants that passed the highest vocabulary thresholds set for the test, which implied being capable of producing words included within the 5,000 most frequent ones in a general corpus and also those in the *University Word List* (Xue & Nation, 1984). Nevertheless, such an assumption could not be attested statistically for the reasons explained in the final section.

Excerpts [4], [5] and [6] exemplify some of the most frequent business MWUs employed by the members of the experimental group as shown in the pre-test. Most of them belonged in the group of complaints and information requests, for instance:

[4] Dear Mr. Artois, **I am writing to you with reference to** (…)

[5] **We have not received any reply yet**. Unfortunately, in September we ordered from your company (…)

[6] **Unless this issue is resolved promptly**, then unfortunately, **we will be forced to cancel all current orders** from your company (…)

4.2. Post-test results

4.2.1. General results

The post-test was identical to the one presented above, as it consisted in writing a full individual letter that would be included in the final exam, once the experimental group had been exposed to the corpus materials and had taken part in the sessions programmed for the experiment. Preliminary results clearly display a tendency for the experimental group to make greater use of those MWUs that were explored during the aforementioned sessions, as well as other typical expressions of business English that could fit into the categories proposed in the working taxonomy, that is, 48.19% of the MWUs extracted from the experimental learner subcorpus fell within the groups included in such taxonomy, as shown below. On the contrary, the members of the control group resorted less frequently to MWUs that could be deemed representative of the business English variety, finding 20.15% of these expressions that would fall within the categories proposed in the working taxonomy among the entire set of MWUs identified automatically by *Sketch Engine* (Kilgarriff et al., 2014).

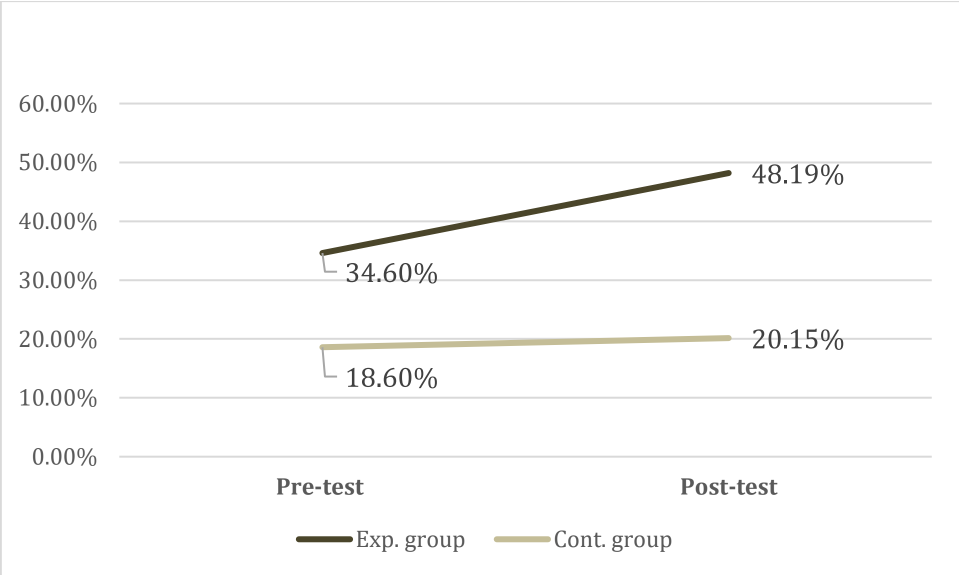

Figure 3. Pre- and post-test results

As a matter of fact, the progression observed on the part of the experimental group if compared with the control one is quite relevant (see Figure 3 above), since the former increased their usage of business MWUs by 13.59 points. On the contrary, the results in the post-test produced by the latter (who had only been exposed to class materials and taken part in the traditional sessions devoted to business letter writing) were much poorer, moving from 18.6% to 20.15% business MWU usage.

As it could be expected and as found by Hiltunen and Mäkinen (2014) with regard to language competence levels, the better results obtained by the informants in the experimental group in Laufer and Nation's (1999) productive vocabulary test might have also conditioned their greater progression, although, as already stated, such connection could not be statistically corroborated due to a lack of balance between the samples. In any case, and also judging by the results shown below, it appears that the implementation of a corpus-based methodology did have a positive impact on the students' overall performance, whose progression from the pre-test to the post-test phase was six times higher (in percentage points) for the experimental group than for the control one.

Some of the most typical MWUs used in the post-test, in this case by the control group, are shown in fragments [7], [8] and [9], which illustrate a great abundance of expressions which fit into the categories *opening and closing formulas* or *complaints*, due to the nature of the task itself:

[7] **I am writing to inform you** that you have not read the last two letters I sent you (…)

[8] Unfortunately, **we have not received any kind of reply** from you to make the adjustment (…)

[9] **I look forward to hearing from you** by return (…)

4.2.2. Results per categories

Another factor that might account for the positive influence of corpus-based instruction on the acquisition of business MWUs is the greater variety of these units (in terms of their function and meaning) that was observed in the experimental subcorpus. Whereas the informants in the experimental group resort to MWUs to express regret, to request, inform or to complain about something, amongst others, the control group sticks to a lesser amount of functional categories like *complaining*, *closing* or *replying*.

As regards the functional categories defined in our corpus of business MWUs (see subsection 3.2.), some of them were not present in the analysis of the results, as the task assigned to all the informants consisted in writing a letter of complaint, hence the absence of MWUs used to thank the addressee or to make an offer. On the contrary, the most frequently used expressions fit into the category *complaining/warning action*, precisely because of that fact.

Before continuing, a remark must also be made in this respect. The fact that some MWUs are not present in the inventory examined in this study does not imply that they were not used in any of the corpora. For instance, expressions such as *thank you for* (F=2), or *to inform you that* (F=5) could be found in the control subcorpus. However, taking into account that the identification of MWUs was carried out automatically by comparison with a reference learner corpus (the *BAWE*), the data associated to expressions like those above were not deemed relevant enough for their identification as solid MWUs according to the algorithm, which, at a time, might also point at their lack of representativeness in the focus/study corpus (formed by the letters written by the control group).

	Salutation	*Requesting*	*Replying*	*Regretting*	*Opening/ introducing*	*Informing*	*Complaining/ warning action*	*Closing*
Experimental subcorpus	1.2%	6.02%	6.02%	3.61%	28.91%	6.02%	36.14%	12.04%
Control subcorpus	1.6%	-	10.4%	-	9.6%	-	76%	12%

Table 3. Results obtained per category (% MWUs extracted per category)

As illustrated by Table 3, the difference between both groups was not merely quantitative, but also qualitative, since the number of categories which the MWUs used by the experimental group fall into is higher, as well as the distribution of such units throughout categories. Although the most numerous category is *complaining/ warning action*, the amount of items falling into it within the control subcorpus is twice as high as it is in the experimental one (76% vs. 36.14%, respectively). Actually, two thirds of the MWUs retrieved from the control subcorpus fall into this category, which points at a worse distribution of these elements and at a poorer performance on the part of this group, both in terms of function and meaning. The results for replying and closing expressions is similar for both groups, although the categories *requesting*, *regretting* and *informing* remain empty as regards the control group.

4.2.3. Statistical significance tests

In order to triangulate the data displayed in the above sections, a statistical significance test was performed on both data sets. To begin with, the samples were analysed taking into consideration the informants' individual achievement,[8] not the global results per group used in previous subsections. Kolmogorov and Smirnov test (Kolmogorov, 1933; Smirnov, 1948) was implemented in the first place using the software SPSS (IBM Corp., 2021) with the aim of determining whether the distribution of the sample could be deemed normal or not, finding an asymptotic significance in the experimental and the control group's results whose *p*-value was established at 0.039 and 0.008, respectively (below the threshold level, at 0.05, as commonly agreed). Hence, none of the results obtained for either group displayed normal distribution values and had to be processed applying non-parametric statistical tests.

Non-parametric Mann-Whitney U test (Mann & Whitney, 1947) was thus selected due to the fact that it yielded the most conclusive information about the difference observed between both groups of informants when looking into their individual performance. Firstly, the initial results showed an asymptotic significance ≤ 0.001 (*p*-value) – far below the threshold level (at 0.05) –, thus substantiating our previous findings when considering the group's data as a whole, since the difference observed between both groups in this case (which reflects their performance as individuals) could also be deemed statistically significant precisely because of such a low *p*-value.

Secondly, concerning the specific data associated to individual achievement (as illustrated in Table 4 below), it could also be noted that the highest proportion of informants employing correct business MWUs within the experimental group

[8] Individual achievement was computed by registering the amount of MWUs that each of the informants used correctly in the letters they wrote for the test individually. This task had to be performed manually, as the software was not capable of extracting them automatically through corpus comparison due to the small size of each text (roughly 300 words per letter on average).

represented 27.77% of the total. This group managed to use 15 of these lexical elements properly in the texts produced, as opposed to the control sample, where the most numerous group constituted roughly 24% of the total and only used 9 of these expressions accurately per document. Similarly, the second most numerous set in both samples presents better results in the experimental group (22% informants), managing to use 10 correct MWUs, unlike the control group, where 18.1% of the informants use 8 MWUs properly. On a similar note, the third ranking set in the experimental group (11% of the informants, as shown in columns 3, 4 and 7 below) also demonstrate greater accuracy in MWU use, using 16, 14 and 11 MWUs, respectively, although, in this case, the proportion of informants in the control group which ranks third is higher (15%) and the results are partially similar (11 MWUs).

EXPERIMENTAL GROUP										
% INFORMANTS	5.55%	11%	27.77%	11%	5.55%	11%	22%	5.55%		
# MWUs	17	16	15	14	13	11	10	8		
MID-RANGE	36.06									
CONTROL GROUP										
% INFORMANTS	3.03%	6.06%	6.06%	6.06%	15.1%	12.12%	24.2%	18.1%	6.06%	3.03%
# MWUs	16	14	13	12	11	10	9	8	7	6
MID-RANGE	20.52									

Table 4. Individual achievement per group

Thirdly, in order to complement these observations, the mid-range[9] obtained with Mann-Whitney U test (Mann & Whitney, 1947) for each group (as shown in the fourth and eighth row of Table 4) also points at the better performance of the experimental group when compared against the control one. As for the former, the test assigns a mean rank value of 36.06, almost 16 points above the control group. This difference directly relates to the higher observation values recorded in Table 4, as the experimental group employed a higher number of proper MWUs in their texts than the control one as a whole.

[9] When this value gets higher in a group, we can infer a direct relation between this value and the observation values for the same group, therefore, if the mid-range is higher, the group's performance, in this case, is also higher.

5. CONCLUSION

The present research has been aimed at designing a corpus-based methodology to enhance the process of acquisition of MWUs in business letter writing. To that end and given the scarcity of corpus materials available, a corpus of such expressions was gathered based on a collection of over 1,200 business letters taken from business English textbooks.

Subsequently, an experimental DDL method was designed and implemented on one of the groups which the 55 informants involved in this study were divided into, basically consisting in working hands-on with an *ad hoc* business corpus. Finally, all the informants were given the task of writing a letter of complaint and the texts produced were processed automatically to extract the most relevant MWUs found in them for the sake of comparison.

Overall, the results seem to signal the positive effect of the implementation of a corpus-based methodology in the teaching of business MWUs for various reasons. Firstly, from a global perspective, the experimental group managed to use a larger proportion of business MWUs in the texts they produced, excelling the control group by 28 percentage points in the post-test task. Secondly, the degree of progression from the pre-test to the post-test phase was considerably higher for the experimental group, moving from 34.6% to 48.19% MWU usage, as opposed to the control group, who only improved by roughly 2 percentage points between both phases. Thirdly, not only did we record quantitative differences as regards global achievement, but also qualitative ones, since the experimental group managed to cover a larger amount of semantic/functional MWU categories than the control group, who displayed a clear tendency to focus mostly on MWUs expressing complaints or warning to take action. Lastly, having triangulated the general results by implementing non-parametric statistical tests on the data associated to individual performance, it was also attested that there exists a statistically significant difference between the results obtained by both groups of informants.

Nevertheless, it was not possible to measure the actual influence of such variables as vocabulary level or L1 background on the informants' achievement, given the lack of balance of the samples, which was directly related to the difficulty to maintain informant engagement, as attendance to the DDL sessions was voluntary and not all the subjects attended them all. Therefore, this limitation should be addressed in future research implementing similar methodologies, which could answer interesting questions related to the observations presented herein like the way in which the students' L1 background might condition the acquisition and usage of certain structures or the degree to which the greater progression of one group of informants might be directly conditioned by their language competence level.

In spite of such limitations, the pedagogical implications derived from this experiment point at the convenience of implementing corpus-based methodologies in the ESP classroom. Actually, the use of corpus-informed language in ESP allows students to approach the learning of a specific language from a motivating point of view. Students can be guided to achieve autonomy in language learning (Gavioli, 2005) while exploring corpora of texts frequently used in their professional area. In this particular case and regardless of their authentic or semi-authentic nature, all the texts employed in this experiment display the characteristics of the genre concerned herein.

In addition, the incorporation of activities using corpora to produce professional texts promotes language awareness and close contact with the professional field of work or study, offering the students the possibility to face in context the wide array of formulaic language used, as it is the case of business letter writing. The results of this study could thus encourage practitioners to implement the use of corpora as part of the practical part of the syllabuses to explore and produce samples of texts belonging to various business written genres, as is the case.

To conclude, we would like to briefly tackle the issue of Artificial Intelligence (AI) and the role it could play in the production of business correspondence, whose fixed nature might make it particularly prone to automatic text generation. There is no doubt that the power and ability of AIs to create human-like text is increasing at a very fast pace and the speed at which business letters could be written by applications like Chat GPT cannot be matched by a human being. Yet, the subjects involved in this experiment are students of a university degree whose professional profile is strongly related to the area of Second Language Teaching, which implies knowing the language and its features in depth in order to become professional language instructors in the uture. Therefore, using such means as Chat GPT to do tasks like the one presented above might not be beneficial for them in the long run, as they may have to design and teach ESP courses and should have developed the expertise to do so autonomously in advance. As regards the DDL experience in itself, close monitoring should be required to avoid such practices. Additionally, using AI content detectors [10] would also be recommendable to ensure the authenticity of the texts produced by the informants.

Nonetheless, and still along the lines of the use of generative AI in the ESP classroom, an experiment of this sort might benefit greatly from the work by Crosthwaite and Baisa (2023), who explore the expansion of DDL studies through the introduction of AI-generated materials used as a complement to actual corpora, since certain limitations that are intrinsic to the latter (such as the complexity in the use of corpus tools or the scarcity of available corpora in certain fields) could somehow be overcome by implementing a mixed, less conservative DDL methodology. In this particular case, and given the difficulty in finding corpus materials that could be employed to teach business English, employing AI-generated texts might be of help and also contribute to raise awareness amongst students about the advantages of using tools such as Chat GPT in a constructive and beneficial manner.

[Paper submitted 3 Jan 2024]
[Revised version received 28 Mar 2024]
[Revised version accepted for publication 5 Apr 2024]

[10] https://copyleaks.com/ai-content-detector

References

Anthony, L. (2022). *AntConc* (Version 4.2.0) [Computer software]. Waseda University. https://www.laurenceanthony.net/software

Argyroulis, V. (2022). Investigating student motivation in the use of corpus concordancing in ESP learning at university level. *ESP Today*, *10*(1), 71–98. https://doi.org/10.18485/esptoday.2022.10.1.4

Ashley, A. (2003). *Oxford handbook of commercial correspondence: Student's book*. Oxford University Press.

Aston, G. (1997). Involving learners in developing learning methods: Exploiting text corpora in self-access. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (204–214). Longman.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, *26*(3), 263–286. https://doi.org/10.1016/j.esp.2006.08.003

Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, *60*(3), 534–572. https://doi.org/10.1111/j.1467-9922.2010.00566.x

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348–393. https://doi.org/10.1111/lang.12224

Brieger, N. (2011). W*riting*. HarperCollins Publishers.

Chen, M., & Flowerdew, J. (2018). Introducing data-driven learning to PhD students for research writing purposes: A territory-wide project in Hong Kong. *English for Specific Purposes*, *50*, 97–112. https://doi.org/10.1016/j.esp.2017.11.004

Clerehan, R., Kett, G., & Gedge, R. (2003). Web-based tools and instruction for developing IT students' written communication skills. In A. Ellis, D. McGilvery, C. Spratt, & D. Weaver (Eds.), *Proceedings of exploring educational technologies* (16–17). Monash University.

Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, *2*(3), 223–235. https://doi.org/10.1093/applin/II.3.223

Crosthwaite, P., & Baisa, V. (2023). Generative AI and the end of corpus-assisted data-driven learning?: Not so fast! *Applied Corpus Linguistics*, *3*(3), Article 100066. https://doi.org/10.1016/j.acorp.2023.100066

Curado Fuentes, A. (2002). Exploitation and assessment of a business English corpus through language learning tasks. *ICAME Journal*, *26*, 5–32.

Elkasović, S., & Jelčić Čolakovac, J. (2023). The teaching of ESP lexis revisited: The effect of authentic text adaptation on vocabulary retention. *ESP Today*, *11*(1), 75–95. https://doi.org/10.18485/esptoday.2023.11.1.4

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, *20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Erman, B., Forsberg, F., & Lewis, M. (2016). Formulaic language in advanced second language acquisition and use. In K. Hyltenstam (Ed.), *Advanced proficiency and exceptional ability in second languages* (111–148). Walter de Gruyter. https://doi.org/10.1515/9781614515173-007

Flowerdew, L. (2012). Exploiting a corpus of business letters from a phraseological, functional perspective. *ReCALL*, *24*(2), 152–168. https://doi.org/10.1017/S0958344012000043

Forsberg, F., & Bartning, I. (2010). Can linguistic features discriminate between the CEFR-levels?: A pilot study on written L2 French. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (133–158). European Second Language Association.

Forsberg, F. (2021). Formulaicity. In N. Tracy-Ventura & M. Paquot (Eds.), *The Routledge handbook of second language acquisition and corpora* (370–381). Routledge.

Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996-2005. *Journal of English Linguistics*, *36*(1), 5–38. https://doi.org/10.1177/0075424207311247

Gavioli, L. (2005). *Exploring corpora for ESP learning*. John Benjamins. https://doi.org/10.1075/scl.21

Glabasova, D., Brezina, V., & McEnery, T. (2017). Exploring learner language through corpora: Comparing and interpreting corpus frequency information. *Language Learning*, *67*(1), 130–154. https://doi.org/10.1111/lang.12226

Granger, S. (2001). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (145–160). Oxford University Press. https://doi.org/10.1093/oso/9780198294252.003.007

Hadley, G. (2002). An introduction to data-driven learning. *RELC Journal*, *33*(2), 99–124. https://doi.org/10.1177/003368820203300205

Hiltunen, T., & Mäkinen, M. (2014). Formulaic language in economics papers: Comparing novice and published writing. In M. Gotti & D. Giannoni (Eds.), *Corpus analysis for descriptive and pedagogical purposes: ESP perspectives* (347–368). Peter Lang.

Hutchins, W. J. (1977). On the problem of “aboutness” in document analysis. *Journal of Informatics*, *1*(1), 17–35.

IBM Corp. (2021). *IBM SPSS Statistics for Windows* (Version 28.0.) [Computer software]. IBM Corp.

Karpenko-Seccombe, T. (2018). Practical concordancing for upper-intermediate and advanced academic writing: Ready-to-use teaching and learning materials. *Journal of English for Academic Purposes*, *36*, 135–141. https://doi.org/10.1016/j.jeap.2018.10.001

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, *1*(1), 7–36. https://doi.org/10.1007/s40607-014-0009-9

Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione [On the empirical determination of a distribution law]. *Giornale dell'Istituto Italiano degli Attuari*, *4*, 83–91.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33–51. https://doi.org/10.1177/026553229901600103

Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, *40*(5), 721–753. https://doi.org/10.1093/applin/amy012

Lougheed, L. (2002). *Business correspondence: A guide to everyday writing* (2nd ed.). Pearson Education.

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, *18*(1), 50–60. https://psycnet.apa.org/doi/10.1214/aoms/1177730491

Marín, M. J. (2015). Measuring precision in legal term mining: A corpus-based validation of single and multi-word term recognition methods. *ESP World*, *46*, 1–23.

Paquot, M. (2018). Phraseological competence: A missing component in university entrance language tests?: Insights from a study of EFL learners' use of statistical collocations. *Language Assessment Quarterly*, *15*(1), 29–43. https://doi.org/10.1080/15434303.2017.1405421

Paquot, M. (2019). The phraseological dimension in interlanguage complexity research. *Second Language Research*, *35*(1), 121–145. https://doi.org/10.1177/0267658317694221

Parizoska, J., & Rajh, I. (2017). Idiom variation in business English textbooks: A corpus-based study. *ESP Today*, *5*(1), 46–67. https://doi.org/10.18485/esptoday.2017.5.1.3

Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, *9*, 131–156. https://doi.org/10.1075/ijcl.9.1.07par

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (191–226). Longman.

Pérez-Paredes, P., Sánchez-Tornel, M., & Alcaraz Calero, J. M. (2012). Learners' search patterns during corpus-based focus-on-form activities: A study on hands-on concordancing. *International Journal of Corpus Linguistics*, *17*(4), 482–515. https://doi.org/10.1075/ijcl.17.4.02par

Scott, M. (2001). Mapping key words to problem and solution. In M. Scott & G. Thompson (Eds.), *Patterns of text: In honour of Michael Hoey* (109–128). John Benjamins. https://doi.org/10.1075/z.107.07sco

Scott, M. (2022). *WordSmith Tools* (Version 8) [Computer software]. Lexical Analysis Software.

Sharma, R. C., & Mohan, K. (2017). *Business correspondence and report writing: A practical approach to business and technical communication* (5th ed.). McGraw Hill.

Shi, J. (2013). The application of data-driven learning in business English teaching for MBA students in China: A critical assessment. *International Journal of English Linguistics*, *3*(1), 26–30. https://doi.org/10.5539/ijel.v3n1p26

Shiyab, S. M., & Halimi, S. (2015). *Writing business letters across languages: A guide to writing clear and concise business letters for translation purposes*. Cambridge Scholars Publishing.

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

Sinclair, J. M. (1998). The lexical item. In E. Weigand (Ed.), *Contrastive lexical semantics* (1–24). John Benjamins. https://doi.org/10.1075/cilt.171.02sin

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, *19*(2), 279–281. https://doi.org/10.1214/aoms/1177730256

Someya, Y. (2000, 9 August). *Online business letter corpus KWIC concordancer and an experiment in data-driven learning/writing* [Paper presentation]. The 3rd International Conference of the Association for Business Communication, Doshisha University, Kyoto, Japan.

Soto-Almela, J. (2024). Constructing persuasion in tourism promotion websites: A corpus-assisted study of hyphenated adjectives in English. *ESP Today*, *12*(1), 71–92. https://doi.org/10.18485/esptoday.2024.12.1.4

Sympson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487–512. https://doi.org/10.1093/applin/amp058

Taylor, S. (2012). *Model business letters, emails and other business documents*. FT Press.

Thurstun, J., & Candlin, C. N. (1998). Concordancing and the teaching of vocabulary of academic English. *English for Specific Purposes*, *17*(3), 267–280. https://doi.org/10.1016/S0889-4906(97)00013-6

Todd, R. W. (2001). Induction from self-selected concordances and self-correction. *System*, *29*(1), 91–102. https://doi.org/10.1016/S0346-251X(00)00047-6

Ueno, S., & Takeuchi, O. (2023). Effective corpus use in second language learning: A meta-analytic approach. *Applied Corpus Linguistics*, *3*(3), Article 100076. https://doi.org/10.1016/j.acorp.2023.100076

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Wulff, S. (2018). Acquisition of formulaic language from a usage-based perspective. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (17–37). Routledge.

Xia, D., Sulzer, M., & Pae, H. K. (2023). Phrase-frames in business emails: A contrast between learners of business English and working professionals. *Text & Talk*. https://doi.org/10.1515/text-2022-0027

Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, *3*(2), 215–229.

Mª JOSÉ MARÍN-PÉREZ, PhD, works as a lecturer and researcher at the English Department of the University of Murcia, Spain. Her major research interests are Corpus Linguistics, ESP and corpus-based discourse analysis. She has published her work in prestigious journals such as *Corpora*, *Terminology* and *ESP Today*.

MARÍA PILAR AGUADO JIMÉNEZ, PhD, is a tenured lecturer at the University of Murcia. She teaches Advanced English and ESP, especially business and legal English, to students from the Degree in English Studies taught at this university. Her research interests are ESP (business English), Corpus Linguistics and ELT. She has published with John Benjamins, also in *CALL* and *System*, among other journals.