http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

Digital Lexicography in ESP: Enhancing Vocabulary Acquisition and Reading Comprehension in Students of Commerce Degree

Beatriz Chaves-Yuste, *Department of English Studies, Complutense University of Madrid, Madrid, Spain (bchave01@ucm.es) (https://orcid.org/0000-0002-4442-9138)*
and
Manuel Márquez-Cruz, *Department of Linguistics, Arabic, Hebrew, Basque and East Asian Studies, Complutense University of Madrid, Madrid, Spain (manmarqu@ucm.es) (https://orcid.org/0000-0001-9040-169X)*

Abstract: This study investigates the impact of Sketch Engine and Lexonomy on vocabulary acquisition and reading comprehension in English for Specific Purposes (ESP), focusing on international commerce among undergraduate students. Conducted over thirteen weeks, the research involved 90 first-year Commerce students from a public university in Madrid. A pre-experimental design was implemented, dividing participants into control and experimental groups. The control group used traditional analogue dictionaries, while the experimental group utilized digital tools to compile specialized dictionaries. Performance was assessed through pre- and post-intervention reading comprehension tests and a questionnaire assessed the use of lexicographical resources by ESP undergraduate students. On the one hand, results revealed that the use of these two lexicographical digital tools implies significant improvements regarding vocabulary acquisition and reading comprehension. On the other hand, digital tools emerged as the preferred resources, emphasizing translation, contextual meaning, and voice search functionalities, though challenges such as navigation difficulties and interpreting phonetic transcriptions were noted. The findings underscore the pedagogical potential of digital lexicography to enhance ESP instruction, foster autonomous learning, and equip students with essential professional skills. This research provides a foundation for integrating innovative methodologies into higher education, contributing to more effective and equitable teaching practices.

Keywords: SKETCH ENGINE, LEXONOMY, DIDACTIC LEXICOGRAPHY, ENGLISH FOR SPECIFIC PURPOSES, READING COMPREHENSION, VOCABULARY ACQUISITION

Opsomming: Digitale leksikografie in ESD: Die verbetering van woordeskatverwerwing en leesbegrip by handelstudente. In hierdie studie word die impak van Sketch Engine en Lexonomy op woordeskatverwerwing en leesbegrip in Engels vir Spesifieke

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

Doeleindes (ESD) ondersoek. Daar word spesifiek gefokus op internasionale handel onder voorgraadse studente. Die navorsing wat oor dertien weke uitgevoer is, het 90 eerstejaar-handelstudente van 'n openbare universiteit in Madrid betrek. 'n Pre-eksperimentele ontwerp wat deelnemers in kontrole- en eksperimentele groepe verdeel, is geïmplementeer. Die kontrolegroep het tradisionele analoogwoordeboeke gebruik, terwyl die eksperimentele groep digitale hulpmiddels gebruik het om gespesialiseerde woordeboeke saam te stel. Prestasie is aan die hand van pre- en postintervensieleesbegripstoetse beoordeel en met behulp van 'n vraelys is die gebruik van leksikografiese hulpbronne deur ESD-voorgraadse studente beoordeel. Aan die een kant het die resultate aangedui dat die gebruik van hierdie twee leksikografiese digitale hulpmiddels beduidende verbeterings rakende woordeskatverwerwing en leesbegrip inhou. Aan die ander kant het digitale hulpmiddels na vore getree as die voorkeurhulpbronne met die fokus op vertaling, kontekstuele betekenis, en stemsoektogfunksies. Uitdagings soos navigasieprobleme en die interpretasie van fonetiese transkripsies is egter ook aangeteken. Die bevindings beklemtoon die opvoedkundige potensiaal van die digitale leksikografie om ESD-onderrig te verbeter, outonome leer te bevorder en studente met noodsaaklike professionele vaardighede toe te rus. Hierdie navorsing vorm 'n basis vir die integrering van innovorende metodologieë in die hoër onderwys, wat sodoende bydra tot meer effektiewe en billike onderrigpraktyke.

Sleutelwoorde: SKETCH ENGINE, LEXONOMY, DIDAKTIESE LEKSIKOGRAFIE, ENGELS VIR SPESIFIEKE DOELEINDES, LEESBEGRIP, WOORDESKATVERWERWING

1. Introduction

Lexicography is a millennia-old discipline which, in its practical aspect, has been focused on the compilation of comprehensive or partial vocabularies of one or several languages. In its theoretical aspect, it has focused on the study of the construction of these vocabularies throughout history from a prescriptive, descriptive and historical perspective, or on the methodologies used in the compilation of such works. In this context, the dictionary has emerged as the flagship product, though early lexicographic manifestations in Europe and Asia often took the form of word lists or glosses added to specific works (cf. Tarp and Gouws 2023: 429-430). These glosses facilitated the comprehension of terms that might prove complex for a community, due to the antiquity of the vocabulary, or its technical, religious, scientific, or literary nature, among others (Boisson et al. 1991: 261-262; Béjoint 2016: 8). Such was the case with the Homeric glosses written by ancient scribes in the works of Homer (Tarp 2019: 11), or with the *Cangjiepian*, a textbook dated around 220 B.C. used to teach children Chinese writing, which could only be read with the help of specialized annotations (Yong and Peng 2008: 16). These glosses and annotations were later compiled into independent books, whose diffusion reached significant levels thanks to the technological revolution of the printing press, and more recently, through their publication in digital format.

From the points outlined above, it is evident that one of the primary purposes of a language dictionary — among others — is pedagogical. This pedagogical function is even more apparent in dictionaries specifically designed to

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

enhance the process of language learning, whether for the native language or a second or foreign language (Márquez Cruz 2024: 64). This field is known as pedagogical lexicography, and the works derived from its practice are typically categorized, depending on the type of user, into school dictionaries and learner's dictionaries. These works aim to meet the cognitive and communicative needs of learners, either of their native language (school dictionaries) or of a second or foreign language (learner's dictionaries). The lexicographic structures of both types of dictionaries are determined by the quantity, quality, and organization of data provided in the paratexts and nomenclature, in the headwords (macrostructure), in the lexicographic entries (microstructure), in the relationships between different parts and elements of the dictionary (mediostructure), and in the images and/or multimedia elements (iconostructure; cf. Camacho 2017: 2). Additionally, the vocabulary used in the definitions and the typography of the textual conglomerate that makes up the lexicographic article should also be considered. Regarding the type of information provided in these dictionary structures, Heuberger (2016: 30-40) identified a series of features that should be present in learner's dictionaries, which we also believe should be part of school dictionaries: (i) the use of controlled defining vocabularies, (ii) definitions, (iii) examples, (iv) grammatical information, (v) usage information (pragmatics), (vi) collocations, (vii) pronunciation, (viii) data accessibility, (ix) preliminary texts, and (x) appendices (final materials).

Regarding school dictionaries, it is crucial to examine the outer texts, both the preliminary and concluding sections of the dictionary in conjunction with its macrostructure, microstructure, mediostructure, iconostructure, and target users. Many of these dictionaries are published — in analogue format — in lightweight, manageable volumes (especially when the target user group is primary or secondary school students). They use italics, different font sizes, and various colours to differentiate content within the dictionary entries. They also feature a selective and appropriate lexicon for the school vocabulary, including non-standard terms, such as some neologisms. In addition, they employ controlled vocabularies (between one thousand and two thousand words) to ensure that the definitions are simple and provide ample examples, as well as illustrations, grammatical indications, synonym and antonym relationships, lexical families, semantic fields, idiomatic expressions, and collocations. Finally, they usually contain grammatical appendices, and the prefaces often provide guidance on dictionary usage, including explanations of the various components of dictionary entries based on word type, abbreviations, etc. for users who may not be familiar with them.

Learner's dictionaries differ from school dictionaries — those created for language learning by native speakers — in that they tend to be slightly larger in size. They use different colours and font sizes to distinguish the various types of data provided in dictionary entries. The order of definitions generally follows usage frequency, with detailed descriptions of phraseology, idioms, and collocations, and synonyms and antonyms. This information is followed by the use of descriptors and guide words, syntagmatic or sentence-based definitions through

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

equivalence or examples (Climent 2005: 363-366), phonetic information, grammatical information, usage labels, grammatical appendices, and annexes with varied information on culture, economics, geography, or encyclopaedic content (currencies, weights, places, famous figures, etc.). The iconostructure is usually more extensive in bilingual dictionaries than in monolingual ones in order to visually convey the relationship between the target language and the mother tongue.

The described features provide a concise theoretical framework on the essence of the dictionary models used in language learning classrooms. This may seem unusual given the learning context of the present study, which focuses on English for Specific Purposes (henceforth ESP), specifically English in the field of Commerce. We should argue that the lexicographic proposal we present should be underpinned by theoretical principles and practical developments specific to specialized lexicography, terminology, and terminography. It is true that the creation of a commerce dictionary should adhere to this branch of lexicography and that, therefore, we would need to describe a theoretical framework more aligned with specialized dictionaries. However, the primary objective of our lexicographic proposal is to enhance the acquisition of grammatical and lexical knowledge of the English language, using commerce as a thematic area, but without intending to create a dictionary whose purpose is to provide definitions of commerce-specific terms. In other words, the goal is not to create a dictionary of commerce in English, but to improve linguistic knowledge of English, especially, the reading comprehension, through the lexicographic coding of lexical and grammatical structures that are characteristic of, but not exclusive to, the English language in the context of commerce. For this purpose, first-year Commerce students at a public university in Madrid developed a monolingual English dictionary, with its nomenclature based on the compilation and analysis of a scientific corpus focused on international commerce.

2. Corpus Linguistics

Of the many different definitions of Corpus Linguistics (hereafter, CL), one of the most current and, in our opinion, one that clearly and precisely captures the essence of this very eclectic discipline is the definition provided by Rojo (2021: 48). He (ibid.) described it as an approach to the study of linguistic phenomena with an empirical orientation, based on the detailed analysis of large amounts of data (corpora). As Rojo (2021: 48) argued, CL is characterized by its empirical nature, deriving data from naturally produced texts collected in the form of text corpora. In this task, the use of computers is essential to store and manage the corpora, employing both qualitative and quantitative analytical techniques, and conducting systematic and exhaustive analyses of the data provided by corpora. It is worth noting that working with computers has not always been a must for corpus analysis, as Shaxrizoda (2024: 104-105) pointed out, given that corpus-based research existed prior to the use of computers for analysing linguistic data.

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

Regarding the concept of a corpus, it is an organized collection of oral and/or written texts in a language, compiled to exemplify its use. The coding of the corpus should adhere to a standard that allows for the retrieval of information using different systems based on various linguistic criteria. Therefore, when compiling a corpus, it is essential to consider the language(s) or variety(ies) to ensure representativeness, as well as the text's mode (oral or written), type (e.g., book, magazine, advertisement, etc.), domain (academic or popular), location, and date (Sinclair 2005: 8). The size of the corpus will depend on several factors, including its intended use; however, whenever possible, it should consist of complete documents to avoid any bias.

Corpus-based research follows two main approaches: starting from corpus data to develop theories (corpus-driven) or consulting corpora to verify theories (corpus-based). In either case, computational tools are essential for compiling and analysing corpora automatically or semi-automatically, reducing the noise caused by irrelevant or questionable-quality information in the corpus (Costa et al. 2015: 74). Tools like Babouk, BootCat, Corpógrafo, and Terminus (Toledo and Camero 2021: 20) are used for terminological extraction and corpus compilation and analysis. Among these, BootCat stands out for extracting large amounts of digital textual material using seed words representative of a specific semantic field. BootCat randomly combines these seed words to generate sequences called tuples, which are then used as search strings in Google to retrieve web pages. The retrieved HTML documents are cleaned and converted into a plain text (Flores 2022: 174). An extension of BootCat is WebBootCat, a tool based on the seed words technique used by Sketch Engine.

3. Integrating corpus compilation and digital lexicography in ESP vocabulary acquisition

With the aim of enhancing the effectiveness of the vocabulary acquisition process in the domain of ESP — specifically, Business English — through lexicographic tasks, two tools were made available to the students: Sketch Engine (https://www.sketchengine.eu/) and Lexonomy (https://www.lexonomy.eu/). The former was employed to compile and manage a corpus that served as the foundational resource for the dictionary developed by the students. The latter was used, first, to design the macrostructure and microstructure of the dictionary and, subsequently, to input lexicographic data, publish the entries, and share them among all participants.

The corpus was compiled from nine business-themed English texts. The Word Sketch tool was used to explore collocations and word combinations; Wordlist, to identify the most frequent lexical items; Keywords, to extract core terminology specific to the commercial domain; and Concordances, to analyse the contextual usage of high-frequency keywords and obtain authentic usage examples.

Regarding Lexonomy, the microstructure of the dictionary entries was

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

designed according to the following core lexicographic information: lemma, part of speech, British and American English pronunciation, definition, and examples of use.

The theoretical principles of corpus linguistics and the use of corpus compilers and analysers, as well as Dictionary Writing Systems,[1] have already been successfully implemented in various teaching practices aimed at promoting incidental lexical learning in foreign languages classrooms. Chaves-Yuste and Márquez-Cruz (2024: 183; 2022: 364) demonstrated that Lexonomy, used to create a cooperative digital dictionary, helped to improve reading comprehension, vocabulary and linguistic competence of secondary education students. They demonstrated that Lexonomy was an efficient tool in a b-learning context and facilitated vocabulary acquisition. However, in any of these studies they used it in combination with Sketch Engine to compile the corpus they were working with.

Srdanović (2020: 1059-1060) implemented a classroom project in the field of Tourism in which instructors and students collaboratively utilized Sketch Engine and Lexonomy to create a bilingual Japanese-Croatian dictionary. The study showed two main findings: (i) the use of these advanced tools fosters students' proficiency in leveraging technologies for linguistic analysis and lexicographic resource creation; (ii) the pressing need to train professionals capable of applying Japanese language skills within the tourism sector (Mavrić et al. 2023: 51). Furthermore, the benefits of these tools extend beyond language acquisition. They impart crucial knowledge and skills to dictionary users, enabling them to maximize the utility of these resources. As argued by Caruso (2024: 139), when students engage in the process of creating dictionaries in one or more second languages, they achieve multifaceted gains. By compiling corpora using Sketch Engine and designing dictionary structures with Lexonomy, students not only enhance their passive and active knowledge of the target language but also deepen their understanding of the microstructural organization of lexicographic entries in learner dictionaries.

However, it is essential to acknowledge that these tools comprise certain limitations, and their implementation should be carefully evaluated. First, Sketch Engine requires the purchase of a license, which may limit its accessibility for some educational institutions. Second, while Lexonomy is freely available, it has constraints concerning data storage capacity, which could impact the scalability of certain projects (Kovalevskaitė et al. 2020: 246). Thus, the use of Sketch Engine and Lexonomy presents a valuable opportunity to integrate technology into linguistic and lexicographic education, equipping students with practical and theoretical expertise. Nonetheless, educators must weigh the tools' benefits against their financial and technical limitations to ensure optimal application in educational contexts. Hence, this work aims to explore the benefits of the compilation of a digital learner's dictionary using Sketch Engine and Lexonomy in students' acquisition of commerce vocabulary. The objective is twofold: On the one hand, to measure the effect of the compilation of a corpus and a dictionary on students' acquisition of specific vocabulary of commerce (especially, inter-

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

national negotiation) which helps students develop their reading comprehension of business and trade texts, and on the other hand, to know the real use of the available lexicographic tools for ESP students who need to foster English for international commerce firstly, for academic purposes, and secondly, for their future professional endeavours. Thus, the research questions raised are:

1. To what extent do the corpus compiler software Sketch Engine and the Dictionary Writing System Lexonomy influence commerce students' acquisition of specific commerce vocabulary and development of their reading comprehension?
2. Which is the real use of commerce students of the available lexicographic tools in their English for commerce learning process?

Based on the scientific literature and the proposed research questions, the following hypotheses are formulated:

— Hypothesis 1 (H1): The use of the corpus compiler software Sketch Engine and the Dictionary Writing System Lexonomy positively impact on the acquisition of commerce vocabulary and on the development of reading comprehension by commerce undergraduate students.
— Hypothesis 2 (H2): Students use more digital than analogue lexicographic tools in the English for commerce learning process.

This study aims to provide a preliminary contribution to the scientific understanding of vocabulary acquisition through the creation and use of lexicographic tools by students. We believe that through an incidental, intentional, and lifelong learning experience, students can acquire the target specialized vocabulary necessary to develop their professional skills successfully.

4. Design

This research follows a quantitative approach, gathering data through a quasi-experimental design. On the one hand, it collects quantitative data by measuring academic performance before and after an intervention in which the creation of either an analogue or digital dictionary was employed in two groups of university students randomly assigned as a control group and an experimental group. On the other hand, it also gathers quantitative data concerning the use and evaluation of lexicographic tools, particularly dictionaries, for the learning of English for professional purposes. The research spanned thirteen weeks, from February 6 to May 10 of the 2023–2024 academic year.

Initially, the academic scores obtained in the PRE reading comprehension test in the field of commerce were collected, administered to all the students before the intervention began. After thirteen weeks of instruction a similar test was administered (POST), providing a text which gathered some of the terms collected in the different dictionaries. During the instruction period, the stu-

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

dents of the experimental group worked in heterogeneous cooperative teams and compiled and used dictionaries using Sketch Engine and Lexonomy. Meanwhile, the control group worked with analogue traditional glossaries. The dependent variable in this study is the acquisition of a specialized professional lexicon, measured through the scores of these two tests (PRE and POST), administered before and after compiling the various dictionaries in both the control group and the experimental group. The independent variable is the tool used and the procedure followed, specifically the use of digital lexicographic tools such as Sketch Engine and Lexonomy.

4.1 Participants

The sample was selected based on accessibility, using a non-probabilistic purposive sampling of 90 first-year students enrolled in the Commerce degree at a public university of Madrid. These students were recruited through purposing samplings since one of the researchers was the instructor of two pre-existing groups (the control group and the experimental group) which were naturally split up. The criteria established for selecting the sample were: being enrolled in the first year of the Commerce degree, attending classes regularly, completing all the tests and providing signed consent to participate in the research. Of all participants, 54.4% were male (n=49) and 45.6% were female (n=41). The average age of the participants was 19.58 years, with a standard deviation (SD) of 1.83 years. None of the students had prior experience using Sketch Engine or Lexonomy. However, all participants demonstrated familiarity with technological devices and exhibited a moderate level of digital competence, particularly when operating their mobile phones.

4.2 Instruments

In order to assess the usefulness of these digital lexicographic tools in the development of our students' professional skills, two custom-designed tests were used for data collection. The first test was administered in February (February 6), prior to the experimental use of digital lexicographic tools in the experimental group, and the second test in June (June 19), after the classes had ended. The second test was administered after the conclusion of the semester and at the onset of the examination period. During this five-week interval, students were assessed in various subjects within their degree program, completed internships at local companies, and presented their end-of-year assignments. Therefore, despite the time gap, no significant external factors are considered to have influenced the results obtained in the post-test. Both tests were two reading comprehension exercises, part four of the official Business English Certificate (BEC) Vantage (level B2, according to the Common European Framework of Reference for Languages, Council of Europe, 2001) test (University of Cambridge ESOL

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

Examinations 2015). Students needed to read a text and choose the best word to fill each gap (fifteen multiple choice items with four options each). All students took the tests, which were scored from 0 to 10 (0.66 points each correct answer), with the following scale: 0-4.99 = fail, 5-6.99 = pass, 7-8.99 = good, and 9-10 = excellent. After the intervention, a questionnaire, based on Knežević et al.'s (2021), containing 51 Likert scale questions was administered to assess students' evaluations regarding the actual use of lexicographic tools, particularly dictionaries, in the learning of ESP. The content validity of this questionnaire was analysed by displaying the content of this instrument to a panel of four experts in lexicography and foreign language instruction. They validated its content by guaranteeing that the items addressed all related aspects of the objectives of the study (Roebianto et al. 2023) and they agreed on their coherence and relevance. The KMO value of this questionnaire was 0.641 and the Bartlett's Test of Sphericity value was significant (p=0.001), so factor analysis was appropriate. The first fifteen components explained 78.126% of the accumulated variance, which is a valid and acceptable percentage since it exceeds the 60% threshold. This implies that the questionnaire captured a significant amount of information and that the initial eigenvalues of the extracted squared loadings corroborated that the retained components of the questionnaire were relevant. Its reliability was assessed using the rank-biserial correlation (r) with a reliability coefficient of 0.575, which indicated a moderate effect size.

4.3 Procedure

The study was first introduced to the students, and informed consent was obtained, allowing them the option to withdraw from the study at any time. The didactic experience was then conducted from February through mid-May, with a total of 32 hours of instruction, scheduled from 1:00 p.m. to 3:00 p.m. To avoid bias, the same instructor taught both the control and experimental groups. The tests were administered in the regular classroom under proper conditions of sound, lighting, and ventilation. The participation rate was 100%, meaning that all 90 students who took the PRE test also completed the POST test and the questionnaire on dictionary usage. The data collected were treated following the ethical guidelines of the Declaration of Helsinki, ensuring confidentiality.

4.4 Data analysis

Quantitative data, both for descriptive and inferential analysis, were examined using SPSS software, version 29 for Windows (IBM 2016). First, the normality of the variables was checked using the Kolmogorov-Smirnov test, considering the sample size, which indicated that neither test met the normality criteria (PRE p=0.001; POST p=0.001). Second, descriptive statistics such as mean and standard deviation were used for each variable, and non-parametric inferential

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

statistics such as the Mann-Whitney U test (for independent samples) were employed to compare the PRE and POST test results in the control and experimental groups, thus verifying the research objective and hypothesis. Third, the PRE and POST test results were compared within the two groups employing the Wilcoxon test. Fourth, the effect size of the questionnaire was calculated and interpreted using ranks-biserial correlation (r), and its control validity was explained through a factor analysis. All analyses were conducted with a significance level of $p < 0.05$. Finally, we did a descriptive analysis of the means and standard deviation obtained in the different items of the questionnaire to analyse the answers provided by students.

5. Results

The findings presented here suggest a potential causal link between the use of teaching dictionaries and the enhancement of specialized lexicon knowledge among students of Commerce. However, it is important to interpret these results with caution, acknowledging possible limitations and alternative explanations. The descriptive analysis is presented in Table 1, showing the mean and standard deviation of both tests, as well as the inferential analysis results that address the research objectives and hypotheses. In the PRE test, the average scores in both groups exhibited comparable levels of lexicon knowledge, with no significant differences (p=0.35; r=0.13). However, in the POST test, the scores of the control group were significantly lower than the average scores obtained by the experimental group, with a difference of 4.68 points. The Mann-Whitney U test indicated significant differences between the scores (p=0.001; r=0.49), which implies that the experimental group benefited more from the intervention, namely, the creation and use of digital learner's dictionaries created by the students. While these results appear to support the alternative hypothesis and reject the null hypothesis, we should be cautious when attributing the improvement solely to the intervention since other factors, such as their individual learning preferences, may have influenced the intervention. Notwithstanding, the findings suggest that the use of digital learner's dictionaries may have contributed to improve lexicon acquisition and professional skills development in the field of international trade.

Table 1: Descriptive results of control and experimental groups

	Control group		Experimental group	
Variables	M	SD	M	SD
PRE test	7.78	1.63	7.62	1.50
POST test	4.36	2.72	9.04	1.41

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

These results corroborate the effectiveness of creating and utilizing digital lexicographic tools to succeed in using English in professional contexts with sufficient fluency and accuracy, enabling students to perform effectively in the labour activities demanded by the commercial sector.

To complement the analyses and address the effect of confounding variables such as gender or age, the non-parametric Mann-Whitney U test was employed. As observed in Table 2, neither the age nor the gender significantly influenced the academic performance in relation to dictionary use. This strengthens the claim that the observed differences were likely due to the intervention. Both men and women do not differ in academic performance in creating and using analogic dictionaries or digital dictionaries.

Table 2: Complementary inferential results by gender and age

Variables	Genre		Age	
	U	*P*	*U*	*P*
Analogue dictionaries	2.81	0.356	2.38	0.919
Sketch Engine + Lexonomy digital dictionary	2.82	0.207	2.02	0.553

Finally, an additional quantitative study was also conducted to analyse the actual use of dictionaries in students' learning process of English for commerce. The high reliability of the administered questionnaire (Cronbach's alpha= 0.946) adds credibility to these findings. This high score confirms that the instrument is both valid and reliable for the purposes of this study. A descriptive analysis was subsequently conducted, focusing on the mean and standard deviation of the variables under study. Particular attention was given to the types of dictionaries used, the reasons for their consultation, the nature of the queries, and the difficulties encountered which reveal nuanced patterns in dictionary usage.

As illustrated in the accompanying figure (Figure 1), digital dictionaries, related applications and Spanish–English bilingual dictionaries emerged as the most frequently used resources, which reflect broader trends in technology adoption rather than the unique advantage of digital learner's dictionaries. Conversely, paper dictionaries, monolingual dictionaries, and specialized bilingual dictionaries for professional purposes were the least utilized. For a better interpretation of the figures, the most frequent answers have been presented in green whereas the most infrequent ones are coloured in red.

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

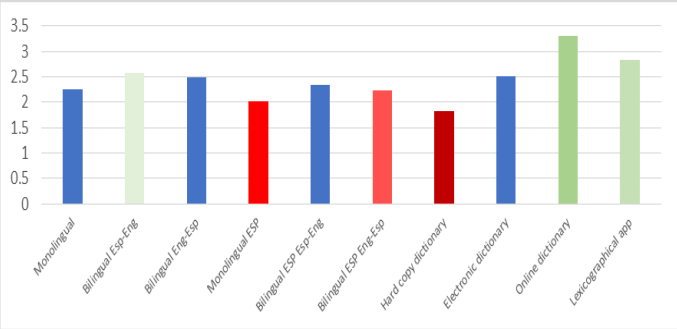

Figure 1: Types of dictionaries

Regarding the reasons for dictionary use displayed in Figure 2, the most common motivations included determining the meaning of unfamiliar words, completing activities that explicitly required dictionary use, and verifying the accuracy of an inferred meaning. However, it is infrequent to use dictionaries for speaking and listening tasks in English or for improving pronunciation skills. This suggests that their impact may be domain-specific rather than comprehensive.

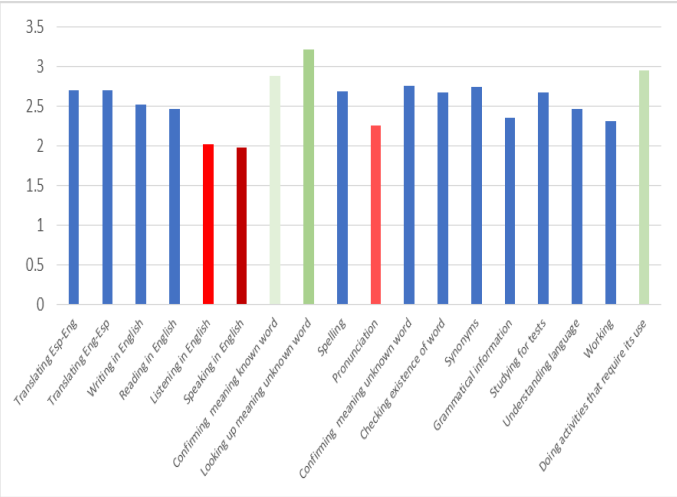

Figure 2: Reasons for dictionary use

As can be observed in Figure 3, in terms of the type of information sought, the most frequent queries were related to accessing the meaning of terms within specific contexts, finding equivalent translations, and using voice search functions. Features such as appendices, images, and usage guides were notably less popular among users.

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

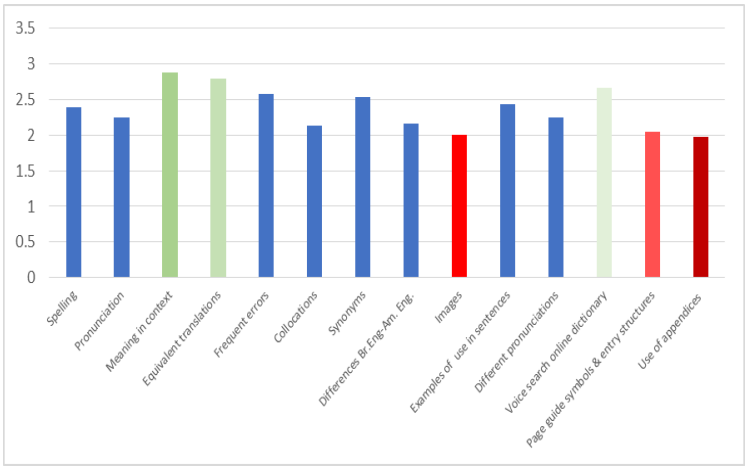

Figure 3: Types of information sought

Finally, regarding the primary difficulties reported in using dictionaries summarized in Figure 4, participants included challenges in navigating the tools, misunderstanding the information provided, and the need to consult multiple references depending on specific needs. These difficulties highlight areas for improvement in integrating digital dictionaries into pedagogical strategies. The least frequent issues included unfamiliarity with abbreviations, difficulty assessing the quality of a dictionary, and challenges in interpreting phonetic transcriptions.

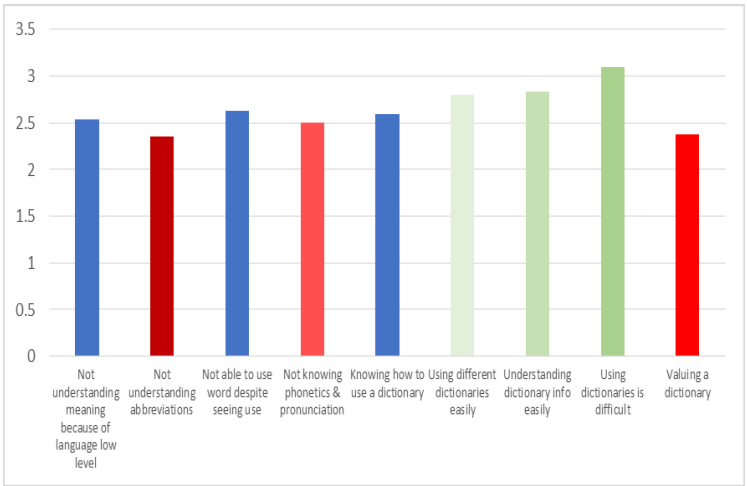

Figure 4: Difficulties encountered when using a dictionary

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

6. Discussion

This study investigates whether incorporating digital lexicographic tools, specifically the corpora compiler Sketch Engine and the DWS Lexonomy, enhances vocabulary acquisition and reading comprehension of scientific texts in the field of international commerce among undergraduate students studying English for Specific Purposes . Conducted in a face-to-face educational context, the research involved 90 students randomly assigned to either a control or an experimental group. Over an equal duration of instruction, the control group relied on traditional teaching methods, including the creation and use of analogue paper dictionaries or glossaries, while the experimental group utilized the aforementioned digital lexicographic tools. To assess the impact of these instructional approaches, participants completed two standardized reading comprehension tests (BEC Vantage, Cambridge Exam). These tests were based on business and international commerce texts, incorporating the vocabulary featured in their respective dictionaries. The findings indicate significant improvements in the experimental group's reading comprehension, attributed to their enhanced vocabulary acquisition. The target vocabulary was acquired through both incidental and intentional learning approaches. As a result, their reading comprehension improved substantially, particularly in the use of specialized vocabulary — an essential skill for success in their professional careers within the commercial sector. Moreover, the practical component of the study enabled students to not only create their own dictionaries but also to navigate and utilize lexicographic resources with proficiency, collaboratively managing entries produced by their peers. These results validate Hypothesis 1, confirming that the integration of Sketch Engine and Lexonomy in ESP instruction fosters vocabulary development and, consequently, reading comprehension skills. The study underscores the potential for these tools to optimize ESP teaching and learning at the university level and in other educational contexts, such as Vocational Education and Training (VET), where specialized English proficiency is essential for professional success. To ensure that gender and age did not bias the results, an inferential analysis was conducted, revealing no significant differences between the groups. These findings strengthen the reliability of the study's conclusions, supporting the idea that digital lexicographic resources can be broadly implemented without demographic limitations. Descriptive analyses reveal a compelling need to optimize university students' academic performance through a balanced approach combining traditional teaching and digital resources. These results align with Bartolomé-Díaz and Frontini (2020: 67), who examined the use of Sketch Engine and Lexonomy in creating a bilingual French–Spanish architectural dictionary. Their findings highlighted the innovative potential of digital lexicography while emphasizing the need for further development to achieve full operational capability. This study supports prior research emphasizing the critical role of vocabulary acquisition in enhancing academic performance. For instance, Shahjahan et al. (2021: 1145) argued that well-prepared graduates exhibit stronger academic and professional competen-

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

cies, while Atik and Çelik (2021: 124-125) demonstrated that academic achievement yields broader personal and societal benefits, including motivation, engagement, and life-long learning. To this end, universities must prioritize pedagogical innovation, as suggested by Van Laar (Van Laar et al. 2020: 10), by integrating active methodologies and digital tools into the curriculum. Such practices position students as active participants in their learning processes, fostering motivation and engagement (Snezhko et al. 2022: 88). The study further validates Hypothesis 2, revealing that students favour digital lexicographic tools over analogue methods in their ESP learning. This finding aligns with contemporary research (such as Alshammary and Alhalafawy 2022: 2294; Jia et al. 2024: 13), which highlights the efficacy of digital tools in enhancing vocabulary acquisition and academic performance. Digital resources emphasize active, student-centred learning, personalized instruction, and innovation (Strelan et al. 2020: 16). The increasing adoption of digital ecosystems in higher education, as advocated by Nguyen and Tuamsuk (2022: 2), necessitates the alignment of teaching practices with frameworks such as DigCompEdu (Redecker and Punie 2017: 20). These frameworks emphasize the development of educators' digital competence and provide practical tools to integrate digital strategies effectively. For students, the use of such resources not only enhances learning outcomes but also equips them with essential 21st-century skills, promoting lifelong learning (Molnár et al. 2022: 274).

This study has certain limitations. The sample size restricts the generalizability of findings, necessitating cautious interpretation. Additionally, while efforts were made to control for sociodemographic variables, participants were selected based on convenience and accessibility rather than through randomized sampling. Despite these constraints, the study provides valuable insights into the pedagogical application of digital lexicographic tools, offering a foundation for future research. Future studies could extend these findings to online university settings, other ESP contexts, and additional languages or disciplines. Longitudinal research would also be valuable, providing a deeper understanding of the long-term impacts of digital lexicography on vocabulary acquisition and reading comprehension. Expanding the scope of analysis to include other variables, such as professional performance, could further elucidate the broader implications of this approach. Investigating the relationship between digital resource use and other professional competencies, such as oral communication and writing skills, could further clarify their role in holistic language development. By building on the current findings and addressing existing limitations, future studies can continue to advance the integration of digital lexicography in ESP instruction, fostering more effective and equitable educational practices in a rapidly evolving academic and professional landscape.

7. Conclusion

This study demonstrates the potential of digital lexicographic tools, such as Sketch Engine and Lexonomy, to revolutionize vocabulary acquisition and reading

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

comprehension in ESP education. By placing students at the centre of the learning process and equipping them with specialized language skills, these tools prepare undergraduates to meet the demands of their professional fields. This research provides educators with a roadmap for enhancing both student learning outcomes and professional preparedness. The results emphasize the necessity of adopting innovative teaching methodologies that leverage digital resources. Such tools not only optimize students' linguistic competence but also foster their ability to meet the demands of professional environments requiring specialized English skills. Specifically, the creation and use of digital learning dictionaries empower students to develop autonomy in learning, encouraging active participation and critical engagement with language content. This pedagogical approach aligns with contemporary trends in higher education, which prioritize student-centred and technology-enhanced learning. Furthermore, the study reveals the importance of addressing common challenges faced by students in using digital dictionaries, such as navigation difficulties and the need for targeted guidance on interpreting lexicographic data. These insights suggest that educators should incorporate explicit instruction on the effective use of such tools, tailoring support to diverse learning styles and levels of digital literacy. The findings also highlight the role of dictionary use in promoting incidental and intentional vocabulary learning. By integrating tasks that require students to infer word meanings, verify accuracy, and apply vocabulary in context, educators can create a learning environment that mirrors real-world professional scenarios. Such practices not only strengthen students' lexical competence but also enhance their confidence in using English for specialized purposes.

The significant improvements observed in the experimental group's reading comprehension scores provide robust evidence that the use of Sketch Engine and Lexonomy supports the development of key linguistic and professional skills. These results validate the hypothesis that digital dictionaries are instrumental in equipping students with the fluency and accuracy required for success in the commercial sector. The descriptive and inferential analyses further illustrate the prevalence of digital dictionaries over traditional analogue resources in the learning process. Students reported a strong preference for tools that facilitate contextual understanding, translation, and voice search functions, reflecting a shift towards more dynamic, interactive modes of language learning. However, challenges such as tool navigation, and the interpretation of phonetic transcriptions indicate areas where additional instructional support is needed. Importantly, the findings reveal no significant differences in outcomes based on gender or age, suggesting that the benefits of digital lexicography are broadly applicable across diverse student demographics. This universality underscores the potential of these tools to enhance educational equity and inclusivity in ESP instruction.

Despite the positive results of this intervention, further studies are recommended to confirm these results across diverse educational contexts. In addition, additional factors influencing the effectiveness of digital lexicographic tools,

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

such as the students' different learning styles or affective factors that may influence their learning process could also be analysed. Moreover, apart from the quantitative methodology employed in this study, further studies would also benefit from qualitative methods, such as semi-structured interviews with students or open-ended surveys. They would provide new insights about how students experience and struggle with digital lexicographic tools and which challenges need to be met. Nevertheless, these results provide strong evidence supporting the efficacy of digital learner's dictionaries in enhancing specialized lexicon knowledge taking into account the domain-specific limitations of this study.

As digital innovation continues to reshape higher education, integrating such resources represents a critical step toward fostering autonomous, meaningful learning experiences that align with the needs of contemporary society. Furthermore, the administered questionnaire provided valuable insights about the real use of lexicographic tools by ESP Spanish undergraduate students, enabling targeted interventions to address these gaps in classroom settings and thereby enhance students' learning processes.

Endnote

1. These systems are referred to by various names. Tiberius and Heylen (2022) document terms such as *dictionary editing systems*, *dictionary compilation software*, *lexicography software or dictionary production software*, *lexicographic workbench*, *dictionary management system or lexicographer's workbench*, *dictionary editing tool*, *dictionary building software*, or simply *editorial system*. Similarly, Rubio et al. (2021: 342) highlighted additional terminology found in the scientific literature: *dictionary building platform* (Mangeot and Chalvin 2006: 1666), *dictionary content management system* (Alegría et al. 2006: 2), *dictionary editor and browser* or *DEB platform* (Horák and Rambousek 2007: 129), and *dictionary writing and publishing systems* (Měchura 2017: 1).

References

Alegría, Iñaki, Xabier Arregi, Xabier Artola, Mikel Astiz and Leonel Ruiz Miyares. 2006. A Dictionary Content Management System. Corino, Elisa, Carla Marello and Cristina Onesti (Eds.). 2006. *Atti del XII Congresso Internazionale di Lessicografia, Turin, Italy, 6–9 September 2006. Vol. 1:* 105-109. Alessandria: Edizioni dell'Orso.

Alshammary, F.M. and W.S. Alhalafawy. 2022. Sustaining Enhancement of Learning Outcomes across Digital Platforms during the COVID-19 Pandemic: A Systematic Review. *Journal of Positive School Psychology* 6(9): 2279-2301. https://journalppw.com/index.php/jpsp/article/view/12650

Atik, S. and O.T. Çelik. 2021. Analysis of the Relationships between Academic Motivation, Engagement, Burnout and Academic Achievement with Structural Equation Modelling. *International Journal of Contemporary Educational Research* 8(2): 118-130. https://doi.org/10.33200/ijcer.826088

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

Bartolomé-Díaz, Zaida and Francesca Frontini. 2020. Building a Domain-specific Bilingual Lexicon Resource with Sketchengine and Lexonomy: Taking Ownership of the Issues. Kernerman, Ilan, Simon Krek, John P. McCrae, Jorge Gracia, Sina Ahmadi and Besim Kabashi (Eds.). 2020. *Proceedings of the 2020 Globalex Workshop on Linked Lexicography, Marseille, France, 11–16 May 2020*: 62-68. Paris: European Language Resources Association.

Béjoint, H. 2016. Dictionaries for General Users: History and Development; Current Issues. Durkin, P. (Ed.). 2016. *The Oxford Handbook of Lexicography:* 7-24. Oxford: Oxford University Press.

Boisson, C., P. Kirtchuk and H. Béjoint. 1991. Aux origines de la lexicographie: les premiers dictionnaires monolingues et bilingues. *International Journal of Lexicography* 4(4): 261-315. https://doi.org/10.1093/ijl/4.4.261

Camacho, J. 2017. Aproximación al concepto de medioestructura lexicográfica y su influencia en la tipología lexicográfica: el diccionario de aprendizaje especializado y consulta. *Romanica Olomucensia* 29(1): 1-16. https://doi:org/10.5507/ro.2017.001

Caruso, V. 2024. Between Collocation and Colligation: An Experiment in Collaborative Lexicography. *International Journal of Lexicography* 37(1): 75-94. https://doi.org/10.1093/ijl/ecad026

Caruso, Valeria, Alessandra Chervino and Giulia Daniele. 2022. Disseminating Dictionary Skills with e-lex Tools. Klosa-Kückelhaus, Annette, Stefan Engelberg, Christine Möhrs and Petra Storjohann (Eds.). 2022. *Dictionaries and Society. Book of Abstracts of the XX EURALEX International Congress, Mannheim, Germany*, *12–16 July 2022*: 138-139. Mannheim: IDS-Verlag.

Chaves-Yuste, B. and M. Márquez-Cruz. 2022. Lexicografía digital y adquisición de léxico en inglés: una experiencia b-learning en los tiempos de Covid-19. Puebla-Martínez, B., R. Vinader and N. Navarro (Coord.). 2022. *Covid-19, multiverso de disciplinas. Una mirada desde la salud, la educación y la comunicación*: 350-367. Madrid: Dykinson.

Chaves-Yuste, B. and M. Márquez-Cruz. 2024. A Cooperative Digital Dictionary to Improve English for General Purposes Reading Comprehension in Secondary Education. Bobkina, J. and E. Domínguez Romero (Eds.). 2024. *Transformative Trends in Language Education. Pioneering Autonomous Learning Technologies*: 171-188. Berlin: Peter Lang. https://doi.org/10.3726/b21998

Climent, J. 2005. *El uso del diccionario bilingüe castellano/valenciano como instrumento didáctico en los ciclos de enseñanza secundaria obligatoria y bachillerato.* Unpublished PhD Thesis. Alicante: Universidad de Alicante.

Costa, Hernani, Gloria Corpas, Miriam Seghiri and Ruslan Mitkov. 2015. iCorpora: Compiling, Managing and Exploring Multilingual Data. Corpas, G., M. Seghiri, R. Gutiérrez Florido and M. Urbano (Eds.). 2015. *Nuevos horizontes en los Estudios de Traducción e Interpretación. Libro de resúmenes de AIET17, Málaga, Spain, 29–31 January 2015*: 74-76. Geneva: Tradulex.

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.

Flores, E. 2022. Compilación de un corpus de especialidad con BootCaT para la traducción jurídica italiano / español: el caso de un contrato de agencia. *Onomázein. Revista de lingüística, filología y traducción* 57: 167-214. https://doi.org/10.7764/onomazein.57.09

Heuberger, R. 2016. Learners' Dictionaries: History and Development; Current Issues. Durkin, P. (Ed.). 2016. *The Oxford Handbook of Lexicography*: 25-43. Oxford: Oxford University Press.

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

Horák, Aleš and Adam Rambousek. 2007. Dictionary Management System for the DEB Development Platform. Sharp, Bernadette and Michael Zock (Eds). 2007. *Proceedings of the 4th International Workshop on Natural Language Processing and Cognitive Science, NLPCS 2007, Funchal, Portugal, 12–13 June 2007*: 129-138. Funchal: Insticc Press.

Jia, W., L. Zhang, A. Pack, Y. Guan and B. Zou. 2024. Digital Game-based Learning's Effectiveness on EFL Learners' Receptive and Productive Vocabulary Knowledge. *Language Learning & Technology* 28(1): 1-21.
https://hdl.handle.net/10125/73554

Knežević, L., S. Halupka-Rešetar, I. Miškeljin and M. Milić. 2021. Millennials as Dictionary Users: A Study of Dictionary Use Habits of Serbian EFL Students. *SAGE Open* 11(2): 1-12.
http://doi.org/10.1177/21582440211008422

Kovalevskaitė, J., L. Boizou, A. Bielinskienė, L. Jancaitė and E. Rimkutė. 2020. The First Corpus-driven Lexical Database of Lithuanian as L2. *Frontiers in Artificial Intelligence and Applications* 328: *Human Language Technologies — The Baltic Perspective:* 245-252.
https://doi.org/10.3233/FAIA200630

Mangeot, Mathieu and Antoine Chalvin. 2006. Dictionary Building with the Jibiki Platform: The GDEF Case. Calzolari, N. et al. 2006. *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06), Genoa, Italy, 22–28 May 2006*: 1666-1669. Genoa: European Language Resources Association.
http://www.lrec-conf.org/proceedings/lrec2006/pdf/105_pdf.pdf

Márquez Cruz, M. 2024. *Modelos lexicográficos digitales en la lexicografía didáctica.* Madrid: Guillermo Escolar Editor.

Mavrić, M., I. Srdanović and D. Špica. 2023. Analiza kulturno specifičnih termina na temu tradicionalnih japanskih festivala. Mikulaco, I. (Ed.). 2023. *Jezik, prijevod i međukulturna komunikacija*: 50-61. Pula: University of Pula.

Měchura, Michal Boleslav. 2017. Introducing Lexonomy: An Open-source Dictionary Writing and Publishing System. Kosem, Iztok, Carole Tiberius, Miloš Jakubíček, Jelena Kallas, Simon Krek and Vít Baisa (Eds.). 2017. *Electronic Lexicography in the 21st Century: Lexicography from Scratch. Proceedings of the eLex 2017 Conference, Leiden, Netherlands, 19–21 September 2017*: 662-679. Brno: Lexical Computing CZ s.r.o.

Molnár, György, Beáta Orosz and Katalin Nagy. 2022. Current Issues and Possible IT Solutions for Digital Competence Development. Turčáni, Milan, Zoltán Balogh, Michal Munk, Martin Magdin, Ľubomír Benko and Jan Francisti (Eds). 2022. *14th International Scientific Conference on Distance Learning in Applied Informatics, Conference Proceedings, Štúrovo, Slovakia, 2–4 May 2022:* 267-276. Alphen aan den Rijn: Wolters Kluwer.

Nguyen, L.T. and K. Tuamsuk. 2022. Digital Learning Ecosystem at Educational Institutions: A Content Analysis of Scholarly Discourse. *Cogent Education* 9(1): 1-17.
https://doi.org/10.1080/2331186X.2022.2111033

Redecker, C. and Y. Punie (Eds.). 2017. *European Framework for the Digital Competence of Educators: DigCompEdu*. Seville: Publications Office of the European Union.
https://doi.org/10.2760/159770

Roebianto, A., S.I., Savitri, I. Aulia, A.R.I.E. Suciyana and L. Mubarokah. 2023. Content Validity: Definition and Procedure of Content Validation in Psychological Research. *TPM — Testing, Psychometrics, Methodology in Applied Psychology* 30(1): 5-18.
https://doi.org/10.4473/TPM30.1.1

http://lexikos.journals.ac.za; https://doi.org/10.5788/35-1-2004 (Article)

Rojo, G. 2021. *Introducción a la lingüística de corpus en español.* New York: Routledge.

Rubio, R.Y., J.E. Bonilla and J.A. Bernal. 2021. Dictionary Writing Systems y otras herramientas informáticas para la elaboración, administración y publicación de diccionarios. *Lingüística y Literatura* 42(80): 340-360. https://doi.org/10.17533/udea.lyl.n80a20

Shahjahan, M., K.R. Ahmed, A. Al Hadrami, R. Islam, S. Hossain and S. Khan. 2021. Factors Influencing Poor Academic Performance among Urban University Students in Bangladesh. *International Journal of Evaluation and Research in Education (IJERE)* 10(4): 1140-1148. http://doi.org/10.11591/ijere.v10i4.21158

Shaxrizoda, Q. 2024. The History of Corpus Linguistics and Basics. *Best Journal of Innovation in Science, Research and Development*: 104-111.

Sinclair, J. 2005. Corpus and Text — Basic Principles. Wynne, M. (Ed.). 2005. *Developing Linguistic Corpora: A Guide to Good Practice*: 1-16. Oxford: Oxbow Books.

Snezhko, Z., D. Babaskin, E. Vanina, R. Rogulin and Z. Egorova. 2022. Motivation for Mobile Learning: Teacher Engagement and Built-in Mechanisms. *International Journal of Interactive Mobile Technologies (iJIM)* 16(01): 78-93. https://doi.org/10.3991/ijim.v16i01.26321

Srdanović, I. 2020. From Specialized Web Corpora of Tourism to a Learner's Dictionary. *Rasprave Instituta za hrvatski jezik i jezikoslovlje* 46(2): 1059-1083. https://doi.org/10.31724/rihjj.46.2.31

Strelan, P., A. Osborn and E. Palmer. 2020. The Flipped Classroom: A Meta-analysis of Effects on Student Performance across Disciplines and Education Levels. *Educational Research Review* 30: 100314. https://doi.org/10.1016/j.edurev.2020.100314

Tarp, S. 2019. La ventana al futuro: despidiéndose de los diccionarios para abrazar la lexicografía. *RILEX. Revista sobre investigaciones léxicas* 2(2): 5-36. https://doi.org/10.17561/rilex.v2.n2.1

Tarp, S. and R.H. Gouws. 2023. A Necessary Redefinition of Lexicography in the Digital Age: Glossography, Dictionography and Implications for the Future. *Lexikos* 33(1): 425-447. http://doi.org/10.5788/33-1-1826

Tiberius, C. and K. Heylen. 2022. ELEXIS Pathfinder to Computational Lexicography for Developers and Computational Linguists. Version 1.0.0. DARIAH-Campus. [Training module]. https://campus.dariah.eu/resources/pathfinders/elexis-pathfinder-to-computional-lexicography-for-developers-and-computational-linguists

Toledo, M.C. and A. Camero. 2021. Compilación semiautomática con Sketch Engine de un corpus ad hoc comparable bilingüe (inglés–español) de contratos de distribución exclusiva (DISTRIBUCOR). Sartor, E. (Ed.). 2021. *Los corpus especializados en la lingüística aplicada: traducción y enseñanza*: 15-41. Mantua: Pliegos Hispánicos, Universitas Studiorum S.r.l. – Casa Editrice.

University of Cambridge ESOL Examinations. 2015. *Business Vantage. Business English Certificate 5.* Cambridge: Cambridge University Press.

Van Laar, E., A.J.A.M. van Deursen, J.A.G.M. van Dijk and J. de Haan. 2020. Determinants of 21st-Century Skills and 21st-Century Digital Skills for Workers: A Systematic Literature Review. *SAGE Open* 10(1): 1-14. https://doi.org/10.1177/2158244019900176

Yong, H. and J. Peng. 2008. *Chinese Lexicography: A History from 1046 BC to AD 1911.* Oxford/New York: Oxford University Press.