The journal of El-Ryssala for studies and research in humanities

Volume : 09 / N° : 04/ 2020. pp 244-268

PISSN : 2543-3938 - EISSN : 2602-7771

The Effectiveness of Using Data-driven Learning Activities in Teaching English Near-synonyms for Algerian EFL Learners: A Pilot Study

Imen CHAALAL[1,*]

[1] **École Normale Supérieure- Messaoud Zeghar Sétif, Research Laboratory on Language Didactics (RLLD) (Algeria), i.chaalal@ens-setif.dz**

Received: 06/01/2024 **Accepted**: 04/12/2024

Abstract:

Working with an online DDL course (Crosthwaite, 2020), this pilot study sets out to examine whether using corpus linguistics tools would be useful for the development of Algerian sophomore EFL learners' use of English near-synonyms. Drawing on this course, pre/posttests, students' pre questionnaire and post semi-structured interviews were carried out. Findings reveal that in the pretest, near-synonyms were used interchangeably, regardless the context they occurred in. Conversely, after treatment, positive results were obtained, in which learners could produce accurate co-occurrences of those collocates across different disciplines. The paper concludes that DDL activities helped learners differentiate between near-synonyms and built positive attitudes among learners towards learning these items using the DDL approach.

Keywords: *Algerian EFL learners; Collocations; Data-driven learning; Near-synonyms; Sketch Engine*

*Imen CHAALAL

I. INTRODUCTION

Vocabulary has always been a major stumbling block for English as Foreign/Second Language (EFL/ESL) learners, and one of its crucial challenges is the use of synonymous words among tertiary level students. EFL/ESL students often misuse, overuse, or underuse synonyms and fail to distinguish between near-synonyms when they are used in various contexts. These struggles are observed among Algerian EFL learners who might not be different from other foreign language learners, in this regard.

Near-synonyms are words which share semantic similarities, yet differ in their contextual usage. Because of variations in their meaning and usage, they pose problems for learners when they need to enrich their language repertoire and enhance their writing skills. The escape to thesaurus, dictionaries, or translation engines is not always adequate to study near-synonyms; therefore, using corpus evidence to identify and extract them has its implications for EFL learning/teaching. This paper shows how the issue of near-synonyms' use can be tackled through the integration of the new Data-driven Learning (DDL) approach.

In the last decades, a plethora of corpora of different types have become very pertinent pedagogical tools in language teaching and learning and the development in Corpus Linguistics (CL) has led to the emergence of the DDL approach as well as different useful CL tools. Since their introduction, corpora have changed the analyses of attested examples of language use in real contexts. They offered lexicographers, linguistics researchers and educators "the opportunity to draw informed, objective, data-driven conclusions about the type and frequency of linguistic features that occur together, occur often, and occur in specific contexts" (Crosthwaite, 2020, p.1), and when the application of corpora found its way in teaching language learners, the term DDL, coined by Johns (1991), has come to the surface. It became, henceforth, effortless for language learners to access large databases of naturally occurring language and to investigate specific language patterns. Moreover, recent-day tools, such as Sketch Engine for Language Learning (SkELL) (Baisa & Suchomel, 2014) are considered to be valuable corpus tools that go beyond exposing words in real contexts and word frequency counts. Thanks to their improved functionality, automatic identification and attractive visualizations, these corpus interfaces are offering complex statistics and collocation functions, which were not available in the past; they simplified the access to information and made it much easier and quickly accessible.

Unlike traditional DDL paper-based format, in which students are provided with printed concordance input, guided by their teachers, DDL most recent outgrowth is witnessed in the use of online DDL interfaces. The Short Private Online Course (SPOC) by Queensland University on Edx platform (Crosthwaite, 2020) is a fine example of CL autonomous training. Crosthwaite (2020, p.2) explained that "given the current shift towards the adoption of blended/flexible learning in many contexts, taking DDL training online is a potentially fruitful method for developing DDL ".

Hence, thanks to the development of corpus tools and the escalating interest in similar online courses, the affordances of DDL seem to be more fruitful for EFL learners; this combination is believed to offer unprecedented opportunities for EFL learners resulting in a considerable influence on language pedagogy. Drawing on studies of DDL, the study examines the effectiveness of using this approach in teaching English near-synonyms for Algerian EFL learners, and presumably calls for considering DDL as an exceptional mainstream language teaching approach that should be integrated in Algerian EFL classes in the future .

This article, then, aims at revealing the pedagogical potential of DDL in EFL learning/teaching. It explores the effectiveness of using DDL activities (Crosthwaite, 2020) among Algerian EFL learners; the online course was employed to help learners improve their near-synonymy use, and, consequently, foster their language learning process. The study describes the application of this DDL course on a group of sophomore Algerian students who are destined to become teachers of English in High and Middle School levels .

Through this pilot study, we aim to sense the improvement in the near-synonyms' use before and after treatment and describe the students' experience in implementing this course in EFL classes. Details on the incorporation of DDL through this online course in teaching synonyms are then given, and an examination on how near-synonyms are handled by learners across different disciplines is suggested; in addition, the attitudes of learners towards this implementation are summarized in semi-structured interviews. Before doing so, a review of literature summarizing the previous studies on DDL applications in language pedagogy and near-synonymy use in EFL classes is presented. The paper concludes with the potential of DDL in teaching synonyms and discusses the value of DDL as a pedagogical tool and innovative technology that could be integrated in EFL classes.

1. Literature Review

Over the past three years, integrating CL tools and methods in language pedagogy has become evident in many educational contexts all over the world, and the proponents of this trend have proposed its application in various related domains. The vast array of topics that have been tackled with regard to corpus applications ranged from: pedagogical applications (Leech, 1997; Frankenberg-Garcia, 2005), classroom applications (O'Keeffe, McCarthy, & Carter, 2007), teacher education (O'Keeffe and Farr, 2003; Farr, 2022), etc..

With the spread of this trend, i.e., by considering corpora as valuable tools in language teaching and learning, their applications in language classrooms have become prevalent and indispensable. In her article, "Corpora and Language Teaching", Römer (2008, p.1) distinguished between, 'direct' and 'indirect uses' of corpora in language education. She (Römer, 2008) explained that, on the one hand, the indirect applications of corpora imply that the data derived from corpora result in hands-on approaches for researchers and materials writers; the effects then are sensed in the teaching syllabus, teaching materials and reference works. On the other hand, the direct applications of corpora are more teacher-and learner-focused; they entail hands-on approaches for teachers and learners (DDL), as corpus evidence are exploited in learner and teacher interactions.

1.1. Data-driven Learning

The use of corpora for pedagogical purposes is known as DDL. According to Gilquin and Granger (2010, p.1), "DDL consists in the application of tools and techniques of corpus linguistics for pedagogical purposes." The DDL approach was first introduced by Johns (1991) at the University of Birmingham. It involves using a large amount of authentic language data to investigate specific language patterns, and it is based on corpus concordance tools (in Key Word in Context (KWIC) form), in order to help students improve the different language skills. With the help of concordance functions, frequency lists and collocates, students would be able to retrieve from the corpus the usage of specific patterns in particular contexts, and uncover, for example, the lexical, syntactic and semantic properties of the words under study .

Johns (1991b, p.30 in Boulton, 2009a) described DDL as "the attempt to cut out the middleman as far as possible and to give the learner direct access to the data." The "middleman" refers of course to the teacher, but the computer is not seen as "a surrogate teacher or tutor, but as a rather special type of informant" (p.1). DDL thus implies that learners are exposed to large amounts of authentic language data in which they discover the language by themselves and be typical "language detectives", in Johns' terms (1997, p.101) "every student is a Sherlock Holmes". For Boulton (2009a, p.2), learners "are at the centre of the process, taking increased responsibility for their own learning rather than being taught rules in a more passive mode".

Following Johns's claims (1991, p.2) that "research is too serious to be left to the researchers", the role of the teacher has shifted significantly from the conveyer of information

to a facilitator of language learning; i.e., he has become "the director and coordinator of student-initiated research" (Johns, 1991, p.3). For Römer (2011), instead of depending on the researcher as 'mediator and provider of corpus-based materials' there is now an interaction between the learner and the corpus, and in a more controlled way, an interaction between the learner, the teacher, and the corpus. Römer (2011, p. 211) stated: "language learners and teachers get their hands on corpora and concordance tools themselves and find out about language patterning and the behavior of words and phrases in an "autonomous" way (Bernardini, 2002, p.165)." Corpora, then, are considered to be invaluable sources of "serendipitous" learning activities; i.e., learners pursue their own enquiries and pinpoint any interesting observations individually (Bernardini, 2000). Römer (2011, p.212) put forward: "in Bernardini's view, the learner plays the role of a "traveller instead of a researcher" (Bernardini, 2000a, p.131), and was less "interested in the starting or end point of a learning experience" than in what the learner experiences in between, on her or his journey (Bernardini, 2000a, p.142)".

Such learning activities promote learners' autonomy when "learners observe corpus data, make hypotheses and formulate rules in order to gain insights into language (inductive approach) or check the validity of rules from their grammar or textbook (deductive approach)" (Gilquin and Granger, 2010, p.1). Through using DDL, learners are presented with authentic examples of language use and asked to draw conclusions from the context (Hunston, 2002), instead of memorizing "abstract "rules" of grammar and inauthentic samples of discourse from textbooks (Crosthwaite, 2020). In this view, DDL type of learning fosters the learners' language learning process and prepares them to be independent, self-corrective, and constructive learners.

1.2. The Effectiveness of DDL in EFL Teaching

With a clear understanding of what DDL is, its pedagogical value is highlighted. The effectiveness of DDL in language pedagogy in general and EFL in particular has recently been widely brought up to the forefront when the advancements in technology and user-friendly corpus interfaces improved the potentiality of corpora. The research publications of Boulton (2009a, 2017); Gilquin and Granger (2010); Chambers (2022) are some examples.

A number of advantages claimed for DDL were suggested in Boulton (2009a). In his comparison between the DDL approach and the traditional rule-based approach, Boulton (2009a) explained that DDL is more advantageous at various levels. First, DDL reduces the teachers' intuition, as many language rules are intuition-based. Second, DDL provides a full picture of language that traditional rules fail to provide. Third, DDL provides concrete examples of language use, as its relevant parts can be identified and their form and frequency of occurrence are revealed; in other words, DDL helps learners not only in finding answers to general language queries, but also in discovering specific registers and texts' types. In their introductory chapter, Leńko-Szymańska and Boulton (2015, p.10) suggested that "corpus data is not used to elicit a rule or pattern but to exemplify it." Fourth, DDL brings new roles to the classroom since learners are "asked to abandon deeply rooted norms of classroom behaviour" (Bernardini, 2001a, p.23) since learners are usually told what to do by their teachers and refuse to assume their responsibility for their own learning. Fifth, the process of discovery learning guaranteed by the acquired corpus skills makes learning more efficient and refreshed, and, thus, trains learners to be more autonomous not only in language learning, but also in ICTs. Learners at this point are able to analyze language and interpret corpus queries appropriately on their own and can even extend corpus use to other fields of study. Boulton (2017, p.1) emphasized that "the underlying rationale […] is that such contact provides the massive contextualised exposure needed for language learning, but in a more controlled way than purely haphazard

exposure via regular reading or listening, thus promoting or enhancing noticing, language awareness, autonomy, and ultimately producing 'better learners.'''

The earlier studies that have been conducted with respect to DDL applications succeeded to provide foundational evidence of the value of DDL in language pedagogy, and highlighted the need for additional classroom applications and continued research in the field. Boulton's influential contributions (e.g., Boulton, 2009b; Boulton, 2010) that have been carried out for over fifteen years seemed to be the most prominent ones in flourishing the DDL trend. Nevertheless, prior to his concentration, several research works (e.g., Aston (1997) and Bernardini (2000), as reviewed in Boulton's Timeline (2017)) were carried out. Other good examples include: Bennett (2010); Flowerdew (2012); Leńko-Szymańska & Boulton.(2015)

Moreover, a series of most recent studies, such as Crosthwaite (2019, 2020); Boulton and Vyatkina (2021); Pinto et al. (2023) have covered different pedagogy related domains. Pinto et al. (2023), for instance, discussed the latest applications of DDL and language corpora for learning different languages, such as English, Spanish and Portuguese by international learners. The authors provided efficient examples of lesson plans that might help teachers prepare DDL activities for their learners. Another example is Crosthwaite's (2019) exploration of the affordances of DDL for pre-tertiary learners, a category of learners that has been overlooked since the beginning of DDL integration in language pedagogy. In his assessment of the feasibility of this approach with young learners, Crosthwaite (2019) provided empirical outcomes and suggested DDL activities that seem to be beneficial in pre-tertiary classrooms.

Though the DDL approach evidently has its place in the classroom, to the best of our knowledge, the primary limitation is the underdevelopment of this approach in tackling lexicogrammar elements. Therefore, the difficulty in integrating DDL into the inductive learning of lexicogrammar must also be taken into consideration.

1.3 .Synonyms and Near-synonyms

The term synonymy is described as a semantic relation between word pairs (Palmer, 1981). In the Cambridge Online Dictionary[2], the state of being synonymous is described as "having the same or almost the same meaning as another word or phrase in the same language or the fact that words or phrases are synonymous." However, in very rare occasions, two words can substitute each other without a change in meaning. In the Cambridge Online Dictionary, it is stated that "there is no such thing as perfect synonymy, but the meanings of many pairs of words are close enough for this term to be used in the context of a crossword." The words, 'peculiar', 'different' and 'unique', for instance, are typically used in different ways and convey different connotations. Thus, the firm belief that full equivalence exists between such items can be rather misleading for learners, as synonyms' use differ in various contexts .

Two main types of synonyms are found in English: absolute synonyms and near-synonyms. On the one hand, absolute synonyms are the words that can be used interchangeably in any context, without altering the meaning or connotation of the message transmitted (Jackson & Amvela, 2000). On the other hand, near-synonyms are words which share semantic similarities, yet differ in their contextual usage. Jackson (1988, p.67) described them as "words that can substitute for each other in a wide range of contexts but not necessarily absolutely because they differ in a number of ways".

The distinctions of near-synonyms can be made on the basis of a number of criteria,

[2] https://dictionary.cambridge.org

including: dialects, degree of formality, connotations, collocations and grammatical patterns. First, near-synonyms are different in terms of various dialects. British and American English synonyms, such as 'lift' and 'elevator', 'biscuit' and 'cookies', are well-known examples. Second, in terms of stylistic differences, i.e., their degree of formality, synonyms such as 'request' and 'appeal' are more formal than 'ask', 'beg', and 'plead'(Phoocharoensil,2010). A third criterion is the connotative meaning of near-synonyms, which can be positive, negative or neutral; words such as: 'skinny' and 'slim' hold different meanings in negative and neutral contexts. Another aspect is collocations. The co-occurrence between near-synonyms and other words surrounding them is restricted by collocation; for instance, the verb 'pay' usually collocates with the noun 'attention' and 'commit' collocates with the noun 'crime' (McCarthy and O'Dell, 2008). Lastly, grammatical patterns constitute another criterion. In terms of syntactic differences, near-synonyms, despite their shared core meaning, they have different syntactic structures. Near-synonymous phrases such as 'consist of' and 'be made up of', which denote the same literal meaning, do not have the same grammatical structures (Petcharat and Phoocharoensil, 2017) .

A number of recent publications provided interesting insights on the application of the corpus-based approach in the study of near-synonyms. We can find: Kim's (2020) analysis of Korean EFL learners' trajectories of discriminating four synonymous pairs: 'demand' and 'request'; 'mend' and 'repair'; 'outcome' and 'consequence'; 'happen' and 'take place'; Alanazi's (2022) study on the collocational behaviors and the semantic preferences of the near-synonymous verbs 'affect' vs. 'impact'; and Kruawong and Phoocharoensil's (2022) differentiation between three synonymous verbs 'teach', 'educate', and 'instruct' in terms of their collocational patterns. The implications of such lines of research have established the pedagogic value of DDL in EFL contexts. It becomes obvious, then, that the inherent characteristics of near-synonyms, which are definitely challenging for EFL learners, can be successfully tackled using the DDL approach .

On account of the advantages brought up by this approach in extracting near-synonyms, this study aims at investigating its usefulness in Algerian EFL classrooms. Working with SPOC on DDL (Crosthwaite, 2020), we aim at examining the development of Algerian EFL learners' use of English near-synonyms using SkELL, and explore how they are handled by Algerian learners across different disciplines, when examining two disciplinary corpora (BAWE and MICUSP). For that reason, we have followed the corpus-based teaching pattern "3Is model: illustration-interaction-induction", as proposed by Carter and McCarthy (1995 in Crosthwaite, 2020). In parallel, some DDL activities were designed and practiced to answer three main questions :

1 .Does the teaching of near-synonyms using SPOC on DDL increase EFL learners' awareness of near-synonyms differentiations ?

2 .Does the teaching of near-synonyms using SPOC on DDL improve the learners' awareness of collocational patterns?

3. How are EFL learner's attitudes impacted by SPOC on DDL focus on near- synonyms' use?

II. METHODOLOGY

This pilot study is part of a one year quantitative/qualitative project investigating the effectiveness of using DDL to improve Algerian EFL students' writing skill. The paper explores the students' use of online DDL course to develop one key element of vocabulary learning, near-synonymy distinction, which tightly correlates with writing competency. Therefore, the focus will be on the analysis of students' performance before and after using DDL and describe their experience in implementing this course in their written expression classes.

1. Research Design

The participants were Algerian sophomore students of English at the department of foreign languages at the Teacher Training School of Sétif. They were seven students (6 female and 1 male, aged19-20 years of age) destined to become teachers of English in High and Middle schools. Students learned English for more than seven years. The fusion of the two levels, i.e., High and Middle School levels is believed to be adequate for students of the two groups have the same English language proficiency level and they are receiving the same training in combined English classes. The selection of the participants came after the teacher's invitation for a volunteering participation in this project; while the invitation targeted 73 students, only 7 students showed obvious interest to take part in the pilot study. Working on this sample seems to be feasible since students have never been involved in corpus-based studies or heard about this approach before. So, to avoid any possible risks, we assumed that dealing with a limited number of students would be more efficient in such first encounters with DDL.

2. Research Instruments

The main research tools employed to meet the objectives of this study involve a vocabulary test, a student pre questionnaire and semi-structured interviews. The vocabulary pre/posttest aims at testing the students' appropriate use of near-synonyms before and after receiving treatment. It is used because it provides us with qualitative data about the participants' tendencies while dealing with near-synonyms. The students' pre questionnaire strives to find out how students deal with their uncertainties while selecting the right synonyms, and what language resources do they use while doing so. Finally, semi-structured interviews with our students give us further information about their feedback on their experience of using DDL (See appendices A and B).

3. Study Design

The study of the application of DDL took place over a two months period during the academic year 2022-2023. Before introducing CL, in general, and the online course, in particular, to our students, near-synonyms designed activities were administered as a pretest. In addition, a pre questionnaire was given to the same students. After one week, each student was signed in on Edx platform[3] and examined superficially the website and course modules before starting the treatment. Next, students underwent four live online teaching sessions with their English language instructor, the researcher. The online corpus-based instruction took place on Google Meet (once session per week, each of which took two hours). Choosing online sessions over in-class or laboratory instruction is encouraged by the shift towards blended learning, the trend nowadays in the students' institution. While the two first online sessions were devoted to employing CL tools and SkELL functions, the two others were restricted to teaching synonyms' use through DDL activities on SkELL and near-synonymy discrimination in disciplinary corpora (BAWE and MICUSP). At this time, participants were equipped with the necessary corpus functions and tools and encouraged to use them as a resource for identifying near-synonyms as well as their differentiation across different genres, in terms of formality, connotations, collocations and semantics preferences. By the end of the online teaching, the same test was administered again to the same group of students so as to examine their development in solving the synonyms' issues they have faced in their first encounter. As a final step, short semi-

[3] https://learning.edge.edx.org/course/course-v1:UQx+SLATx+2019/home

structured interviews took place, in which students gave their feedback on their experience with DDL.

4. The Collection of Data through Near-Synonymy Test

The corpus data representing near-synonymous were derived from two main sources: online learners' dictionaries and corpus consultation. The Oxford Learner's Dictionary[4], Merriam Webster Dictionary[5] and Cambridge Dictionary were the references chosen by students for the identification of near-synonyms. Students found them useful resources, in that they offer clear definitions and fine examples of the words under investigation. However, because synonyms may not always be explicitly displayed in dictionary entries, students may not understand the right meanings of words and may even fail to think about the existing relationships between them. At this point, we assumed that the corpus data would enable students to explore what the differences really are. We believe that finding solutions to synonyms' issues with corpus data analysis is a good strategy to raise the students' interest and help them reach conclusions autonomously. Hence, the BAWE Corpus[6] supported by SkELL[7] was the second source of data collection, in addition to the Michigan Corpus of Upper-Level Student Papers (MICUSP)[8], which was intended for the comparison of synonyms' use across different genres and text types. Moreover, SkELL was used for most of the instruction of near-synonyms and the design of relevant DDL activities. The tool is well-known for its several flexible functions, namely, Concordance, Keywords, Word Sketch, Collocations, Thesaurus, Text Type Analysis, etc.

In the corpus-based instruction, the researcher opted for designing DDL activities of the same synonymous items given in the pretest, in which students examined the words' meanings, shared semantic similarities, and contextual usage. During the instruction, students carefully examined the target near-synonyms from the provided concordance instances and completed the tasks individually. In order to avoid any difficulties that might take place, the teacher interfered only when she felt it was necessary to monitor them about the use of the appropriate SkELL function and how to complete the tasks. Following the "3Is model", students would undertake the tasks and end up generalizing the rules by themselves. They were supposed to be able to discover about collocations, connotations, and genres of the synonyms found in the DDL analysis and realize the usefulness of their discovery learning process.

4.1. Materials and Tasks

In this section, the materials and tasks administered to students are presented. The target words, comprising three-word categories: verbs, nouns and adjectives, were mostly selected on the basis of their appearance the BAWE and MICUSP corpora. With the occasional help of dictionary consultation (in the pretest) and corpus tools (after treatment) students examined the target words in a series of three main activities, as summarized below .

[4] https://www.oxfordlearnersdictionaries.com
[5] https://www.merriam-webster.com
[6] The BAWE Corpus (6,968,089 Words) is a British English corpus of student writing at UK universities. Built at the University of Reading, the corpus covers four disciplinary areas (Arts and Humanities, Social Sciences, Life Sciences and Physical Sciences) and four levels of study (undergraduate and taught Masters' level)
[7] https://app.sketchengine.eu

[8] The MICUSP is a collection of papers with a grade of A from the University of Michigan written for upper undergraduate and early graduate courses. https://elicorpora.info/main

4.1.1. Pretest

In Task One, students were tested on their understanding of target words' meanings, and were asked to select from six triple words the best synonyms items: purpose, goal, objective; consequence, result, outcome; appropriate, proper, suitable; normal, regular, common; advise, recommend, suggest; far, distant, remote. The task given provided information about students' knowledge of synonyms' meanings and their ability to differentiate between them .

Task Two was essentially a sentence construction test, in which students' ability to produce correct sentences using six pairs of synonyms was under investigation. The target synonyms were: assess and evaluate; explain and clarify; acquire and obtain; alternative and option; chance and opportunity; problem and trouble. The aim of the activity was to examine the students' use of collocations and whether they succeed to produce accurate grammatical patterns or not.

Words connotations were the focus of Task Three. The task consisted of four sub-activities. First, on this test, students listed isolated words according to their positive, neutral or negative meanings; nine triple words were suggested. Next, three additional sub-activities were administered to consolidate students' knowledge of words' connotations, but this time the words were contextualized in sentence structures. Learners were asked to identify the words that were more likely to have only negative connotations, select from a set of items those with the most positive connotations to complete full sentences, and determine, by themselves, whether the words in bold have a positive or negative connotation (See Appendix A).

4.1.2. Posttest

Table 1 below reveals DDL activities and the corpus tools used in this study. Five SkELL functions were employed to achieve our objectives: Frequency, Collocations using particularly LogDice measures, Word Sketch, Word Sketch Difference and Thesaurus. In addition, the MICUSP was used to compare between synonyms across disciplines. The same activities designed for the pretest using dictionary consultation and the same instructions provided by the teacher were undertaken using SkELL functions. As shown in the table below, each activity required at least one tool for the analysis of synonyms, yet Frequency and Collocations were principally used for most of the activities.

Table (1): Summary of Near-Synonyms' Distinction and Corpus Tools Used

DDL Activities	Synonymy Distinction	Tools Used
Task 1	Formality	Word Sketch; Frequency
	Synonym Differentiation	Word Sketch Difference; Collocations
	Genre Differentiating	MICUSP; Word Sketch
Task 2	Grammar Patterns, Collocations	Word Sketch; Collocations; Thesaurus
Task 3	Connotations; Collocations;	Collocations; Frequency

Since these activities aim at assessing the effectiveness of using DDL in teaching near-synonyms more than testing the students' knowledge, the analysis shows how the teacher can provide students with the given resources and what kind of input they may find appropriate to

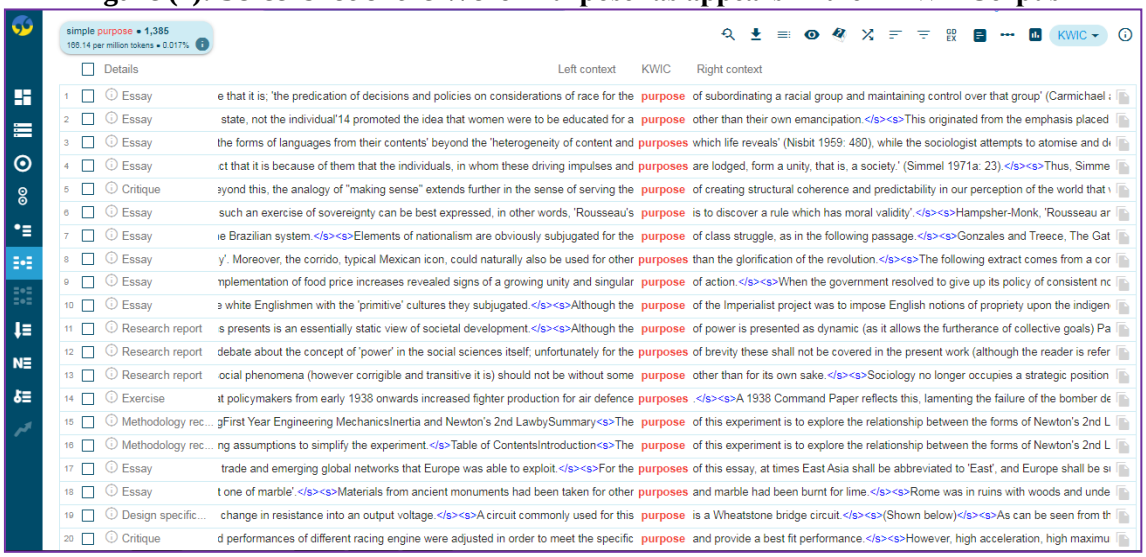
answer the questions. Students followed the DDL approach shown in Table 1 using the "3Is model". In the first stage, students were given for each task a number of inductive corpus-based activities developed by the teacher, in which the students used DDL and employed the most suitable corpus tool for synonyms' distinction. During these tasks, students examined the concordance lines and worked individually; eventually, they shared their discoveries with the teacher and made their own rules about the meaning and usage of the target synonymous words .

Because of space limit, in what follows, only three main DDL activities, illustrating one group of near-synonyms (purpose, goal and objective) on SkELL, are demonstrated. Nevertheless, it should be noted that the same procedure of analysis was undertaken to examine all the target words of the suggested tasks.

a. DDL Activity 1: Frequency

The first DDL activity is illustrated in Figure1 below (screenshot provided). Students were asked to complete the first activity using Concordance and Frequency list to understand the meanings of a set of triple words (e.g., purpose, goal and objective), compare between them and decide if they can be viewed as best near-synonyms. The search queries were carried out with the three words, yet the illustration is restricted, here, to the word 'purpose'. This search item produced 1,385 hits in the BAWE Corpus, appearing 166.14 times per million tokens. In addition, the 'simple query' search provided information about the frequency and genre of the text involved from as shown below (e.g., essay, research report, exercise, etc.).

Figure (1): Screenshot of the Word 'Purpose' as appears in the BAWE Corpus



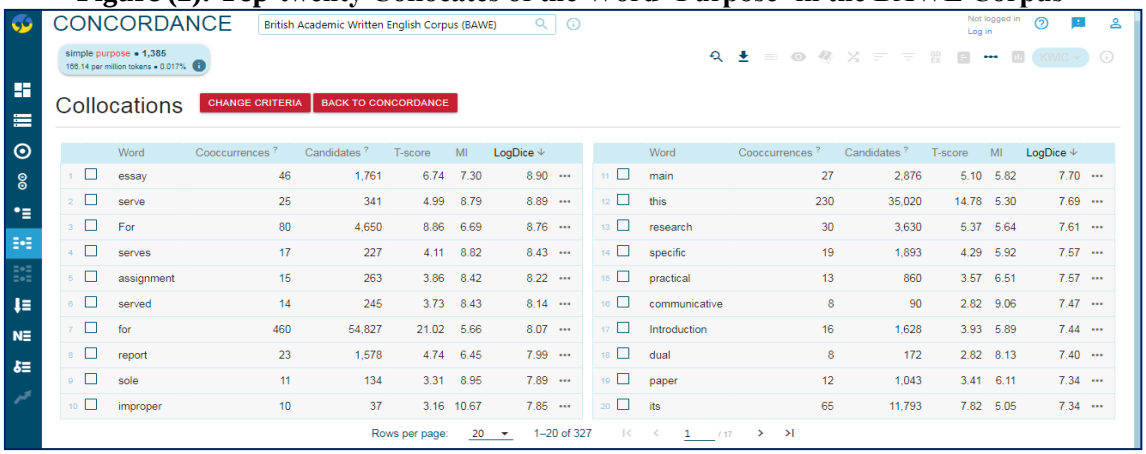
b. DDL Activity 2:

Collocations

A second example of DDL activity is to elicit the collocations and semantic preferences of the target words. The Collocation function gives students hints about particular 'verb and noun', 'adjective and noun' and 'verb and adverbial' combinations. The collocation strength between items is calculated using LogDice measures, which are interpreted through the LogDice score (if LogDice is greater than 6, the words are good collocates). Figure2 shows the top-twenty collocates of the noun 'purpose', including verb and adjective combinations. Thanks to the KWIC

function, collocations are demonstrated in real context, allowing learners to see the words in contextualized sentences.

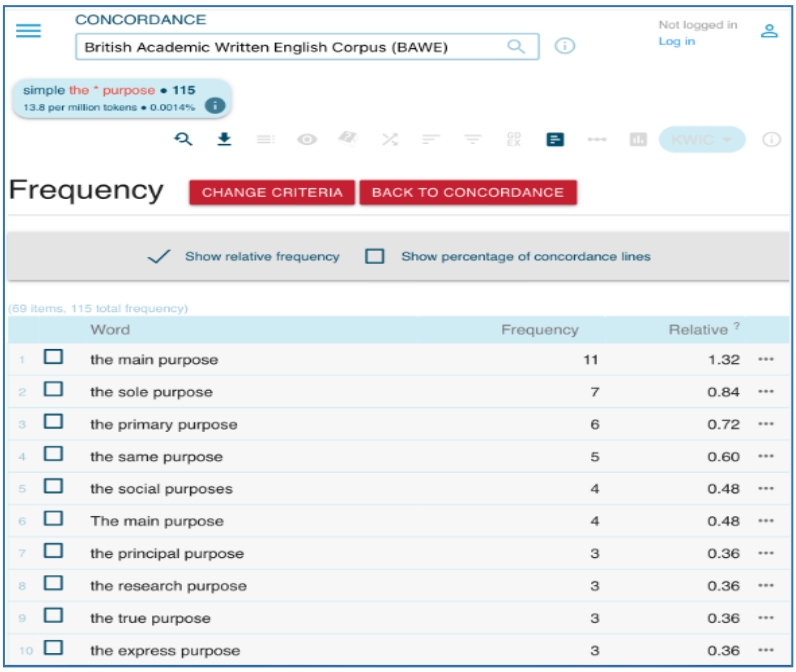
Figure (2): Top-twenty Collocates of the Word 'Purpose' in the BAWE Corpus



Wildcards and Word Sketch Difference

Another way of choosing the best collocate is to use Wildcards on SkELL. Wildcards are very useful tools in helping students discover what word is supposed to fill in a gap in a phrase. A simple search for "the * purpose" along with the KWIC Word Forms allow students to find out the mostly used adjective collocates of the word 'purpose'. The results of this query show that the adjective 'main' is the best collocate for the word 'purpose'.

Figure (3): Top-ten Adjective Collocates of the Word 'Purpose' in the BAWE Corpus using Wildcards



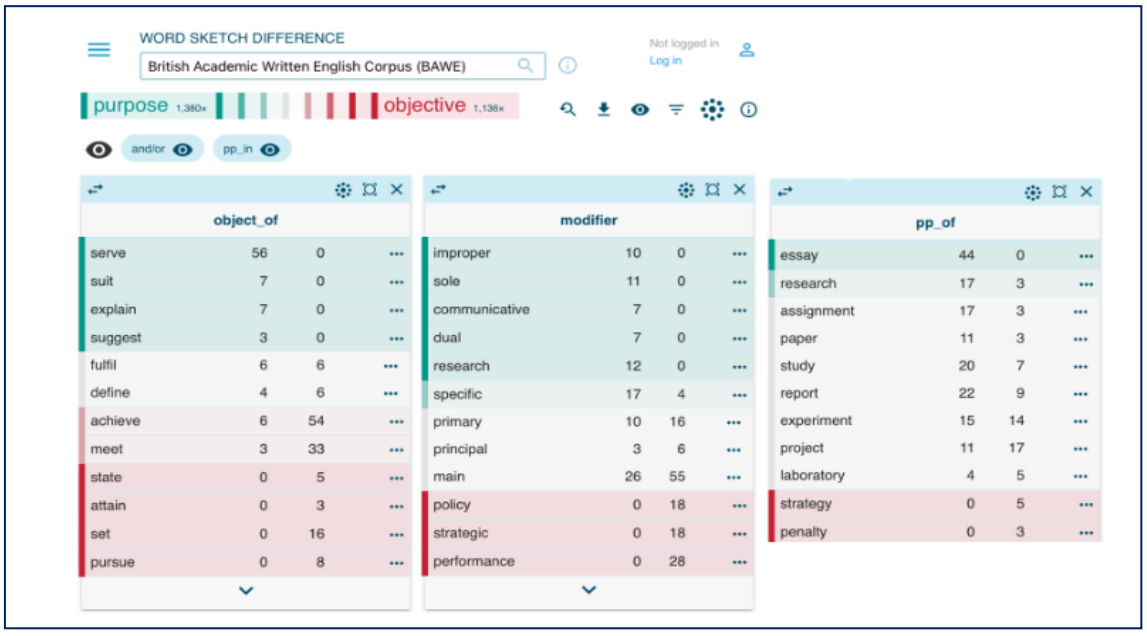
Word Sketch Difference is an extra useful function to generate synonyms and compare between them. The results obtained from comparing the lemmas 'purpose and goal' and 'purpose and objective' revealed various shared collocates. It is shown in Figure 4 and Figure 5 below that, on the one hand, 'fulfil', 'define', 'meet', 'ultimate' and 'specific' are the shared lexical collocates of 'purpose' and 'goal'. On the other hand, the words which collocate most with

'purpose' and 'objective' are 'achieve', 'primary', 'assignment', etc. Obviously, when words with many shared collocates are found, they are said to be most similar in meaning and are considered to be near-synonyms.

Figure (4): Shared Collocates between 'Purpose' and 'Goal' in the BAWE Corpus

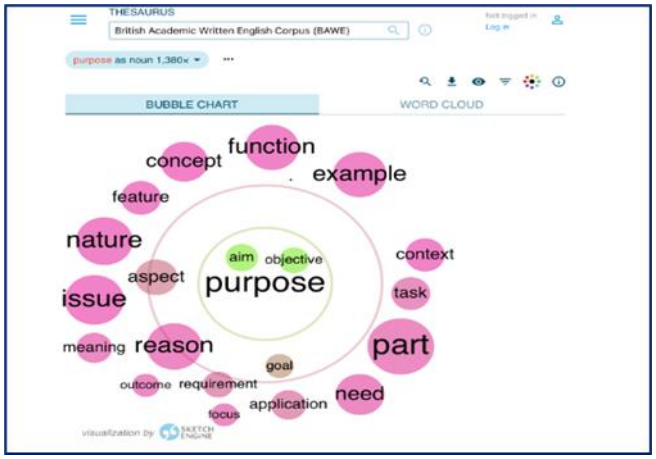


Figure (5): Shared Collocates between 'Purpose' and 'Objective' in the BAWE Corpus



Moreover, for a better visualization of 'purpose's near-synonyms, SkELL's Bubble Chart (Figure 6), supported by the Thesaurus functionality, provides interesting details. Students could discover that the word 'aim' is more similar to 'purpose' than 'goal' is, while 'goal' is more frequently used than 'aim' and 'achieve'.

Figure (6): Bubble Chart of the Word 'Purpose' in the BAWE Corpus



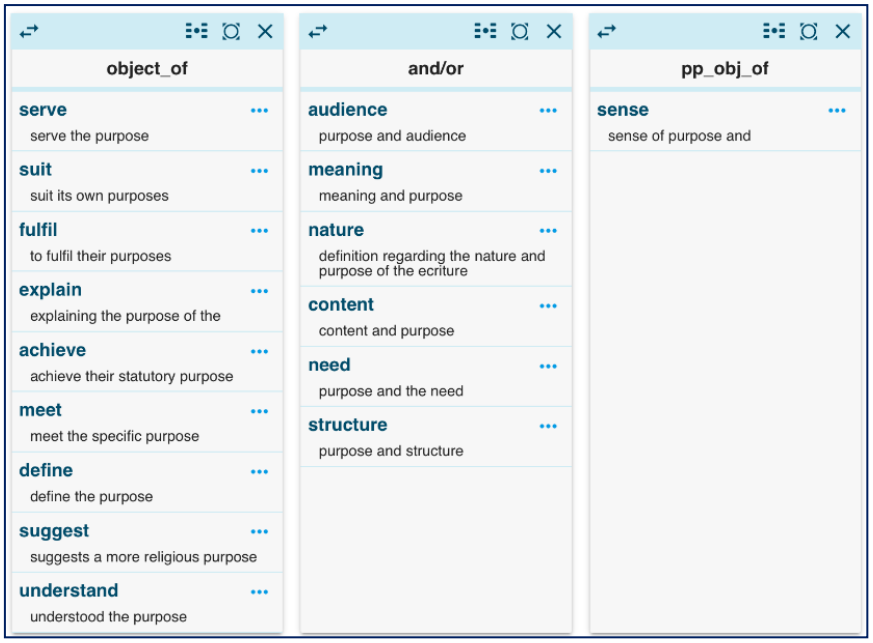
Word Choice and Collocations

Another suggested activity is selecting the best word choice for collocation patterns. The best realization of accurate collocates is to ask students to check if the sentence (e.g., We need to achieve purposes.) is correct, where 'achieve' is the verb choice error. After searching the query, students found out that the verb collocate 'achieve purposes' displays only five hits from the corpus, suggesting its infrequent co-occurrence. So, to find a better collocate, students need to discover what the most appropriate answer could be. Using the Collocations function, students search, then, for the word 'purposes' in the corpus. By looking for the immediate word before it, the advanced bar allows them to change the collocation range to option (-1); the results of this query generated the verb collocate 'serve'. Students can, thus, discover the semantic preferences of words when they read the full sentence displayed either in KWIC or sentence format. Similar activities confirm that 'purpose' cannot be used interchangeably with 'objective' and 'goal' since the verb 'achieve' collocates rather strongly with 'goals' and 'objectives' in the BAWE corpus (LogDice equals 9.65 and 8.88, respectively). It is evident, hence, that the three nouns share more overlapping semantically-related themes and collocations; however, we find out that the verbs 'achieve' and 'meet' really stick together with 'goal' and 'objective' more than that with 'purpose'. Also, there are some words that do not stick together, whatsoever; for example, 'pursue' or 'attain' do not combine with 'purpose.'

Word Sketch

Lastly, further details on the grammar pattern of 'purpose' can be obtained from the WordSketch function. The function could also be used to reveal which verb collocates are most appropriate with the word 'purpose', when it is the object of the sentence. The list of verb collocates is displayed on the left of Figure 7.

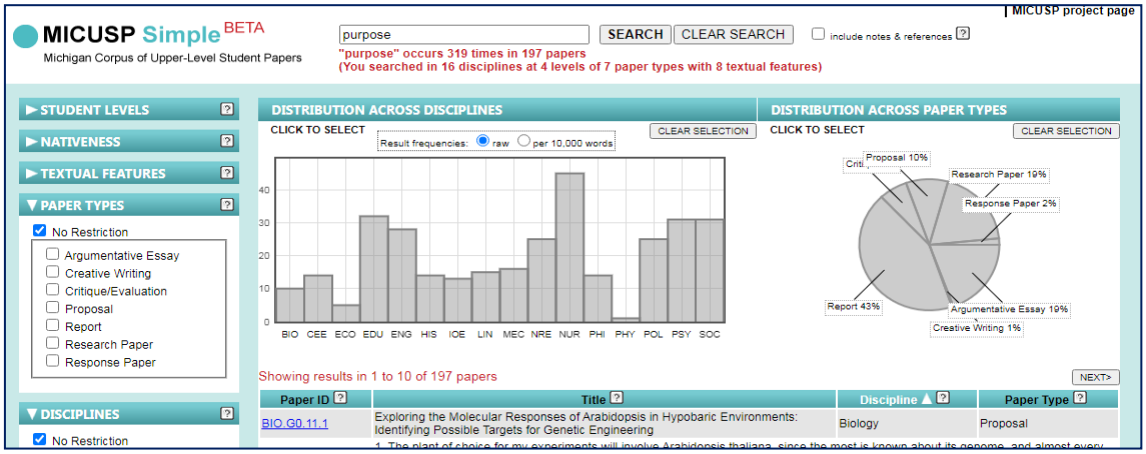
Figure (7): Verb Collocates of 'Purpose' as an Object of the Sentence in the BAWE Corpus



c. DDL Activity 3: Disciplinary Corpora

Synonymy use across disciplines varies, so exploring disciplinary corpora is a good way to discriminate between them. In addition to the possible use of the function Text Types in the BAWE corpus, for academic English, we intended to explore another disciplinary corpus, the MICUSP, to examine the use of near-synonyms in other disciplinary fields, such as the scientific and technical ones. The first search query of the word 'purpose' in the MICUSP shows that 'purpose' appears 319 times in 197 papers. The bar chart in Figure 8 reveals that the word's most frequent occurrence is found in the field of Nursing (NUR), appearing 45 times in 17 papers, followed by Education (EDU) with 32 occurrences in 15 papers.

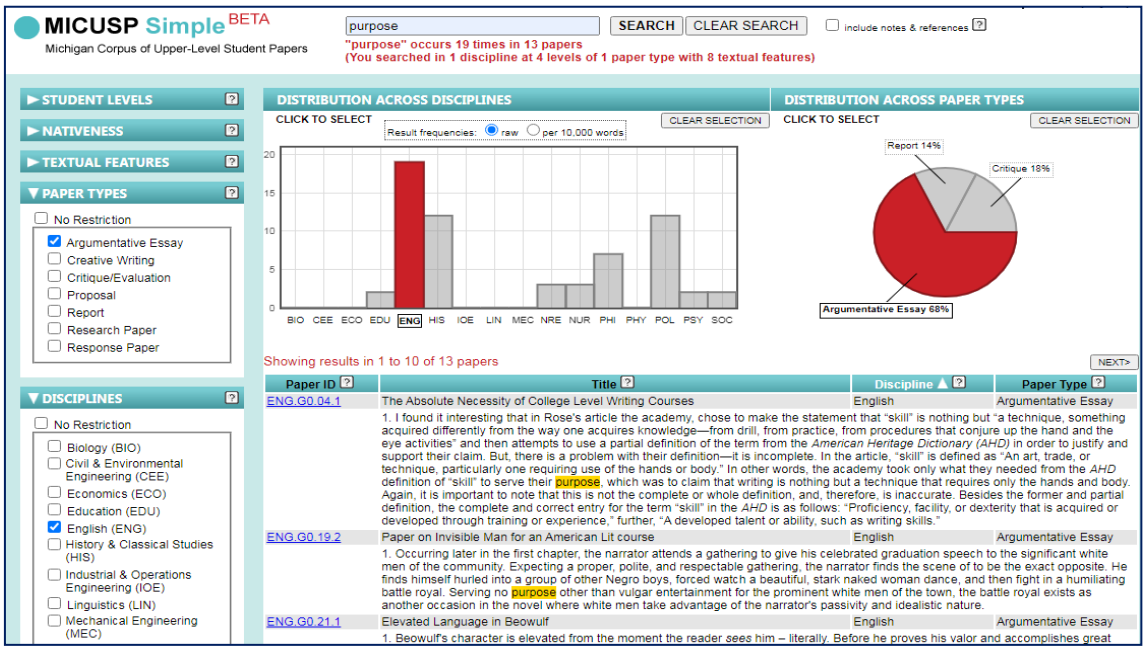
Figure (8): Screenshot of the Word 'Purpose' in the MICUSP



Concerning EDU text types, the analysis shows that 'purpose' is commonly used in report (56%) followed by proposal (19%); however, in NUR, 'purpose' is found more often in reports and research papers (58% and 22% respectively). Such information can be considered useful for students when comparing between genres and text types of different research fields. In our study, the yielded results help students figure out that some synonyms are more frequent in

scientific fields than in the humanities. Since we are dealing with an English language major, we are more interested in obtaining information about these items from English (ENG) in specific types of texts (e.g., argumentative essays, creative writing, or even in a combination of them). Hence, by comparing the words (purpose, goal, and objective), in the MICUSP, we found out that 'purpose' occurs 19 times in 13 papers. It is displayed in argumentative essays (68%) more than in reports (14%) and critique (18 %.). Interestingly, as shown in Figure 8, in comparison to all fields of research, 'purpose' is mostly used in ENG argumentative essays.

Figure (9): Screenshot of the Word 'Purpose' in the MICUP Restricted Search ENG/Argumentative Essay



Further queries' comparisons revealed that 'goal' appears in ENG (8 times), following Philosophy (10 times). While 'goal' in ENG argumentative essays covers 50% of the chart, 'objective' is very rare in this type of texts; it appears only 2 times in 2 papers (18%). The results obtained from similar comparisons help students select the most appropriate synonymous word in specific types of texts; in this case, students can elicit that 'purpose' is widely used in English argumentative essays more than 'goal' or 'objective.'

All in all, it can be said that using disciplinary corpora is fairly successful and the affordances of DDL can bring fruitful results for EFL learners' investigation of discipline-specific vocabulary.

5. The Collection of Data through a Pre-Questionnaire and End Of Course Semi-Structured Interviews

A pre-questionnaire and a semi-structured interview (Appendix B) were employed in this pilot study to collect facts about the students' attitudes towards the impact of integrating DDL activities on their near-synonymy learning. Before the pre-test, a 22-question pre-questionnaire was completed by seven sophomore students (6 female and 1 male, aged 19-20 years of age) destined to become teachers of English in High and Middle schools. It was derived from SPOC on DDL (Crosthwaite, 2020), though some modifications were carried out so that they meet the objectives of this study. The questionnaire was submitted online via Google forms and was divided into three main sections. While section one was devoted to the students' personal

information, section two was concerned with the students' tendencies of using computers. Section three, the focus of this study, aimed at providing information about the language resources used by students and finding out how students deal with their uncertainties while selecting near-synonyms. Then, after completing the post-test, short semi-structured interviews were carried out with only four students, as the three others were absent the day of the interviews. The interviews aimed at eliciting the students' attitudes towards their experience of learning near-synonyms through DDL activities. The questions administered revolved around whether students have become confident when they draw on DDL in their learning, whether they are going to change their learning references' habits to DDL method, and what was the most useful function they thought it may help them with their learning of near-synonyms.

III. RESULTS

In order to examine the usefulness DDL activities in teaching near-synonyms for Algerian EFL learners and obtain a complete picture of the students' learning growth, a quantitative-qualitative data analysis was carried out. First, in an attempt to answer Question 1 and 2, a paired samples t-test was employed to compare the means of pretest and posttest. The obtained results provide us with information about the students' increased awareness of near-synonyms differentiation and their corresponding collocational patterns. Second, the results obtained from the pre questionnaire were analyzed and triangulated with the semi-structured interview results to answer Question 3, which explores the students' attitudes towards the impact of integrating the DDL course on near-synonymy learning. The yielded results are summarized in what follows.

1. Pre and Posttest Results

Table 2 and 3 below summarize the students' raw scores of the pre and posttests, focusing on the learners' performance. The first column presents the students names[9] and the three other columns present their completed tasks' scores.

Table (2): Students' Pretest Raw Scores

Pretest Results				
Students	Task1	Task2	Task3	Total Score
Merriam	2	2	11	15
Adam	3	3	10	16
Luna	5	5	19	29
Naya	4	2	14	20
Mila	1	2	14	17
Ayla	2	2	14	18
Sara	1	2	14	17

Table (3): Students' Posttest Raw Scores

Posttest Results				
Students	Task1	Task2	Task3	Total Score
Miriam	4	5	17	26
Adam	5	6	17	28
Luna	5	6	24	35
Naya	4	4	21	29
Mila	3	5	21	29
Ayla	4	5	21	30

[9] For the anonymity of the participants, the following pseudonyms, chosen by the researcher, were given: Merriam, Adam, Luna, Naya, Mila, Ayla, and Sara.

Sara	3	5	21	29

The tables show that students' scores vary significantly in both pre and posttest. Despite the fact that more than half of the participants' scores were less than the average (18) in the pretest, interestingly, one student, Luna, obtained a higher score (29); she further excelled in the posttest and obtained (35). All the other participants' scores were relatively higher after treatment: 5 scores between (26) and (29) and 1 score between (30) and.(36)

Table (4): Paired Samples t-test Results

Students	Pretest Scores	Posttest Scores	Difference	Dev (Diff-M)	Sq.Dev
Merriam	15	26	11	0.43	0.18
Adam	16	28	12	1.43	2.04
Luna	29	35	6	-4.57	20.9
Naya	20	29	9	-1.57	2.47
Mila	17	29	12	-1.43	2.04
Ayla	18	30	12	-1.43	2.04
Sara	17	29	12	-1.43	2.04
Mean (M)	18.85	29.42	10.57		S: 31.71

The descriptive statistics reported in table 4 revealed that for this set of scores, there is a difference between the pretest (M=18.85) and posttest (M=29.42) scores that is not likely due to chance. The results show that there was a statistically significant difference between the pretest and the posttest mean scores; t=12.17, p=.00002<.05. The result is, hence, significant at p< .05, and the null hypothesis H0 was rejected in favor of the alternative hypothesis H1. We can, therefore, say that the corpus-based instruction was effective and that students succeeded to differentiate between near-synonyms and realize their collocation patterns more than they did before starting the instruction.

2. Pre Questionnaire Results

In what follows, the most relevant results obtained from the analysis of the pre-questionnaire are highlighted .

First, findings revealed the students' modest interest in using computers despite their frequent use of the Internet; the majority of them (57.14%) are often online via smartphones. While the majority (57.14%) somewhat like using computers, only (28.57%) often use computers for school work and doing homework, and (42.85%) do never use them for their personal purposes. These results justify the students' reliance on new technology devices and modern applications rather than on computers.

Second, in our attempt to learn about the students' referencing habits, students were incited to answer open-ended questions about what do they do when they are uncertain about their use of English vocabulary, in general, and synonymy word choice, in particular. Merriam said that she resorts to checking the word on the Internet or simply just ignore it and try to focus on conveying the general meaning. Ayla tended to simply look the word up in the dictionary and try to find another best equivalent. In fact, students' answers revealed that when searching a word, they always check the words in dictionaries (57.14%), use often Google search engine (42.85%), and opt sometimes for online translation websites (42.85%); however, it was reported that their search on the Internet is not very well organized, as they do not use specific websites .

Third, when reporting about their dictionary use, 57.14% of participants used often online dictionaries, without any regular use of the same dictionary (i.e., each time, a different dictionary is used). In another question about how often do students guess the meaning of words instead of confirming their word choice with a reference resource, 71.42% of students said they often guess, 14.28% always guess and 14.28%sometimes do so. Moreover, in a more practical question, students were asked if they were supposed to choose between two synonyms in the sentence "We need to achieve a purpose/goal." what would be their word choice. 57.14% of students said they would search it online, 14.28% claimed they would try to guess and 28.57% would look the words up in the dictionary. Searching for words online seems to be the most favorite option as it is more accessible (quick and easy) .

Finally, concerning corpus knowledge, all participants (100%) reported they had never heard of CL and language corpora before, nor they are knowledgeable about their usefulness in vocabulary learning.

3. Semi-structured Interview Results

The results indicated that students had positive attitudes towards learning near-synonymy through SPOC on DDL. Some of our students' various insights are listed below.

First, the analysis of the interview revealed that the participants' experience with DDL was novel and interesting. Merriam, who seemed very motivated, said that unlike the traditional English learning classes, DDL activities were more enjoyable and fun. She explained that the attractive visualizations of the various words, in Bubble Charts, for instance, were remarkably helpful for a better understanding of the relationship between synonyms. She was even very encouraged to use DDL, not only for the purpose of vocabulary learning, but also for her overall English academic practices in the future and decided to recommend it to her classmates .

For Naya, she liked the idea that in search of words' meaning, it would be possible to use corpus interfaces instead of looking up in dictionaries or textbooks. Her answer implied that that the easy and simplified access to information made it much easier for learners to accept corpora as valuable tools .

Luna claimed that using concordance lines was very beneficial as they allowed her discover collocation patterns that she did not know before. In her view, seeing many examples of near-synonyms in real contexts is a unique feature that cannot be found in dictionaries or any other references. In this regard, the four students agreed on the usefulness of SkELL's concordance function. Luna added that she could even correct her own mistakes, when she used the concordance; consequently, she developed a better understanding of synonyms' meaning and collocation patterns in a straightforward manner .

Regarding their general feedback on the tasks given, Luna, Naya, and Mila found learning synonyms through these tasks rather challenging and time consuming since they needed more time to be familiar with the functions and practice their activities in their own speed. Finally, concerning the instruction process, Mila had positive opinions about the nature of DDL activities and appreciated the teacher's occasional intervention in guiding them all the way through their queries; despite this daunting task and technical problems, she looked forward for a development in her learning autonomy in the future.

To sum up, the pre-questionnaire and interview results revealed that students had discovered new computer software and learning methods that aided them build positive attitudes towards learning vocabulary, in general, and near-synonymy, in more particular .

IV. DISCUSSION

It is evident from the statistical results obtained that using online DDL activities for learning near-synonyms among Algerian EFL learners is useful. The results of the pretest indicated a serious problem in differentiating between near-synonyms by students, especially in Task One. Apart from Luna's scores, low scores of students indicated that synonyms' distinction, understanding connotations and using the right collocations were the most difficult tasks. So, in an attempt to answer question 1 of this study, it can be said that, after treatment, and owing to the use of SkELL functions, students could identify which words are similar. For example, the word 'appropriate', which most of the students failed to decide on its best synonym, either 'proper', 'suitable', or both of them, was easily understood after using the Thesaurus. When students could read the frequency and similarity scores of the target words displayed in a list, they compared between them and completed their answers. This procedure was entirely automatic, without the teacher's involvement. The challenges involved with these specific near-synonyms were explained in Petcharat and Phoocharoensil's work (2017). By confirming the contribution of corpora to the study of near-synonyms, in their study, it was revealed that the three synonyms share the same core meaning, but differ in their detailed meanings, degrees of formality, collocations, and grammatical patterns. Consequently, the three words are not used interchangeably in all contexts. Moreover, it was found that some additional data, which were non-existent in the dictionaries employed, could be provided thanks to the corpus evidence .

In addition, the significant differences between the means' scores of pre and posttest implied that at the outset of the study, students showed poor understanding of words' connotations, a problem that was fairly overcome when the target words were contextualized in full sentences and displayed in concordance lines; these latter played a vital role in helping learners see authentic language use in real context. This study proves that even when dictionaries are available, the use of concordancers is crucial for the selection of the right near-synonym. In his comparison between dictionary information and corpus data information, McGee (2018, p.315) proved that such reliance on corpus data solves word meaning issues, as "it enables us to expand our interest away from purely paradigmatic or narrowly semantic considerations of a word to the syntagmatic environment of the word. Rather than attempting to 'understand' a word in isolation, one can observe its environment and see how it is used, rather than merely consider what it means".'

Furthermore, as for collocations' use, after being introduced to a variety of combinations, students observed how the meanings of collocates change according to the context of their use and genre. In order to answer research question 2, students searched for the 'verbs' and 'adjectives' that frequently collocate, for instance, with the three synonymous nouns (consequence, result and outcome); they figured out that 'consequence' and 'outcome' are found more in formal genres (academic contexts). Student's productions enhanced consequently, after treatment. Similar findings were obtained in Phoocharoensil's examination of the Corpus of Contemporary American English (COCA) (2020), in which he discovered that these target synonyms are used with the highest frequency in academic texts, suggesting its association with a high degree of formality, while frequencies are the lowest in informal genres,

more precisely, 'consequence' and 'outcome' are lowest in frequency in TV and movie subtitles, and 'result' in fiction .

In this study, the tests' improved results indicated that before introducing learners to DDL activities, near-synonyms were used interchangeably, regardless the context they occurred in; this is mainly due to the learners' frequent exposure to these terminological variations in the same dictionary entries. However, after treatment, learners could produce accurate co-occurrences of those same collocates. That is to say, once learners became aware of their inadequate prior knowledge, they were able to correct their erroneous use of these synonyms and acquired new knowledge of collocations. These findings are consistent with the students' responses in the interview section, in which they confirmed the effectiveness of DDL activities in using near-synonyms. As the design of the activities was influenced by Crosthwaite's online course (2020), the findings converge also with his participants' satisfactory opinions on the usefulness of corpus for learning about vocabulary. According to Crosthwaite (2020, p.20), "resolving the errors techniques and activities described should be of use to potential DDL practitioners interested in using corpora for error resolution or as a reference resource for students of academic writing".

In summary, this paper argues that the success of this corpus-based instruction is ascribed to both the integration of online DDL course and the students' acquired skill of discovery learning. This relates to the use of online DDL interfaces, namely, SPOC (Crosthwaite, 2020) and its contribution to the development of students' autonomous training. It seems evident that the students' misuse of near-synonyms and their inappropriate perception of their common interchangeability have changed. They became more aware of their use in different genres and text types and, hence, developed their synonymy knowledge. In addition, such instruction has raised the students' awareness of the importance of self-correction errors in solving synonymy issues. Students developed their ability to identify near-synonyms' errors areas, focus on the collocations patterns and semantic preferences of each synonym and, as a result, became self-corrective learners. This discovery learning strongly corroborates with the "learner is a language detective" metaphor (Johns, 1997). Moreover, providing learners with the necessary corpus tools, concordance and Word Sketch, in particular, allowed them to see real examples in an attractive platform. Such practices are believed to be fruitful for future corpus-based instructions relevant to vocabulary learning.

V. CONCLUSION AND RECOMMENDATIONS

This pilot study demonstrated the effectiveness of using DDL in teaching near-synonyms for Algerian EFL learners. It examined the use of a potentially effective online interface (SPOC on DDL) and discussed the strengths of integrating SkELL and disciplinary corpora to foster students' use of near-synonyms. The study explored also the students' attitudes towards incorporating DDL activities in EFL classes. Results showed that SPOC was a useful tool that aided students more efficiently in acquiring near-synonyms and developing their vocabulary knowledge, accordingly. The findings led us to conclude that DDL can be considered as a valuable tool that should be implemented in Algerian EFL classes. Moreover, we believe that the affordances brought up by the concordance tools and various language corpora would help in the design of corpus instructional materials to raise students' awareness of vocabulary issues. Nevertheless, in this pilot study, the selected sample might be viewed as untypical; therefore, ideally, these findings should be replicated when dealing with large-scale studies. Further research on DDL

studies should be developed in the Algerian context, and studies that target other areas, such as grammar use, mechanisms, pragmatics development, idioms use, etc, are also recommended.

REFERENCES

Alanazi, Z. (2022). Corpus-based analysis of near-synonymous verbs. *Asian-Pacific Journal of Second and Foreign Language Education*, *7*(1), 1-25, doi:10.1186/s40862-022-00138-5

Baisa, V & Suchomel, V. (2014). Skell: Web interface for English language learning. In *Eighth Workshop on Recent Advances in Slavonic Natural Language Processing*, (pp.63-70). Brno: Tribun EU.

Bennett, G. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers*. Michigan: ELT .

Bernardini, S. (2000). Systematising serendipity: Proposals for concordancing large corpora with language learners. In Lou Burnard and Tony McEnery (ed.), R*ethinking language pedagogy from a corpus perspective*, (pp.225–234). Germany: Peter Lang .

Boulton, A. (2009a). Data-driven learning: Reasonable fears and rational reassurance. *Indian Journal of Applied Linguistics*, 35(1):1-28. Retrieved from: https://www.researchgate.net/publication/29598965_Data-driven_learning_Reasonable_fears_and_rational_reassurance

Boulton, A. (2009b). Testing the limits of data-driven learning: Language proficiency and training. *RECALL*, *21*(1), 37-54. doi:10.1017/S0958344009000068

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

Boulton, A. (2017). Corpora in language teaching and learning: Research timeline. *Language Teaching*, *50*(4), 483-506. doi: 10.1017/S0261444817000167

Boulton, A & Vyatkina, N. (2021). Thirty year of data-driven learning: Taking stock and charting new directions over time. *Language Learning and Technology*, *25* (3), 60-89 https://www.lltjournal.org/item/10125-73450 /

Chambers, A. (2022). What is data-driven learning?. In Anne O'Keeffe and Michael, J. McCarthy, (ed.), *The Routledge handbook of corpus linguistics*, (pp. 416-429). London: Routledge. doi: 10.4324/9780367076399-29

Crosthwaite, P. (2020). Taking DDL online: Designing, implementing and evaluating a SPOC on data-driven learning for tertiary L2 writing. *Australian Review of Applied Linguistics*, *43*(2), 169-195. doi:/10.1075/aral.00031.cro.

Crosthwaite, P. (2019). *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners*. London, Routledge. doi:10.4324/9780429425899

Farr, F. (2022). How can corpora be used in teacher education?. In O'Keeffe, Anne and Michael J. McCarthy, (ed.), *The Routledge handbook of corpus linguistics*, (pp.456-468). London: Routledge. doi: 10.4324/9780367076399-29

Flowerdew, L. (2012). *Corpora and language education*. UK: Palgrave MacMillan.

Frankenberg-Garcia, A. (2005). Pedagogical uses of monolingual and parallel concordances. *ELT Journal*, *59*(3), 189-198, doi:10.1093/elt/cci038

Gilquin, G &Granger, S. (2010). How can DDL be used in language teaching?. *DRAFT, F.N.R.S.* (pp. 1-12). Centre for English Corpus Linguistics Université Catholique de Louvain.

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

Jackson, H. (1988). *Words and their Meaning: Learning about Language*. NewYork: Routledge.

Jackson, H & Zé Amvela, E. (2000). *Words, meaning, and vocabulary*. London: Cassell .

Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. In Tim. F. Johns and Philip King (ed.), *Classroom Concordancing ELR Journal*, 4, 1-16 .

Johns, T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In Anne. Wichmann, Steven Fligelstone, Tony McEnery and Gerry Knowles (ed.), *Teaching and language corpora*, (pp.100-115). London: Longman.

Kim, M. (2020). A qualitative analysis of EFL learners' discrimination of near-synonyms in a data-driven learning task. *English Teaching*, 75(3), 25-47. Available at https://eric.ed.gov/?id=EJ1274664

Kruawong, T & Phoocharoensil, S. (2022). A genre and collocational analysis of the near-synonyms teach, educate and instruct: A corpus-based approach. *TEFLIN Journal: A Publication on the Teaching & Learning of English*, *33*(1), 75-97. Available at https://eric.ed.gov/?id=EJ1354662

Leech, G. (1997). Teaching and language corpora: A convergence. In Anne Wichmann, Steven Fligelstone, Tony McEnery and Gerry Knowles (ed.), *Teaching and language corpora*, (pp.1-23). London: Longman

Leńko-Szymańska, A & Boulton, A. (2015). Data-driven learning in language pedagogy. In Agnieszka Leńko-Szymańska and Alex Boulton (ed.), M*ultiple affordances of language corpora for data-driven learning*, (pp.1-14). Amsterdam: John Benjamins. https://benjamins.com/catalog/scl.69

McCarthy, M & O'Dell, F. (2008). *English collocations in use: Advanced*. Cambridge: Cambridge University Press.

McGee, I. (2018). Corpus linguistics and the classroom: Avenues for innovation. In Rahma. Al-Mahrooqui and Denman Christopher (ed.), *English education in Oman: Current scenarios and future trajectories*, 15, (pp.313-322). Singapore: Springer. Available at: https://doi.org/10.1007/978-981-13-0265-7_18

Merriam-webster Dictionary. https://www.merriam-webster.com

O'Keeffe, A &Farr, F. (2003). Using language corpora in initial teacher education: Pedagogic issues and practical applications. *TESOL Quarterly*, *37* (3), 389-418. Available at: https://www.jstor.org/stable/3588397

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

Online Cambridge Dictionary. https://dictionary.cambridge.org/dictionary/english/synonymy (consulted 16/06/2023)

Oxford Learners Dictionaries. https://www.oxfordlearnersdictionaries.com

Palmer, F. (1981). *Semantics*. Cambridge: Cambridge University Press.

Petcharat, N & Phoocharoensil, S. (2017). A corpus-based study of English synonyms: Appropriate, proper, and suitable. *LEARN Journal: Language Education and Acquisition Research Network*, *10*(2), 10-24. Available at: https://eric.ed.gov/?id=EJ1229630

Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, *3*(10), 227-245.

Phoocharoensil, S. (2020). A genre and collocational analysis of consequence, result, and outcome. *3L: Southeast Asian Journal of English Language Studies*, 26(3), 1-16. doi:10.17576/3L-2020-2603-01

Pinto, P. T., Crosthwaite, P., de Carvalho, C. T., Spinelli, F., Serpa, T., Garcia, W., & Ottaiano, A. O. (Ed) (ed.). (2023). *Using language data to learn about language: A teachers' guide to classroom corpus use*. University of Queensland. doi/10.14264/3bbe92d

Römer, U. (2008). Corpora and language teaching. In Lüdeling Anke and Merja Kytö (ed.), *Corpus linguistics: An international handbook*, (pp. 112-130). Berlin: Mouton de Gruyter .

Römer, U. (2011). Corpus Research applications in second language teaching. *Annual Review of Applied Linguistics*, *31*, 205–225. doi/10.1017/S0267190511000055

SkELL. Availabe at: https://app.sketchengine.eu . Consulted on July, 7th, 2023.

SPOC on DDL. Available at: https://learning.edge.edx.org/course/course-v1:UQx+SLATx+2019/home Consulted on Octover, 7th, 2022.

APPENDICES

Appendix I: Pretest Activities[10]

Definition: Near-synonyms are words which share certain semantic similarities, yet differ in their contextual usage (e.g., acquire vs. obtain; evaluate vs. judge).

Task one:
In each of the following items, think about the meaning of the *italicized* word. Then, highlight the right synonym of that word below. Use a dictionary to justify your answer.

1. *Appropriate,* Proper, Suitable
2. *Normal,* Regular, Common
3. *Advise,* Recommend, Suggest
4. *Purpose,* Goal, Objective
5. *Far,* Distant, Remote
6. *Consequence,* Result, Outcome

Task Two:
For each item, think about the meaning of the pairs below and use them in a sentence.

1. Alternative and Option
2. Chance and Opportunity
3. Acquire and Obtain
4. Assess and Evaluate
5. Explain and Clarify
6. Problem and Trouble

Task Three:

1. **Directions:**

List the following items into their positive, neutral or negative meanings.

1. Curious, Nosy, Interested
2. Use, Exploit, Employ
3. Saving, Stingy, Thrifty
4. Stubborn, Steadfast, Tenacious
5. Confident, Courageous, Conceited,
6. Peculiar, Different, Unique,
7. Selective, Picky, Meticulous,
8. Old, Vintage, Decrepit
9. Happy, Manic, Elated

2. **Directions:**

Read the sentences below. Identify the word(s) that have a negative connotation.

1. Bedford is a gritty neighborhood, but the rent is affordable.
2. On my flight to Los Angeles, I sat next to this geezer. He was very old.

[10] Synonymy activities were derived from: https://www.yourdictionary.com/articles/connotative-words-examples

3. Every morning my neighbor takes his mutt to the park. It always barks loudly when leaving the building.
4. You need to be pushy when you are looking for a job.
5. Bob is bullheaded sometimes, but he always gets the job done.

3. Directions:

Choose the word with the most positive connotation to complete the sentences below.

1. I love spending time with my aunt. She has a very (childlike, youthful, childish) outlook on life.
2. I enjoy finding ways to make the most of my budget by being (frugal, miserly, cheap).
3. I am (confident, arrogant, egotistical) that I will be selected for a role in the school play.
4. I realize that you're feeling (nosy, prying, inquisitive), but I'd rather not discuss the matter.
5. My cousin is very (blathering, jabbering, talkative).

4. Directions:

For each of the sentences below, determine if the bold word has a positive or negative connotation.

1. How long have you been dieting? Maybe too long. You look so **skinny**.
2. I can tell you put a lot of work into your book report. It is **extraordinary**.
3. I hope to look like my sister when I grow up. She is absolutely **gorgeous**.
4. I know that Eddie and I are the same age, but he is just so **juvenile**.
5. It's certainly **unusual** to use ComicSans font on a resume.

Appendix II: Pre Questionnaire "Improving Writing through Corpora: Data-driven Learning" (Crosthwaite, 2020)

Personal Information

Gender

Male ☐

Female ☐

Age

18 years old ☐

19 years old ☐

20 years old ☐

Computer Use Tendencies

Overall, do you like to use a computer?

Yes ☐ **No** ☐ **Somewhat** ☐

Do you have Internet access at home?

Yes ☐ **No** ☐

How often are you online via your cell phone/Smartphone?

Never ☐ **Always** ☐

Sometimes ☐ **Often** ☐

How often do you use a computer for school work? (Writing an essay, doing homework)

Never ☐ **Rarely** ☐ **Sometimes** ☐

Often ☐ **Always** ☐

How often do you use a computer for personal purposes (e.g., emailing)?

Never ☐ **Rarely** ☐ **Sometimes** ☐

Often ☐ **Always** ☐

Synonymy Use

When you are unsure about English vocabulary to express your idea, what do you do? ……

How often do you use major Internet search engines to find an answer (Google/Yahoo)?

Never ☐ **Rarely** ☐ **Sometimes** ☐

Often ☐ **Always** ☐

Do you use online translation sites to help you write in English?

Yes ☐ **No** ☐

How often do you use these search engines in English?

Never ☐ **Rarely** ☐ **Sometimes** ☐

Often ☐ **Always** ☐

Which sites do you use? **……**

When you write in English how often do you use an electronic dictionary (Dictionary apps)?

Never ☐ **Rarely** ☐ **Sometimes** ☐

Often ☐ **Always** ☐

Which electronic dictionary do you use?**……**

How often do you use online dictionaries?

Never ☐ **Rarely** ☐ **metimes** ☐

Often ☐ **Always** ☐

Which sites do you use?

……

When you are unsure about English synonyms, do you use a vocabulary book or online reference vocabulary website?

Yes ☐ ☐ **No**

What other resources do you use to write in English?

……

How often do you search for words while you are writing?

Never ☐ **Rarely** ☐ **Sometimes** ☐

Often ☐ **Always** ☐

When you are unsure about some English patterns, how often do you guess instead of confirming your word choice with a reference resource?

Never ☐ **Rarely** ☐ **ometimes** ☐

Often ☐ **Always** ☐

If you were supposed to choose between two synonyms in the sentence “We need to achieve a purpose/goal.”, what would you do to decide on the most appropriate word choice?

Guess ☐

Look in the dictionary ☐

Check a grammar book ☐

Search online ☐

Search in Arabic ☐

Other ☐

Are you knowledgeable about language corpora and corpus linguistics?

Yes ☐ **No** ☐

Appendix III: **Semi-structured Interviews**

1. Can you briefly tell me about your experience with using SPOC on DDL?
2. Why do feel that way?
3. Did you find learning through SPOC on DDL easy or difficult?
4. What was the most useful function you thought it can help you with learning near-synonyms/vocabulary?
5. Why you do you think so?
6. How did you become aware of the usefulness of DDL?
7. Do you think you become more confident when you draw on DDL in your learning ?
8. Are you going to change your learning references' habits to DDL method?