

Using Corpus-Based Instruction in ESL Students' Collocation Learning

Rania Hussien¹

1. Department of English Language and Literature, Beirut Arab University, LEBANON.

Citation: Hussein, R. (2025). Using Corpus-Based Instruction in ESL Students' Collocation Learning. *Gloria: International Multidisciplinary Studies*, 1(3), 1-17. <https://gloria-leb.org/UsingCorpus-BasedInstruction.htm>

Abstract

One of the most effective ways to provide solutions for writing problems is to use corpora in foreign language instruction. This can be done through establishing corpus-based teaching strategies or by using corpus tools to support the material. Learners of second languages particularly university students encounter significant obstacles in their pursuit of mastering writing skills. Utilizing English collocations in writing is one of these difficulties because it's so challenging to understand them, and it can be difficult to know which words collocate. Therefore, it could be hard for second language (L2) English learners to effectively employ English collocations in their writing because of the vast diversity of collocations and their various forms. Despite these challenges, these collocations are important for achieving fluency, precision, effectiveness, and coherence in writing. This paper explores the role of corpora in collocation learning, reviews existing literature and focuses on teaching students how to use the Corpus of Contemporary American English (COCA) more skillfully in order to support their acquisition of various English collocations and ensure that they use them correctly in their writing. Additionally, this study provides examples of search concordances and procedures that could be used on COCA to provide more accurate and representative collocation results.

Keywords: corpus-based instruction, corpora, collocation, COCA, concordances, writing, English language learners

الملخص

يُعتبر استخدام المدونات اللغوية في تعليم اللغة الأجنبية كلغة ثانية أحد أكثر الطرق فعالية لحل المشاكل اللغوية. ويمكن القيام بذلك من خلال إنشاء استراتيجيات التدريس القائم على المدونات اللغوية. تشكل المشاكل التي يواجهها متعلمو اللغة الانكليزية خصوصًا طلاب الجامعات عقبات كبيرة في سعيهم لإتقان مهارات الكتابة حيث يعد استخدام التراكيب اللفظية الإنجليزية في الكتابة أحد هذه الصعوبات لأنه من المهم جدًا فهمها، وقد يكون من الصعب أيضًا معرفة الكلمات المترابطة وكيفية استخدامها وتطبيقها. لذلك، قد يكون من الصعب على متعلمي اللغة الإنجليزية كلغة ثانية أن يستخدموا التراكيب اللفظية الإنجليزية بشكل فعال في كتاباتهم بسبب التنوع الكبير للتعابير المطابقة وأشكالها المختلفة. بالرغم من هذه التحديات، فإن هذه التراكيب اللغوية مهمة وضرورية لتحقيق الطلاقة، والدقة، والفعالية، والتماسك في الكتابة. تستكشف هذه الدراسة دور أهمية المدونات اللغوية في تعلم التراكيب اللفظية، وتستعرض الدراسات الموجودة والمتعلقة بهذا الموضوع. كما أنها تركز على تعليم الطلاب كيفية استخدام مجموعة اللغة الإنجليزية الأمريكية المعاصرة (COCA) بمهارة أكبر كأداة تعليمية من أجل دعم اكتسابهم لمجموعات اللغة الإنجليزية والتأكد من استخدامها بشكل صحيح في كتاباتهم. بالإضافة إلى ذلك، تقدم هذه الدراسة أمثلة على التطابقات النصية التي يمكن استخدامها في تطبيق (COCA) من أجل توفير نتائج أكثر دقة وإتقانًا في كتابة المفردات اللغوية.

Introduction

The emergence of online corpora in applied linguistics has fundamentally altered the perspectives and methodologies of the majority of modern scholars. This shift has created new research opportunities for linguists, allowing them to identify trends, evaluate theories, and make judgments based on substantial factual data. Wang and Wang (2018) have highlighted in their study the importance of corpora in the language learning environment, especially as technological improvements have made these resources more easily available and user-friendly. As more advanced technology is available, corpora are being utilized more frequently as a reference tool used by language teachers and learners since they provide insights that were previously hard to gain (Wichmann & Fligelstone, 2014).

Because of its extensive collection of texts from various genres, Mark Davis' 520 million-word Corpus of Contemporary American English (COCA) is the richest and free accessible online corpus. It can be effectively utilized to provide answers for a variety of pedagogical issues pertaining to English collocations that confront language instructors and students (Davies, 2010). When it comes to writing, English language learners struggle with the use of collocations. University students frequently face difficulties in employing collocations correctly or utilizing collocations that are too informal. Through the utilization of COCA's extensive data, educators and learners can acquire a more profound comprehension of collocational patterns, leading to an improvement in language competency and writing abilities.

Recent studies have shown that accessing corpora directly allows language learners to engage with the text in the (concordance) by observing, exploring word structures, collocations, and linguistic patterns. Ivanova, Skitina, Shabanova, and Poliakova (2023) showed in their study that students who participated in corpus-based learning greatly improved their academic reading, writing, and speaking abilities as well as their retention of the ESP vocabulary.

Moreover, learners can draw generalizations regarding syntax and grammatical characteristics (Karpenko-Seccombe, 2018). This is especially crucial in an EFL classroom where students' direct access to corpora fosters lexical awareness and helps them become familiar with the different contexts in which the words and phrases occur (Xue, 2021). This paper aims to review the existing literature, investigates the function of corpora in language learning, and focuses on teaching students how to use the Corpus of Contemporary American English (COCA) more proficiently to support their learning many collocations in English and make sure they use them correctly in their academic writing.

Problem Statement

Collocations can be difficult for many ESL students to employ in their writing, which affects both their language competency and writing ability. Collocations are idiomatic statements made up of a series of words that occur together primarily for arbitrary reasons, which makes them challenging to master. Collocations are a vocabulary component that naturally occurs to native language speakers; learners typically aren't aware that these expressions exist. One of the main markers of unfamiliarity is the inability of EFL/ESL learners to employ collocations appropriately (Demir,2017). Moreover, traditional instruction methods often fail to provide students with suitable tools that help students use these collocations effectively. As a result, students frequently struggle to correctly understand and use collocations. They frequently find them difficult to comprehend and sometimes tend to pair words that do not collocate, which could produce a speech or writing that is not natural. This study seeks to determine whether participating in corpus-based activities can result in demonstrable gains in students' writing competency as well as how a corpus-based approach can affect students' capacity to use

collocations in their writing. In addition, it seeks to determine whether using corpus-based activities can lead to any improvement of students' writing and shows how this approach helps students overcome the challenges related to their understanding of collocations, thus enhancing their coherence and effectiveness in writing. To achieve this, the study will rely on secondary resources rather than primary ones, drawing insights from existing literature and data to analyze the potential impacts of a corpus-based approach.

By reviewing the existing literature, the following questions are addressed:

1. What does the existing research reveal about the impact of using corpus-based approach on the students' ability to employ collocations in their writing?
2. How have corpus-based activities been shown to influence students' writing proficiency in previous studies?

Examining the existing literature will help identify the findings, tools and strategies that should be applied in classroom and their effects on writing collocations. It also provides a comprehensive understanding of the current research and offers recommendations for future research.

Objectives of the Study

In relation of the research questions above, the objectives of the research are:

1. To review previous literature and show the influence of a corpus-based approach on ESL students' ability to effectively and appropriately employ collocations in their writing.
2. To explore and analyze how corpus-based approach affects students' writing proficiency after engaging in corpus-based activities focused on collocation learning through different studies.

Significance of the Study

Teaching collocations using a corpus-based approach is crucial for both teachers and students. By giving students access to authentic instances of collocations in everyday situations, this method improves their comprehension of language usage. Students' comprehension of proper use of word combinations is enhanced by this exposure, which also helps them write more proficiently. Additionally, as students can use corpora to freely explore and resolve linguistic concerns, participating in corpus-based activities promotes greater autonomy in learning. Moreover, by providing students with contextualized learning opportunities, this approach helps them better grasp and remember collocations in a variety of contexts.

For educators, the corpus-based approach gives access to real and current teaching resources, which enhances the engagement and interest of their classes. It facilitates data-driven education through computer-based instruction, which enables teachers to build more effective classes by focusing their pedagogical choices on actual language use data and using corpus analysis to find prevalent collocation patterns and errors. This method also helps teachers better meet the requirements of their students by providing more targeted and efficient guidance.

Literature Review

Corpora and Language Learning

According to Anderson and Corbert (2017), a corpus is the collection of extensive spoken or written electronic text collections. It is when a computer software that is capable of analyzing the relevant linguistic constructs can simply manage the machine-readable texts. A thorough examination might reveal patterns in the normal and widespread uses of language. A corpus's size might range from

millions to billions of words, and it might include a variety of genres that students find interesting to investigate. A word, phrase, tag, lemma, or even an extensive lexical or grammatical structure can be found through examples given by a concordance. According to Queens University Library (2024), a concordance is “a comprehensive index of the words used in a text or a body of texts. Ordinarily it will not only index but also cite all passages in which a given word occurs.”

With a concordancing tool, researchers can see every instance of a word in a corpus in its surrounding context. They may observe real language use in a variety of circumstances due to the wide range of texts included in these corpora, which aids in their understanding of minute variations in usage, meaning, and form.

Corpora may offer to students some useful resources like the genre's fundamental lexical, grammatical, and organizational elements (Flowerdew, 2014). In addition to examining the real-world examples given, students can examine linguistic patterns by using a corpus and a concordancer (Alsaifi, 2022). Corpora use the online computer concordances that show word collocations through which students can read multiple instances of language usage in various settings. Collocational frequencies provide learners with surroundings and data for the word they are searching for. The context that a corpus provides for examples is another benefit. By examining the examples, students can get a sense of how terminology should be utilized. Through inference, students can determine the meaning of a word. Through the examination of authentic instances, students can identify common patterns in vocabulary, grammar, and syntax, resulting in a more comprehensive comprehension of the language. Additionally, using corpora promotes a data-driven method of teaching languages where students take an active role in their own education. They are able to investigate certain linguistic aspects, put theories to the test, and make judgments based on empirical evidence, all of which improve their linguistic proficiency and develop critical thinking abilities.

Some researchers have supported Corpus-based data-driven learning approach (DDL). This approach was first suggested by Johns (1991), who encourages inductive, self-discovery-oriented learning. The basis of DDL activities is "noticing" and "discovering," which are key components of "Noticing Hypothesis" and "Discovery Learning," two significant theories of language acquisition (Boulton & Cobb, 2017). Several academics and educators later embraced Johns' (1991) notion of "identify, classify, and generalize" to highlight the student's participation in self-detection and to inspire them to use the rich real language resources. It is believed that this self-discovery strategy works well for helping students correct their linguistic errors. By employing concordances to investigate corpus examples, students participate in exploratory learning. The concordance can be utilized to improve students' writing including vocabulary, collocations and grammar when it is selected carefully. The most common Corpus tools used in English are:

1. The Brown Corpus: It is one of the most widely used English corpus tools. It includes one million words from a wide range of American English sources.
2. The British National Corpus: It is a general corpus which contains 100 million words of Modern British English.
3. The American National Corpus (ANC): It is a general corpus containing 22 million words of American English.
4. The International Corpus of English (ICE): It is a specialized corpus that includes examples of spoken and written English from all across the world.
5. The Corpus of Contemporary American English (COCA): It comprises more than one billion words from a variety of sources, including spoken language, academic journals, novels, newspapers, and magazines.

This study's demonstration of how corpus tools, like COCA, help ESL learners enhance their writing abilities is one of its distinctive contributions. Students who studied collocations using concordance lines, for instance, were able to write more precise and contextually relevant phrases in their writings, progressing from simple formulations to more complex expressions (Mansour,2017). Students are empowered to actively participate in their education through innovative teaching techniques like guided discovery projects, which involve having students find patterns in real-world linguistic data. Furthermore, by incorporating corpus tools into peer-editing sessions, students may compare their writing to actual language usage, which promotes collaborative learning and improves linguistic accuracy.

The advantages of corpus-based learning transcend teaching writing; it may also be successfully used to improve speaking and listening comprehension in other ESL learning domains. For example, corpora can enhance students' listening comprehension by assisting them in identifying popular collocations and natural intonation patterns in spoken English (Djalilova, 2024). To help students speak more fluently and clearly, teachers can also employ spoken corpora to teach conversational skills like taking turns or using discourse markers. Teachers can offer a more comprehensive and integrated language learning experience by integrating corpus-based tools into multiple areas of ESL training, giving students the tools they need for communication in the real world.

Collocations and Corpora

Celce-Murcia and Olshtain (2000) refer to collocation as words that come together or “chunks that native speakers can access for comprehension or production” (p.83). Similarly, O'Dell and McCarthy (2008) define collocations as “a combination of two or more which frequently occur together” (p.6).

In terms of their grammatical characteristics, collocations are typically categorized in the following ways:

- a. Adjectives and nouns: a major setback, a critical moment, an intense debate
- b. Nouns and verbs: leadership emerges, opinions diverge, challenges persist
- c. Verbs and nouns: address a concern, achieve a goal, experience a change
- d. Nouns and nouns: coffee cup, phone call, book cover
- e. Verbs and expressions with prepositions: fall in love, sink into despair, tremble with fear
- f. Verbs and adverbs: sing beautifully, laugh heartily, wait patiently
- g. Adverbs and adjectives: deeply grateful, incredibly talented, remarkably quiet

A wide variety of benefits of corpora for vocabulary language learning have been reported in the literature. Concordance lines are an effective tool for vocabulary education, as evidenced by empirical studies on the subject. Numerous research looked into how corpus tools affected EFL learners' acquisition of verb-noun collocations and how concordances can be used to successfully identify verb-noun collocations in context.

Significance of collocations Acquisition in Writing

Collocation instruction is necessary in ESL courses and mastery of collocations is crucial for learners. It is widely recognized that mastery of collocations contributes to the development of language proficiency, communicative competence, and fluency (Bui, 2021). Therefore, learning collocations is essential in second or foreign language acquisition.

Further research showed that errors made by ESL learners are a result of a lack of collocations, which affects their fluency in writing. Collocational competency gives students a useful writing tool.

Numerous academics have argued that teaching collocations in EFL writing classes is crucial in this regard. Lewis (2000) asserted that a greater portion of written language consists of collocations. He stated that “fluency is achievable in case the acquisition of a substantial store of fixed or semi-fixed prefabricated items is accomplished” (Lewis,1997, p. 15). In this sense, it suggests that mastery of a foreign language would not be attainable without a clear emphasis on collocations. They must therefore be instructed and provided with a lot of significance because they are a vital category of vocabulary and a requirement for writing skill.

Using COCA Search for Collocations

Users have several possibilities to check various collocation types with COCA search tools. It could be used to search adjective-noun collocations, verb-verb collocations, verb-preposition collocations, and adverb-adjective collocations using the following link: www.english-corpora.org/coca/ . Since the current study aims to show search mechanisms of COCA, certain strings can be used by students which might foster their collocations’ acquisition and in fixing errors in their writing.

Figure 1 shows possible search display option that can be conducted to check different types of collocations using the *collocates option(POS) list*.

The screenshot shows the COCA search interface. At the top, there's a blue header with the COCA logo and the text 'Corpus of Contemporary American English'. Below this are two tabs: 'SEARCH' and 'FREQUENCY'. The main search area has a text input field containing 'makeVERB', a 'Word/phrase' dropdown menu set to 'verb.ALL', and a 'Collocates' dropdown menu set to 'noun.ALL'. Below these are two columns of results, each with a list of text types (TV/MOVIES, BLOG, WEB-GENL, SPOKEN, FICTION, MAGAZINE, NEWSPAPER, ACADEMIC) and a frequency count. The first column is labeled '1' and the second is labeled '2'. Both columns show the same list of text types and frequency counts.

Figure 1: The Collocates Display Option for Collocation Search on COCA

According to Evert (2008) a collocation is a “combination of two words that exhibit a tendency to occur near each other in natural language” (p.4). Through COCA corpus tool, students can find matching strings. In Figure 2, the search results are shown with examples of words which can collocate with “make” whether with nouns, verbs, adjectives or adverbs.

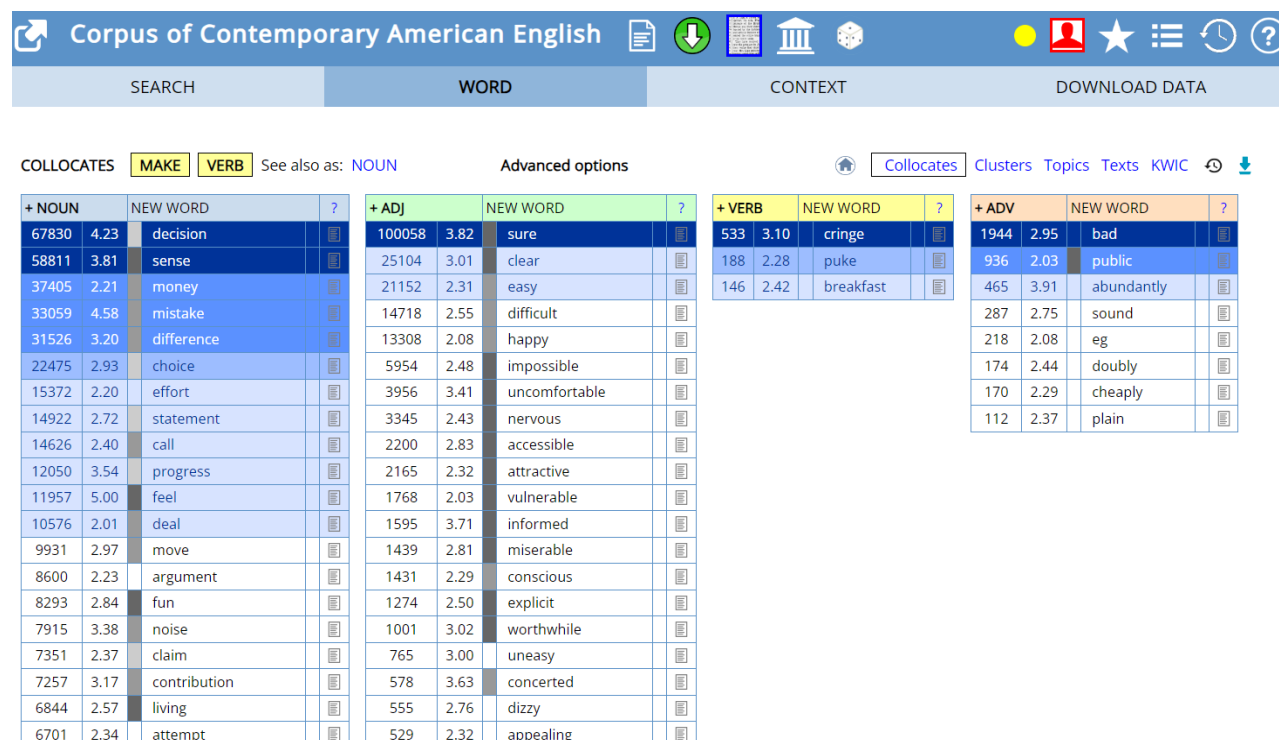


Figure 2: COCA Display for “make” Collocations

Additionally, students can benefit from the word collocates as shown in Figure 3.

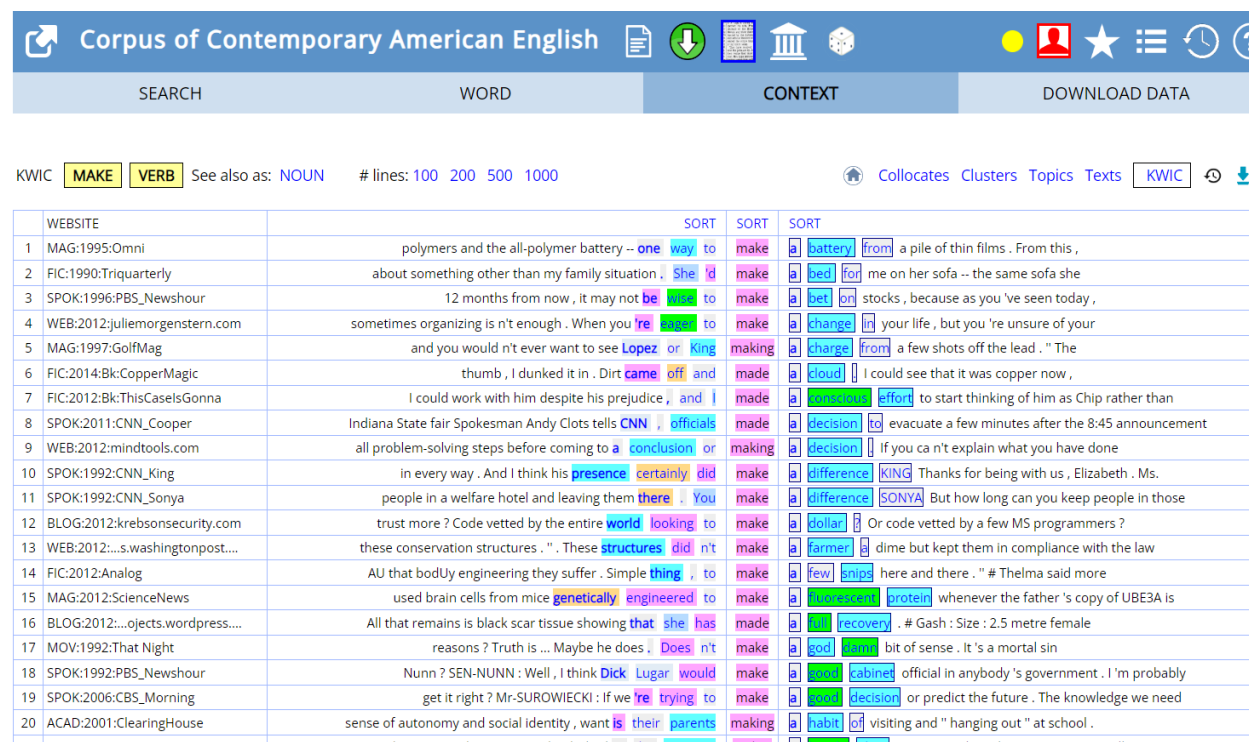


Figure 3: A Screenshot from COCA for “make” and its Collocations

Moreover, students can use the following concordance lines to detect the correct usage of each of the collocations. Figure 4 helps students identify how these words are used in context.

The screenshot displays the COCA interface with the following elements:

- Header:** Corpus of Contemporary American English, with icons for search, download, print, and other functions.
- Navigation:** SEARCH, FREQUENCY, CONTEXT, and ACADEMIC (AVL) tabs.
- Search Results:** A list of 16 entries, each with a year, source, and concordance lines. The concordance lines show the word 'make' followed by 'decision' in various contexts.
- Controls:** A 'SELECT ENTRIES BELOW' button, an 'ADD TO' button, a 'NEW' button, a 'SAVE LIST' button, and a 'SEE LISTS' button.
- Additional Tools:** A 'SAVE' button, a 'TRANSLATE' button, an 'ANALYZE' button, and a 'HELP' button.

Year	Source	Concordance Lines
2018	MAG Business Insider	, " our modest request is that she be given an additional day to make decision . " # Ford has accused Kavanaugh of sexually assaulting her di
2018	ACAD Kutafin University Law Review	is why in the adjudication stage, only the Trial panel is empowered to make decision on the application of supervisory and educational mea
2016	SPOK CNN: Anderson Cooper	It might not be even total convention. TUCKER# I'm not going to make decision base on my discussion with each of the candidates or their s
2016	SPOK CBS: Face The Nation	that there's a whole group of people you hire who then go and make decision that may never get to you. How important is hiring in terms o
2015	NEWS Denver	a decision. Samour said the jury can't rely on outside sources to make decision in this case and again rejected this request. Moving on to th
2014	SPOK CBS: Face The Nation	decision on that? RUBIO# Well, you know, I'll have to make decision in 2016 either way, because I'm up for re-election in the Senate and
2014	SPOK CBS: Face The Nation	president is a political leader, and it's not forbidden for him to make decision that are sensible politics for his party and his governance -- gc
2013	SPOK ABC: 20/20	saying that I wasn't, had a part of it because I did make decision to go. But, you know, was this girl that, like,
2012	SPOK NBC: MeetPress	the people in America having the freedom... MR-GREGORY: But... FMR-SEN-SANTORUM:... to make decision for themselves. He has been for
2012	WEB gowers.wordpress.com	when there is nothing -- neither abstract nor extract, so the reader should make decision based only on the title and authors' names. This is
2012	WEB ...ontporchrepublic.com	you give them too much freedom that you need to have an enlightened elite make decision for them. That is essentially what Salyer is sayin
2012	WEB mayoclinic.com	provoke anxiety and confusion. Concern that the cancer may not be life-threatening can make decision making complicated. # The number
2012	WEB abcnews.go.com	pm # How in the hell is a guy in a coma supposed to make decision ? Ron Paul and his supporters are idiots! # Posted by: PauletteB Septeml
2012	WEB mnn.com	human DNA to have social sex partners. It is in human DNA to make decision or choose. Humans do not care to exercise the capacity to do
2012	WEB pearsonified.com	and let other share their experiences or point of views but certainly you should make decision not because of other people. If somethings v
2012	BLOG skeptoid.com	for a neuroendocrine tumor, " the length of time before you have to make decision about your treatment is not as quick as someone whose

Figure 4: A Screenshot from COCA Concordances for “make” and “decision” Collocations

Previous Studies on Integrating Corpora Tools for Collocation Teaching

The literature has documented a broad variety of advantages of corpora for vocabulary language learning. The results of empirical studies on the use of corpora in collocation education also show the effectiveness of using concordance lines. Several studies examined how corpus tools affected EFL learners' acquisition of collocations.

Altun (2021) studied how successful corpus-based pedagogy is when compared to traditional vocabulary teaching techniques, especially those that involve the use of dictionaries. At a public institution, the study involved two advanced groups of 44 L2 learners who did not major in English. While the experimental group studied 16 pre-selected formal academic words and their strong and weak collocations with a corpus (COCA, the corpus of contemporary American English), the comparison group used advanced learner's dictionaries to study the same collocations. The corpus-based strategy had a higher impact on the acquisition of strong collocations at a slightly higher rate, according to the corpus group's mean scores (Corpus = 45.91, Dictionary = 44.06). The dictionary use group did better at identifying weak collocations.

According to a study conducted by Charles (2012), there is the viability and benefits of teaching EAP writing through the development and analysis of student-made, discipline-specific corpora. Fifty students from multidisciplinary classes of advanced-level students participated in the study. Data were collected from initial and final questionnaires. More than 90% of students reported that creating their own corpora was simple. Moreover, working with their own corpora was a source of enthusiasm for the majority of students, as about 90% of them agreed that it improved their writing and that they planned to utilize it in the future.

Alruwaili (2020) examined the effectiveness of using corpora for learning verb-noun collocations in the classroom. The study used information gathered from 51 students enrolled in a Saudi Arabian university's foundation year general English course at the intermediate level. During the course of the five-week study, participants received three sessions of training on how to read, analyze, and use a corpus tool called AntConc. Results revealed that the participants could independently use the DDL (Data-driven Learning) approach in the same method as they had been trained, indicating the effectiveness of the training. By using concordance lines, the students were also able to recognize the common verb patterns.

Kartal and Yangineksi (2018) in their pre- and post-test quasi-experimental research design, investigated how using corpus tools affected the development and learning of verb-noun collocations by university EFL student teachers. The participants were sixty first-year student teachers from a Turkish state university's English Language Teaching Department and were divided into experimental and control groups. The majority of student teachers reported that collocation learning was greatly aided by concordancing technologies and it is necessary to incorporate corpus consultation into EFL classes.

Chan and Liou (2005) explored the impact of employing five web-based practice units on verb-noun collocations in English. Using the design of a web-based Chinese-English bilingual concordancer (keyword retrieval tool), this study examines the impact of using five web-based practice units on English verb-noun collocations. A total of thirty-two college EFL students in Taiwan participated in the study by completing two post-tests, a background questionnaire, and an evaluation questionnaire. The findings revealed that although learners regressed later, they achieved a considerable improvement in collocation right after the online exercise. Even yet, the ultimate performance exceeded the students' initial level. It was discovered that learners with varying prior knowledge of verb-noun collocations and types of collocations were not equally responsive to the practice effects. The majority of participants found the concordancer and the online educational units to be effective.

However, a few studies highlight the difficulties that students may run into while using concordance-based exercises. According to the participants in a study by Asik, Vural, and Akpinar (2016), there was an unsatisfactory increase in their understanding of word frequency, idioms, and vocabulary learning methodologies. Negative views were also linked to the technical issues with the program and being time-consuming in nature while using the corpus to perform assignments. In the same way, Lu and Dang's (2022) research showed that learners appeared to find the materials' terminology difficult to use in instruction.

According to the studies mentioned above, despite the challenges, using a corpus-based approach when combined with technology-based teaching approaches can help students learn collocations more easily, improve their competency, and aid in the process of correcting errors. Therefore, using it in English language classes could be groundbreaking.

Methodology

This study employs a literature review design to analyze the effect of corpus-based approach on teaching collocations in ESL classrooms. The relevant data for this review of literature were gathered from the following databases: Scopus, Web of Science and EBSCO. The Keywords that were used are “concordances”, “collocation acquisition”, “corpora”, “corpus-based approach”. Some relevant articles from the early 2000s until 2024 were carefully screened and included in the literature review. The main emphasis of the inclusion criteria was based on peer-reviewed studies that clearly investigated the application of corpus-based approaches in collocation instruction within ESL settings which offered comprehensive methods and documented measurable results. Studies were included in

relation to the research questions on whether the selected studies examined the use of corpora for teaching collocations, involved participants who were learners or teachers in ESL or EFL contexts, and explored measurable outcomes such as improved collocation use or writing quality after being engaged in corpus-based activities.

The exclusion criteria included irrelevant studies, book reviews, general reviews, and meta-analyses. These standards assured the chosen studies' effectiveness, reliability, and relevance. In screening studies, full-text articles were thoroughly examined after abstracts were first reviewed for relevance. Studies that used experimental or quasi-experimental designs, employed a variety of corpora, and addressed the use of concordances and data-driven learning techniques received particular consideration. The included studies closely aligned with the research questions of the study and provided high-quality reliable data for the review.

Data Collection and Findings

The following table organizes and presents the title, author, purpose, methodology, participants, corpus used and findings of the five studies included in the literature review.

Table 1: Research Studies on Corpus Approach

1	Title	The Learning Effect of Corpora on Strong and Weak Collocations: Implications for Corpus-Based Assessment of Collocation Competence
	Author(s)	Hatice Altun
	Purpose	To shed light on how successful corpus-based pedagogy is when compared to traditional vocabulary teaching techniques, especially those that involve the use of dictionaries
	Participants	Forty- four participants- two advanced groups of L2 learners at a public university who did not major in English
	Corpus/corpora used	Corpus of Contemporary American English (COCA)
	Methodology	The comparison group examined the identical collocations using advanced learner's dictionaries, while the experimental group researched 16 pre-chosen formal academic words and their both strong and weak collocations using a corpus (COCA, the corpus of contemporary American English). Data was collected using the Oxford Placement Test, pretest, posttest, and exercises created for particular collocation areas of instruction.
	Findings	The corpus group's mean scores (Corpus = 45.91, Dictionary = 44.06) show that the corpus-based approach performed significantly better and had more of an impact on the understanding of strong collocations. Interestingly, the dictionary use group (Corpus=54.08, Dictionary=57.18) acquired weak collocations more effectively.
2	Title	Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building
	Author(s)	Maggie Charles

	Purpose	To examine the feasibility and benefits of teaching EAP writing in a way that allows students to create and analyze their own unique, discipline-specific corpora
	Participants	50 advanced non-native speakers (NNS) of English participated; 62% were Oxford University PhD students and 30% were master's students
	Corpus/corpora used	DIY corpus
	Methodology	The first two sessions mostly devoted to introducing corpus lines and developing the corpora, while the study of a particular rhetorical purpose using discourse and corpus work was the main focus of the final four sessions. Students were asked how often they used corpus, whether they used it for editing or drafting their writing.
	Findings	The results showed that about 90% agreed that their corpus helped them improve their collocations in writing. 88% of participants stated that using the program was simple, and 83% said that using their corpus to find answers to their own questions was easy.
3	Title	Data-driven learning approach for teaching verb-noun collocations in an English foreign language context
	Author(s)	Awatif Katib Alruwaili
	Purpose	To address how students actually employ corpora to learn verb-noun collocations in the classroom
	Participants	51 students enrolled in a Saudi Arabian university's freshman year general English course at the intermediate level.
	Corpus/corpora used	AntConc
	Methodology	The training and testing phases made up the two parts of the experiment. Every session lasted an hour and took place in a computer lab. The training sessions concentrated on using AntConc, a concordancer, which is a corpus tool. The use of corpus linguistics in language analysis was briefly introduced throughout these three training sessions. The participants received instruction in using AntConc to look at broad verb patterns using Sinclair's model during the training sessions. Software trackers were used to monitor the participants during the training and testing sessions. Activity sheets and tracking logs were used to gather data.
	Findings	The main findings revealed that the participants could independently use the DDL (Data-driven Learning) technique in the same manner as they had been trained, indicating the effectiveness of the training. By using concordance lines, the students were also able to recognize the common verb patterns.
4	Title	The Effects of Using Corpus Tools on EFL Student Teachers' Learning and Production of Verb-Noun Collocations
	Author(s)	Galip Kartal Gonca Yangineksi
	Purpose	To examine the effects of using corpus tools on university EFL student teachers' learning and production of verb-noun collocations

	Participants	Sixty first-year student teachers in a Turkish state university's English language teaching department
	Corpus/corpora used	Corpus of Contemporary American English (COCA)
	Methodology	Two classes participated in the study's experimental and control groups according to a pre- and post-test quasi-experimental research design. There was use of both qualitative and quantitative research approaches. A Collocation Knowledge Test was used to collect quantitative data, and it was given as a pre- and post-test. Additionally, the development of collocations was tracked by examining the writings of student instructors both before and after the adoption. Also, a survey was used to get participant feedback regarding the usage of corpus tools.
	Findings	The majority of student teachers reported that collocation learning was greatly aided by concordancing technologies. There has been discussion of the results of exposing students to real data and integrating corpus consultation into EFL lectures. With extremely low p-values (both less than 0.05), showing statistical significance, the pretest and group factors both have significant effects on the production of collocation scores.
5	Title	Effects of Web-based Concordancing Instruction on EFL Students' Learning of Verb – Noun Collocations
	Author(s)	Tun-pei Chan and Hsien-Chin Liou
	Purpose	To explore the impact of five online practice units on verb-noun collocations in English by constructing an internet-based bilingual concordancer for Chinese and English (keyword retrieval program)
	Participants	Thirty-two college EFL students in a university in Taiwan.
	Corpus/corpora used	Web-based bilingual concordance
	Methodology	First, 36 items from the pre-test, the immediate post-test, and the delayed post-test are the three test groups were used. They were utilized to gauge students' understanding of V–N lexical collocation at various stages of the learning process. Second, an introductory questionnaire and an assessment questionnaire in Chinese were created to gauge students' perceptions both before and after the online exercise. The 14-item Background Questionnaire was created to investigate how students learn vocabulary and whether they prefer inductive or deductive instruction.
	Findings	The findings showed that although learners relapsed later, they achieved considerable improvement in collocation right after the online exercise. Even yet, the ultimate performance exceeded the students' initial level. It was discovered that learners with varying prior knowledge of verb-noun collocations and types of collocations were not equally responsive to the practice effects. The majority of participants found the concordancer and the online educational units to be acceptable. The paired t-test comparison between the post-test and pre-test scores revealed significant improvements in students' collocation knowledge

		($p = .000$, $p < .05$). Their collocation awareness was retained in the delayed post-test after two and a half months, with no significant difference ($p = .309$).
--	--	---

Discussion

The two research questions that are addressed are answered based on the literature reviewed. Research Question One: What does the existing research reveal about the impact of using corpus-based approach on the students' ability to employ collocations in their writing?

The research findings in Table 1 show how corpus-based approaches can help L2 English learners acquire collocations more effectively. According to Altun's (2021) research, traditional dictionary approaches are not as efficient at teaching strong collocations as corpus-based education; nevertheless, dictionaries appear to be more useful when teaching weak collocations. The advantages of letting students generate and evaluate their own discipline-specific corpora are demonstrated by Charles's (2012) research, which found that 90% of participants reported increases in their writing abilities. In addition, it was revealed that the corpus helped improve students' collocations in writing, with 88% reporting ease of use, and 83% stating its effectiveness while answering their own questions. This aligns directly with research question 1 by showing how corpus-based activities enable students to identify and apply appropriate collocations in a specific context.

Research Question Two: How have corpus-based activities been shown to influence students' writing proficiency in previous studies?

As shown in Table 1, several studies confirm that corpus-based activities positively influence writing proficiency by improving students' collocation skills. The research conducted by Alruwaili (2020) provides more evidence for the effectiveness of corpus tools, as it demonstrates that students who were trained with AntConc were able to use the Data-Driven Learning (DDL) method on their own to identify common verb patterns. Moreover, the research conducted by Yangineksi and Kartal (2018) discovered that corpus tools greatly improved students' understanding of verb-noun collocations. Resulting in low p -values (both less than 0.05), the pretest and group factors both had significant effects on the increase in collocation scores. This supports the second research question by showing how corpus-based activities, particularly those involving DDL, enhance writing proficiency. Additionally, the study by Chan and Liou (2005) showed that students achieved significant immediate improvement in collocation learning after some training on Web-based bilingual concordancer. Chan's and Liou's paired t -test showed that collocation knowledge had significantly improved ($p = .000$), and that awareness had been retained after two and a half months ($p = .309$). Their paired t -test results further illustrate the lasting impact of corpus-based learning on writing proficiency.

These studies highlight how corpus-based learning can help develop collocation competency. They also show how different approaches can work depending on the kind of collocation and the particular tools employed.

Conclusion

In this study, an overview of the application of computer corpora and concordances in collocations learning and teaching has been provided. As stated before, the corpus-based approach provides a strong and practical method for improving collocation learning among L2 English learners, especially when the Corpus of Contemporary American English (COCA) is used. With the usage of corpus tools, students are able to observe and understand the contextual usage of a variety of collocations by having direct access to real-world language use. This approach tackles the major

obstacles that college students must overcome in order to become proficient writers, especially when it comes to English collocations. However, teachers and students may encounter some difficulties in the classroom because of their lack of experience with corpora and their inability to develop and construct corpus-based activities. When selecting concordance lines for their students, educators need to choose them appropriately (Ma & Mei, 2021).

Research on corpus-based language learning is a relatively recent field. Consequently, one of the most significant fields of study in computer-assisted language acquisition is the use of corpora and concordances (Ballance, 2017). To better comprehend the possible function that computer corpora and concordances may play in L2 language acquisition, more research is required. It is essential that educational institutions and language teachers make investments in professional development and training related to corpus-based tools and methodologies in light of these findings. They need to develop cooperative workshops or practice where educators may exchange corpus-based learning tools, best practices, and experiences. This collaborative support network can assist teachers in overcoming obstacles and gaining useful knowledge. The gap between the theoretical advantages and real-world implementation of these technologies can be closed by providing educators with the skills and knowledge they need to successfully incorporate corpus-based techniques into their instruction. Additionally, improving the ease of use of corpus tools' interfaces can promote their usage and increase accessibility for educators and learners equally. In order to offer a more comprehensive approach to language acquisition, curricula design should contain explicit instructions on how to use corpus tools with other teaching strategies like task-based language teaching or project-based learning. Moreover, policy makers should provide funds for pilot programs to encourage the adoption of corpus-based methodologies and certify educators who use innovative methods. Collaboration between educators, linguists, and technology experts will be vital for establishing most of the possibilities of corpora in improving L2 learners' proficiency-particularly in grasping collocations as the field of corpus-based language learning continues to develop.

References

- Alruwaili, A. K. (2020). Data-driven Learning Approach for Teaching Verb-Noun Collocations in an English Foreign Language Context. *Asiatic: IIUM Journal of English Language and Literature*, 14 (1), 138- 159.
- Alsahafi, M. (2022). Corpora and concordancing in second language vocabulary learning. *Theory and Practice in Language Studies*, 12(12), 2651-2657.
- Altun, H. (2021). The learning effect of corpora on strong and weak collocations: implications for corpus-based assessment of collocation competence. *International Journal of Assessment Tools in Education*, 8(3), 509-526.
- Anderson, W., & Corbert, J. (2017). *Exploring English with Online Corpora*. Palgrave.
- Asik, A., Vural, A. S., & Akpınar, K. D. (2016). Lexical awareness and development through data driven learning: attitudes and beliefs of EFL learners. *Journal of education and training studies*, 4(3), 67-96.
- Ballance, O. J. (2017). Pedagogical models of concordance use: Correlations between concordance user preferences. *Computer Assisted Language Learning*, 30 (3-4), 259-283.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language learning*, 67(2), 348-393.
- Bui, T. L. (2021). The role of collocations in the English teaching and learning. *International Journal of TESOL & Education*, 1(2), 99-109.

- Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and context in language teaching: A guide for language teachers*. Cambridge University Press.
- Chan, T. P., & Liou, H. C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer assisted language learning*, 18(3), 231-251.
- Charles, M. (2012). 'Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes*, 31(2), 93-102.
- Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and linguistic computing*, 25(4), 447-464.
- Demir, C. (2017). Lexical collocations in English: A comparative study of native and non-native scholars of English. *Journal of Language and Linguistic Studies*, 13(1), 75-87.
- Djalilova, A. N. (2024). Use Of Spoken Corpora in Language Learning and Teaching. *Texas Journal of Philology, Culture and History*, 28, 4-7.
- Evert, S. (2008). Corpora and collocations. *Corpus linguistics. An international handbook*, 2, 1212-1248.
- Flowerdew, L. (2014). Corpus-based analyses in EAP. In *Academic discourse* (pp. 105-124). Routledge.
- Ivanova, E. G., Skitina, N. A., Shabanova, V. P., & Poliakova, N. V. (2023). Corpus-Based Teaching of Specific and Academic Vocabulary for ESL Students. *Journal of Higher Education Theory and Practice*, 23(19), 151-157.
- Johns, T. (1991). Should you be persuaded: Two examples of data driven learning. *ELR Journal*, 4, 1-20.
- Karpenko-Seccombe, T. (2018). Practical concordancing for upper-intermediate and advanced academic writing: Ready-to-use teaching and learning materials. *Journal of English for Academic Purposes*, 36, 135-141.
- Kartal, G., & Yangineksi, G. (2018). The effects of using corpus tools on EFL student teachers' learning and production of verb-noun collocations. *PASAA*, 55(1), 100-125.
- Lewis, M. (1997). *Implementing the Lexical Approach: Putting Theory into Practice* (1st ed.). Heinle ELT.
- Lewis, M. (2000). *Teaching collocation: Further developments in the lexical approach*. Hove, England: Thompson – Heinle Publications.
- Lu, C., & Dang, T. N. Y. (2022). Vocabulary in EAP learning materials: What can we learn from teachers, learners, and corpora? *System*, 106, 102791.
- Ma, Q., & Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *Journal of China Computer-Assisted Language Learning*, 1(1), 177-190.
- Mansour, D. (2017, June). Using COCA to foster students' use of English collocations in academic writing. In *Proceedings of the 3rd International Conference on Higher Education Advances (HEAd'17)* (pp. 600–607). Editorial Universitat Politècnica de València, València, Spain.
DOI: <https://doi.org/10.4995/HEAd17.2017.5301>
- O'Dell, F., & McCarthy, M. (2008). *English Collocations in Advanced Use*. Cambridge University Press.
- Queens University Library (2024, October 5). *English language and literature*. Retrieved from <https://guides.library.queensu.ca/english/concordances-quotations>
- Wang, S., & Wang, D. (2018). Reform of College Students' Foreign Language Vocabulary Teaching and Effectiveness Analysis of Students' Vocabulary Competence Based on Corpus Teaching Method. *Educational Sciences: Theory & Practice*, 18(6).
- Wichmann, A., & Fligelstone, S. (2014). *Teaching and language corpora*. Routledge.
- Xue, L. (2021). Using data-driven learning activities to improve lexical awareness in intermediate EFL learners. *Cogent Education*, 8(1), 1996867.

Appendix A: Corpus-Based Activity Collocations Use

Activity 1: Read the following concordance lines (Figure 3) and then answer the following questions.

1. What are the recognized collocates?
2. What part of speech comes after the word “make”?
3. Write a synonym for the collocation in lines 7,9 and 12.
4. Write one sentence using one of the collocations.

Activity 2:

Fill in the blanks with “do” or “make” using the concordance lines.

- a. The instructions you gave me don't_____ sense.
- b. On the weekend, I usually _____ chores around the house.
- c. She _____ extra cash by freelancing online.
- d. It's important to _____ decisions carefully.
- e. Scientists need to_____ their work thoroughly.
- f. We hope to _____ more projects with your company in the future.

Now, write 2 sentences using 2 collocations with make.

- a. _____
- b. _____

Activity 3:

- a. Search COCA www.english-corpora.org/coca/
- b. Click on Find matching strings.
- c. Click results for words in context.
- d. Write five sentences using “make”.

1. _____
2. _____
3. _____
4. _____
5. _____

Dr. Rania Hussein is a skilled English language instructor with extensive experience in teaching at both school and university levels. She is passionate about educational research and is recognized for her strong communication and leadership skills. She holds a Ph.D. in English Linguistics from Beirut Arab University. Her work reflects a deep commitment to enhancing language education through innovative methods.

