

## The Learning Effect of Corpora on Strong and Weak Collocations: Implications for Corpus-Based Assessment of Collocation Competence

Hatice Altun <sup>1,\*</sup>

<sup>1</sup>Pamukkale University, The School of Foreign Languages, Denizli, Turkey

### ARTICLE HISTORY

Received: Dec. 22, 2020

Revised: Apr. 26, 2021

Accepted: May 20, 2021

### Keywords:

Strong and weak collocations,  
Concordancers,  
Traditional learning,  
Lexical assessment.

**Abstract:** Although corpora and corpus linguistics have been applied for quite long in foreign and second language settings, there is still limited understanding about how EFL learners use corpus tools along with dictionaries to enhance their collocation knowledge. This study aims to gain insight into the effectiveness of corpus-based pedagogy in comparison with the conventional vocabulary teaching methods, particularly using dictionaries. The study was conducted with two non-English major advanced groups of L2 learners at a public university. The experimental group studied 16 pre-selected formal academic words and their strong and weak collocations with corpus (COCA, the corpus of contemporary American English), while the comparison group studied the same collocations using advanced learner's dictionaries. The instruments for collecting data included the Oxford placement test, pretest, posttest, and exercises devised for particular teaching points of collocations. Results of Repeated Measures ANOVA tests showed no significant difference between the two experimental groups. However, the corpus-based approach showed more impact on the reception of strong collocations acquired by the corpus group at a slightly better performance rate, as evidenced by the group's mean scores (Corpus=45.91, Dictionary= 44.06). Interestingly, the acquisition of weak collocations was better for the dictionary use group (Corpus=54.08, Dictionary= 57.18). The paper thus offers some implications for teaching and assessing collocation knowledge and makes suggestions that EFL practitioners should create variations in instructional methodologies through gaining awareness of the increasing availability of innovative technologies. Further research on collocations' assessment has also been suggested.

## 1. INTRODUCTION

Numerous studies and publications have emphasized the contribution of corpora to the language learning environment, and corpora are being used more frequently as a reference tool for language teachers and learners as a result of the growing availability of advanced technology. Initially, corpora were used mainly for the production of dictionaries and language textbooks (e.g., Barlow & Burdine, 2006; Gilquin et al., 2007; Sinclair, 2001; Thurston & Candlin, 1997). The common use of corpora in material development is reported as a result of the effectiveness of the published materials in foreign and second language classrooms. The use of authentic language samples from corpora serves as a comprehensible input for language learning settings,

---

\*CONTACT: Hatice ALTUN ✉ [haticealtun@gmail.com](mailto:haticealtun@gmail.com) 📍 Pamukkale University, The School of Foreign Languages, 20070, Denizli, Turkey

---

particularly in foreign language (EFL) classrooms where it is rather challenging to expose language learners to various uses and contexts of a word studied. Corpora were also used as a source of linguistic research on lexical studies, grammar, discourse analysis, pragmatics, and linguistics (e.g., Benesch, 2001; Biber et al., 1999; Cortes, 2002; Flowerdew & Peacock, 2001).

More recently, direct access to corpora by learners comprises the subject of a number of studies (e.g., Bernardini, 2002; Boulton & Cobb, 2017; Chambers, 2005; Chambers & O'Sullivan, 2004; Gaskell & Cobb, 2004; Gilmore, 2009; Yoon & Hirvela, 2004). In all these corpora studies, language learners interact with the text in the concordancer to observe, speculate, and explore language patterns, word forms, and collocations. Learners can make generalizations about grammatical features, syntax, agreement, and stylistics thanks to this inductive learning approach. This is particularly important in the EFL contexts where students usually receive most of their language education through another medium but English. Learner's direct access to corpora promotes lexical consciousness, through which students familiarize themselves with the various contexts of the lexical items. For instance, if the students create a list of vocabulary and prepositions used in context, the concordance lines help students to understand that the same lexical items can be used in multiple contexts. This process can promote students' guessing ability by demonstrating the various uses of language items studied (Johns, 1991).

A new path for corpus use is applying corpus linguistic methods and tools in the design and validation process of language teaching and assessments. Some recent studies particularly focus on the potential benefits of exploiting a learner corpus for testing and assessment of L2 proficiency in writing and also speaking (Callies, 2016; Callies & Götz, 2015). Although corpus studies with the testing focus are still at an early stage, they contribute a lot to the research on the assessment of L2 proficiency (Deshors et al., 2016; McCarthy, 2013). This study is also expected to offer some potential beneficial implementations for assessing L2 vocabulary proficiency, particularly in the context of the *Common European Framework of Reference for Languages* (CEFR).

Another flourishing interest area among language teachers and researchers is formulaic expressions and idiomatic language use (Biber et al., 2004; Wray, 2002, 2008). It is considered that mastery of formulaic expressions is essential to acquire lexical competence and an idiomatic control of language (Ellis, 2002, 2003). The phenomenon of collocations occupies a focal point in the scheme of formulaic language research (Firth, 1957; Lewis, 1993; Lewis, 1997, 2000; Liu, 2010; Nesselhauf, 2003). The study of collocations is of great interest in language teaching because language learners are considered to benefit from the naturally occurring word combinations to gain a more natural phraseology of L2. Thus, instead of memorizing long chunks of words, the learners would be able to produce some of the collocation combinations and would also develop some understanding of linguistic features and processes which affect the way collocations are formed (Walker, 2011). Recently emerging awareness on the importance of corpus consultation, especially corpus concordancing, in the study of collocations has led to the penning a number of studies devoted to this issue (Breyer, 2009; Chan & Liou, 2005; Cheng et al., 2003; Durrant & Schmitt, 2009; Lee & Swales, 2006; Liu, 2010). Nevertheless, despite a plethora of research articles and projects comparing the effectiveness of traditional methods and dictionaries to corpora (Basal, 2019; Çelik, 2011; Daskalovska, 2015; Lai & Chen, 2015), corpus-based language teaching focusing on learners' corpus consultation about different collocation types (i.e.; strong vs. weak collocations) is still on all fours, and more effort is needed to draw up-and-coming implications for EFL contexts.

This paper attempts to contribute to the above-stated niche as a way to teach collocations. More specifically, it aims to see if concordancing exercises, which rely on collocation competence, can enhance the nature of vocabulary learning. This experimental study, therefore, aims to

explore the potential benefits of hands-on concordancing over dictionary use in-class activities for teaching strong and weak collocations over five weeks.

## **2. LITERATURE REVIEW**

### **2.1. Corpora and Language Learning**

A corpus is the accumulation of vast spoken or written electronic text archives (Anderson & Corbert, 2017). The texts are machine-readable and can easily be manipulated by software that can analyze the linguistic constructs in question. A careful analysis can provide insights into how language is used typically and commonly. The size of a corpus can change from millions to billions of words, and it may contain several genres which learners found useful to explore. A concordancing program enables researchers to view all of the occurrences of a particular word in its immediate environment in a corpus. The immediate environment contains several words before and after the search word itself. The full concordance lines indicate the larger text in which examples occur (*ibid.*). Concordancing allows the researchers to perform basic qualitative and quantitative analysis to show all aspects of the nature of the word as well as its frequency in a specific context (Flowerdew, 1996).

Corpora may provide learners with valuable tools such as basic lexical, grammatical, and organizational details for the genre (Tribble, 2001). With a corpus and a concordancer, learners not only see the authentic examples provided but also have the opportunity to study language patterns (Biber et al., 1999). Corpora display word collocations via the concordancing program. Learners can see preceding and subsequent data for the term they are searching for by looking at collocational frequencies. Another advantage of a corpus is the context it brings in examples (Biber et al., 2004). Learners can appreciate the sense in which terms should be used by looking at the examples. By making inferences, students can be able to figure out what a word means. Corpora may also foster an atmosphere conducive to inductive learning (Flowerdew, 2009). This gives students power over their language learning. In this sense, foreign language students take on the position of linguistic researchers, analyzing data and coming up with their own rules and conclusions.

Some scholars and language teachers (Johns, 1991; Tribble, 2001) have strongly supported the use of corpora instead of dictionaries and traditional activities to develop competencies in various skills on account of the fact that concordances are argued to promote learners' analytical thinking skills and autonomy. Adherents of corpora have also argued that traditional learning tools, including dictionaries, are tedious and tiring and also nonproductive tools, particularly for vocabulary learning. The inauthentic examples and the vague language use in dictionaries prevent learners from realizing various authentic contexts of words (Tribble & Johns, 1990). However, according to Cobb (2003), in spite of all the burdensome and time-consuming effects of dictionaries, many language learners still depend upon the dictionaries to learn vocabulary. Additionally, in their meta-analysis, Lee et al. (2018) argue that corpora can increase vocabulary gains considerably, particularly in in-depth vocabulary knowledge of collocations. Also, some collocations which are even difficult to be recognized by the native speakers can easily be taught through concordances.

As opposed to the importance credited to the exploitation of corpora in language teaching, however, total reliance on it may be problematic in that corpora may pose some challenges and obstacles for some learners. First of all, all learners may not have positive attitudes towards inductive discovery learning (Flowerdew, 2009). According to Flowerdew (2009), corpus use is typically correlated with an inductive approach, which may not be suitable for all students due to their differing cognitive styles. This style of learning can benefit field-dependent students who enjoy discussions based on the application of rules from examples (*ibid.*). Field-independent learners, on the other hand, who prefer simple rule instruction will not find it

useful. Cobb (1998) also raises another practical question about corpora exploitation. Lexical information is massive and maybe potentially confusing to the learners. While words occur in a wide range of contexts, many of the words in the concordance lines are unfamiliar, and the contexts are short, incomplete, and do not indicate a coherent and unified context (Cobb 1998). As a result, the teacher's function as a facilitator is essential to overcome the challenge caused by the context (Flowerdew, 2009).

## 2.2. Collocations and Corpora

Collocations are words that appear together in a text more often than their individual frequencies or than would be predicted by chance (Halliday, 1966). Collocating words predict each other, i.e., when one part of a collocating pair is detected, the odds of discovering the other part improve (Hoey, 1991; Jones & Sinclair, 1974). However, there is no set definition of what word combinations are considered as collocations among language educators. The controversy often stems from the disagreement over how structurally fixed and meaningfully transparent a word combination should be to be considered a collocation. Yet most educators agree that collocations are word forms with restricted structural variations and vary from free words, and to alleviate the problem of this arbitrariness, some scholars offered a scale with subcategories, such as 'strong,' 'medium strength,' (Crowther et al., 2002) or 'strong,' 'weak' and 'fixed' (O'Dell & McCarthy, 2008). For this study, the researcher exploited this scale of collocations and focused particularly on strong and weak collocations to be studied by advanced L2 learners.

Language users need to develop collocational links for an efficient lexical network. However, Nesselhauf (2003) and Altenberg and Granger (2001) argue that even advanced English learners have issues with the correct use of collocations. In the EFL settings, developing collocational competence is rather challenging due to the arbitrary nature of collocations. Collocational mistakes are usually the most dominant ones in EFL learners' outputs (Gui & H., 2002; Hsu & Chiu, 2008). Koç (2006) also discovered that one of the main problems with Turkish EFL learners is the lack of collocational competence. Learners tend to learn vocabulary as isolated units rather than as formulaic sequences of words in combination with each other. Furthermore, Prodromou (2003) contends that collocations, either fixed or more flexible, are formed after many years of habitual use by the native speakers of a language. Collocations offer 'chunks' of English that are part of formulaic language ready to be used; therefore, the automation of collocations enables 'native speakers' to express themselves fluently. Second language learners, however, lack this automation and, thus, are more prone to using unnatural phraseologies. In order to achieve automaticity in collocational use, second language learners should be aware that they need to develop an ability to comprehend and produce collocations as unanalyzed chunks (Prodromou, 2003).

Since mastering collocations is rather challenging (Wray, 2000), a large body of study has concentrated on learner mistakes and the primary challenges second language learners encounter while studying collocation norms (Howarth, 1998; Liu, 2010; Nesselhauf, 2003, 2005). It is also well known that second language learners tend to rely on weak collocations, which are non-restricted word combinations (e.g., nice memories, a good meal, bad friends) (Hasselgren, 1994; Nesselhauf, 2005). Considering that word frequency is one of the imperative determiners in making lexical choices (Foster & Chamber, 1973), it is not surprising that high-frequency weak collocations are processed quickly. So the element of familiarity plays a vital role to clutch for the words learners feel safe with, and even advanced learners systematically overgeneralize these 'lexical teddy bears' – "core words – learnt early, widely useable, and above all safe (because they do not show up as errors)" (Hasselgren, 1994, p. 250). On the other hand, strong collocations – low frequency, more clear-cut lexical combinations – take a longer time to learn and are less likely to be used by second language learners (Conzett, 2000). However, strong collocates are expected to facilitate the processing of the following noun

because they prime the subsequent noun and make it more restricted than the same word preceded by a weak collocate (e.g., auburn hair vs. brown hair, inclement weather vs. bad weather) (Hoey, 2005). This inherent paradoxical nature of strong and weak collocations poses an additional challenge to learning collocations. Therefore, the contradictory effects of strong and weak collocations – learners’ reliance on weak collocations but their being less predictable or strong collocations’ facilitating effect of the subsequent word but being difficult to be processed – on gaining collocational competence need to be studied more in EFL settings. There are, however, few studies focusing on learning above mentioned collocations through concordancing (Conzett, 2000).

Only a few studies have looked into the effects of using concordancers to teach EFL students. Some notable ones are as follows. In Sun and Wang’s (2003) study, the efficacy of inductive and deductive teaching approaches on EFL students was investigated. Participants used an online monolingual concordancer to research collocations of various difficulty levels. The inductive group benefited substantially more than the deductive group after the posttest. There was no significant difference between the learners’ performance affected by the teaching method, inductive or deductive, in terms of tricky collocations. However, the inductive approach was more effective in teaching easier collocations with the help of corpora.

Daskalovska (2015), in another notable study, explored the influence of concordance on 44 first-year English language and literature learners’ adverb-noun collocation knowledge. The experimental group outperformed the control group, who studied the collocations through traditional exercises and dictionaries. She underpinned the valuable contribution of concordance use on the collocational production of ELT learners. One last research study worth mentioning is Nesselhauf’s (2003) groundwork. She conducted an exploratory study on verb-object-noun collocations in a corpus of academic essays written by non-native speakers of English. He concluded that although rote learning and behaviorism are discredited, a number of collocations need to be taught and learned explicitly; in this case, the criteria for the selection of collocations to be taught can be determined based on the acceptability and frequency of collocations in any special register of interest to the learner.

Nesselhauf (2005) suggests three criteria to select collocations to be taught to advanced level students: frequency, difficulty, and degree of disruption. Frequency is the number of occurrences of a collocation set in a certain text that students need to study. Collocations with high-frequency and wide-range collocations are deemed worthy of teaching in some studies (Hill, 2000; Hill et al., 2000), and in others, collocations with medium or weak strength (Hill, 2000). Degree of difficulty, i.e., degree of susceptibility to deviation, is the second criterion in the model in which two types of difficulty are explained: absolute difficulty and relative difficulty. Both of them serve as a rating scale for the learnability of a collocation. Deviation in the model means using unnatural phraseology or ungrammatical word combinations. The third criterion is the degree of disruption, i.e., the extent to which a deviant expression confuses the reader or listener and obstructs the quality of meaning to be conveyed or even disrupts the communication. Nesselhauf (2005) admits that this criterion, the disruption criterion, is rather challenging to measure because it is hard to express the degree of disruption in numbers. Moreover, the fuzziness of the idea of disruption (e.g., according to whom and according to what situation) makes the criterion challenging to justify.

Collocations, as a necessary form of vocabulary awareness, have caused learning problems for EFL learners, according to the studies described above (Liu, 2010; Nesselhauf, 2003). The selection of collocations to be taught does not seem to be applicable to all proficiency levels. In the case of advanced levels, learners strive for high proficiency; thus, learners’ needs should be considered as a criterion as well as other dimensions related to collocations. Furthermore, depending on collocation instructions, various forms of collocations seem to behave differently.



Drawing on the criterion of frequency and degree of difficulty in Nesselhauf's (2005) model, the current study, therefore, aims at exploring advanced level Turkish EFL learners' learning processes of strong and weak collocations (Crowther et al., 2002; O'Dell & McCarthy, 2008) with the help of a corpus, i.e., the corpus of American English, COCA- (Davies, 2008). The study looks into the causes of individual treatment differences (with or without concordancers) and various collocation types in order to fill in the gaps identified in the previous research survey.

### 3. METHOD

#### 3.1. Design

This study addresses the possible aftereffects of hands-on concordancing exercises on advanced level Turkish EFL learners' learning strong and weak collocations in comparison to traditional dictionary use. The study has employed a pretest and posttest design, with 44 participants in two groups, a control and an experimental. Both groups took part in the treatment sessions between pretest and posttest. The control group studied the selected collocations through dictionaries and the experimental group via a corpus. The dependent variable of the study is learners' achievement on a collocation test developed by the researcher. The independent variables are two groups who study using concordancing activities and an online dictionary, and the type of collocations taught: strong and weak. Instruction was delivered to both groups through explicit classroom teaching based on the activities prepared by the researcher. The participant groups showed differences as to whether they used dictionaries or concordances during the treatment. The experimental group explored concordance lines of COCA to make meaningful deductions about the collocations to be learned. The control group studied the same vocabulary using the traditional advanced learners' dictionary.

#### 3.2. Research Questions

The following research questions have been addressed:

1. Do concordancing exercises have any impact on L2 learners' collocation competence in comparison with traditional dictionary use?
2. Does the reception level of strong and weak collocations reveal a significant difference in advanced L2 learners?

#### 3.3. The Hypotheses

It was hypothesized that:

H<sub>0</sub>: Statistically no significant difference will be observed in students' posttest scores across the two groups after a period of explicit vocabulary teaching.

H<sub>0</sub>: Statistically no significant difference will be observed in students' performances in posttest scores with regard to strong and weak collocations across the two groups after a period of explicit vocabulary teaching.

The first research question was investigated by assessing the performance of the experimental group against the control group. The second research question was explored by comparing the potential development of the two groups with regard to the strong and weak collocations.

#### 3.4. Participants

All the participants were EFL learners enrolled in an academic writing class at a public university, and they were all native speakers of Turkish. These 51 students took the course of academic writing, during which the collocation treatment was administered for five weeks. Eight participants did not take the posttest; therefore, they were excluded from the data. In total, 44 students took part in all phases of the study. All participants had an upper-intermediate or advanced level of English language proficiency. Their proficiency level was checked using an

Oxford placement test as part of the study. The mean score was 43.86 (SD, 3.968), which was classified as B2 level-upper intermediate by CEFR. According to their own assessment, their computer skills ranged from basic to intermediate and more advanced. None of the participants had any previous knowledge of corpus linguistics.

### 3.5. Data Collection Instrument and Target Structures

A multi-faceted protocol was adopted during the vocabulary collection and test creation phases. Drawing on the criterion of frequency and degree of difficulty in Nesselhauf's model (2005, see literature review 2.2 for the detailed account of the model), the researcher has identified several strong and weak collocations from the teaching materials used in classes and exams in order to meet the advanced EFL learners' needs. The lexical items with a medium degree of difficulty but the relatively low frequency, or vice versa, received a fair amount of attention while preparing the list of collocations to be taught, and they were tested later on in the study. The relative degree of difficulty is measured by comparing the number of deviant expressions of collocation with its overall number in a particular text.

The collocations were selected from among those identified as important because they were considered to be of help to the advanced learners of English in their written and spoken English outputs. Additionally, the researcher focused on collocations that are not immediately obvious (e.g., *adhere to standards*, *auburn hair*, and *broad accent*), considering that those collocations would be helpful for their language exams given in the school and also for the standard exams such as TOEFL, IELTS, and GRE that they might need to take according to their future aspirations.

All collocations tested were adjective-noun bigrams. According to corpus studies, the most common grammatical element in academic texts is nouns (300,000 nouns per million words) (Biber & Conrad, 1999; Biber & Gray, 2016; Biber et al., 1999). The other two most common grammatical functions are adjectives and prepositions (Biber & Gray, 2011). Due to their frequency in the teaching and testing materials, only adj+noun collocations were included as the items to be used in the treatment. Additionally, it is considered that students would encounter adj+noun collocations in most of the high stake tests as well, so these combinations seemed like the most appropriate choice from among the other collocation types.

The selected collocations were divided into two categories. The first category is defined as strong collocations, in which the words are very closely associated, e.g., *mitigating circumstances* or *factors* (see the literature review for the detailed information about strong and weak collocation types). The second one is that of weak collocations in which words collocate with a range of other words. For example, *broad* collocates with a broad range of different nouns, e.g., *broad avenue*, *accent*, *view*. It is also considered that, in terms of their fixedness and idiomaticity, the weak and strong collocations form a continuum, with stronger ones at one end and weaker ones at the other (Conzett, 2000). Most collocations lie somewhere between the two.

The strength of collocations was operationalized through Mutual Information (MI) scores calculated for selected adjective-noun bigrams in the COCA. From among the other association measurements (AM) like T- scores and Log Dice, only MI scores were used as a reference to calculate the probability of co-occurrence of the collocations for some reasons. First, T-scores are considered to be the best indicator for lexical PP-verb collocations among all association measures (Hoffmann et al., 2008) so it was not the best alternative to measure adj+noun combinations' strengths. Although another AM, Log Dice, has been introduced as an alternative to MI scores, it has not been explored enough in language learning research yet (Gablasova et al., 2017). Therefore, MI scores seemed to be the most relevant measure to give information about the bond of probability between the adjectives and nouns used in the current study.

---

Additionally, MI scores are one of the most frequent and reliable measurement tools recommended in the literature to calculate the strength of collocations (Hunston, 2002; Hunston & Laviosa, 2000; Walter, 2012).

MI scores calculate the extent to which specific words co-occur compared to the number of times they appear separately, and they strongly rely on frequencies. Therefore, in order to make sure about the strength of collocations, MI scores were checked using COCA and BNC (British National Corpus). In total, 16 collocations were identified: 8 strong and eight weak ones. Weak collocations were defined as adjective-noun bigrams with an MI score lower than 3, and strong collocations were defined as adjective-noun bigrams with an MI score higher than 8. These cutting edges were recommended by Hunston (2002, p.71). It is generally accepted that MI scores lower than 3 suggest an insignificant likelihood of co-occurrence between the node and its collocate. Therefore, the MI score over 8 would show a highly significant relation of the probability between the searched items.

The classroom exercises were designed to explore the collocates of the pre-determined 16 words. Due to the semantic unrelatedness of these 16-word collocations, exercises focused on discrete items in a rather structured way in a multiple-choice test. In tandem with Nesselhauf (2003), the researcher adopted an explicit teaching method while studying the collocations with learners. The five-week teaching material comprised matching, gap filling, paraphrasing, error correction, and production type of exercises, which allowed learners to explore the selected words and their collocations. The materials also sought to assess learners' ability to adapt their vocabulary information to new contexts. These exercises were studied as part of the academic writing course for almost half an hour every week.

The collocational knowledge test utilized in the study was also designed and developed by the researcher and was used to evaluate students' collocation competence. The multiple-choice test format was chosen for the receptive collocational test, given the objectivity of scoring it allows. The test instructed participants to determine the correct collocate of the highlighted 16 words.

All the distractors were chosen from among the pseudo-collocates, weakly collocated or unrelated items in the lists of COCA and BNC in relation to the search item. For each item, the strong collocates were defined after a thorough search on both corpora. Those collocates that has the highest frequency rate were chosen as the correct answer. Then all the distractors' frequency and strengths were checked in order to make sure that the correct answer is the best option. The piloting of the collocational knowledge test was conducted with 20 ELT students at a different public university and with three English teachers. All the necessary items and distractors' developments were done based on the results obtained from piloting. Cronbach alpha was .815 for the collocation test, which indicated a high internal consistency. One week before and after the five-week experiment, pretests and posttests, which were basically the same test, were administered.

### **3.6. Treatment**

The two participant groups in the experiment were assigned according to lists provided for the academic writing course; that is to say, section one was the first group, and section two the second. The first group (G1), called the corpus group, studied the words and their collocations with concordance and corpus-based activities, but the other group (G2), the control (dictionary) group, used traditional dictionaries while studying the same words. There were five sections in each course, and at each session, four collocation combinations were studied. The activities were completed in half an hour under the guidance of the course instructor. After the administration of the pretest, an introductory lesson in which the collocations and their particular uses were taught by the course instructor was conducted in the first part of the experiment. The corpus group received additional information on the utilization and searching



with a concordancer. The dictionary group dwelled upon the exploitation of dictionaries while studying collocations during the introduction week.

The main part of the experiment, the five-week teaching treatment, was unique to the groups. Corpus group (G1) delved into the corpus queries with COCA, one of the largest corpora in the world with one billion words from eight different genres. Although COCA’s web page offers several linguistic search opportunities, the learners were only asked to use the frequency counts for the collocation search. All the learners in the corpus group performed the classroom task, which required searching through COCA using their own computers during the class period. The control group (G2) used several advanced learners’ dictionaries to do the same collocation searches. They completed the same tasks with the corpus group, and they were not introduced concordances. After the five-week treatment, the posttest was administered to evaluate participants’ performance in collocation learning. Participants’ test scores in each group were accumulated to conduct the necessary analyses.

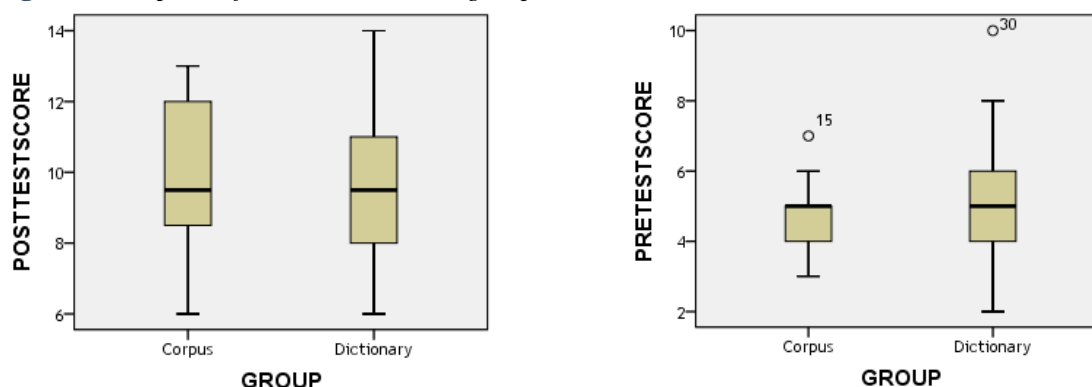
### 3.7. Data Analysis

A one-way ANOVA with repeated measures was used to explore the effect of the two treatments on advanced L2 learners’ collocational competence in two stages. The test type and the collocation types were taken as within-measures of the study. To assess the assumption of a one-way ANOVA, the researcher first checked the normality condition of the data set before making a decision on which statistical method should be used, using skewness and kurtosis indexes along with the Shapiro-Wilk test. Table 1 shows that all skewness and kurtosis values of the data were between -1.96 and +1.96, a threshold recommended by Ghasemi and Zahediasl (2012). This suggests the normal distribution of the data sets of the study. The result of the Shapiro-Wilk test also showed that all four data sets satisfied the normality condition,  $p > .05$ ; therefore, the null hypothesis fails to be rejected. As is shown in the boxplots (see Figure 1), there seems to be one outlier in each group, so they were excluded from the data to conduct the analysis.

**Table 1.** Descriptive statistics of two groups in pre and posttest.

| Groups          | N  |          | Mean | SD    | Skewness | Kurtosis | Shapiro-Wilk |
|-----------------|----|----------|------|-------|----------|----------|--------------|
| Corpus (G1)     | 20 | Pretest  | 4.70 | 1.081 | .117     | -.212    | .919         |
| Dictionary (G2) | 24 |          | 5.13 | 1.849 | .521     | .918     | .953         |
| Corpus (G1)     | 20 | Posttest | 9.70 | 2.130 | -.072    | -.749    | .936         |
| Dictionary (G2) | 24 |          | 9.33 | 2.297 | .046     | -.580    | .942         |

**Figure 1.** Boxplots of test scores across groups.



Since the within-subject test time variable has only two levels, the test for sphericity could not be applied. After the assumptions were met to conduct a repeated-measures ANOVA design, the test was run to pursue the analysis.

#### 4. RESULTS

Both groups performed at a similar rate, according to the means of the pretest results ( $G1_M = 5.56 / G2_M = 4.97$ ). These findings showed that there were no major variations in pre-learning histories between the groups prior to the pretest. The means of the posttest findings, on the other hand, showed a positive variance in favor of G1, which explored lexical items using corpora and concordance-based exercises ( $G1 = 67.24 / G2 = 64.81$ ). Table 2 below presents the summary of descriptive statistics of the pre and posttest results of the collocation test.

**Table 2.** Descriptive statistics from RM ANOVA for pre and posttest of collocation test, M (means) - SD (Standard Deviation).

| Group   | M/SD | Pretest | Posttest |
|---------|------|---------|----------|
| G1 (20) | M    | 32.74%  | 67.24%   |
|         | SD   | 7.269   | 7.269    |
| G2 (24) | M    | 35.17%  | 64.81%   |
|         | SD   | 8.923   | 8.924    |

A repeated measures one-way ANOVA was used to see whether the observed difference in the means of the posttest findings for the two groups was statistically meaningful. The findings, as presented in Table 3, showed that the variance in posttest results favoring the corpus group (G1) was not significant ( $F(1, 42) = .955, p = .334$ ). In response to the first research question about whether corpus activities create a significant difference between the two groups' collocational performance, it can be concluded that the null hypothesis cannot be rejected based on this result. In other words, the groups performed in a parallel manner on both collocation tests. Although the increase rate (of means) in both groups was quite large (nearly 30 points), there was not a significant difference between the groups. However, it can still be commented that regardless of the collocation type, the groups' learning performance during the practice period in the context of collocations was positive.

**Table 3.** Test of within-subjects effects from RM ANOVA for test results.

|              | Sum of Squares | df    | F       | Sig. | Partial Eta Squared | Observed Power |
|--------------|----------------|-------|---------|------|---------------------|----------------|
| Pre/posttest | 22313.204      | 1.000 | 165.243 | .000 | .797                | 1.000          |
| WithinGroups | 128.989        | 1.000 | .955    | .334 | .022                | .15            |
| Error        | 5671.370       | 42    |         |      |                     |                |

Table 4 shows descriptive statistics for the pre and posttest results in terms of collocation form. When the pretest outcomes of both strong and weak collocations were compared, the means of both classes were found to be reasonably similar ( $G1$  strong collocation (sc) = 47.68/ weak collocation (wc) = 47.68), ( $G2$  sc = 49.62 / wc = 50.41).  $G1$  showed better performance in weak collocation items, but  $G2$  revealed equally better performance in strong collocation items in the pretest. The means of posttest results of both groups as compared to those of pretest results indicated a decline in terms of strong collocation items. Despite this stated decline, the study revealed that the decline in the experimental group was less than the control group ( $G1$  sc = 47.68 / 45.91,  $G2$  sc = 49.62 / 44.06). Both groups, on the other hand, revealed better

performance in the weak collocation items in the posttest in comparison to the pretest results, but the results showed positive variance in terms of G2 this time (G1 wc = 54.08, G2 wc = 57.18).

**Table 4.** Descriptive statistics from RM ANOVA for pre and posttest with regard to strong and weak collocations, *M* (means) - *SD* (Standard Deviation).

| Group   | M-SD | Pretest Strong C | Pretest Weak C | Posttest Strong C | Posttest Weak C |
|---------|------|------------------|----------------|-------------------|-----------------|
| G1 (20) | M    | 47.68%           | 52.37%         | 45.91%            | 54.08%          |
|         | SD   | 20.82            | 20.86          | 14.96             | 14.96           |
| G2 (24) | M    | 49.62%           | 50.41%         | 44.06%            | 57.18%          |
|         | SD   | 17.58            | 17.54          | 12.96             | 12.54           |

A repeated-measures ANOVA analysis was run to see if the variance in the means of the posttest on strong collocation items was statistically significant, and the results (see Table 5) obtained from the test revealed that the difference in the strong collocation items observed in favor of corpus G1 was not statistically significant ( $F_{1,42} = .421, p = >.05$ ).

**Table 5.** Test of within-subjects effects from RM ANOVA for strong collocations.

|                  | Sum of Squares | <i>df</i> | F     | Sig. | Partial Eta Squared | Observed Power <sup>a</sup> |
|------------------|----------------|-----------|-------|------|---------------------|-----------------------------|
| Collocation type | 323.789        | 1.000     | 1.733 | .195 | .040                | .251                        |
| WithinGroups     | 78.660         | 1.000     | .421  | .520 | .010                | .097                        |
| Error            | 7848.761       | 42        |       |      |                     |                             |

The test results of the repeated measures analysis conducted for weak collocation items within the tests revealed that the difference in the weak collocation items observed in favor of dictionary G2 was not statistically significant ( $F_{1,42} = .361, p = >.05$ ), either, (see Table 6).

**Table 6.** Test of within-subjects effects from RM ANOVA for weak collocations.

|                  | Sum of Squares | <i>df</i> | F     | Sig. | Partial Eta Squared | Observed Power <sup>a</sup> |
|------------------|----------------|-----------|-------|------|---------------------|-----------------------------|
| Collocation type | 440.116        | 1.000     | 2.701 | .108 | .060                | .362                        |
| WithinGroup      | 139.426        | 1.000     | .856  | .360 | .020                | .148                        |
| Error            | 6843.968       | 42        |       |      |                     |                             |

The second research question explores the acquisition level of strong and weak collocations in both groups. From this analysis, it can be concluded that we cannot reject the null hypothesis, which claims that there was no significant difference between groups in terms of their competence with regard to collocation types. The experimental group showed slightly better performance, as evidenced by the groups' mean scores (G1 sc = 45.91, G2 sc = 44.06). However, the acquisition of weak collocations was slightly better for the dictionary use group despite not being evidenced by the statistical result (G1 wc = 54.08, G2 wc = 57.18). That is to say; it was found from the study that the corpus-based approach might have created some impact by chance on the reception of strong collocations.

## 5. DISCUSSION

The current study aimed to determine the more effective way of teaching strong and weak adjective-noun collocations using either concordancing tools or traditional learning tools of dictionaries. The results of the study did not support the hypothesis that corpus-based treatment would be better in teaching collocations, unlike some other studies which provided some profound effects in favor of corpus use in the literature (Chan & Liou, 2005; Daskalovska, 2015; Tsai, 2019). However, in terms of the collocation learning after the treatment, it can be argued that if enough time and effort spent on the study of collocations, L2 learners improve their lexical competence through guided teaching (Flowerdew, 2009). Although there is no significant difference between the experimental group and the control group in terms of learning collocations, the overall performance of both groups improved considerably. Particularly, in terms of the experimental group's experience, it might be weakly assumed that minimal training about how to use concordancing tools enabled learners to use concordance software well enough to conduct independent searches. In that regard, it can be argued that the study might offer some insight into the contemporarily debated research topic of whether teacher-prepared concordance lines or students' use of concordances on their own should be a more efficient way of teaching. It can be inferred from the study that learners' independent and direct use of corpus and concordancing tools have the potential to help learners to have control over their learning and thus boost their self-autonomy (Gaskell & Cobb, 2004; Sun & Wang, 2003).

Statistically not significant, but the relative success of the experimental group can be associated with the novelty effect of corpora, i.e., Hawthorne effect (Levitt & List, 2011). The students had no prior knowledge and experience of using corpus and concordancing activities. They were aware that they were studying a new and engaging tool to study collocations and expected to perform better. Therefore, this novelty effect might have contributed to their relative success in the posttest. Additionally, the rich input provided by concordance lines allowed students to engage actively in target collocations and to expose themselves repeatedly to the collocations. Lee, Warschauer, and Lee's (2018) meta-analysis demonstrates that corpus use improves in-depth vocabulary knowledge more than definitional knowledge or productive useability. In that sense, with regard to the relative success of the experimental group, it can cautiously be argued that corpus tools provide students with easy and ample access to explore the several aspects of a lexical item. Students' active involvement and spending time on the environment of a word increases the thought process, which may lead to more successful vocabulary gains. On the other hand, for the control group, limited access to the example uses and what is involved in better exploring a word did not require deep processing of the input about a word combination. Therefore, they might have scored slightly less in the posttest.

When the results are explored closely, it can be observed that there are interesting points with regard to the developments in different collocation types. Although the results are not statistically significant, it may be assumed that the experimental group's performance on strong collocation type could be associated with the instruction provided for this group through the corpus considering the previous literature about inductive learning (Sun & Wang, 2003). Strong collocations by nature are less frequent but more fixed collocations in comparison to weak ones. As was hypothesized in the literature (Hoey, 2005), strong collocations make the preceding nouns more marked; thus, it takes lesser time to process them than when a noun is preceded by a weaker collocate. Corpus, in that regard, might have allowed the experimental group to observe and explore ample and authentic use of target strong collocations. It seems quite likely that collocations observed in corpus provide cues on which learners can draw easily. Yet of course further research should be conducted to make strong arguments about it. Students' spending time on the collocates increases the thought process which may facilitate learning challenging strong collocations. Students might have found online concordancing motivational

and engaging while focusing on strong collocations. The control group, on the other hand, continued to rely on lexical teddy bears, i.e., weak collocations in our case, as indicated in the literature (Hasselgren, 1994; Siyanova & Schmitt, 2008). Even though they were advanced learners, once again, learners' dependence on the familiar was revealed through their overgeneralized use of the weak collocations.

When we examine the results from a pedagogical perspective, we can offer a combined methodology of corpora and dictionaries to teach collocations. Although today's language learners are '*digital natives*' (Prensky, 2001), who have sophisticated skills to use digital technologies and also developed new cognitive capacities adaptable to these new technologies, it is evident that some paper-based traditional teaching methodologies still apply to some learners' cognitive styles. As Flowerdew (2009) cautioned us, some field-independent students may not enjoy the inductive learning approach that corpus use adopts. Some students in the experimental group might, in this sense, not have had a positive attitude towards inductive discovery learning on account of their cognitive tendency.

There is consensus in the literature that teaching instruction should guarantee learners to develop an extensive repertoire of formulaic sequences – in our case, particularly collocations (Wray, 2002). The findings presented here seem to support this proposition with regards to collocation learning. The current study was conducted by comparing two instructional methodologies while teaching two different types of collocations and the results of the pre and posttests demonstrated that language teachers should combine concordancing activities with dictionary tasks in order to address various learning needs and styles. Web-based activities can also offer new possibilities to supplement the existing teaching materials.

### **5.1. Implications for Corpus Use in EFL Classes and Exploitation of Corpus for Testing**

Two directions of pedagogical implications can be extrapolated from the present study. First, L2 learners are in need of hybrid teaching tools such as web-based tools and dictionaries to compensate for the limitations of each tool when they are used exclusively in an EFL setting. Dictionaries have been in good use for a long time in language classrooms. But a corpus is a relatively new tool for learners and teachers in particular EFL settings. Therefore, corpus tools should be introduced to both teachers and students in order to gain advantages of using corpus-based teaching/learning activities to address the needs of today's digital-native students. However, total reliance on corpus can pose several challenges on students, as warned by Flowerdew (2009). Since corpus use is based on inductive discovery learning, field-independent students might not benefit from corpus use as much as field-dependent learners. At this point, dictionary use with clear instructions would be more fruitful for the setting. The training sessions for the corpus group had three steps: 1) explicitly describing and teaching several corpora and the concordance, 2) demonstrating how the concordancers and collocation search is conducted, 3) having students hands-on practices in a flexible time frame. So if all the students were given a similar training, they would all make most of the use of corpus tools in vocabulary learning. The scope of this study is focused on vocabulary learning; however, corpus tools could be exploited in teaching many language skills such as writing and speaking. Thus, corpus tool, as a new type of learning aid, mediates language learning when appropriate training is provided for students.

The second direction of implications can focus on exploiting corpus as a testing aid. Teaching collocations is a rather challenging task due to the inherently complex nature of collocations. Choosing collocations to be taught is another task that poses difficulties for teachers. For this study, the researcher chose several adjective-noun collocations with various difficulty and frequency levels. An additional challenge is caused by the paradoxical nature of strong and weak collocations exploited. Participants' errors could provide some insight for teachers about what to focus on and how to improve the lacking information regarding the collocation type.



---

Every teacher might want to build their own learner corpus in order to custom their learners' needs and test their particular proficiency. So, this emerging research field, i.e., exploiting corpus for language testing and assessment, relies on learner corpora which comprise learners' outputs. Learners' errors provide valuable insight for teachers while preparing tests to assess proficiency levels in different constructs of language in the context of CEFR. Using learner corpus improves test content and also decreases the subjectivity of human raters whose holistic ratings are inevitably affected by their value judgment. Thus corpus-driven assessment also helps to validate human raters' claims (See Callies & Götz, 2015 for further research).

## 6. CONCLUSION

The current study is an attempt to delve into the area of learning collocations using different tools, i.e., corpora and dictionaries. Corpus and concordance programs are powerful tools in EFL settings. According to the results of the study, potential differences in learners' performance on collocation tests and their improvement in learning collocations cannot be attributed only to the corpus-based approach. Dictionaries still contribute to the language learning environment; therefore, a combined approach could be a better choice in studying collocations. Many researchers are strong proponents of corpus use in language teaching, yet some reservations about the benefits of corpus exploitations in language classrooms could still be valid in terms of learners' learning styles and needs (Cook, 1998; Widdowson, 2000). Therefore, traditional teaching materials like dictionaries should be supplemented with concordance programs to improve educational settings to respond to the various needs of language learners.

Although the results were not significant, relatively higher mean scores could still be considered to mean that corpus-based pedagogies may be more suitable for today's generations, who were grown up as computer and Internet literates and thus demand faster and cheaper technologies. In that regard, corpora can be a solution to some problems about vocabulary learning in language classrooms. Concordance-based activities provide learners with a chance to conduct research by allowing them to take on their own learning responsibilities (Johns, 1991) and expose them to authentic language (Biber, 2004). To a certain extent, the results of this study also support the literature postulating that corpus-based vocabulary learning exercises have a positive impact in improving lexical competence (Biber, 2004; Cobb, 1997, 2003). The findings of the study are compatible with the corresponding research in the research field (Cobb, 1997; Anđ, 2006).

However, it should be noted that the researcher is fully aware of the fact that a deeper and more detailed analysis would be necessary regarding the linguistics and psycholinguistics factors that affect the intrinsic difficulty of collocations. Therefore, the results should be regarded with caution. The short period of research time and lack of student training about corpus use were among the several limitations of the current study. Additionally, the sample size was not enough to draw generalizable results. The two groups were divided unevenly due to outliers. The number of participants in the corpus group was fewer in number, which could have impacted the results to be statistically significant. Unfortunately, the small sample size did not allow the researcher to draw reliable conclusions about whether the exploitation of corpus or dictionary could improve collocation learning. Mainly because of the sample size for the type of collocations (weak and strong), the researcher did not have enough statistical power to compute the within-effects between the variables. A follow-up qualitative research study could give us some detailed information about the learners' particular vocabulary choice concerning weak and strong collocations. It is necessary to conduct a more longitudinal study with a larger sample size in different settings to explore the effect of corpus use on the collocational competence of advanced students. For further research, a study based on learner corpora would give a more satisfying insight as to why participants made certain errors in collocational pairs

and how these lexical misselections can contribute to L2 vocabulary gain. Further development in computer technology will definitely spawn more efficient tools for incorporating corpus exploitation in L2 vocabulary learning, which will merit further empirical research.

### Declaration of Conflicting Interests and Ethics

The author declares no conflict of interest. This research study complies with research publishing ethics (Ethics Committee Approval: 178.233.42.148). The scientific and legal responsibility for manuscripts published in IJATE belongs to the author.

### ORCID

Hatice Altun  <https://orcid.org/0000-0003-4096-4018>

## 7. REFERENCES

- Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of make in native and non-native student writing. *Applied Linguistics*, 22(2), 173-194.
- Anderson, W., & Corbert, J. (2017). *Exploring English with Online Corpora*. Palgrave.
- Barlow, M., & Burdine, S. (2006). *Phrasal verbs*. Athelstan Pubns.
- Basal, A. (2019). Learning collocations: Effects of online tools on teaching English adjective - noun collocations [Article]. *British Journal of Educational Technology*, 50(1), 342-356. <https://doi.org/10.1111/bjet.12562>
- Benesch, S. (2001). *Critical English for academic purposes: Theory, politics, and practice*. . Lawrence Erlbaum Associates.
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165–182). Rodopi.
- Biber, D., & Conrad, S. (1999). Lexical Bundles in Conversations and Academic Prose. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: studies in honour of Stig Johansson*. Rodopi.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25, 371-405.
- Biber, D., & Gray, B. (2011). Grammatical change in the noun phrase: the influence of written language use. *English Language and Linguistics*, 15(2), 223-250. <https://doi.org/10.1017/S1360674311000025>
- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English : linguistic change in writing*. Cambridge University Press.
- Biber, D., Johansson, S., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Boulton, A., & Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis. *Language Learning*, 67(2), 348-393. <https://doi.org/10.1111/lang.12224>
- Breyer, Y. (2009). Learning and teaching with corpora: reflections by student teachers. *Computer Assisted Language Learning*, 22(2), 153-172.
- Callies, M. (2016). Towards a process-oriented approach to comparing EFL and ESL varieties [Article]. *International Journal of Learner Corpus Research (IJLCR)*, 2(2), 229-251. <https://doi.org/10.1075/ijlcr.2.2.05cal>
- Callies, M., & Götz, S. (2015). *Learner Corpora in Language Testing and Assessment* (Vol. 70). John Benjamins Publishing Company.
- Çelik, S. (2011). Developing Collocational Competence Through Web Based Concordance Activities. *Novitas Royal research on Youth and Language*, 5(2), 273-286.
- Chambers, A. (2005). Integrating corpus consultation in language studies *Language Learning and Technology*, 9 (2), 111–125.
- Chambers, A., & O’Sullivan, I. (2004). Corpus consultation and advanced learners’ writing skills in French. *ReCALL* 16(1), 158–172.

- 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. <https://doi.org/10.1080/09588220500185769>
- Cheng, W., Warren, M., & Xu, X. (2003). The language learner as language researcher: putting corpus linguistics on the timetable. *System*, 31 (2), 173–186.
- Cobb, T. (2003). Analyzing Late Interlanguage with Learner Corpora: Québec Replications of Three European Studies. *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 59(3), 393-424. <https://doi.org/10.3138/cmlr.59.3.393>
- Conzett, J. (2000). Integrating collocation into a reading and writing course. In M. Michael Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 70-87). Language Teaching Publications.
- Cook, G. (1998). The uses of reality: A reply to Ronald Carter. *ELT Journal*, 52, 57-64.
- Cortes, V. (2002). Lexical bundles in Freshman composition. In R. Reppen, S. M. Fitzmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 131-145). Benjamins.
- Crowther, J., Dignen, S., & Lea, D. E. (2002). Oxford collocations dictionary for students of English. In *Oxford collocations dictionary for students of English*.
- Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, 28(2), 130-144. <https://doi.org/10.1080/09588221.2013.803982>
- Davies, M. (2008). *The Corpus of Contemporary American English: 520 million words, 1990-present*. <http://corpus.byu.edu/coca/>.
- Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL*, 47, 157-177.
- Ellis, N. C. (2002). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition*, 24, 297-339.
- Ellis, N. C. (Ed.). (2003). *Constructions, chunking, and connectionism: the emergence of second language structure*. Blackwell.
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. In *Studies in Linguistic Analysis* (pp. 1-32). Oxford.
- Flowerdew, J., & Peacock, M. (2001). Issues in EAP: A preliminary perspective. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 8-24). Cambridge University Press.
- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14, 393-417. <http://dx.doi.org/10.1075/ijcl.14.3.05flo>
- Foster, K., & Chamber, S. M. (1973). Lexical Access and Naming Time. *Journal of Verbal Learning and Verbal Behavior*, 12, 627-635.
- Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence: Collocations in Corpus-Based Language Learning Research. *Language Learning*, 67(S1), 155-179. <https://doi.org/10.1111/lang.12225>
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors?. *System*, 32(3), 301–319.
- Gilmore, A. (2009). Using online corpora to develop students' writing skills. *ELT Journal*, 63(4), 363-372.
- Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, 6(4), 319-335.
- Gui, S., & H., Y. (2002). *Chinese learner English corpus (CLEC)*. Shanghai Foreign Language Education Press.

- Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, & R. H. Robins (Eds.), *In Memory of J.R. Firth* (pp. 148-162). Longman.
- Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4(2), 237-258.
- Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching Collocation* (pp. 47-69).
- Hill, J., Lewis, M., & Lewis, M. (2000). Classroom strategies, activities and exercises. In M. Lewis (Ed.), *Teaching Collocation* (pp. 88-177).
- Hoey, M. (1991). *Patterns of Lexis in Text*. Oxford University Press.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.
- Hoffmann, S., Evert, S., Smith, N., Lee, D., & Prytz, Y. B. (2008). *Corpus linguistics with BNCweb: a practical guide*. Peter Lang.
- Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications*. (pp. 161-168). Clarendon Press.
- Hsu, J. T., & Chiu, C. Y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *The Asian EFL Journal*, 10(1), 181-204.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge University Press.
- Hunston, S., & Laviosa, S. (2000). *Corpus Linguistics*. School of English CELS.
- Johns, T. (Ed.). (1991). *Should you be persuaded: two samples of data-driven learning materials* (Vol. 4).
- Jones, S., & Sinclair, J. M. (1974). English lexical collocations. A study in computational linguistics. *Cahiers de lexicologie*, 24, 15-61.
- Koç, G. (2006). *Developing Collocational Awareness* [MA Thesis]. Bilkent University.
- Lai, S.-L., & Chen, H.-J. H. (2015). Dictionaries vs concordancers: actual practice of the two different tools in EFL writing. *Computer Assisted Language Learning*, 28(4), 341-363. <https://doi.org/10.1080/09588221.2013.839567>
- Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25, 56-75.
- Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, 40(5), 721–753. <https://doi.org/10.1093/applin/amy012>
- Levitt, S. D., & List, J. A. (2011). Was there really a Hawthorne effect at the Hawthorne plant? An analysis of the original illumination experiments". *American Economic Journal: Applied Economics*, 3, 224-238. <https://doi.org/doi:10.1257/app.3.1.224>
- Lewis, M. (1993). *The Lexical Approach: the state of ELT and a way forward*. Language Teaching Publications
- Lewis, M. (1997). *Implementing the Lexical Approach* Language Teaching Publications.
- Lewis, M. (Ed.). (2000). *Teaching collocations: further developments in the lexical approach*. Thomson.
- Liu, D. (2010). Going Beyond Patterns: Involving Cognitive Analysis in the Learning of Collocations. *TESOL quarterly*, 44(1), 4-30. <https://doi.org/10.5054/tq.2010.214046>
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223-242.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. J. Benjamins Pub. Co.
- O'Dell, F., & McCarthy, M. (2008). English collocations in use: Advanced. In Prensky, M. (2001). *Digital Natives, Digital Immigrants*.
- Prodromou, L. (2003). *Collocations*. Macmillan Publishers.

- Sinclair, J. (2001). Preface In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small Corpus Studies and ELT* (pp. vii-xv). Benjamins.
- Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, 64(3), 429-458.
- Sun, Y.-C., & Wang, L.-Y. (2003). Concordancers in the EFL Classroom: Cognitive Approaches and Collocation Difficulty [doi: 10.1076/call.16.1.83.15528]. *Computer Assisted Language Learning*, 16(1), 83-94. <https://doi.org/10.1076/call.16.1.83.15528>
- Thurstun, J., & Candlin, C. N. (1997). *Exploring academic English: A workbook for student essay writing*. National Centre for English Language Teaching and Research.
- Tribble, C. (2001). *Corpora and language teaching: Adjusting the gaze*. Universite' Catholique de Louvain.
- Tribble, C., & Johns, G. (1990). *Concordances in the classroom*. Longman.
- Tsai, K.-J. (2019). Corpora and dictionaries as learning aids: inductive versus deductive approaches to constructing vocabulary knowledge. *Computer Assisted Language Learning*, 32(8), 805-826. <https://doi.org/10.1080/09588221.2018.1527366>
- Walker, C. P. (2011). A Corpus-Based Study of the Linguistic Features and Processes Which Influence the Way Collocations Are Formed: Some Implications for the Learning of Collocations. *TESOL quarterly*, 45(2), 291-312. <https://doi.org/10.5054/tq.2011.247710>
- Walter, E. (2012). Using a corpus to write dictionaries In A. O'Keefe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* Routledge.
- Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied Linguistics*, 21, 3-25.
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21(4), 463-489.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge University Press.
- Wray, A. (2008). *Formulaic Language: pushing the boundaries*. Oxford University Press.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257-283.