View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by ScholarSpace at University of Hawai'i at Manoa

Language Learning & Technology
ISSN 1094-3501

October 2017, Volume 21, Issue 3
pp. 153–171

ARTICLE

Using corpora to develop learners' collocational competence

Shuangling Li, *Southwestern University of Finance and Economics*

Abstract

This article investigates the role of direct corpus use in learners' collocational competence in academic writing. An experiment was conducted between two groups of Chinese postgraduates who had no previous knowledge of corpora. It was embedded in a regular 4-month linguistics course in the students' programmes, where a corpus-assisted method was used for the experimental group and a traditional, or rule-based, method was used for the control group. The English essays written by these two groups of learners from different time periods (before, immediately after, and two months after the course) were analysed regarding the learners' collocational use—in particular, verb-preposition collocations. The results reveal that while both groups showed improvements in their academic writing, the students in the experimental group displayed a significant improvement in the use of collocations, including a higher rate of accuracy, or naturalness, and an increased use of academic collocations and fixed phraseological items. It is thus concluded that the knowledge and use of corpora can help students raise their awareness of habitual collocational use and develop their collocational competence. This supports the positive role of direct corpus application in an EFL context.

Keywords: *Corpus-Assisted Learning, Collocational Competence, Verb-Preposition Collocations*

Language(s) Learned in this Study: *English*

APA Citation: Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology*, *21*(3), 153–171. Retrieved from http://llt.msu.edu/issues/october2017/li.pdf

Introduction

There is an increasing consensus that a corpus, typically a large collection of naturally-occurring texts, is of great value to language teaching. It provides both a rich source of attested language and an authentic learning context for EFL learners who do not usually have as much exposure to the target language as native speakers (McEnery & Hardie, 2012; O'Keeffe & McCarthy, 2010; Sinclair, 1991, 2004). As a result, a significant number of corpus-based studies have been conducted to aid in the teaching of collocations. For instance, many of them analyse collocations used by learners or by native speakers, providing useful insights into classroom teaching (Fellbaum, 2007; Marco, 2011; Namvar, Nor, Ibrahim, & Mustafa, 2012; Nesselhauf, 2003, 2005). A few focus on creating corpus-informed pedagogical materials such as corpus-based collocation dictionaries and academic collocation lists (Ackermann & Chen, 2013; Bahardoust, 2013; Durrant, 2009; Huang, Chen, Tsao, & Wible, 2015; McGee, 2012).

While the above-mentioned areas attract an increasing amount of attention, direct corpus application in the teaching of collocations in the EFL context, particularly on a long-term basis, is still rare. More than a decade ago, Granger (2004) pointed out that "the number of concrete corpus-informed achievements [was] not proportional to the number of publications advocating the use of corpora to inform pedagogical practice" (p. 136); this may still be true today. The number of actual applications of corpora in classrooms is in no way proportional to the number of corpus-based studies of collocation. As Leńko-Szymańska and Boulton (2015) noted recently, "the direct uses of corpora in language teaching are treated rather marginally" (p. 3). The main reason for such scarce empirical use, for example in China, is that many English teachers are not equipped with sufficient knowledge or techniques to be able to use corpora in

Copyright © 2017 Shuangling Li

teaching. A great many teachers are still skeptical of the role of corpus-assisted teaching in contrast to traditional methods.

Against such a background, the current research incorporates direct corpus use into a 4-month course in a university setting in China and investigates its role in the development of learners' collocational competence. More specifically, this research explores the following questions:

1. Is it feasible to incorporate direct corpus use into a regular curriculum? And if it is, is the long-term effect positive?
2. Is the corpus-assisted method more effective than a traditional or rule-based method in the teaching of collocations?
3. If so, what role does the corpus-assisted method play in learners' development of collocational competence, particularly regarding the frequency of collocations, used and misused, and the use of academic collocations and fixed phraseological items?

It is hoped that by addressing these questions, new insights can be provided into corpus applications in the modern era of EFL teaching.

Corpus Application in the Teaching of Collocations

There is no lack of corpus-based studies informing the teaching of collocations, but many of them focus on an indirect application of corpora in classroom settings, for example in designing corpus-informed materials such as collocation dictionaries or lists to be used in classrooms (Ackermann & Chen, 2013; Durrant, 2009; McGee, 2012). On the other hand, it is rare to observe direct uses of corpora in a course to develop learners' collocational competence—probably because it presents teachers with several practical obstacles. For instance, it may be argued that the attested data in a corpus are not simple enough for learners, especially those with lower-level proficiency, or that a great amount of time is often spent simply in familiarising students with corpus use (see Leńko-Szymańska & Boulton, 2015).

Despite such difficulties, direct corpus application in classrooms offers multiple advantages, including access to authentic language, learner autonomy in the practice of concordance analysis, and opportunities for learners to explore how language really behaves and to raise their awareness of natural collocational use. As Johns (1990) suggests, a central point of data-driven learning (DDL) is that we should "cut out the middleman as far as possible and […] give the learner direct access to the data" (p. 18).

Luckily, there are a number of research projects that have experimented with mediated or direct uses of corpora in the teaching of collocations. For instance, Vyatkina (2016) showed that paper-based DDL materials were more effective than traditional methods for teaching new collocations to students at lower-proficiency levels. This is one particularly revealing study that focused on the role of corpus use in learning verb-preposition collocations, but this research mainly involved one-time DDL interventions. Wu, Witten, and Franken (2010) evaluated a system that used a web-derived corpus with several student participants and illustrated how it was useful for expanding their collocational knowledge in writing. Their study is worth mentioning, because students had direct access to a pre-processed and filtered collection of concordances in the revision process of collocational use in their writing. However, this study was primarily a test of effectiveness of that system, which awaits application in classroom teaching. Similarly, while Reynolds (2016) demonstrated that the adoption of a web-based collocational concordancer largely increased learners' accuracy in collocational use, the study only partially incorporated direct corpus use in a writing course, namely when students self-edited their essays for verb-noun errors. Other studies, such as Çelik (2011), Daskalovska (2015), and Huang (2014), also found that corpus application was beneficial in the teaching of collocations, but they mostly involved short-term experiments where a corpus was not truly embedded in a regular course. This gap justifies the current research to explore the feasibility of incorporating direct corpus use into a regular curriculum to teach collocations, especially on a relatively long-term basis.

Verb-Preposition Collocations

Rationales

This study investigates verb collocations because EFL learners, even advanced ones, tend to have frequent problems with them. For instance, Källkvist (1998) suggested that awkward collocations used by advanced Swedish learners of English are often related to the use of verbs. Both Nesselhauf (2005) and Marco (2011) asserted that verb collocation is the major source for questionable or deviant combinations in EFL learners' language use. In addition, Wang and Shaw (2008) summarised the results from other studies on collocational errors and concluded that EFL learners made the most errors in the collocational use of verbs, followed by prepositions and determiners.

There are generally three types of collocations that involve the use of verbs: V + N, Adv + V, and V + Prep (see Benson, Benson, & Ilsen, 1986, p. ix; Lewis, 2000; Wu et al., 2010). Many previous studies have been dedicated to investigating verb-noun collocations (e.g., Ebrahimi-Bazzaz, Samad, bin Ismail, & Noordin, 2014; Marco, 2011; Nesselhauf, 2003, 2005; Zinkgräf, 2008), but little attention has been given to the use of verb-preposition collocations (see Vyatkina, 2016; Wong, 2014). This gap also serves as one of the key motivations for the present study.

Two other reasons highlight the importance of investigating verb-preposition collocations. First, this type of collocation occurs relatively frequently in learners' writing. For example, Namvar et al. (2012) examined nine types of collocations (e.g., V + N, Adj + N, N + N, V + Prep, and Adj + Prep) in learners' writing, and found that the occurrence of verb-preposition collocations actually ranks second, immediately after verb-noun collocations. Second, it will be rather rewarding to focus on verb-preposition collocations in teaching. The study conducted by Wu et al. (2010) supports this point. They utilised lexical data from a web-derived corpus to expand learners' collocational knowledge, and showed that the learners can perform particularly well (100% correct) on verb-preposition collocations after looking at a collection of natural collocational use (pp. 97–99).

Identifying Verb-Preposition Collocations

Although the term *verb-preposition collocation* has emerged in a number of studies, an ideal or agreed definition of it has yet to be provided. Since the current research does not attempt to propose a theoretically rigorous definition, it follows three main criteria to identify verb-preposition collocations: frequency, span of combinations, and pragmatic function.

First, one of the main preconditions for a verb-preposition combination to be viewed as a collocation in this study is its frequent occurrence in language use. Combinations that are infrequent do not fully merit the label *collocation* and are also less important in an EFL learning context (see Barnbrook, Mason, & Krishnamurthy, 2013; Handl, 2008; Howarth, 1996). As a consequence, combinations that occur less than three times per hundred million words in a general reference corpus, such as the British National Corpus (BNC), were not considered in this study.[1]

Second, the current analysis focuses on the collocations where the preposition occurs within a span of three words of the verb. In other words, both adjacent collocations (V + Prep) and discontinuous collocations (V + ? + Prep and V + ? + ? + Prep) are examined. Discontinuous collocations are considered so as to increase the opportunity to identify verb-preposition collocations. For example, a few verb-preposition collocations such as *put emphasis on* and *pay attention to* may have words inserted between the verb and the preposition. However, combinations of a verb and a preposition that occur outside the span of three words are not considered in this study, taking account of the feasibility of the analysis.

Third, the pragmatic function is used as a criterion to distinguish between verb-preposition collocations and verb-particle or verb-adverb constructions. Although a few studies (e.g., Aarts, 1989; Farrell, 2005; Keizer, 2009) included verb-particle and verb-adverb combinations under the umbrella term verb-preposition constructions for reasons such as efficiency, this research makes a distinction between these combinations

because it is believed that the usage of a verb-preposition combination and that of a verb-particle or verb-adverb combination are essentially different, both semantically and pragmatically (Gries, 2003; Kim & Baldwin, 2010; Treffers-Daller, 2011). To this end, this study refers to Jackendoff (2002, p. 69–70), who proposed two primary rules to disambiguate these constructions. When the verb in the combination is intransitive and the non-verbal element serves as the only complement (e.g., *George grew up* or *Fred freaked out*), this combination is a verb-particle construction. Conversely, if the non-verbal element is not the only complement (e.g., *Bill ran up the street* or *companies need to communicate with customers*), this combination is a verb-preposition construction. When the verb in the combination is transitive and the non-verbal element can appear on either side of the object (e.g., *Pat put out the garbage* or *Pat put the garbage out*), this combination is a verb-particle construction. In contrast, if the non-verbal element can only appear on the left side of the object or complement (e.g., *language teaching will benefit from a DDL approach*), the combination is a verb-preposition construction.

Methodology

Overview of Research Procedure

In this research, an experiment was conducted to investigate the role of direct corpus use in the development of learners' collocational competence in English academic writing. Both the experimental and control groups completed a course in linguistics, with the former using a corpus-assisted approach and the latter a more traditional approach. English essays written by these two groups of learners from different time periods (before, immediately after, and two months after the course) were collected and analysed in terms of the use of verb-preposition collocations. In the following sections, detailed information about the participants, the courses they took, the data used for the current analysis, and the procedure for retrieving verb-preposition collocations is discussed.

Participants and Courses

The participants in this experiment were 60 Chinese postgraduate students who had no previous knowledge of corpora. They were all majoring in English at three top universities in Chengdu. They were either in the first or second year of postgraduate study, and their ages ranged from 21 to 26. Before the experiment, they were assessed by writing an English essay of around 2,000 words on any topic related to linguistics (for the assessment criteria, see Appendix A). Based on their writing performance, 30 students were assigned to the experimental group and the other 30 to the control group. This initial assessment was to ensure that the average performance of the two groups in writing was similar.

Next, the two groups completed a 15-week course on linguistics using, respectively, a corpus-assisted method and a *traditional* method (i.e., no introduction or use of a corpus, usually with a rule-based and teacher-fronted style). The main reason for choosing linguistics is that this course is often essential for English major postgraduate programmes in China. Very few universities in China provide general English language courses to English major postgraduate students, so it is usually the case that the linguistics course will aim to develop students' ability in language analysis as well as their English language proficiency. The course introduced several important language phenomena, including words, words and meaning, words and grammar, phraseology, collocation, discourse, and genre analysis. Were the experiment based on this course successful, the result would indicate that it is feasible to integrate direct corpus use into a regular curriculum so as to develop learners' collocational competence or English language competence in general.

More specifically, the experimental group was taught the course using mainly the academic part of the BNC, along with Corpus of Contemporary American English (COCA).[2] Each session incorporated direct corpus use. The teacher adopted an inductive approach and helped the students first to understand new language phenomena (e.g., collocation and phraseology) and explore their usage (for an example worksheet, see Appendix B). The control group, on the other hand, was taught this course using a traditional or rule-based method with the same teacher. These students had access to dictionaries and carried out similar activities with a more teacher-centred approach and no corpus use. Appendix C provides more details of

the course contents and activities for the two groups.

For this experiment, written informed consent was obtained from all the participants who kindly allowed their essays to be used for research purposes. They were informed that these essays would be a means to monitor their progress in English academic writing, but they did not know that their collocational competence would be the main focus of this research. Withholding this piece of information prevented students from paying extra attention to collocational use while writing their essays.

Data for the Current Analysis

As mentioned above, each participant wrote three essays in different time periods (before, at the end of, and two months after the course). For each essay, the students were instructed to write on any topic related to linguistics for around 2,000 words and were given one week to write the essay after class.[3] They were also allowed access to any tools or materials they used in the course (e.g., the experimental group was allowed to use the corpora while the control group had access to dictionaries and their learning materials).

The essays written by two students in the control group were not considered for further analysis, because they did not participate in all sessions of their course, which to some extent invalidated the analysis of their essays. Therefore, the current analysis focused on 174 essays in total (90 by the experimental group and 84 by the control group).

These 174 essays were later processed to anonymize participants' personal information and then tagged in terms of part of speech (POS), constituting the corpus for the current investigation.[4] This corpus is referred to as the Corpus of Student Essays (CSE), consisting of 375,672 tokens. The CSE was further divided into six subcorpora to distinguish texts from different groups and time periods (see Figure 1). The six subcorpora were analysed using WordSmith Tools 6 (Scott, 2015).

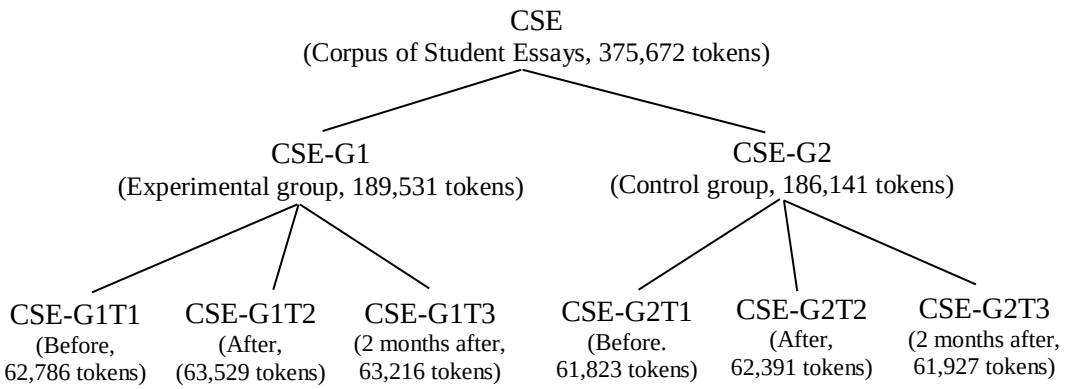

Figure 1. The construction of the corpus for the current analysis

Retrieving Verb-preposition Collocations

Based on the criteria discussed earlier for identifying verb-preposition collocations, two steps were followed to retrieve the target items. First, all the instances in the POS-tagged corpus that contained the combinations of a verb and a preposition within a span of three words were obtained. Second, these retrieved instances were examined carefully to separate those that did not contain verb-preposition collocations. Two native speakers of English were also involved in this identification process. Based on a random sample of 1,000 concordance lines from the retrieved instances, an inter-coder reliability analysis was performed using the *Kappa* statistic to determine consistency, and a substantial agreement was found among the two coders: $Kappa = 0.76$, $p < .001$, 95% CI [0.72, 0.80].

To summarise, six cases were not considered in this study, as shown in Table 1. The first case concerned the most important criterion for identifying verb-preposition collocation: frequency. Combinations that

occurred less than three times in the BNC were not considered in the subsequent analysis. The second and third cases related to the mis-tagging of data—for example, a mis-tagged verb or preposition in a combination. The fourth case involved combinations of the verb *to be* and a preposition (e.g., *is about* and *are in*). They were disregarded because they were not very revealing with regard to learners' collocational use. In the fifth case, the preposition in the combination was actually part of another fixed phrase (e.g., *violate on purpose* and *arranged at the same time*). Finally, verb-particle and verb-adverb combinations were distinguished from verb-preposition collocations by considering the two criteria provided by Jackendoff (2002) and the feedback from the two coders.

Table 1. *Cases not Considered as Verb-Preposition Collocations*

Category	Examples
Infrequent combination	elders especially have their own sons **accompanying beside** their bed. … what she has done is **affirmed through** the high salary …
Mis-tagged verb	Some scholars also did **researches about** the translation of … … and to give them **lectures about** the society …
Mis-tagged preposition	… list several examples to **illustrate that** the violation of the … The language learning process may **benefit if** the emphasis of study is …
Be verbs	The third part **is about** comparison and analysis … … that millions of bicycles **are in** use in Guangzhou …
Preposition as a part of another fixed phrase	… the Cooperative Principle is **violated on** purpose in daily life … his job and your work were **arranged at** the same time …
Verb-particle or verb-adverb combinations	Apart from the terms **mentioned above**, sometimes other researchers… The new definition of this term **turned** the situation **around** and showed …

Results

General Overview of Participants' Collocational Use

A general overview of each learner's collocational use can be observed from three aspects: the total number of tokens of verb-preposition collocations, the variety of collocations used, and the frequency of misused collocations. Table 2 presents an overview concerning the experimental and control groups in different time periods. These three aspects of collocational use were examined in terms of the mean value, that is, the average tokens, the types of collocations, and the misuse of collocations associated with each group (also illustrated in Figure 2).

Table 2. *Overview of Collocational Use by the Experimental and Control Groups*

		T1 (before)		T2 (after)		T3 (2 months after)	
Aspect	**Group**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Token	Experimental	41.27	2.02	75.53	2.47	73.57	2.40
	Control	42.14	2.49	57.29	2.84	58.46	2.61
Type	Experimental	19.47	1.84	33.67	2.37	34.07	2.02
	Control	20.14	1.81	26.89	2.62	25.57	2.46
Misuse	Experimental	12.40	2.14	4.50	1.26	4.00	1.46
	Control	11.86	2.23	9.54	2.18	9.68	1.87

Note. For the experimental group, n = 30; for the control group, n = 28.

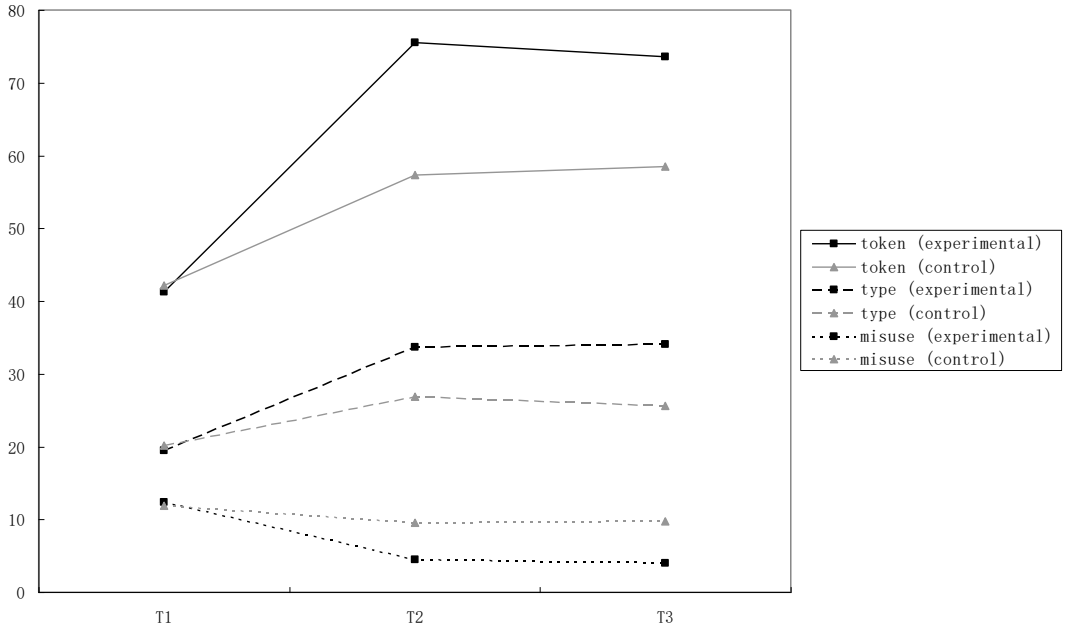

Figure 2. An overview of learners' collocational use over time

It can be inferred from Table 2 that before the course (Time 1), the average number of tokens, types, and misuse of verb-preposition collocations associated with the two groups were similar. Levene's test of equality of variances also indicated that the variance between the two groups in Time 1 was not statistically different, whether for tokens ($F_{(1, 56)} = 0.01$, $p = .949$), types ($F_{(1, 56)} = 0.23$, $p = .633$), or misuse ($F_{(1, 56)} = 0.03$, $p = .862$).

Regarding the average frequencies of misused collocations, it is worth pointing out that although the mean values of misuse associated with the two groups seemed small, the error rate was actually not low. The experimental group had an average error rate of 30.0% (12.40/41.27), and the control group had a rate of 28.1% (11.86/42.14). Therefore, the misuse of verb-preposition collocations needed to be given adequate attention.

After taking the course, it seemed that both groups exhibited a similar trend in the development of collocational use, using more tokens and types of verb-preposition collocations and producing fewer misused collocations over time. The repeated-measures ANOVA revealed that the increased use of tokens and types of collocations were statistically significant for both groups (tokens, $F_{(1, 56)} = 2115.91$, $p < .001$, $\eta^2 = .98$; types, $F_{(1, 56)} = 420.72$, $p < .001$, $\eta^2 = .94$) and that the decrease in misuse was also significant for all participants ($F_{(1, 56)} = 123.47$, $p < .001$, $\eta^2 = .82$).

However, it was also noticeable that for these three aspects of collocational use, any increase or decrease associated with the experimental group was more striking than that of the control group (Figure 2). First, there was a much more dramatic increase in the average number of verb-preposition collocations used by the experimental group than by the control group. The ANOVA performed on the tokens of collocations indicated a statistically significant difference between the two groups ($F_{(1, 56)} = 757.12$, $p < .001$, $\eta^2 = .93$), both from Time 1 to Time 2 ($p < .001$) and from Time 1 to Time 3 ($p < .001$), though there was no significant difference from Time 2 to Time 3. Second, while both groups tended to use more types of verb-preposition collocations after the course, the increase of average collocation types associated with the experimental group was far more than that of the control group. This difference between groups was statistically significant ($F_{(1, 56)} = 270.08$, $p < .001$, $\eta^2 = .83$), but pairwise comparisons only indicated significant differences between Time 1 and Time 2 ($p < .001$) and between Time 1 and Time 3 ($p < .001$); not between Time 2 and 3 ($p = .337$). Third, there was a much more dramatic decrease in the frequency of misused

collocations by the experimental group than by the control group (Figure 2), with a significant effect ($F_{(1, 56)} = 138.34$, $p < .001$, $\eta^2 = .71$). Follow-up comparisons again revealed that pairwise differences were significant between Time 1 and Time 2 ($p < .001$) and between Time 1 and Time 3 ($p < .001$); not between Time 2 and Time 3.

This overview of results indicates that both groups benefitted from the course in terms of improvement in their collocational use. The experimental group, in particular, showed a significant improvement in terms of the variety and accuracy of collocational use. The following sections will discuss in detail how the knowledge and use of corpora affect learners' collocational competence.

Participants' Misuse of Collocations

Main Types of Misuse

How, or when, learners misuse collocations is often considered an intriguing question by English teachers and researchers who aim to develop learners' collocational competence. Based on the essays written by the participants before the course, it was found that their misuse of collocations could be generally categorised into four types, as illustrated in Table 3. The inter-coder reliability test for tagging the misuse showed that there was a very high level of agreement among the two coders: *Kappa* = 0.87, $p < .001$, 95% CI [0.84, 0.91].

Table 3. *Major Types of Misused Collocations in Learners' Essays Before the Course*

Category	%	Examples
Spelling	17.6	Regarding the structure, this paper is ***devided into** five sections. In this construction, a noun is ***preceeded by** one or several adjectives …
Transitivity of verbs	12.8	… in that case, the word nearly **penetrate *into** people's everyday life **Concerning *on** this term, previous studies have often regarded it as …
Collocate (e.g., misuse of preposition)	62.5	daily communications generally **comply *in** the Cooperative Principle He has now been **dedicated *on** providing humanitarian …
Other awkward use	7.1	The utility of those words could be ***distinguished as** metaphorical use… The above analysis of verbs, which ***follow with** this linguistic feature, …

*Note. * indicates the location of misuse.*

The first type of misuse involved misspellings of the verb in a collocation. These misspellings often resulted from students' confusions about the form of a few verbs (e.g., **devide* for *divide*, **preceed* for *precede*, **seperate* for *separate*, or **indentify* for *identify*). Surprisingly, many of these errors were made by more than one student in their writing. In other words, this type of misuse was not rare in participants' essays, despite the fact that it could have been the easiest type to avoid in writing.

The second type of misuse related to the transitivity of verbs. Although it was not frequent in the current study, Moehkardi (2002) asserts that the problems of verb transitivity in the use of verb collocations can be overwhelming. In the examples in Table 3, verbs like *penetrate* and *concern* do not need a preposition to take an object. Another case related to this misuse arose when a verb could be either transitive or intransitive, depending on the context. Take the collocation *enter into* and the verb *enter* for instance. Nouns such as *agreement* and *contract* often follow *enter into*, exhibiting senses like commencement, participation, or consideration. In contrast, nouns such as *room* and *house* often follow the verb *enter*, reflecting its more physical association. Some participants were not fully aware of these usages and constructed inappropriate expressions such as **enter into the building* and **enter into the school*.

The third type of misuse concerned a misuse of the preposition in a collocation, and was the most frequent type of misuse associated with the participants in this study before the course (62.5% among all types of

misuse). For example, the following is a list of such misused verb-preposition collocations, which occurred relatively frequently in the participants' essays:

> *adapt to* (*adapt *for*); *be confined to* (*be confined *in*); *be engaged in* (*be engaged *with*); *be related to* (*be related *with*); *comply with* (*comply *in*); *concentrate on* (*concentrate *with*); *dedicate to* (*dedicate *on*); *derive from* (*derive *with*); *differ from* (*differ *with*); *distinguish between x and y* (*distinguish *from x and y*); *immigrate to* (*immigrate *in*); *suffer from* (*suffer *with*)

It was found that the top three misused collocations of this type contained the prepositions *to*, *with*, and *from* (an error rate of 21.7%, 17.3%, and 15.5%, respectively). This result could be partially explained by the high frequency of the prepositions *to* and *with* themselves, or it could be an indication that these collocations were relatively more difficult for the participants to acquire than others (see Jalali & Shojaei, 2012, p. 89–90). Either way, it can be argued that explicit teaching of these frequently misused collocations in classrooms is worthwhile in order to reduce learners' use of awkward collocations. On the other hand, the collocations that contained the preposition *by* were less frequently misused among all the combinations (an error rate of 1.9%). It is possible that the preposition *by* was easier for these learners to use than other prepositions since it mainly reflected a sense of agency or passive voice. This speculation was also supported by the study conducted by Zhou, Rong, and Huang (2014, p. 1439–1440) which showed that the preposition *by* had a higher precision rate than many other prepositions in Chinese learners' writing.

The fourth type of misuse concerned the remaining awkward collocations used in participants' essays, often a result of misunderstanding the entire collocation. For instance, in the first example in Table 3, the student used the combination *distinguished as* to express a meaning that can be more appropriately realised by fixed collocations such as *recognised as* or *considered as*. Similarly, in the second example, the combination *follow with* could have been substituted with *fit*, *exhibit*, or *show*, which would have been more suitable in that context.

Misuse of Collocations over Time

This section examines the occurrences of the aforementioned four types of misused collocations in students' essays over time so as to indicate the development of their collocational competence in terms of accuracy. Table 4 presents such an overview, showing the average frequencies for each type of misuse found in the different time periods.

Table 4. *Misuse of Collocations by the Two Groups over Time*

		T1 (before)		T2 (after)		T3 (2 months after)	
Category of Misuse	**Group**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Spelling	Experimental	2.17	1.04	1.33	0.87	0.97	0.75
	Control	2.11	1.18	1.39	0.90	1.57	1.15
Transitivity	Experimental	1.59	1.02	0.93	0.73	0.73	0.68
	Control	1.52	0.91	0.93	1.00	0.61	0.67
Preposition	Experimental	7.73	1.92	2.03	0.95	1.93	1.06
	Control	7.42	2.26	6.21	2.47	6.71	1.89
Other awkward	Experimental	0.91	1.01	0.30	0.53	0.40	0.55
	Control	0.81	0.76	1.00	1.10	0.82	0.85

Note. The numbers in the Mean column (M) indicate the average frequencies for one certain misuse associated with the experimental (or the control) group in a certain time period. For example, the number in the top left of this table (2.17) indicates the average frequency of misuse in spelling associated with the experimental group at Time 1.

Concerning the first two types of misuse, Table 4 shows that after the course, both groups reduced the

number of misused collocations in their essays (spelling, $F_{(1, 56)} = 10.43$, $p < .001$, $\eta^2 = .28$; transitivity, $F_{(1, 56)} = 18.06$, $p < .001$, $\eta^2 = .40$). However, there was no statistically significant difference between groups (spelling, $F_{(1, 56)} = 2.57$, $p = .114$, $\eta^2 = .04$; transitivity, $F_{(1, 56)} = 0.18$, $p = .674$, $\eta^2 = .00$). This finding thus highlights the importance of explicit instructions of any kind in the development of learners' collocational competence, especially with regard to spelling and transitivity.

It was evident that after the course, the final two types of misuse appeared much less frequently in the essays written by the participants in the experimental group, with a significant effect (preposition, $F_{(1, 29)} = 128.11$, $p < .001$, $\eta^2 = .90$; awkward, $F_{(1, 29)} = 4.07$, $p = .028$, $\eta^2 = .23$). In contrast, these two types of misuse still occurred frequently in the essays by the control group (no significant effect for the decrease of preposition, $F_{(1, 27)} = 2.00$, $p = .156$, $\eta^2 = .13$; no significant effect for the decrease of awkward $F_{(1, 27)} = 0.21$, $p = .813$, $\eta^2 = .02$). This contrast between groups was also statistically significant: significant from Time 1 to Time 2 and from Time 1 to Time 3 for preposition ($F_{(1, 56)} = 97.99$, $p < .001$, $\eta^2 = .64$) and similarly significant from Time 1 to Time 2 and from Time 1 to Time 3 for awkward ($F_{(1, 56)} = 8.20$, $p = .006$, $\eta^2 = .13$). This suggests that the corpus-assisted learning method may have been more useful than the traditional method in terms of helping students reduce their misuse of collocations, particularly those related to the use of prepositions.

Collocations Frequently Used by the Participants

Apart from the above-mentioned misused collocations, participants used many other collocations reasonably well in their essays. Table 5 lists the 15 most frequently used verb-preposition collocations, both before and after the course.

Table 5. *Frequently Used Collocations by the Two Groups Before and After the Course*

Group	T1 (before)	T2 (after)
Experimental	*accepted by*, *affected by*, *argue about*, *base on*, *borrow from*, *caused by*, *depend on*, *discuss with*, *elaborate on*, *exclude from*, *help with*, *influenced by*, *distract from*, *negotiate with*, *originate from*	*associate with*, *collocate with*, *concentrate on*, *consider as*, *defined as*, *depend on*, *distinguish between* a *and* b, *divide into*, *pay attention to*, *proposed by*, *provide with*, *put emphasis on*, *suggested by*, *regard as*, *relate to*
Control	*affected by*, *base on*, *borrow from*, caused *by*, *depend on*, *divide into*, *fight for*, *graduate from*, *influenced by*, *know about*, *learn from*, *negotiate with*, *prevent from*, *substitute for*, *wait for*	*created by*, *depend on*, *discuss with*, *experiment on*, *help with*, *influenced by*, *know about*, *learn from*, *lecture on*, *originate from*, *present to*, *proved by*, *regard as*, *specialise in*, *written by*

First, among these frequent collocations, some were used much less frequently after the course, while a few new collocations were used—often, they were academic collocations (i.e., combinations that occurred significantly more frequently in academic discourse; see the definition of *academic collocations* in Ackermann & Chen, 2013; Durrant, 2009). For example, before the course, both groups used many general collocations, such as *borrow from*, *know about*, and *learn from*, which occur relatively less frequently in academic writing. After completing the course, it seems that these general collocations were used less frequently. On the other hand, both groups made use of more academic collocations: *defined as*, *proposed by*, and *suggested by* in the essays from the experimental group; *proved by* and *written by* in the essays from the control group. This trend was particularly evident concerning the experimental group (see Table 6). It was clear that in both Time 2 and Time 3, the essays from the experimental group contained more academic collocations than those from the control group. This difference between groups was also statistically significant ($F_{(1, 56)} = 579.51$, $p < .001$, $\eta^2 = .91$; significant both from Time 1 to Time 2 and from Time 1 to Time 3).

Table 6. *Academic Collocations Used by the Two Groups over Time*

		T1 (before)		T2 (after)		T3 (2 months after)	
Aspect	**Group**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Token	Experimental	12.97	1.17	28.77	2.20	31.30	1.27
	Control	13.25	1.30	23.07	1.41	19.93	1.60
Example	Experimental	*argue about, base on, caused by*		*associate with, collocate with, proposed by, defined as, suggested by*		*regard as, consider as, relate to, associate with, conducted by*	
	Control	*base on, depend on, divide into*		*created by, discuss with, proved by, written by*		*proved by, experiment on, adopted by*	

Note. The identification of academic collocations refers to the lists provided by previous researchers (e.g., Ackermann & Chen, 2013; Durrant, 2009; Huang et al., 2015).

Second, it was also noticed from Table 5 that some verb-preposition collocations used in participants' essays formed a part of longer phrases (e.g., *paid to* could be considered as a part of *attention was paid to*; *put on* as a part of *put emphasis on*). Interestingly, it seemed that such collocations were used relatively frequently by the experimental group after the course, and warranted examination over time. Table 7 shows the average tokens and frequent examples of such collocations used by the two groups in different time periods. It was found that both groups used more such collocations after the course, which suggests that the knowledge learned from the course may have facilitated use of phraseological expressions, among the experimental group in particular. This group difference was also statistically significant ($F_{(1, 56)} = 335.97$, $p < .001$, $\eta^2 = .86$; significant both from Time 1 to Time 2 and from Time 1 to Time 3).

Table 7. *Collocations as Part of Fixed Phraseologies Used by the Two Groups over Time*

		T1 (before)		T2 (after)		T3 (2 months after)	
Aspect	**Group**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Token	Experimental	19.57	2.09	35.73	2.37	36.10	2.10
	Control	20.39	2.04	27.82	2.19	26.64	2.30
Example	Experimental	*pay price for ...* *argue with ... about ...*		*attention being paid to* *put emphasis on …* *as has been depicted by*		*distinguish between ... and ...* *pay attention to …*	
	Control	*give way to ...* *keep ... from* V-*ing*		*complain to ... about ...* *pay attention to …*		*attach importance to ...* *talk to ... about ...*	

Note. The discussion of phraseological expressions takes into account three main criteria: frequency, syntagmatic fixedness, and semantic non-compositionality (see Barnbrook et al., 2013; Handl, 2008; Howarth, 1996; Sinclair, 1991).

Discussion

The experiment showed that while both the experimental and control groups benefitted from the course over time, the participants in the experimental group greatly improved in terms of their collocational competence. More specifically, the knowledge and use of corpora seemed to contribute to three aspects of learners' collocational competence: decreased use of awkward collocations, increased use of academic collocations, and a higher degree of phraseological features exhibited in their writing.

Awkward Collocations

Awkward collocations are often considered as *markers* of non-native use in writing because EFL learners, regardless of their language levels, frequently produce sentences with awkward collocations (Marco, 2011; Nesselhauf, 2003, 2005). Various reasons may be given for this, including influence from learners' L1 and cultural background (Farghal & Obiedat, 1995; Namvar et al., 2012), insufficient knowledge about the usage of these collocations (Koosha & Jafarpour, 2006; Namvar & Ibrahim, 2014), and learners' personal experience and language competence (Ebrahimi-Bazzaz et al., 2014; Ganji, 2012). Fortunately, many of these aspects can be effectively addressed with autonomous implicit learning—or in particular, explicit teaching. A number of studies (e.g., Kennedy, 2003; Sonbul & Schmitt, 2013; Zaferanieh & Behrooznia, 2011) suggest that learners who receive explicit instruction on collocations perform better than those who receive no instructions or implicit instruction through mere exposure.

The findings from the current research also support the prominent role of explicit instruction in the teaching of collocations. The participants in both groups experienced an improvement in their collocational use after explicit instruction. The improvement associated with the control group suggests that any sort of explicit teaching may be useful to develop learners' collocational competence.

More importantly, the current analysis revealed that the incorporation of the corpus-assisted approach into teaching was much more useful than the traditional or rule-based explicit instructions. After taking the corpus-assisted course, it was evident that the participants in the experimental group used fewer awkward collocations than those in the control group. This result affirms the positive role of a corpus in teaching. The corpus, as a collection of authentic language texts, provides a rich source of natural language for EFL learners and raises learners' awareness of the naturalness of collocational use in writing (Flowerdew, 2012; Sinclair, 1991). Further, it enables the participants to develop a learning habit to check their use of collocations or other phraseological expressions with attested language data (Flowerdew, 2015; Reppen, 2010).

Academic Collocations

The second finding was that corpus use contributed to an increase in learners' use of academic collocations, a finding that has not been frequently addressed in previous research. A large number of studies have aimed to provide a corpus-based analysis of academic collocations in specific disciplines or across disciplines, showing the relationship between collocations and genre (Biber, Conrad, & Cortes, 2004; Marco, 2000; Ordem & Bada, 2016; Ward, 2007); others have sought to create lists of academic collocations for classroom use (Ackermann & Chen, 2013; Durrant, 2009; Huang et al., 2015). More studies are still needed to highlight the connection between corpus use and learners' choice of academic collocations.

The current analysis contributes to the existing literature in showing that the experimental group used academic collocations more frequently than the control group after the course (see Table 6). It is inferred that corpus use has given the participants constant exposure to academic discourse and opportunities to notice, either intentionally or subconsciously, features of such discourse. In other words, direct access to a corpus of academic texts may enable learners to notice the style of the given discourse and foster an awareness of using academic collocations in their writing. Therefore, this finding implies that the teaching of academic collocations could be carried out with a combination of methods. In addition to methods suggested by previous studies, such as making use of corpus-informed academic collocation lists and collocation dictionaries in classrooms, it is also beneficial to give learners direct access to a corpus so that they can observe and investigate collocational use themselves.

Phraseological Features

The results from the current study also suggest that the corpus-assisted learning course had a positive influence on the use of more fixed phraseologies. It was found that after the course, the essays written by the experimental group contained far more collocations that were a part of fixed phraseologies than those written by the control group. This indicates how direct corpus use may change the extent to which learners

use phraseological items or the extent to which their writing is phraseological. This change echoes the argument from previous studies that the integration of corpora into teaching can help learners detect lexico-grammatical patterning in given texts (McEnery & Hardie, 2012; O'Keeffe & McCarthy, 2010; Sinclair, 1991). Opportunities to analyse words through concordance tools may help learners to realise how language itself tends to be phraseological, and how words do not just respond to the open choice principle, but more importantly, the idiom principle (see Sinclair, 1991, 2004). Inductive learning with corpora is, perhaps, more important and effective in the long term than having teachers simply highlight the phraseological features of language use.

Conclusion

This research empirically applied direct corpus use in classroom settings in China on a relatively long-term basis and explored its role in learners' development of collocational competence in academic writing. An experiment was conducted between two groups of Chinese postgraduate students, using a corpus-assisted approach for the experimental group and a traditional one for the control group. The analysis focused on the essays written by these participants in three different time periods: before, immediately after, and two months after the course. This experiment first suggests that it is feasible and positive to incorporate direct corpus use into a regular course, such as linguistics. The results further show that the experimental group experienced a significantly greater improvement in collocational use than the control group. The improvement is reflected in many aspects, such as the increase in the total number of collocations used in their writing, the increased variety of collocations, the increased accuracy in collocational use, and the increased use of academic collocations.

These findings support the view that corpus-assisted learning can greatly contribute to the development of learners' collocational competence. Traditional explicit instruction on collocational use may be useful to learners to some extent, but corpus-assisted instruction is particularly effective in improving learners' use of collocations. What a corpus offers learners is an authentic learning context and the opportunity to investigate language use themselves. This exposure to attested language data raises learners' awareness of using collocations in a more natural or near-native way. Additionally, the method of inductive learning facilitates noticing of habitual collocations, which reduces learners' tendency to form awkward collocations. Given all these advantages, it would be beneficial for more researchers and teachers to investigate direct corpus applications in classroom settings. Since this study mainly focuses on verb-preposition collocations to investigate the role of corpus use, it would be useful for future research to explore other types of commonly-used collocations by learners or other aspects of learners' collocational competence, providing a more comprehensive picture of how the corpus-assisted approach will greatly facilitate language learning in the modern era.

Acknowledgements

I would like to thank all the enthusiastic participants for making this research possible. I am extremely grateful to the *LLT* editors, the anonymous reviewers, and my colleagues for their valuable feedback and insightful suggestions. Needless to say, any remaining errors or misconceptions are all mine.

Notes

1. The BNC can be accessed through the BNCweb.
2. The teachers and students accessed the COCA through Davies's BYU interface.
3. These essays were checked for plagiarism before the submission, using PaperPass.
4. This study used TreeTagger, developed by members of BFSU Corpus Research Group.

References

Aarts, B. (1989). Verb-preposition constructions and small clauses in English. *Journal of Linguistics*, *25*, 277–290. doi: 10.1017/S0022226700014109

Ackermann, K., & Chen, Y. H. (2013). Developing the Academic Collocation List (ACL): A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, *12*(4), 235–247. doi: 10.1016/j.jeap.2013.08.002

Bahardoust, M. (2013). Grammatical collocation in writing production of EFL learners. *The Iranian EFL Journal*, *9*(1), 266–279.

Barnbrook, G., Mason, O., & Krishnamurthy, R. (2013). *Collocation: Applications and implications*. London, UK: Palgrave Macmillan.

Benson, M., Benson, E., & Ilsen, R. F. (1986). *The BBI combinatory dictionary of English: A guide to word combinations*. Amsterdam, Netherlands: John Benjamins. doi: 10.1075/z.bbi

Biber, D., Conrad, S., & Cortes, V. (2004). 'If you look at...': Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371–405. doi: 10.1093/applin/25.3.371

Briguglio, C. (2007). Assessing the writing skills of entry-level undergraduate business students to enhance their writing development during tertiary studies. In S. Frankland (Ed.), *Enhancing teaching and learning through assessment* (pp. 16–23). Dordrecht, Netherlands: Springer. doi: 10.1007/978-1-4020-6226-1

Çelik, S. (2011). Developing collocational competence through web based concordance activities. *Novitas-ROYAL Research on Youth and Language*, *5*(2), 273–286. Retrieved from http://www.novitasroyal.org/Vol_5_2/CelikS.pdf

Coffin, C., Curry, M., Goodman, S., Hewings, A., Lillis, T., & Swann, J. (2003). *Teaching academic writing: A toolkit for higher education*. London, UK: Routledge.

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, *28*(2), 130–144. doi: 10.1080/09588221.2013.803982

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, *28*(3), 157–169. doi: 10.1016/j.esp.2009.02.002

Ebrahimi-Bazzaz, F., Samad, A. A., bin Ismail, I. A., & Noordin, N. (2014). Verb-noun collocation proficiency and academic years. *International Journal of Applied Linguistics and English Literature*, *3*(1), 152–162. doi: 10.7575/aiac.ijalel.v.3n.1p.152

Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Review of Applied Linguistics*, *28*(4), 315–331. doi: 10.1515/iral.1995.33.4.315

Farrell, P. (2005). English verb-preposition constructions: Constituency and order. *Language*, *81*(1), 96–137. doi: 10.1353/lan.2005.0017

Fellbaum, C. (2007). *Idioms and collocations: Corpus-based linguistic and lexicographic studies*. London, UK: Continuum.

Flowerdew, L. (2012). *Corpora and language education*. New York, NY: Palgrave MacMillan.

Flowerdew, L. (2015). Data-driven learning and language learning theories. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15–36). Amsterdam, Netherlands: John Benjamins. doi: 10.1075/scl.69.02flo

Ganji, M. (2012). On the effect of gender and years of instruction on Iranian EFL learners' collocational competence. *English Language Teaching*, *5*(2), 123–133. doi: 10.5539/elt.v5n2p123

Granger, S. (2004). Computer learner corpus research: Current status and future prospects. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp.123–146). Amsterdam, Netherlands: Rodopi.

Gries, S. T. (2003). *Multifactorial analysis in corpus linguistics: A study of particle placement*. London, UK: Continuum.

Hamp-Lyons, L., & Heasley, B. (2006). *Study writing: A course in written English for academic purposes*. Cambridge, UK: Cambridge University Press.

Handl, S. (2008). Essential collocations for learners of English: The role of collocational direction and weight. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 43–66). Amsterdam, Netherlands: John Benjamins.

Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Tübingen, Germany: Niemeyer.

Huang, P., Chen, C., Tsao, N., & Wible, D. (2015). The development of a corpus-based tool for exploring domain-specific collocational knowledge in English. *Taiwan Journal of TESOL*, *12*(2), 117–141. Retrieved from http://www.tjtesol.org/attachments/article/395/07_TJTESOL-266-0930%20(2).pdf

Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL*, *26*(2), 163–183. doi: 10.1017/S0958344014000020

Jackendoff, R. (2002). English particle constructions, the lexicon, and the autonomy of syntax. In N. Dehé, R. Jackendoff, A. McIntyre, & S. Urban (Eds.), *Verb-particle explorations* (pp. 67–94). Berlin, Germany: Mouton de Gruyter.

Jalali, H., & Shojaei, M. (2012). Persian EFL students' developmental versus fossilized prepositional errors. *The Reading Matrix*, *12*(1), 80–97. Retrieved from http://www.readingmatrix.com/articles/april_2012/jalali_shojaei.pdf

Johns, T. (1990). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, *10*, 14–34.

Källkvist, M. (1998). Lexical infelicity in English: The case of nouns and verbs. In K. Haastrup & A. Viberg (Eds.), *Perspectives on lexical acquisition in a second language* (pp. 149–174). Lund, Sweden: Lund University Press.

Keizer, E. (2009). Verb-preposition constructions in FDG. *Lingua*, *119*, 1186–1211. doi: 10.1016/j.lingua.2007.12.007

Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *TESOL Quarterly*, *37*(3), 467–487. doi: 10.2307/3588400

Kim, S. N., & Baldwin, T. (2010). How to pick out token instances of English verb-particle constructions. *Language Resources & Evaluation*, *44*, 97–113. doi: 10.1007/s10579-009-9099-7

Knoch, U. (2011). Rating scales for diagnostic assessment of writing: What should they look like and where should the criteria come from? *Assessing Writing*, *16*(2), 81–96. doi: 10.1016/j.asw.2011.02.003

Koosha, M., & Jafarpour, A. A. (2006). Data-driven learning and teaching collocations of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal*, *8*(4), 192–209. Retrieved from http://asian-efl-journal.com/December_2006_EBook.pdf

Leńko-Szymańska, A., & Boulton, A. (Eds.). (2015). *Multiple affordances of language corpora for data-driven learning*. Amsterdam, Netherlands: John Benjamins. doi: 10.1075/scl.69

Lewis, M. (2000). *Teaching collocation: Further developments in the lexical approach*. Hove, UK: Language Teaching Publications.

Marco, M. J. L. (2000). Collocational frameworks in medical research papers: A genre-based study. *English for Specific Purposes*, *19*, 63–86. doi: 10.1016/S0889-4906(98)00013-1

Marco, M. J. L. (2011). Exploring atypical verb + noun combinations in learner technical writing. *International Journal of English Studies*, *11*(2), 77–95. doi: 10.6018/ijes.11.2.149651

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory, and practice*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511981395

McGee, I. (2012). Collocation dictionaries as inductive learning resources in data-driven learning: An analysis and evaluation. *International Journal of Lexicography*, *25*(3), 319–361. doi: 10.1093/ijl/ecr040

Moehkardi, R. (2002). Grammatical and lexical English collocations: Some possible problems to Indonesian learners of English. *Humaniora, 14*(1), 53–62. doi: 10.22146/jh.v14i1.745

Namvar, F., & Ibrahim, N. (2014). Construction of collocations in the writing of postgraduate students. *International Journal of Arts & Sciences, 7*(2), 487–497. Retrieved from http://www.universitypublications.net/ijas/0702/pdf/P4G76.pdf

Namvar, F., Nor, N. F. M., Ibrahim, N., & Mustafa, J. (2012). Analysis of collocations in the Iranian postgraduate students' writings. 3L: *The Southeast Asian Journal of English Language Studies*, *18*(1), 11–22. Retrieved from http://ejournals.ukm.my/3l/article/download/945/864

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223–242. doi: 10.1093/applin/24.2.223

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam, Netherlands: John Benjamins. doi: 10.1075/scl.14

Ordem, E., & Bada, E. (2016). Lexical collocations (verb + noun) across written academic genres in English. *European Journal of Education Studies*, *3*(1), 20–36. doi: 10.5281/zenodo.51567

O'Keeffe, A., & McCarthy, M. (2010). *The Routledge handbook of corpus linguistics*. London, UK: Routledge.

Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge, UK: Cambridge University Press.

Reynolds, B. (2016). Action research: Applying a bilingual parallel corpus collocational concordancer to Taiwanese medical school EFL academic writing. *RELC Journal*, *47*(2), 213–227. doi: 10.1177/0033688215619518

Scott, M. (2015). *WordSmith tools manual (version 6.0)*. Stroud, UK: Lexical Analysis Software Ltd. Retrieved from http://lexically.net/downloads/version6/wordsmith6.pdf

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

Sinclair, J. (2004). *Trust the text: Language, corpus, and discourse*. London, UK: Routledge.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, *63*(1), 121–159. doi: 10.1111/j.1467-9922.2012.00730.x

Treffers-Daller, J. (2011). Grammatical collocations and verb-particle constructions in Brussels French: A corpus-linguistic approach to transfer. *International Journal of Bilingualism*, *16*(1), 53–82. doi: 10.1177/1367006911403213

Vyatkina, N. (2016). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, *28*(2), 207–226. doi: 10.1017/S0958344015000269

Wang, Y., & Shaw, P. (2008). Transfer and universality: Collocation use in advanced Chinese and Swedish learner English. *ICAME Journal*, *32*, 201–232. Retrieved from http://clu.uni.no/icame/ij32/ij32_201_232.pdf

Ward, J. (2007). Collocation and technicality in EAP engineering. *Journal of English for Academic Purposes*, *6*(1), 18–35. doi: 10.1016/j.jeap.2006.10.001

Wong, M. (2014). Verb-preposition constructions in Hong Kong English: A cognitive semantic account. *Linguistics*, *52*(3), 603–635. doi: 10.1515/ling-2014-0001

Wu, S., Witten, I., & Franken, M. (2010). Utilizing lexical data from a web-derived corpus to expand productive collocation knowledge. *ReCALL*, *22*(1), 83–102. doi: 10.1017/S0958344009990218

Zaferanieh, E., & Behrooznia, S. (2011). On the impacts of four collocation instructional methods: Web-based concordancing vs. traditional method, explicit vs. implicit instruction. *Studies in Literature and Language*, *3*(3), 120–126. doi: 10.3968/n

Zinkgräf, M. (2008). V+N miscollocations in the written production of university level students. *Elia: Estudios de lingüística inglesa aplicada*, *8*, 91–116.

Zhou, Y., Rong, Z., & Huang, G. (2014). A correcting model for preposition error in English essays of Chinese student based on hybrid features classification. *Applied Mechanics and Materials, 631–632*, 1435–1441. doi: 10.4028/www.scientific.net/AMM.631-632.1435

Appendix A. Criteria Used for Assessing the Essays Written by the Participants

These criteria are based on Briguglio (2007, p. 19), Coffin et al. (2003, p. 77–80), Hamp-Lyons and Heasley (2006, p. 206), and Knoch (2011, p. 91).

Criteria	Feature	Score
1. Knowledge	Ability to show knowledge and understanding of an area of linguistics	_____ (give a score from 1 to 14)
2. Argument	Ability to present and pursue an argument	_____ (give a score from 1 to 14)
3. Critical Thinking	Ability to discuss and evaluate alternative explanations and arguments	_____ (give a score from 1 to 14)
4. Clarity	Ability to express himself/herself clearly in the essay	_____ (give a score from 1 to 14)
5. Organisation	Ability to organise the essay coherently and cohesively	_____ (give a score from 1 to 14)
6. Academic style	Ability to write the essay in an academic way	_____ (give a score from 1 to 10)
7. Accuracy	Ability to show sophisticated use of English language that is free of errors	_____ (give a score from 1 to 10)
8. Complexity	Ability to show a level of lexical variation and density	_____ (give a score from 1 to 10)
	Total score:	_____

Appendix B. An Example Worksheet Used by the Experimental Group

Introducing collocation in the BNCweb:

1. What is collocation?
   Group activity: Search online, find out its definitions and examples, and present in class
   Hint: Collocation as a concept versus collocation as a methodology
2. The collocation function of the BNCweb: The case of *tea*
   a. Analyse the first page of concordances of *tea* in the BNCweb.
      What are its collocates and why? What criteria do you use?
   b. Use the collocation function.
      What features have you noticed from the collocation list? (Hint: grammatical words)
      Change the span from the default (-3 to +3) to (-1 to -1). What is different in the collocation list?
      Now give restrictions to collocates: any adjective. What are the most frequent words in the list?
3. Practice with the collocation function: The case of *happen*
   Group activity: Find out about its collocates and features of these collocates

Appendix C. Brief Overview of the Course Contents

Week	Course for the Experimental Group (Using a corpus-Assisted Method)	Course for the Control Group (Using a Traditional or Rule-Based Method)
1	Session 1: Introduction Outline for the course Introduction to the BNC and the COCA	Session 1: Introduction Outline for the course Main areas of linguistics
2–3	Session 2: Words Definitions and features of a word Corpus-based inductive activities (e.g., discuss definitions of a word from a corpus-linguistic perspective, observe features of a word using the BNC)	Session 2: Words Definitions and features of a word Teacher-guided activities (e.g., discuss definitions of a word using dictionaries, understand potential features exhibited by a word)
4-5	Session 3: Words and Meaning Part of speech; polysemy Corpus-based inductive activities: analyse features of polysemous words with the BNC: word usage and meaning (using *time*, *light*, and *like* as examples)	Session 3: Words and Meaning Part of speech; polysemy Teacher-guided activities: analyse features of polysemous words with dictionaries (using *time*, *light*, and *like* as examples)
6	Session 4: Words and Grammar Lemma; morpheme and affix Corpus-based inductive activities: observe the relationship between word form and usage with the BNC (using *eye* and *eyes* as examples)	Session 4: Words and Grammar Lexeme; morpheme and affix Teacher-guided activities: discuss the relationship between word form and usage with dictionaries (using *eye* and *eyes* as examples)

7–8	Session 5: Phraseology Criteria for phraseology; various types Corpus-based inductive activities: analyse the use of phraseological items with the BNC (using *of course*, *at the same time*, and *it* + v-link + adjective + *that* as examples)	Session 5: Phraseology Criteria for phraseology; various types Activities: the teacher illustrates the use of phraseological items (using *of course*, *at the same time*, and *it* + v-link + adjective + *that* as examples)
9–11	Session 6: Collocation Definition; types of collocation (including verb-preposition collocations) Corpus-based inductive activities: analyse the use of collocations with the BNC (using *heavy rain*, *learn from*, and *cause* + noun as examples)	Session 6: Collocation Definition; types of collocation (including verb-prep. collocations) Activities: the teacher illustrates the use of collocations (using *heavy rain*, *learn from*, and *cause* + noun as examples)
12–13	Session 7: Discourse Spoken versus written discourse; features of academic discourse (using the BNC and the COCA) Corpus-based inductive activities (e.g., investigate the features exhibited by university essays and job interviews with the AntConc)	Session 7: Discourse Spoken versus written discourse; features of academic discourse (illustration by the teacher) Teacher-guided activities (e.g., discuss the features exhibited by university essays and job interviews)
14–15	Session 8: Genre Analysis Keywords of a discourse Corpus-based inductive activities (e.g., investigate the linguistic features exhibited by a novel, a legal contract, and business emails with the AntConc)	Session 8: Genre Analysis Keywords of a discourse Teacher-guided activities (e.g., discuss the linguistic features exhibited by a novel, a legal contract, and business emails)

Note. Concerning the activities in each session, both groups focused on identical subjects—for example the same words, phrases, collocations, and texts. Three verb-preposition collocations were explicitly shown or taught to the participants: pay for, stop from, and learn from. The first two collocations were used as examples when introducing verb-preposition collocation and the third one was used in class activities.

About the Author

Shuangling Li is an associate professor in the School of Foreign Languages for Business at the Southwestern University of Finance and Economics, China. She obtained her PhD in corpus linguistics at the University of Birmingham. Her main research interests centre on the study of phraseology, corpus linguistics, and English language teaching.

E-mail: lisl@swufe.edu.cn