ATU
PRESS

The Primary and Secondary Effects of Data-Driven Learning on High- and Low-Intermediate Learners' Knowledge of Collocations

Afsaneh Saeedakhtar* (iD)

Assistant Professor in TEFL, University of Mohaghegh Ardabili, Ardabil, Iran

Abstract

The present study scrutinized the role of data-driven learning in recognizing and producing collocations by high- and low-intermediate learners of English. Moreover, the influence of secondary effect (or transfer of training) on learning secondary collocations (collocations not provided in concordancing but embedded implicitly in tasks) was examined. The learners' attitudes towards the effect of concordancing on learning collocations and their attitude change over time were also elicited through a questionnaire. A total of 40 Iranian learners were randomly divided into an experimental and a control group. Each group was further subdivided into high- and low-intermediate learners. The experimental group received a 10-session treatment in which they had access to concordancing to perform the paraphrasing tasks. The control group was taught the same collocations in a traditional explicit way. Results revealed that the high-intermediate learners benefitted from concordancing in both primary and secondary learning of collocations more than the low-intermediate learners. Both high- and low-intermediate groups appreciated the positive role of concordancing in learning collocations and stated that their attitudes towards concordancing changed positively over time.

---

*Corresponding author's email: a.saeedakhtar@ut.ac.ir

INTRODUCTION

With the advent of digital literacy, interest emerged in using technological tools for enhancing long-lasting language learning. To date, numerous studies have employed interactive Web 2.0 synchronous tools such as chat, Skype, video-conferencing, etc. (e.g., Alameen, 2011; Wang, Zou, Wang, & Xing, 2013), asynchronous tools such as email, blog (e.g., Fathi & Nourzadeh, 2019), wiki, Viber, mobile-assisted language learning (e.g., Bikowski & Casal, 2018; Godwin-Jones, 2011; Nami, 2020), as well as Web 1.0 tools including corpus and concordancing (e.g., Daskalovska, 2015; Rezaee, Marefat, Saeedakhtar, 2015; Saeedakhtar, 2013; Saeedakhtar, Bagerin, Abdi, 2020; Vyatkina, 2016a, 2016b; H. Yoon & Jo, 2014) to teach different components of language and have reported significant achievements. Concordancer–a computer retrieval program, used either online or offline, to provide learners with numerous authentic materials embedded with the grammar, vocabulary, and collocations–is an instance of technology-enhanced tools for teaching collocations (e.g., Rezaee et al., 2015; Chang & Sun, 2009), i.e., the co-occurrence of words such as *cast doubt* and *draw attention*.

According to Daskalovska (2015), advanced learners with the ability to self-regulate their learning can benefit more from data-driven learning (DDL), put forward by Johns (1990) as "vast databases of the English text (corpora) with software programs called concordancers, which isolate common patterns in authentic language samples" (Hadley, 2002, p. 106). Given that the high-intermediate learners benefit more from DDL, studies mostly have exposed intermediate or advanced learners to either hands-on, i.e., online (e.g., Geluso & Yamaguchi, 2014; C. Yoon, 2016) or hands-off, i.e., paper-based DDL (e.g., Huang, 2014; Smart, 2014). However, a limited number of studies have used hands-on (e.g., Johns, Lee, & Wang, 2008; Mirzaee, Rahimi Domakani, & Rahimi, 2015) and hands-off DDL (e.g., Vyatkina, 2016b) for low proficiency learners and have reported controversial findings: Some studies (e.g., Vyatkina, 2016a; C. Yoon, 2016)

recommended hands-on DDL for only high proficiency learners while a few studies (e.g., Saeedakhtar et al., 2020; Johns et al., 2008; Mirzaei et al., 2015) concluded that low proficiency learners also benefit from hands-on DDL. Previous studies have not compared the achievement of high- and low proficiency learners when they are exposed to hands-on DDL. Further research is necessary to shed more light on the role of hands-on DDL in low proficiency learners' L2 achievement. Moreover, the missing link among previous studies is that they have measured the "primary effects" (Benati & Lee, 2008) of training and have turned a blind eye to the "secondary effects" (White & DeMill, 2013), also known as the transfer of training–transferring the acquired knowledge and skills to similar situations (Larsen-Freeman, 2013). The primary effects in the present study refer to exposing learners to the target collocations in concordancing and measuring learning those collocations while the secondary effects are operationalized as learning the secondary collocations which have been embedded only implicitly in paper-based input (i.e., paraphrasing tasks) to check the transfer of training. Put differently, the present study intends to measure the extent to which teaching collocations in concordancing can raise learners' consciousness to spot secondary collocations they come across while reading the paraphrasing tasks.

Another gap in the literature is that most studies (e.g., Geluso & Yamaguchi, 2014; Rezaee et al., 2015) have only investigated the learners' perceptions of DDL at the end of the treatment and have not taken into account their probable attitude change over time. A few studies (e.g., Boulton, 2012; Vyatkina, 2016a), however, surveyed learners' 'receptivity' to DDL over time and came up with different results. Boulton (2012) found that 'receptivity' to DDL decreased over time, though not significantly, while Vyatkina (2016a) concluded that learners' inclinations to use DDL arose over time though there was no significant difference between the pre-course and post-course perception questionnaire. The present study intends to address the above-mentioned gaps by comparing the effect of hands-on DDL on high- and low-intermediate learners' collocations. Furthermore, it

examines the primary and secondary effects of concordancing on learning targets and secondary collocations. It also attempts to elicit learners' attitudes towards the role of concordancing in learning L2 collocations in general and their attitude change towards concordancing over time by giving the questionnaire in the first and the last session.

LITERATURE REVIEW

Collocations

The term collocation is defined by Lewis (1997) as "the readily observable phenomenon whereby certain words co-occur in natural text with greater than random frequency" (p. 8). Grabe and Stoller (2002) argue that "L2 learners often read materials with glosses for more difficult terms" (p. 58). Since collocations are the component of language which are comprehended easily by L2 learners in comparison with idioms or phrasal verbs, learners may pay less attention to them and only invest in learning unfamiliar words or idioms.

Previous studies have investigated L2 learners' receptive knowledge (e.g., Barfield, 2007) and production performance (e.g., Rezaee et al., 2015; Saeedakhtar, 2013, Vyatkina, 2016a) of collocations. Some other studies (e.g., Altenberg & Granger, 2001) have also used learner corpora to compare L2 learners' production of collocations and that of native speakers. Results of all these studies have revealed that L2 learners overuse (e.g., Nesselhauf, 2005) or underuse (e.g., Granger, 1998; Howarth, 1998) L2 collocations. Such a finding urged L2 researchers to invest in incorporating collocations into classrooms. One way through which such incorporation can happen is using DDL.

DDL

Johns (1990) borrowed the term DDL as an attempt to "allow learners to inductively discover language structures and patterns through interacting

with concordancing software or with concordance-based instructional materials" (Smart, 2014, p. 184). In DDL, learners behave like "language detectives or researchers" (Geluso & Yamaguchi, 2014, p. 227) who delve into authentic materials to identify L2 patterns by themselves.

Concordancing, mostly associated with DDL (C. Yoon, 2016), is one of the most popular reference tools for learning collocations. It provides opportunities for learners to search for collocation patterns in an authentic context (e.g., Daskalovska, 2015; Rezaee et al., 2015; Wu, Witten, & Franken, 2010). Concordancing "sort[s] and display[s] language data in ways that allow[s] users to discover patterns, test hypotheses, and figure out solutions to language problems at hand" (C. Yoon, 2016, p. 210). As a linguistic tool, concordancing enables learners to overcome their linguistic problems (H. Yoon & Jo, 2014) and as a cognitive tool, it helps learners develop their thinking skills (e.g., O'Sullivan, 2007) and strategies for language learning, and reinforces inductive learning (H. Yoon & Jo, 2014).

The effect of DDL has been explored on different components of language including L2 writing (e.g., Kennedy & Miceli, 2010; Park, 2010; C. Yoon, 2016; H. Yoon & Jo, 2014), academic English writing (e.g., Chang, 2014; Flowerdew, 2015), L2 collocations (e.g., Daskalovska, 2015; Rezaee et al., 2015, Saeedakhtar et al., 2020; Vyatkina, 2016a), L2 linguistic items (e.g., Cotos, 2014; Lin & Lee, 2019), L2 vocabulary (e.g., Karras, 2016; Mirzaei et al., 2015), oral L2 proficiency development (e.g., Huang & Hung, 2010), L2 pronunciation (Qian, Chukharev-Hudilainen, & Levis, 2018), error correction (e.g., Crosthwaite, 2017; Luo & Liao, 2015), and teacher education (e.g., Breyer, 2009; Lenko-Szymanska, 2014). Most of these studies have asserted the positive role of DDL in fostering L2 learning.

DDL can be rendered in two different ways, hands-on and hands-off. Fostering life-long learning, increasing learners' autonomy, and providing learners with authentic material (Boulton, 2012) are among the merits of hands-on DDL. However, there are some drawbacks such as wasting the class time due to the lack of adequate computer literacy and back-up and

requiring adequate training on the part of learners and teachers. Hands-off DDL, on the other hand, overcomes such problems by tailoring the material to the needs of the learners. It can be a good choice for low-level learners (Boulton, 2012) with a minimal amount of training and scaffolding.

To overcome the shortcomings of the hands-on DDL, especially for low-level learners, teachers can resort to two options. Incorporating the teachers' guidance and feedback, technically speaking, "guided induction" (e.g., Flowerdew, 2009; Huang, 2014; Smart, 2014) or "dialogue" (Webb, Jones, Barker, & Schaik, 2004) is one solution. Guided induction "provides a structured, scaffolded framework for inductive learning" (Smart, 2014, p. 187) and dialogue refers to "instructions from teachers, guidance, and feedback" (Webb et al., 2004). The next choice is exposing low-level learners to "prepared, preselected … less autonomous" hands-off DDL (Smart, 2014).

Transfer of Training: Secondary Effect

Transfer of training refers to applying and extending what we have learned in different contexts to similar/new situations (Larsen-Freeman, 2013). Despite the fact that transfer plays an important role in fostering language learning, learners are not usually aware that they are expected to transfer the knowledge achieved to similar situations (Lightner, Benander, & Kramer, 2008). Even learners do not receive the training as to how to benefit from transfer strategies (e.g., Billing, 2007). According to White and DeMil (2013), transfer of training effects can be expected on the structure which has not been addressed in the treatment but requires learners to be engaged in the same processing strategy. Transfer of training in workplace-transfer of skills from school to the workplace-(e.g., Chauhan, Ghosh, Rai, & Shukla, 2016; Govaerts, Kyndt, Vreye, & Dochy, 2017), the role of motivation in promoting the transfer of training (e.g., James, 2012; Perkins & Salomon, 2012), and the extent to which learners transfer knowledge from virtual environment to the real world (Wallet, Sauzeon, Rodrigres, &

N'Kaoua, 2009) have already been investigated.

Studies focused on the transfer of training have investigated the secondary effects on linguistic forms (e.g., Benati & Lee, 2008; Leeser & Demil, 2013; White & DeMil, 2013) or pronunciation (e.g., Qian et al., 2018) and transfer of training on collocations in DDL has remained as a less-attended area that invites further exploration.

The present study investigates secondary effects by measuring the extent to which learners benefit from implicit instruction as a result of the awareness that concordancing has raised. In other words, the underlying strategy of learning collocations in concordancing is inductive discovery learning (e.g., Vyatkina, 2016b). The present study attempts to explore whether the same underlying strategy can be resorted to while being exposed to implicit learning.

Empirical Research

A number of meta-analyses represented large effect sizes (e.g., Boulton & Cobb, 2017; Cobb & Boulton, 2015) and medium effect sizes (e.g., Lee, Warschauer, & Lee, 2019) for the superiority of DDL over traditional ways of teaching. Most studies have investigated the effect of hard hands-on DDL (Gabrielatos, 2005) in which learners benefit from online DDL by themselves without the help of their teacher, soft version of DDL (Mukherjee, 2006) in which learners receive the teacher's scaffolding (e.g., Rezaee et al., 2015; Saeedakhtar, 2013), and hands-off version (e.g., Daskalovska, 2015; Huang, 2014; Smart, 2014; Vyatkina, 2016b) where learners receive paper-based printout DDL on different components of language.

Rezaee et al. (2015) compared the role of hard and soft concordancing in recognizing and producing collocations by intermediate learners of English. Results indicated no significant differences between the two versions. Contrary to the previous studies in which the intermediate and advanced learners have been addressed, Vyatkina (2016b) compared the

influence of hands-off DDL with traditional instruction on learning German verb-preposition collocations by low proficiency learners. Results showed that both conditions were effective in improving lexical items and grammar. Vyatkina, however, did not investigate the performance of the learners in the long term to compare their lasting influence.

In the same vein, Saeedakhtar et al., (2020) recruited low-proficiency Iranian learners of English to compare the role of hands-on and hands-off DDL in learning verb-preposition collocations. To this purpose, 60 female pre-intermediate learners were randomly divided into two experimental groups (i.e., hands-on and hands-off DDL groups) and one control group. In a 10-session treatment, all learners were taught 66 verb-preposition collocations under three different conditions.

The hands-on group was exposed to the collocations on the screen of the computer. They searched for collocations in Ant-Conc software with the help of the teacher. The hands-off group received the same collocations in a different color within the same concordance lines on the paper-based corpus. The control group was taught the same collocations explicitly on the board. All groups were then required to do fill-in-the-blank activities to practice the target collocations. When the treatment came to its end, all groups took an immediate post-test. The experimental groups also were required to respond to an attitude questionnaire to elicit their opinions on DDL. Two weeks later, all groups took a delayed post-test. Results indicated that both experimental groups outperformed the control group significantly in learning verb-preposition collocations on the immediate post-test. However, only the hands-on group maintained the gain on the delayed post-test. Both hands-on and hands-off groups welcomed DDL as an interesting and exciting way for learning collocations.

The superiority of hands-on DDL over hands-off DDL was also verified by other recent studies. Sun and Hu (2020), for instance, compared the effect of hands-on and hands-off DDL on learning English hedges in academic writing. A total of 24 upper-intermediate learners were divided into an experimental and a control group. The experimental group had

access to two 2-hour-session hands-on DDL treatment (i.e., two online corpora) in the lab while the control group received hands-off DDL (i.e., teacher paper-and-pencil tasks adopted from corpus). Results showed the outperformance of the hands-on DDL with a small effect size.

Smith (2020) also compared the role of hands-on and hands-off DDL in vocabulary learning. Around 94 English for Academic Purposes (EAP) learners majoring in Accounting and Finance were divided into an experimental group who had access to hands-on DDL (i.e., DIY corpora) and a control group who had access to hands-off DDL. Results demonstrated that the hands-on group did better than the hands-off group in learning specialist vocabulary items. However, both groups appreciated the role of DDL in vocabulary learning.

A look into literature reveals the gap of previous studies in investigating the moderating role of the level of proficiency in the effectiveness of hands-on DDL. To account for some of the shortcomings of the previous studies, the present study intends to explore the role of hands-on DDL in improving collocations of both high- and low-intermediate learners in the short and long term. It also compares the extent to which high- and low-intermediate learners transfer their collocational awareness to contexts other than concordancing. Furthermore, it elicits the experimental groups' perception of concordancing and examines their perception change over time. The following research questions are formulated accordingly.

1. Would high-intermediate learners benefit from concordancing more than low-intermediate learners to enhance their recognition/production of collocations?
2. Would high-intermediate learners spot secondary collocations in the input and recognize/produce them more than low-intermediate learners as a result of concordancing training?
3. What are high- and low-intermediate learners' attitudes towards using concordancing to learn collocations? Did their perceptions change over time?

METHODOLOGY

Participants

A total of 60 (22 male and 38 female) Iranian University learners of English were recruited based on convenient sampling. They aged 22 to 28 and were bilingual in Azari Turkish (L1) and Persian (L2). They voluntarily participated in an extra-curricular class which was held at the university to teach collocations through a technology-enhanced tool, i.e., concordancing. As they expressed in the questionnaire, none of them were already familiar with corpora and concordancing. All learners took the Cambridge Preliminary English Test (PET). According to their scores on PET, they were randomly divided into high- and low-intermediate learners. They were ranked on the basis of their scores in the proficiency test. Then, 33% (i.e., +1 SD) of high achievers who scored more than the cut-off score were taken as high-intermediate learners, and 33% (i.e., -1 SD) of low achievers as the low-intermediate learners. Out of 60 participants, 20 (9 male and 11 female) were classified as high-intermediate, 20 (7 males and 13 females) as low-intermediate learners, and 20 who were intermediate were excluded.

Both high- and low-intermediate learners were sub-divided randomly into an experimental and a control group, including 10 high- and 10 low-intermediate learners in each.

Instruments

Paraphrasing Tasks

For the present study, 10 paraphrasing tasks designed by Saeedakhtar (2013) on different topics (*Diabetes, Biography of Thomas Hardy, All about dolphins, Christmas customers, the Tooth fairy, Dangers of fast food, American wedding customs, American table manners, Dangers of painkillers,* and *Benefits of dark chocolate*) were used as the treatment tasks in which 10 secondary collocations and only the nodes of 125 target collocations were embedded. Three more paraphrasing tasks were designed

to serve as the production task of the pretest, immediate, and delayed post-tests. Since attempts were made to choose genuine materials for the tasks, 15 random texts on each topic were retrieved from Google search-engine, simplified, and summarized in one page. If the texts included some target collocations, their collocate parts (the words that come on the left or right side of the node) were deleted and their node parts (the head of collocations) were underlined followed by the type of the target collocations and their Persian equivalent in parentheses. Learners were required to paraphrase those sentences by incorporating a collocate. For example, learners were required to paraphrase the following sentence by including *strong tendency* and *reach maturity* (*strong* and *reach* as the collocates and *tendency* and *maturity* as the nodes). For example:

- *Their* <u>tendency</u> *(adj-n* تمایل زیاد*) to enlarge their appetite was limitless.*
- *Dolphins* <u>mature</u> *(maturity: v-n* به بلوغ رسیدن *) at the age of 8-15 years.*

The English word for word translations of the collocations تمایل زیاد *(tamayole ziad), and* به بلوغ رسیدن *(be bolog residan) are *much/high tendency and *arrive to maturity.*

Response:
- *Their* **strong tendency** *to enlarge their appetite was limitless.*
- *Dolphins* **reach maturity** *at the age of 8-15 years.*

Learners were asked to paraphrase the one-page text in which 15 nodes were underlined by including an appropriate collocate which was the English equivalent of the Persian collocation (i.e., *strong tendency, reach maturity*) and embedding 15 lexical and grammatical collocations first on their own, without having access to concordancing. In the next step, they were required to refer to concordancing to check the accuracy of their collocations. By so doing, it was assumed that learners' motivation would be triggered to look for the collocations enthusiastically in concordancing.

However, for accomplishing the paraphrasing tasks on the pretest and post-tests, they had no access to concordancing. The tasks were piloted on 11 high-intermediate and nine low intermediate learners of English. Results of the pilot study demonstrated that even high-intermediate learners fell short of writing correct collocates for the given nodes in the absence of concordancing, although all of them stated that the text of paraphrasing tasks was easy to comprehend.

In addition to the target collocations, 10 secondary collocations were embedded in the paraphrasing tasks, without applying any input-enhancement techniques (e.g., using boldface, italics, underlining, etc.), to measure the extent to which HCG and LCG learn those secondary collocations. For example, in session 10, learners were given the following paraphrasing task to use the target collocation *packed with*; meanwhile, the recognition and production of the secondary verb-preposition collocation *contribute to,* embedded in the text, were measured as the secondary effect on the pretest and post-tests.

> ... Studies show that eating dark chocolate, may contribute to improved cardiovascular health. <u>Full of</u> (*packed: adj-pre.* پراز) natural antioxidants, dark chocolate and cocoa ...

Target Collocations

A total of 125 incongruent collocations, collocations that are different in L1 and L2, (Appendix A) were selected to be presented to the learners in concordancing to measure their primary effect. Moreover, 10 secondary incongruent collocations (five lexical and five grammatical collocations, Appendix A) were embedded in paraphrasing tasks to measure their secondary effect. The nodes were selected randomly from among frequency bands 5, 4, 3, and 2 in the Cobuild dictionary (2003) but the collocates were chosen according to their mutual information (MI) presented in Brigham Young University-British National Corpus (BYU-BNC). MI refers to the

strength of association between the node and collocate (Church & Hanks, 1990) and ranges from 0-17 (Siyanova & Schmitt, 2008); the index 0 means that the node and collocate have been used accidentally while the index 17 shows a strong association between the node and collocate. For example, the collocation *commit crime* enjoys a high MI (i.e., 11.37) while the collocation *see crime* enjoys a lower MI (i.e., 3.36).

In the present study, the collocates that possessed the highest MI in BYU-BNC available at: https://www.english corpora.org/bnc/old/?b=x1&c =bnc&q=17774111&q1=17774113 were selected by setting the *Minimum* choice at *MUT INFO* option (Figure 1). The reason behind extracting MI in this study was exposing learners to collocations that are of importance in real-life situations so that they could transfer the learned collocations to real-life activities.
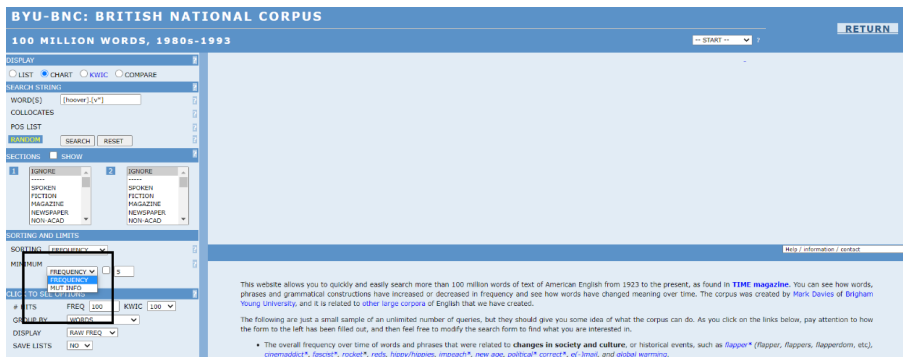

Figure 1: The Screenshot of the Home Page of BYU-BNC

The target collocations included 61 lexical and 64 grammatical collocations. The lexical collocations included 37 verb-noun (e.g., *make a recommendation*), 15 adjective-noun (e.g., *great opportunity*), and nine noun-verb collocations (e.g., *specialty lay*). The grammatical collocations included 16 verb-preposition (e.g., *conserve for*), 15 preposition-noun (e.g., *with determination*), 18 adjective-preposition (e.g., *nervous about*), and 15 noun-preposition (e.g., *reputation for*) collocations. Additionally, 10 secondary collocations other than the target ones were embedded in the text

of paraphrasing tasks to check their secondary effect. They included two verb-noun, three noun-preposition, two adjective-noun, one noun-verb, one verb-preposition, and one adjective-preposition collocations.

Pretest

The researcher-made pretest included a receptive test and a production task. The receptive test is intended to gauge learners' recognition of lexical and grammatical collocations. It included 20 items of 2−3-line sentences in which five collocations were underlined. Learners were required to identify the choice in which the collocation was incorrect. For example, in the following item, learners were required to identify *deep cold* as a miscollocation.

- She did well in the exam, though she suffered from a deep cold. Now some of her
          A   B   C   D
  friends are truly jealous of her.
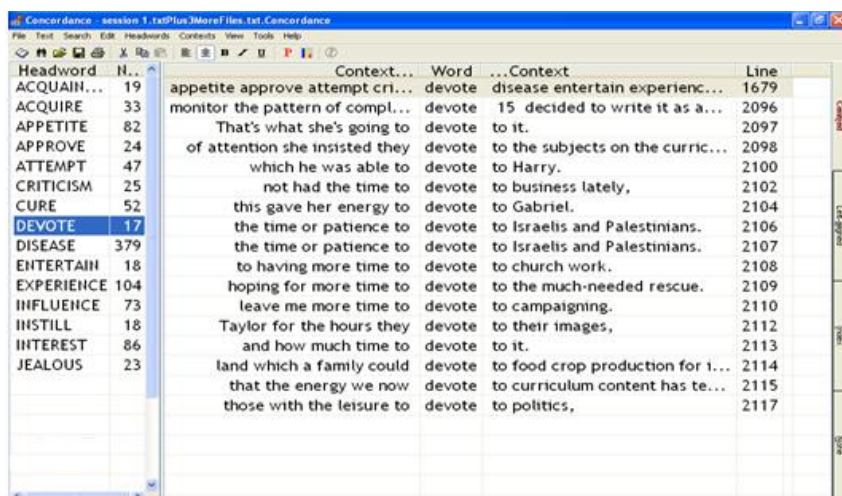   E

   The recognition of 100 random collocations (20 items, five in each) was measured in the receptive test. From the pool of the target collocations, 75 collocations (i.e., 15 items) were selected randomly and the remaining 25 collocations which included non-target collocations were not scored. The Cronbach's alpha reliability index for the receptive test was 0.79.
   The production task, i.e., paraphrasing task (see section 3.2.1), included 24 collocations, of which six were non-target collocations and were not scored. The learners were instructed to paraphrase the text embedding the lexical and grammatical collocations. For ensuring the validity of the receptive and productive pretests, they were piloted prior to the study on a group of 11 high-intermediate and nine low-intermediate learners. The pretests were also evaluated by two experts in TEFL. Following the results of the pilot study and the recommendations of the

experts, a few modifications were made. No time limitation was set for performing the pretest. The pretest was scored by counting the number of correct answers. The maximum score for the receptive and production sections was 15 and 18, respectively.

Concordancing

Concordance 3.3 (2009) at *www.concordancesoftware.co.uk*, not available anymore, was used as the reference tool in the present study. The nodes of 150 collocations along with examples including the collocates for the given nodes were the raw data of concordancing extracted from BYU-BNC. Each session learners had access to only 15 nodes and 17 examples in authentic context for each node in concordancing so that learners would not be overwhelmed with a large pool of information. As shown in Figure 2, the list of the nodes is placed under *Headword*. Learners were instructed to click on those words to have access to the concordance lines, i.e., examples. For the node *devote*, for instance, 17 concordance lines were available in concordancing.

Figure 2: The screenshot of concordancing representing the nodes and concordancing lines

When learners clicked on the concordance lines, another window popped out which showed the collocation in its authentic context (Figure 3).
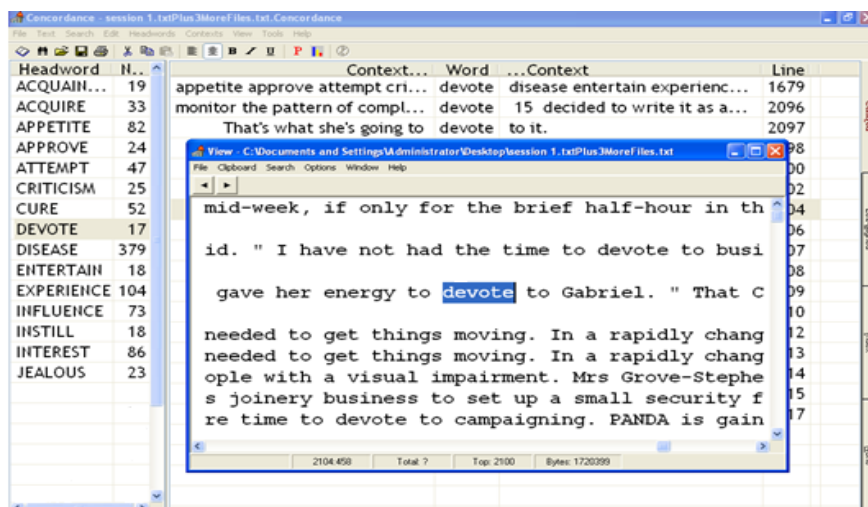
Figure 3: The screenshot of collocations in context

Questionnaire

An eight-item Likert scale researcher-designed questionnaire was given to the experimental group to elicit their opinion about the influence of concordancing on learning lexical and grammatical collocations, their preference to use concordancing on their own or with the teachers' or peer's support, and their attitude change towards concordancing after the 10-session treatment. The same questionnaire was given once in the first session and once in the last session of the treatment to pinpoint any change in their attitude over time. For validity concerns, the questionnaire was also piloted on 11 high-intermediate and nine low-intermediate learners and evaluated by the two TEFL experts.

Procedures

Forty learners took part in an extra-curricular university class to learn

collocations through concordancing. The classes were held twice a week for 13 sessions. Immediately after taking the proficiency test, learners were given the pretest. The production part of the pretest preceded the receptive one to remove any test effect. Based on the results of the proficiency test and pretest, learners were then randomly divided into an experimental group and a control group. Each group was further subdivided into high- and low-intermediate learners.

In session 2, each member of the experimental group individually received a 5-minute instruction in L1 as to how to use concordancing. Attempts were made to ensure that all learners could search for the collocates in concordancing independently. The concordancer was installed on computers in the laboratory. Each session, every member of the experimental group was required to perform the paraphrasing task by themselves, without having access to concordancing, in order to trigger their motivation to compare their production with that of the target one and notice the gap (Swain, 1985); then they were asked to use concordancing individually to fill in the identified gap. They did not receive any teacher or peer support in searching for collocates in concordancing and performing the paraphrasing tasks. There was no time limitation and learners felt free to complete the task at their own pace. At the end of session 2, the experimental groups responded to the questionnaire to elicit their opinions concerning incorporating concordancing into language learning classrooms to learn collocations. The treatment lasted for 10 consecutive sessions.

The control group was required to accomplish the paraphrasing tasks without having access to concordancing. They attended a 10-session speaking class, which was almost the same as their normal communicative classes. Each session learners talked about their (dis)agreement, idea, experience, etc. with the teacher or peers about the main topic of paraphrasing tasks for one hour. The author taught them new vocabulary items, expressions, and collocations explicitly. At the end of the class, learners completed the treatment tasks in half an hour. In session 12, all learners took an immediate post-test−a parallel receptive test and production

task. After the immediate post-test, learners responded to the questionnaire to investigate their attitude change over time. One month later, another parallel test was given to them as the delayed post-test. The post-tests were piloted before the participants took them.

Data Analysis

To analyze the data, three separate one-way ANOVAs were run on the results of the proficiency test, receptive pretest, and production pretest. Then, three separate 2×2 ANCOVAS were run on the results of the receptive post-tests, production post-tests, and transfer of training. The frequency of the learners' responses to the questionnaire was calculated manually.

RESULTS

The first research question addressed the role of concordancing in the recognition/production of collocations by the HCG and LCG. Prior to the study, a one-way ANOVA was run for the results of the proficiency test to ensure the homogeneity of the experimental and control groups. Results showed no statistically significant differences between high-concordancing group (HCG) and high-control group, $F = 2.737$, $p = .192$, and the low-concordancing group (LCG) and low-control group, $F = 2.998$, $p = 1.000$. Two separate one-way ANOVAs were run for the results of the receptive and production pretest. Since the results of the groups' receptive ($F(3, 36) = 2.54$, $p = .60$) and production performance ($F(3, 36) = 1.10$, $p = .35$) yielded no statistically significant difference, they were removed from further analysis. The normality test was run on the results of the proficiency test. The results of the Shapiro-Wilk test of the normality test (Table 1) showed no significant p-value for the participating groups.

Table 1: The results of the Test of Normality

Group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	statistic	df	Sig.
Experimental Group	.244	18	.065	.909	18	.209
Control Group	.169	18	.200	.929	18	.406

In order to adjust for the pre-existing differences among the high- and low-intermediate learners prior to the study, two 2×2 between-groups ANCOVAs were performed for the recognition/production performance of learners. Table 2 shows the descriptive statistics for the recognition of the experimental and control group over time. The comparison of the means of the groups indicated that the HCG outperformed the other participating groups both on the immediate and delayed post-tests in recognizing collocations.

Table 2: Descriptive Statistics for the Recognition of Collocations

Group	Proficiency	N	Pretest		Immediate Post-test		Delayed Post-test	
			M	SD	M	SD	M	SD
Concordancing	High	10	2.73	2.03	9.80	3.88	8.70	2.35
	Low	10	2.33	1.18	4.60	2.91	5.40	2.54
Control	High	10	2.10	1.53	4.50	1.77	5.60	2.17
	Low	10	2.10	1.34	3.30	1.33	3.10	1.19

Results yielded by Wilks' Lambda = .023 indicated that the interaction effect for time (immediate and delayed post-tests) and concordancing was not statistically significant, $F(3, 36) = .005$, $p = .94$. Therefore, learners' performance was similar both on the immediate and delayed post-tests. The main effect for the group was statistically significant, $F(3, 36) = 9.05$, $p = .000$, which indicates that the performance of groups was different from each other. To locate the area of difference, a Scheffe post-hoc analysis was run. Results showed that the experimental group outperformed the control group in recognizing collocations, $p = .001$; however, the difference

between the HCG and the LCG was statistically significant too, $p = .002$. but there was no significant difference between the LCG and high-non-concordancing groups, $p = 1.00$

To address the production performance of the HCG and LCG, another 2×2 between-groups ANCOVA was run. Table 3 represents the descriptive statistics for the production performance of collocations. Results indicated that the HCG performed better than the other groups.

Table 3: Descriptive Statistics for the Production Performance of Collocations

Group	Proficiency	N	Pretest		Immediate Post-test		Delayed Post-test	
			M	SD	M	SD	M	SD
Concordancing	High	10	5.00	1.34	14.70	4.69	15.10	4.84
	Low	10	2.13	1.32	6.50	5.19	5.10	2.99
Control	High	10	5.44	1.50	4.20	1.22	5.20	2.20
	Low	10	2.10	1.34	1.20	1.22	1.20	1.03

Results of Wilks' Lambda = .239 demonstrated that the interaction effect for time (immediate and delayed post-tests) and concordancing did not reach statistical significance, $F(3, 36) = 1.36$, $p = .25$. As in the recognition of collocations, in the production performance, all learners performed the immediate and delayed post-tests similarly. However, the main effect for group was statistically significant, $F(3, 36) = 29.5$, $p = .000$. Another Scheffe post hoc analysis was run to identify the group(s) that produced collocations better than the other ones. Findings showed that the HCG outperformed the control and LCG significantly, $p = .000$. Although the HCG performed better than the high non-concordancing group only on the immediate post-test, such an outperformance did not reach statistical significance, $p = .78$

The second research question attempted to investigate the extent to which the HCG and LCG transferred their collocational awareness to notice other collocations in the input and produce them accurately. Since the proficiency level was suspected to be a source of variation from the very

beginning, another ANCOVA was run to control for the initial differences between the high- and low-intermediate groups. A look at the descriptive statistics for the production performance of collocations (Table 4) indicates that the high-concordancing group achieved the highest mean among the other groups.

Table 4: Descriptive Statistics for Transfer of Training

Group	Proficiency	N	Immediate Post-test		Delayed Post-test	
			Mean	SD	Mean	SD
Concordancing	High	10	4.0	.51	7.0	.67
	Low	10	.00	.00	.00	.00
Control	High	10	3.1	.31	4.3	.52
	Low	10	.10	.31	.40	.51

The results obtained showed that compared with no-concordancing condition, concordancing did not appear to influence transfer of training significantly, $F(3, 36) = .071$, $p = .791$, $\eta^2 = .007$. Interestingly, proficiency as a moderator variable served as a source of variance in the scores obtained as indices of transfer of training. The high-intermediate group which included both +concordancing and –concordancing subgroups outperformed the low proficiency group including both +concordancing and –concordancing, $F(3, 36) = 4.28$, $p = .046$, $\eta^2 = 1.09$. Technically speaking, the interaction effect of concordancing and proficiency was statistically significant with large effect size, $F(3, 36) = 9.15$, $p = .005$, $\eta^2 = .207$.

The last research question intended to elicit learners' attitudes towards using concordancing to learn collocations to estimate their perception change over time (Table 5).

Table 5: The Frequency of HCG and LCG's Responses to the Questionnaire

Questionnaire Items	The First Session		The Last Session	
	HCG Responses (%)	LCG Responses (%)	HCG Responses (%)	LCG Responses (%)
	Yes	Yes	Yes	Yes
1. Do you take concordancing as an effective way of learning collocations?	6 (6%)	4 (40%)	10 (100%)	10 (100%)
2. Do you like to search for the collocations in concordancing alone?	5 (50%)	0	7 (70%)	3 (30%)
3. Do you need any help provided by the teacher in searching for the collocations in concordancing?	4 (40%)	5 (50%)	0	3 (30%)
4. Do you need any help provided by peers in searching for the collocations in concordancing?	6 (60%)	8 (80%)	3 (30%)	4 (40%)
5. If you prefer peer support, do you like them to be at your level of proficiency?	5 (50%)	7 (70%)	3 (30%)	5 (50%)
6. If you prefer peer support, do you like them to be higher than your level of proficiency?	9 (90%)	5 (50%)	10 (100%)	5 (50%)
7. Are you eager enough to take part in other concordancing classes to learn collocations?	10 (100%)	10 (100%)	10 (100%)	10 (100%)
8. Is using concordancing difficult and confusing to you?	8 (80%)	9 (90%)	1 (10%)	2 (20%)

Around 60% of HCG and 40% of LCG took concordancing as an effective way to learn L2 collocations in the first session while these amounts reached 100% in the last session for both groups. About 50% of the HCG and none

of the LCG liked to search for the collocations alone, without any support from the teacher or peers in the first session while it changed to 70% and 30% in the last session, respectively. Representative oral opinions of both groups presented in the first session are given below:

> A learner of HCG:
> *The explanation given by the teacher on how to use concordancing in the first session was enough. The software was user-friendly and I could manage it by myself.*

> A learner of LCG:
> *Since it was our first experience to work on concordancing, I think the teachers' help was necessary. Sometimes I was not sure about the correct collocates. I liked to ask the teacher to make sure.*

Forty percent of the HCG and 50% of the LCG wished to receive teacher scaffolding in benefitting from concordancing in the first session while it changed into 0 and 30% in the last session, respectively. About 60% of HCG and 80% of LCG liked to get peer scaffolding when searching for collocations in the first session while it decreased to 30% and 40%, respectively, in the last session. In the first session, 50% of HCG and 70% of LCG preferred a peer at their own level of proficiency, while 90% of HCG and 50% of LCG welcome more knowledgeable peers. In the last session, their desire for the peers as knowledgeable as themselves decreased to 30% and 50%, respectively and their desire for more knowledgeable peers changed into 100% for the HCG only and there was no change in the percentage of the LCG. All learners of the HCG and LCG liked to take part in such programs in the future both in the first and last sessions. Around 80% of the HCG and 90% of LCG stated that in the first session concordancing was a very time-consuming and difficult task for them to search for the collocates in concordancing by themselves, but in the last session, 10% of HCG and 20% of LCG reported that with the passage of

time they spent less time and energy for benefitting from concordancing and searching for the collocates in concordancing was not that much demanding.

DISCUSSION

The primary objective of the present study was to compare the extent to which high- and low-intermediate learners benefited from hands-on DDL (i.e., concordancing) to learn collocations. In line with previous studies (e.g., Daskalovska, 2015; Geluso & Yamaguchi, 2014; Yoon, 2016), results of the present study demonstrated that the high proficiency learners benefitted from hands-on DDL more than the low proficiency learners. Since high proficiency learners have already mastered the threshold level to take advantage of self-regulated learning, without any "guided induction" (Smart, 2014), they used hands-on DDL easily and improved their receptive and productive knowledge of collocations. Therefore, the level of proficiency might be taken as a moderator variable that can influence the use of hands-on DDL.

Results are also in line with previous studies (Vyatkina, 2016b) which have reported that low proficiency learners may feel confused when working on hands-on DDL. Gremmo and Riley (1995) rightly argued that "when learners in 'high-tech' resource centers are not trained to become competent autonomous learners, the centers risk the same fate as language laboratories suffered decades ago" (p. 160). The LCG might need some degrees of training before working on hands-on DDL. However, results are contrary to some other studies which concluded that low proficiency learners benefitted from hands-on DDL to improve their knowledge of collocations (e.g., Saeedakhtar et al., 2020) and vocabulary (e.g., Mirzaee et al., 2015).

Results of the study might be attributed to the findings of Huang (2014) that concluded that learners failed to recognize the boundary of collocations. The LCG, who had not mastered the threshold level yet, might have experienced "difficulty in recognizing which word is the 'real friend'

accompanying the observed word. Such a mistake in identifying the collocational patterns might also be one of the reasons accounting for the failure of the LCG in benefitting from hands-on DDL. This justification was supported by the results of the questionnaire in which low-intermediate learners needed the teacher's support to ensure that they have recognized collocations correctly. Low-intermediate learners may need the teachers' intervention or "guided induction" (e.g., Flowerdew, 2009; Huang, 2014; Smart, 2014) to overcome their problems in analyzing concordancing. Since in the current study there was no teacher support while searching for the collocates in concordancing, the low proficiency learners did not improve their collocational knowledge. Such a finding is on par with the results of the questionnaire in which LCG was more eager to the teacher intervention.

The secondary objective of the current study was investigating the role of transfer of training in learning secondary collocations. Results demonstrated that the HCG spotted secondary collocations better than the LCG. Results are in line with previous studies (i.e., Leeser & DeMil, 2013; White & DeMill, 2013) which found promising results for the secondary effects of instruction.

Research to date (e.g., Elder & Ellis, 2009; Roehr, 2008) has demonstrated that high proficiency learners benefit more from implicit instruction than low proficiency learners who prefer explicit instruction more. Since the high proficiency learners have a good command of content and meaning, they can easily draw their attention to form (e.g., collocations) more than the low proficiency learners who resort mainly to Van Patten's (2004) "primacy of meaning principle", i.e., learners process meaning before form as a default strategy.

The last aim of the current study was eliciting learners' perception of concordancing in learning collocations and if their perception changed over 10 sessions. Opinions expressed showed that all learners of the experimental group agreed that concordancing was more effective than the traditional ways of learning collocations; they also wished to take part in other concordancing-based classrooms to learn more L2 collocations. Such a

finding is completely compatible with the results of the previous studies (e.g., Geluso & Yamaguchi, 2014; Rezaee et al., 2015; Saeedakhtar et al., 2020; Vyatkina, 2016a) which have reported the positive attitudes of learners towards DDL in fostering collocations.

Interestingly enough, none of the high-intermediate learners felt any need for the teacher's help in benefiting from concordancing. However, the low-intermediate learners believed that some degrees of teacher support was necessary to help them overcome some difficulties they experienced in using concordancing. This finding lends support to Flowerdew's (2009) claim that teachers' intervention is a must in DDL. Such a finding might be rooted in cognitive issues where the low-intermediate learners need the approval of their teacher to learn new elements while the high-intermediate learners are too confident and independent of their teacher.

The experimental group's perceptions of concordancing changed over time. They stated that, in early sessions, concordancing was very difficult and time-consuming for them (especially for the LCG). However, with the passage of time, they spent less time using concordancing and needed less teacher support. Such a finding is compatible with those of previous studies which concluded that in early sessions learners had negative attitudes towards DDL (Vyatkina, 2016a) but their perception changed over time.

CONCLUSION AND IMPLICATIONS

The present study investigated the extent to which level of proficiency can moderate the role of concordancing in learning L2 collocations. It also investigated the role of proficiency level in the transfer of training. Results revealed that level of proficiency can influence the use of concordancing in learning collocations. High-proficient learners outperformed low-proficient learners not only in primary but also in secondary learning. Such a finding lends support to the importance of some degrees of autonomy as an essential prerequisite for discovery learning. High intermediate learners who had

already developed the so-called autonomy or self-regulation did better than other-regulated learners in utilizing hands-on DDL. Learners also had positive attitudes towards hands-on DDL. One of the main implications of the present study for language teachers is exposing only high-intermediate learners to hands-on DDL. If low-intermediate learners are exposed to hands-on DDL, they should benefit from enough teacher support or scaffolding so that they be able to use hands-on DDL effectively. The need for teacher support for low-intermediate learners was also reflected in the results of the questionnaire.

The present study is not free from limitations. The very first shortcoming of the present study is the limited number of participants which may threaten the generalizability of the findings. Another drawback comes from the lack of enough teacher scaffolding to support the low-intermediate learners to benefit from self-directed and autonomous language learning situations. Previous studies have concluded that with enough guided induction (Smart, 2014) even the low proficiency learners can benefit from hands-on DDL.

Future studies are recommended to provide the low-intermediate learners with dialogue (Webb et al., 2004), before and during access to hands-on DDL and then compare the role of hands-on DDL in improving the high- and low-intermediate learners' collocations. Researchers are also recommended to provide learners with opportunities for peer-peer cooperation and interaction while benefitting from hands-on DDL. Furthermore, they are suggested to examine the role of age and individual differences such as aptitude, field-(in)dependence, etc. in the transfer of training.

Disclosure statement

ORCID

Afsaneh Saeedakhtar   ⓘ   http://orcid.org/0000-0003-0979-1198

References

Alameen, G. (2011). Learner digital stories in a Web 2.0 age. *TESOL Journal, 23*, 355−369.

Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics, 22*(2), 173−195.

Barfield, A. (2007). *An exploration of second language collocation knowledge and development.* (Unpublished doctoral dissertation), University of Wales, Swansea.

Benati, A., & Lee, J. F. (2008). *Grammar acquisition and processing instruction: Secondary and cumulative effects.* Bristol: Multilingual Matters.

Bikowski, D., & Casal, J. E. (2018). Interactive digital textbooks and engagement: A learning strategies framework. *Language Learning and Technology, 22*(1), 119−136.

Billing, D. (2007). Teaching for transfer of core/key skills in higher education: Cognitive skills. *Higher Education, 53*(4), 483−516.

Boulton, A. (2012). Hands-on/hands-off: Alternative approaches to data-driven learning. In J. Thomas & A. Boulton (Eds.), *Input, process, and product: Developments in teaching and language corpora* (pp. 152−168). Brno: Masaryk University Press.

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, *67*(2), 348–393.

Breyer, Y. (2009). Learning and teaching with corpora: Reflections by student teachers. *Computer Assisted Language Learning, 22*(2), 153−172.

Chang, J. Y. (2014). The use of general and specialized corpora as reference sources of academic English writing: A case study. *ReCALL, 26*(2), 243−259.

Chang, W., & Sun, Y. (2009). Scaffolding and web concordancers as support for

language learning. *Computer Assisted Language Learning, 22*(4), 283−300.

Chauhan, R., Ghosh, P., Rai, A., & Shukla, D. (2016). The impact of support at the workplace on transfer of training: A study of Indian manufacturing unit. *International Journal of Training and Development, 20*(3), 200–213.

Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics, 16*(1), 22−29.

Cobb, T., & Boulton, A. (2015). Classroom application of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 478−497). Cambridge: Cambridge University Press.

Cotos, E. (2014). Enhancing writing pedagogy with learner corpus data. *ReCALL, 26*(2), 202−224.

Crosthwaite, P. (2017). Retesting the limits of data-driven learning: Feedback and error correction. *Computer Assisted Language Learning*, *30*(6), 447−473.

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning, 28*(2), 130−144.

Fathi, J., & Nourzadeh, S. (2019). Examining the effects of writing instruction through blogging on second language writing performance and anxiety. *Issues in Language Teaching, 8*(1), 63−91.

Flowerdew, L. (2009). Appling corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics, 14*(3), 393−417.

Flowerdew, L. (2015). Corpus based research and pedagogy in EAP: From lexis to genre. *Language Teaching, 48*(1), 99−116.

Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL, 26*(2), 128−146.

Gabrielatos, C. (2005). Corpora and language teaching: Just a fling or wedding bells? *TESL-EJ, 8*(4), 1−37.

Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL, 26*(2), 225−242.

Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning and Technology, 15*(2), 2−11.

Govaerts, N., Kyndt, E., Vreye, S., & Dochy, F. (2017). A supervisors' perspective on their role in transfer of training. *Human Resource Development Quarterly, 28*(4), 515–552.

Grabe, W., & Stoller, F. (2002). *Teaching and researching reading*. London:

Pearson Education.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology, theory, analysis and applications* (pp. 145–160). Oxford: Clarendon Press.

Gremmo, M. J., & Riley, P. (1995). Autonomy, self-direction and self-access in language teaching and learning: The history of an idea. *System, 23*(2), 151–164.

Hadley, G. (2002). An introduction to data-driven learning. *RELC Journal, 33*(2), 99–124.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*(1), 24–44.

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

Huang, H. T. D., & Hung, S. T. A. (2010). Effects of electronic portfolio on EFL oral performance. *Asian EFL Journal, 12*(2), 192−212.

James, M. A. (2012). An investigation of motivation to transfer second language learning. *The Modern Language Journal, 96*(1), 51−69.

Johns, T. (1990). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria, 10(1),* 14–34.

Johns, T. F., Lee, H., & Wang, L. (2008). Integrating corpus-based CALL programs in teaching English through children's literature. *Computer Assisted Language Learning, 21*(5), 483–506.

Karras, J. N. (2016). The effects of data-driven learning upon vocabulary acquisition for secondary international school students in Vietnam. *ReCALL, 28*(2), 166–186.

Kennedy, C., & Miceli, T. (2010). Corpus-assisted creative writing: Introducing intermediate Italian students to a corpus as a reference resource. *Language Learning and Technology, 14*(1), 28−44.

Larsen-Freeman, D. (2013). Transfer of learning transformed. *Language Learning, 63*(1), 107−129.

Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second vocabulary learning: A multilevel meta-analysis. *Applied Linguistics, 40*(5), 721−753.

Leeser, M., & DeMil, A. (2013). Investigating the secondary effects of processing instruction in Spanish: From instruction on accusative clitics to transfer-of-

training effects on dative clitics. *Hispania, 96*(4), 748–762.

Lenko-Szymanska, A. (2014). Is this enough? A qualitative evaluation of the effectiveness of a teacher-training course on the use of corpora in language education. *ReCALL, 26*(2), 260−278.

Lewis, M. (1997). *Implementing the lexical approach: Putting theory into practice*. Hove: Language Teaching Publications.

Lightner, R., Benander, R., & Kramer, E. F. (2008). Faculty and student attitudes about transfer of learning. *A Journal of Scholarly Teaching, 3(1),* 58−66.

Lin, M. H., & Lee, J. Y. (2019). Pedagogical suitability of data-driven learning in EFL grammar classes: A case study of Taiwanese students. *Language Teaching Research, 23*(5), 541−561.

Luo, Q., & Liao, Y. (2015). Using corpora for error correction in EFL learners' writing. *Journal of Language Teaching and Research, 6*(6), 1333−1342.

Mirzaei, A., Rahimi Domakani, M., & Rahimi, S. (2015). Computerized lexis-based instruction in EFL classrooms: Using multi-purpose LexisBOARD to teach L2 vocabulary. *ReCALL, 28*(1), 22−43.

Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art - and beyond. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 5−24). Frankfurt: Peter Lang.

Nami, F. (2020). Towards more effective app-assisted language learning: The essential content and design features of educational applications. *Issues in Language Teaching, 9*(1), 245−278.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company.

O'Sullivan, I. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL, 19*(3), 269–286.

Park, K. (2010). *Enhancing lexicogrammatical performance through corpus-based mediation in L2 academic writing instruction* (Unpublished doctoral dissertation), Pennsylvania State University, University Park, PA.

Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. *Educational Psychologist, 47*(3), 248−258.

Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive high-variability segmental perceptual training: Implementation,

effectiveness, transfer. *Language Learning and Technology, 22*(1), 69–96.

Rezaee, A. A., Marefat, H., & Saeedakhtar, A. (2015). Symmetrical and asymmetrical scaffolding of L2 collocations in the context of concordancing. *Computer Assisted Language Learning, 28(6),* 532−549.

Saeedakhtar, A. (2013). *Symmetrical and asymmetrical scaffolding of high CV and low CV written L2 collocations in web-based concordancing*, (Unpublished doctoral dissertation), University of Tehran, Iran.

Saeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of hands-on and hands-off data-driven learning on low-intermediate learners' verb-preposition collocations. *System, 91*, xx−xx.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review, 64*(3)*,* 429–458.

Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL, 26*(2)*,* 184−201.

Smith, S. (2020). DIY corpora for accounting and finance vocabulary learning. *English for Specific Purposes, 57*, https://doi.org/10.1016/j.esp.2019.08.002.

Sun, X., & Hu, G. (2020). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research,* https://doi.org/10.1177%2F1362168820954459.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235−256). Rowley, MA: Newbury House.

Van Patten, B. (2004). Input processing in second language acquisition. In B. Van Patten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah NJ: Lawrence Erlbaum Associates.

Vyatkina, N. (2016a). Data driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning and Technology, 20*(3), 159−179.

Vyatkina, N. (2016b). Data driven learning for beginners: The case of German verb-preposition collocations. *ReCALL, 28*(2), 207−226.

Wallet, G., Sauzeon, H., Rodrigues, J., & N'Kaoua, B. (2009). Transfer of spatial knowledge from a virtual environment to reality: Impact of route

complexity and subject's strategy on the exploration mode. *Journal of Virtual Reality and Broadcasting, 6*(4), 1860–2037.

Wang, J., Zou, B., Wang, D., & Xing, M. (2013). Students' perception of a wiki platform and the impact of wiki engagement on intercultural communication. *System, 41*(2), 245−256.

White, J. P., & DeMil, A. J. (2013). Transfer-of-training effects in processing instruction: The role of form-related explicit information. *Studies in Second Language Acquisition, 35*(3), 519−544.

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.

Yoon, C. (2016). Concordancers and dictionaries as problem-solving tools for ESL academic writing. *Language Learning and Technology, 20*(1), 209−229.

Yoon, H., & Jo, J. W. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning and Technology, 18*(1), 96−117.

Appendix A

1. The List of Target Lexical Collocations

Verb-noun (n = 37)	MI		
Achieve improvement	11.30	Make prevention	5.55
Bring ruin	5.11	Make recommendation	5.28
Cause damage	8.84	Make requirement	5.03
Develop complication	8.25	Obtain relief	8.65
Develop resistance	8.37	Offer encouragement	8.18
Develop symptoms	9.65	Pay compliment	4.11
Do substitution	3.91	Promote competition	9.95
Earn reputation	10.88	Provide consolation	5.91
establish habit	8.99	Provide consultation	5.91

Establish trust	7,27	Provide explanation	7.85
Exercise caution	8.28	Provide protection	8.46
Gain promotion	9.81	Reach conclusion	9.42
Gain weight	9.09	Reach maturity	11.34
Generate belief	8.30	Run interference	6.85
Grasp opportunity	8.94	Seize control	9.67
Have symbol	3.71	Take option	3.39
Hold expectation	7.83	Wear make-up	12.28
Hold implication	9.94	Win representation	9.19
Level criticism	10.11		

Adjective-noun (n = 15)	MI		
Age-old custom	12.83	Powerful reminder	7.61
Deep impression	7.74	Regular medication	9.69
Deep knowledge	5.38	Serious complication	8.39
Immediate effect	8.57	Strong recommendation	6.82
Great curiosity	5.87	Strong tendency	8.40
Great opportunity	6.47	Urgent need	12.31
Live birth	6.42	Wide variety	11.57
Momentous occasion	12.17		

Noun-verb (n = 9)	MI		
Abuse occur	8.31	Popularity spread	7.49
Attempts make	5.01	Rule operate	7.43
Complications occur	8.75	Specialty lay	7.42
Inclination arise	8.78	Wishes fulfill	11.57
Magic work	6.69		

2. The List of Target Grammatical Collocations

Verb-preposition (n = 16)	MI		
Balance against	3.46	Lead after	3.66

Beg for	5.91	Mourn for	6.31
Blow into	3.86	Prevent from	4.73
Conserve for	6.84	Protect against	10.23
Consist of	5.00	Purchase from	3.02
Equate with	6.11	Replace with	6.40
Feed on	5.70	Seek for	4.88
Interfere with	6.98	Sneak into	7.37
Last for	4.44	Wipe off	7.57

Preposition-noun (n = 15)	MI		
For replacement	4.71	Through distance	3.53
In attendance	5.28	Under influence	6.89
In charge	4.73	Under tuition	4.65
In mirror	5.57	With determination	4.88
In need	4.56	With responsibility	4.84
In socializing	3.38	Without hesitation	10.66
Of deprivation	4.63	Without limit	10.32
Out of curiosity	8.95		

Noun-preposition (n = 15)	MI		
Addiction to	4.63	Improvement over	4.60
Attendance at	6.73	Need for	6.58
Charge of	3.77	Replacement for	4.86
Consolation for	5.06	Reputation for	5.49
Control over	7.30	Resistance to	4.52
Curse upon	7.31	Training in	3.87
Explanation for	5.15	Trust in	4.16
Implication of	4.65		

Adjective-preposition (n = 18)	MI		

Aware of	5.03	Packed with	6.23
Beneficial to	4.49	Popular with	6.05
Confused with	6.09	Recognizable as	5.71
Curious about	8.29	Rude to	4.71
Damaging to	4.96	Scattered throughout	9.88
Harsh toward	7.00	Synonymous with	7.16
Honorable for	6.84	Unhealthy for	4.26
Loaded with	3.08	Urgent for	4.62
Nervous about	7.67	Watchful for	3.67

3. The List of Secondary Collocations (n = 10)

Collocations	Type of collocations
Anxiety grow	Noun-verb
Draw attention	Verb-noun
Contribute to	Verb-preposition
Gain interest	Verb-noun
Hesitated on	Adjective-preposition
In tears	Noun-preposition-
Interference with	Noun-preposition
Long-term abuse	Adjective-noun
Never-ending marriage	Adjective-noun
Taste of	Noun-preposition