Effectiveness of Corpus Consultation in Teaching Verb+Noun Collocations to Advanced ELT Students

Mehmet Akıncı [a], Senem Yıldız [b] *

[a] Şehir University, Turkey
[b] Boğaziçi University, Foreign Language Education Department, Turkey

Abstract

This study examined the effectiveness of data-driven learning (DDL), explicit instruction and these two methods combined in teaching verb+noun (V+N) collocations to advanced Turkish learners of English. Recognition accuracy of V+N collocations was measured along with participants' judgment about the acceptability of these collocations. It also explored the opinions of learners about using corpus in learning V+N collocations. Quantitative data were analyzed via a one-way independent analysis of variance and descriptive statistics. Results revealed a statistically significant difference between the three groups in their recognition accuracy of V+N collocations. While the Explicit-Instruction (EI) Group and the Combined (C) Group scored significantly higher than the Data-Driven Learning (D) Group, the means of the EI-Group and C-Group did not significantly differ. With respect to the judgment about the acceptability of V+N collocations, the EI-Group significantly outperformed the D-Group; but the difference between the EI-Group and the C-Group; or between the C-Group and the D-Group was not statistically significant. The questionnaire data showed that the participants found the instruction with the use of corpus more useful and effective for learning V+N collocations than instruction without.

1. Introduction

The last fifty years have witnessed a multitude of studies conducted on corpus consultation and the advances of digital technology have fueled corpus-based research in educational theory and practice, strengthening the role of corpus in the field of second language acquisition (SLA). The influence of corpus use has been most obvious in vocabulary analysis, training, and theory building (Gardner, 2007). In relation to vocabulary, it has been consulted in discovering the difficulties of learning and teaching collocations whose knowledge promotes target language proficiency considerably (Nation, 2001).

Introduced by Johns (1991) and based on corpus consultation, data-driven learning (DDL) has been regarded as an innovative strategy in English language teaching (ELT). In spite of the research supporting DDL in teaching collocations (Chan & Liou, 2005; Çelik, 2011; Daskalovska, 2015; Huang, 2014; Koosha & Jafarpour, 2006; Rezaee, Marefat & Saeedakhtar, 2014; Sun & Wang, 2003), Nesselhauf (2003, 2005) suggests that explicit attention should be paid to collocations due to the influence of L1 transfer and restrictions on collocations peculiar to the language itself. This study compares the effectiveness of DDL with explicit instruction in teaching verb+noun (V+N) collocations to advanced Turkish learners of English and explored opinions about the use of corpus data in learning and teaching collocations.

1.1. Corpus and data-driven learning

Since it was employed in the collection of Bible and the works of Shakespeare during the Middle Ages (Tribble & Jones, 1990), the definition of corpus has varied remarkably over time. In today's world, it can be defined as a principled (Conrad, 1999) and electronic (Granger, 2002) collection of naturally occurring (Leech, 1997a) written text or transcribed speech coming from various sources; and this collection offers quantitative and qualitative analysis in a fast and reliable way (Kennedy, 1998; Reppen, 2010).

Advances in technology have enabled concordancing to emerge as a trend in various formats which has profoundly affected the ability to process corpora. The Key Word in Context (KWIC) concordance is widely used by researchers to analyze lexical patterns (Koosha & Jafarpour, 2006), linguistic features (Kaur & Hegelheimer, 2005) and the frequency of occurrence of linguistic items (Granger, 2002). With the help of KWIC, learners can observe regularities (Hunston, 2002) that might otherwise unnoticed (Kaltenböck & Mehlmauer-larcher, 2005) in the course of learning the target language.

Although the most striking feature of corpora is that texts as its constituents are naturally-occurring and are therefore authentic, their authenticity has been subjected to negative criticism especially by Widdowson (2000). Basically, Widdowson argues that texts collected in a corpus form a decontextualized language lose their authenticity. Yet authenticating a corpus becomes possible when learners adopt a role of observer and/or a participant interacting with text (Gavioli & Aston, 2001) by making observations and using items which they observe (Chambers & O'Sullivan (2004), which is in line with data-driven learning described by Johns (1991).

1.1.1. Data-driven learning (DDL)

The idea of DDL emerged as an innovative way to implement concordancing materials in the ELT classroom. DDL is based on the principle that language learners should be encouraged to actively and consciously examine and explore naturally occurring language in order to discover patterns on their own. The originator of DDL, Johns, maintains that "research is too serious to be left to the researchers" (Johns,

1991, p. 2); and that language learners are research workers whose learning is driven by access to authentic linguistic data (Johns, 1994). Leech (1997b, p. 3) makes the point that "teaching is a natural extension of research" and claims that a corpus "enables the learner/student to explore, to investigate, to generalize, to test hypotheses; but it does not itself initiate or direct path of learning" (ibid, p. 5). Johansson (2007) perceives the learner as a researcher in accordance with ideas about language learning as an active process: he describes language learning as a process of hypothesis formation and hypothesis testing. In addition, Papp (2007) argues that hypothesis testing is based mainly on a process of comparison between learner's production and the target input. If there is a mismatch, the learner rejects her/his hypothesis and finally discards it. However, if there is a match between the production and the input, the hypothesis gains power and is finally confirmed.

Instead of assuming the role of language expert in charge of teaching and research, teachers in DDL undertake the role of facilitator (Cheng, Warren, & Xu-feng, 2003), assuming that learners need support and guidance from a teacher (Kaltenböck & Mehlmauer-Larcher, 2005). Through a careful selection of concordances, learners are guided to deduce differences in meaning, connotation and grammatical features of given linguistic items and acquire grammatical rules through exploration on the basis of lists with concordances rather than being taught the rule.

Mishan (2004) illustrates how a DDL task overcomes the problem of texts' authenticity that is lost in the process of transplantation. First, a corpus is used in line with the purpose for which it was designed: as a research source; second, learners are involved in goal-oriented tasks where the outcome is not pre-determined as opposed to conventional methods where the result or answer is known before the task is undertaken. Finally, a DDL task authenticates corpus data through the learners' interaction and engagement with it.

The effectiveness of corpus consultation in vocabulary teaching has been examined in a myriad of studies. The first controlled experiment was one by Steven (1991), a small-scale study measuring the effectiveness of consulting printouts from a corpus assembled from ELT course-books. The participants were found to be better at recalling words when cued by concordance lines. This study was extended by Cobb (1997) who conducted a small-scale within-group experimental study with 100 lower-intermediate first-year ESL learners at an Arabic-speaking university. Cobb investigated whether a computer concordance would stimulate and potentially rationalize off-line vocabulary acquisition by presenting target words in several contexts and in doing so promote transferring word knowledge to novel situations. The findings showed that concordances available on screen were more effective than other sources of lexical information. In addition, with the goal of solving the breadth-depth paradox, Cobb (1999) investigated the effectiveness of computerized concordances. The experimental group which utilized a corpus was more successful at supplying new words in gaps in a novel text in addition to gaining definitional knowledge. Analyzing the learning process and the strategies employed by EFL learners using concordancing tools, Sun's (2003) findings indicated that prior

knowledge, cognitive skills, teacher intervention, and skill in using concordances play a significant role. Aiming to reveal the effectiveness of concordance lines on the discovery of language patterns, Boulton (2008) collected data from 113 adult EFL learners consulting concordance lines to learn phrasal verbs. A significant improvement was observed in the post-test results, suggesting that consulting concordance lines was beneficial. In a context of Turkish learners of English, Çelik (2011) tested the effectiveness of DDL in teaching academic words and prepositional phrases. The data were collected from 68 EFL learners at Faculty of Medicine who either studied the target lexical items through web-based concordancing activities (DDL group) or online dictionary use method in a 5-session study. Although the results did not reveal a significant difference between the two groups on post-tests, the DDL group had significantly higher scores in retention tests. Similarly, in a context of Chinese learners majoring in English, Huang (2014) found that paper-based concordance group had a better performance than dictionary use group in using collocational and colligational patterns in their writing. However, some difficulties related to the unfamiliar vocabulary and cut-off sentences in the paper-based corpus were reported in the study.

As can be seen, DDL based on corpus consultation and implicit learning can be beneficial in teaching vocabulary. However, the effectiveness of using corpus in teaching lexical items such as collocations is not obvious, and because such items create a huge deal of difficulty due to L1 transfer and extant restrictions of the language itself, explicit teaching may be required. In the following section, the importance of collocations will be explained along with the challenges encountered by language learners and teachers.

1.2. Teaching and learning collocations

With a wide range of meanings (some of them rather vague) (Nesselhauf, 2005), the term collocation is defined from two main perspectives: one leads to frequency-based approach and the other to a phraseological approach. Stated by Firth (1957) first and improved by Sinclair (1991), the frequency-based one is based on the occurrence of words at a certain distance with the distinction drawn between frequent co-occurrences and the ones that are not regardless of whether there is a syntactic relationship between the elements.

The phraseological approach was strongly influenced by Russian phraseology and adopted by Cowie, Mel'cuk F. J. Hausmann in the fields of lexicography and/or pedagogy (Nesselhauf, 2005). This approach features two types of combinations: formulae, the ones with pragmatic functions; and composites, the ones with syntactic functions (Cowie, 1994). Distinctions between these combinations are based on two criteria: transparency, and commutability or substitutability. While transparency deals with if the elements of a combination or the combination itself have literal or non-literal meaning, commutability or substitutability concerns if the elements can be substituted with other words or not. Nesselhauf (2005, p. 25) defines collocations as "a

type of word combinations in a certain grammatical pattern" and states that this type is used to refer to "an abstract unit of language and its instantiations in texts".

Among a pre-defined set of syntactic relations such as adjective+noun, noun+noun, and noun+verb; verb+(object) noun (V+N) is considered by Howarth (1998), Nesselhauf (2003, 2005) and Chan & Liou (2005) as the most essential one because they make up the communicative core of utterances by representing the propositional core of the fully formed clause; and they constitute dominant EFL weaknesses. In a further classification, V+N collocations include verb object (VO) e.g. *do harm*; verb+preposition+object (VPO) e.g. *deal with a problem*; verb+object+preposition+object (VOPO) e.g. *take smt into consideration*; verb+object+to+infinitive (VO + to + inf) e.g. *force sb to buy*; and verb + adverbial phrase (VA) like *jump out of the window* (Nesselhauf, 2005). In the present study, a phraseological approach is used for the definition of collocations, and Nesselhauf's classification of V+N combinations is adopted.

Many scholars argue that L2 learners have a strong need for the collocation knowledge due to its vitally important role in the production of the target language (Nation, 2001, Wray 2002, 2008). Learners tend to acquire words individually without taking notes of their immediate environment (Skehan, 1998). This leads to the production of unconventional combinations when non-native speakers (NNSs) lack awareness and have little understanding of the phraseological mechanisms of the language (Howarth, 1998). Upon encountering difficulty in producing collocations, L2 learners usually resort to synonymy and/or their L1 to overcome the barriers, but these can lead to mistakes that result in their sounding unnatural. Chang, Chang, Chen, & Liou (2008) propose that most collocation problems of L2 learners are related to L1 interference; and explain this by referring to split categories and direct translation from L1 to L2. A split category occurs when "two words in one language are covered by only one word in another language" (Chang et al. 2008, p.287). For example, the words "win" and "earn" own an overlapping Turkish and because of this shared translation, Turkish learners of English who lack a full understanding of collocation restrictions might produce an inappropriate V+N collocation, such as 'win a living'. Second, direct translation employed by L2 learners can create problems, especially when L1 language patterns interfere with their inferencing process (Natinger & DeCarrico, 1992). The learners might easily produce a miscollocation if one of the components of a collocation can be substituted by a similar word in line with their L1. Nesselhauf (2003, 2005) also argues that non-congruent collocations are more challenging for L2 learners. Furthermore, when there is no congruency between L1 and L2, delexical verbs (e.g. do, make, take, get, have) are observed in collocation mistakes because they have low lexical content and gain meaning from the lexical content of the words they collocate with (O'Keeffe, McCarthy & Carter, 2007). L2 learners also underuse some elements (e.g. boosters) and overuse others to create collocations (e.g. very in an adverbial phrase) to express their ideas (Fan, 2009; Granger, 1998; Siyanova & Schmitt, 2008) due to their inaccurate intuition regarding the existence of collocations; the result is foreignness or oddness in their L2

production (Koosha & Jafarpour, 2006). So as to cross the intermediate plateau and approach to native-speaker norms, L2 learners need to improve their collocation knowledge, as advanced language production is closely related to and therefore necessitates the knowledge and correct use of collocations (Howarth, 1998; Conzett, 2000, Nation 2011).

Numerous studies have been conducted in the last fifteen years to explore efficient ways of teaching collocations to L2 learners using corpora. Sun & Wang (2003) were among the first to explore the use a concordancer to help EFL college students. Their findings indicated that, for teaching easier collocations, an inductive approach based on a corpus was more suitable than a deductive one but there was no significant difference in teaching difficult ones. Chan & Liou (2005) and Ang (2006) found that the corpus consultation can significantly increase learners' awareness about formulaic language and the participants in their study had positive attitudes towards corpus use, supported by the findings of Aşık, Vural and Akpınar (2016). Koosha & Jafarpour (2006) reveal that DDL is highly effective in the process of teaching and learning collocations. They found a relationship between the proficiency levels of learners and the use of collocations and determined that L1 had an effect on Iranian learners' production of English. Additionally, in a 15-session study, Rezaee et al. (2014) looked into the effects of concordancing on the receptive and productive performance of both lexical and grammatical collocations and explored the attitudes of their participants towards the use of concordancing. Concordancing was found to be more effective than traditional methods and was seen to help raise consciousness about the important role of collocations in language learning. In addition, investigating the role of corpus-based activities in learning verb-adverb collocations, Daskalovska (2015) found that participants employing corpus gained more collocation knowledge.

In sum, studies comparing DDL have discovered that DDL is either more effective than explicit instruction or as effective as explicit instruction on teaching collocations. Nesselhauf (2003, 2005), on the other hand, points out that it would be more effective to increase the emphasis on L1 and explicitly draw learners' attention to differences between L1 and L2 during instruction. However, there are very few studies examining the effectiveness of corpus consultation in teaching V+N collocations forming the communicative core of utterances and constituting dominant EFL weaknesses. In addition, opinions of learners exposed to both DDL and explicit instruction have not been explored. Therefore, this study aimed to compare three instruction types: DDL, explicit instruction, and a third method that combines DDL and explicit instruction, by measuring their effectiveness in teaching V+N collocations especially the ones in which L2 learners have difficulty because of the L1 interference.

2. The study

2.1. Research questions

The present study aimed to investigate the following research questions:

1. Are there any differences between the DDL-Group, Explicit Group, and Combined-Group in their recognition accuracy of V+N collocations after the treatment?
2. Are there any differences between the DDL-Group, Explicit Group, and Combined-Group in their judgment on the acceptability of V+N collocations after the treatment?
3. What are the opinions of learners in DDL-Group and Combined-Group about corpus consultation in learning collocations?

We hypothesized that the DDL group would be better than other groups in their recognition accuracy (Hypothesis 1) and judgment on acceptability (Hypothesis 2) of V+N collocations after the treatment in line with the studies showing significant gains of corpus consultation in vocabulary learning (Çelik, 2011; Daskalovska, 2015; Koosha & Jafarpour, 2006; Rezaee et al. 2014; Sun & Wang, 2003). We also hypothesized that learners exposed to corpus consultation would have positive opinions about DDL in learning V+N collocations (Ang, 2006; Aşık et al. 2016; Chan & Liou, 2005).

2.2. Participant characteristics

Participants were 58 Turkish freshman ELT students enrolled in an English composition (EC) course offered by the ELT Program of an English-medium university in Turkey. Their ages ranged from 18 to 20 years. Of the 58 participants, 46 were female and 12 were male. Each had had to pass the university's English proficiency test in order to enroll in their major course of study. The minimum pass mark on this test is taken to be the equivalent of 79 on the TOEFL iBT.

In order to establish a balanced distribution of participants with different proficiency levels in each group, stratified random sampling was employed and participants were divided into three groups, each one of which received one of three instruction types: the first group was the DDL group (D-Group), 20 participants who were instructed using the DDL method; the second was the Explicit Instruction Group (EI-Group), 19 participants who were taught collocations through explicit instruction; and the last was the Combined-Group (C-Group), 19 participants who received instruction that combined DDL and explicit instruction. Data were collected from the C-Group for two reasons. First, collocation knowledge would help students to achieve the goals of the EC course. Second, advanced level learners could benefit from exploring the intricate semantic relationships between words revealed through concordancing (Gardner, 2007).

Participants were enrolled in one of the two different sections of the EC course. Each section was taught by a different instructor and met three hours per week for 13 weeks. The objective of the course was to improve students' ability to produce academic writing in English. Five follow-up laboratory sessions were held and the treatment was given during those sessions by the researcher of the current study.

During the course of the study, the participants who missed a session were excluded from the study. A total of 5 out of 20 participants were lost during the study, all from the EI-Group.

2.3. Data collection procedures

The British National Corpus (BNC) was used to prepare the worksheets for the EI-Group and to design the V+N collocation test and the collocation judgment test that were to be employed in the data collection procedure. Prior to the treatment, nouns whose collocating verbs that the students would work on had been chosen from the Oxford Collocations Dictionary for Students of English (2002), taking into account the possible role of L1 interference, specifically by taking split categories, direct translation and delexical verbs into consideration. Since, at the time of the data collection, there had been no research conducted on collocations that are difficult for Turkish learners of English to learn and produce because of L1 interference, the selection of the nouns was informed by discussions with the course instructors and ten other ELT instructors, all of whom had over five years' experience in the field.

The treatment was given over a period of five weeks, the first of which was allotted to training that was essential for the students to use the corpus (Aşık et al. 2016; Frankenberg-Garcia, 2012) and to be aware of the role of the collocation knowledge and L1 interference. Therefore, the number of the nouns covered in the first session was lower than the other weeks. The five sessions (each lasting 50 minutes) for the D-Group were conducted in a laboratory where each participant had access to a computer. The teacher listed the nouns on the board. The participants then engaged with the corpus to perform a research task (Mishan, 2004). They were asked to use the concordance to find and make a note of the collocating verbs for each noun written on the board: ones they had never used in writing or speaking before; ones they considered highly important; and/or ones that constituted non-congruent collocations. In this way, in their roles of observer and participant interacting with text (Gavioli & Aston, 2001), they became researchers (Johns, 1994). They were also asked to identify discrepancies between their usage of English and native-like patterns (Papp, 2007). The participants exchanged their findings with other learners dealing with the same task (Johns, 1994). The teacher, as a facilitator, guided the learners when they had difficulty (Cheng et al. 2003; Johns, 1994; Kaltenböck & Mehlmauer-Larcher, 2005). In light of the participants' findings, the collocating verbs were written next to the nouns listed on the board, producing a collocation list.

Five sessions (each lasting 50 minutes) were conducted in a conventional classroom for the EI-Group. The teacher, taking full responsibility (Chang & Liou, 2005) and assuming the role of a language expert, wrote on the board the same nouns that were given to the D-Group and then led a discussion on what verbs the participants used with the nouns under investigation. A list of collocations for each noun was created by the teacher in accordance with the list that the D-Group provided through completing a research task. The emphasis was on L1 interference, with the teacher spoon-feeding

the participants by showing the differences between L1 and L2 (Tribble & Jones, 1990). For each week, the researcher prepared a worksheet that consisted of paragraphs from the BNC with a relevant V+N collocation omitted. Participants were assigned to complete the worksheet by referring to the lists on the board (Koosha & Jafarpour, 2006). At the end of the session, the answers were given to the participants, illustrating that the outcome was pre-determined (Mishan, 2004). Combined Group (C-Group) received the same type of instruction as the D-Group did during the first three weeks of the experiment, and during the last two weeks, instruction mirrored what EI-Group received.

2.4. Data collection instruments

To address the research questions of the study, the set of instruments used included a V+N collocation test, a collocation judgment test, a self-evaluation questionnaire and semi-structured interviews. The V+N collocation test developed to address the first research question explored whether there were any differences between the D-Group, the C-Group and the EI-Group in terms of recognition accuracy of collocations after the treatment. The collocation judgment test employed to answer the second research question examined whether there were any differences among the three groups in terms of their intuition about the acceptability of V+N collocations. The self-evaluation questionnaire and semi-structured interviews addressed the third research question, which pertained to the opinions of the participants about corpus consultation in learning collocations.

The V+N collocation test was developed by the researcher to measure to what extent participants had learned collocations. The test was based on the idea of Chan and Liou (2005) and Koosha and Jafarpour (2006) about measuring the knowledge of collocations. The test items were in a gap-fill format, with the first letter of the verb and a Turkish translation of the sentence provided. All sentences were from the BNC, and the translations were made by a professional Turkish interpreter of English and checked by three Turkish-English bilingual language teachers who were raised in an English speaking country. The test had 30 items all of which were sampled from the collocations taught during the 5-week lab sessions so as to assess whether participants had acquired the target V+N collocations as a result of instruction. The content of the items was based on the two criteria that can cause difficulty to Turkish learners of English: non-congruent V+N collocations (15 items), and split categories (15 items) (see an example below for each). Since the test drew on an authentic corpus of the target language use domain, it can be said to have content validity. The test was administered at the end of the experiment ($\alpha = .742$). Each correct answer was given one point (range 0-30).

An example for non-congruent criterion:

(Bir haftalık yağış kuraklığı giderdi.)

-A week of good rains has b_____ the drought. (answer: broken)

An example for split category criterion:

(Bir monopolistin kar etmesi veya sosyal sorumlulukları olan bir organizasyonun zarar etmesi, ekonomik bir etkinliğin işareti olarak kabul edilmez.)

-The ability of a monopolist to e_____ profit, or an organization with social responsibilities to make losses, is hardly to be taken as a signal of economic efficiency. (answer: earn)

The collocation judgment test was developed by the researchers to measure the extent to which participants' intuition about the acceptability of V+N Collocations was accurate. Granger's (1998) notion of acceptability or unacceptability of collocations was adopted. The items were in multiple-choice format, each with three options: 'acceptable', 'unacceptable', and 'I have no idea' (see an example below).

-We've got to <u>strike</u> the appropriate <u>balance</u> between doing what is best for the customer and what is best for us. (answer: acceptable)

a. acceptable

b. unacceptable

c. I have no idea

All sentences were from the BNC. To create unacceptable items, the main verb of each sentence was replaced by a verb that made the sentence unacceptable. The unacceptable item was first examined in the BNC to determine whether it had an occurrence or not; four native-English-speaking students were then consulted on the acceptability of the items. Among the 33 items on this test, all were sampled from the collocations taught during the 5-week lab sessions. Out of the 33 items, 14 were unacceptable and 19 were acceptable. The content of the items was based on four criteria that can cause difficulty for Turkish learners of English: non-congruent V+N collocations (9 items), split categories (16 items), collocations using delexical verbs (2 items), and congruent collocations (6 items). The test was administered at the end of the experiment (α = .367). The option 'I have no idea' was considered wrong. Each correct answer was given 1 point (range 0-33).

A self-evaluation questionnaire with 18 items was constructed by the researchers to probe into the feelings of the participants about using corpus consultation in learning collocations. It was given only to the participants in the D-Group and the C-Group, because they were the ones exposed to corpus consultation. The items were divided into two sections: those dealing with how corpora can help L2 learners, and those that asked for participants' opinions about the use of corpora in the future. All the items were designed on a 5-point Likert scale (ranging from "strongly agree", 5 point to "strongly disagree", 1 point). Having questions based on similar content, both parts were analyzed together. For the two sections, the Cronbach alpha obtained from the responses of the 29 participants in the study was 0.87.

With the goal of getting a better understanding of what participants actually did during the study (Silverman, 2001) through examining their experiences, opinions,

feelings and knowledge of the participants (Patton, 2002); semi-structured interviews were used to probe the opinions of the participants about the use of a corpus as a research tool in general and its ability to facilitate the learning of V+N collocations. Seven participants from the D-Group and six from the C-Group were interviewed about various aspects of corpus consultation. The questions were similar to those Ang (2006) constructed for her study and arranged in a sequence from general to specific. They were designed by the researcher with the aim of promoting retrospective thinking on the goals of the lab sessions, whether the sessions helped to accomplish these goals, the advantages of using a corpus, to what extent the sessions contributed to the participants' English, the participants' learning style preferences and the effect of those preferences on the results, and the assistance they received. Each interview lasted approximately 15 minutes.

3. Results

Prior to the treatment, the participants were randomly assigned to one of three levels of the independent variable in order to ensure that the groups were equal on all aspects. The Kruskal-Wallis test for language proficiency revealed no significant difference across the three groups: $H(2) = .41$, $p > .05$.

This result section is organized according to our three research questions.

3.1. Recognition accuracy of V+N collocations after the treatment

The means and standard deviations for recognition accuracy of V+N collocations across the three groups are given in Table 1.

Table 1: Means and standard deviations for V+N collocation accuracy scores

Groups	Mean	SD	N
EI-Group	15.3	3.7	15
C-Group	13.6	4.6	19
D-Group	10.5	3.9	19
Total	12.9	4.3	53

Table 1 shows that the EI-Group scored the highest and the D-Group scored the lowest on the test of recognition accuracy of V+N collocations. In order to determine whether there was a significant difference among the three groups, a one-way independent ANOVA was conducted. The normality assumption, assessed via a Shapiro-Wilk test, was sustained. The assumption of homogeneity of variance, checked through Levene's test, was sustained, $F(2, 50) = 2, 168$, $p > .05$, as well.

The ANOVA results showed that the recognition accuracy of V+N collocations across the three groups differed significantly, $F(2, 50) = 7, 378$, $p < .05$, partial $\eta^2 =$

.228. Bonferroni post hoc tests revealed a significant difference between the D-Group and the EI-Group, and between the D-Group and the C-Group, but not between the EI-Group and the C-Group.

3.2. Participants' judgment on acceptability of V+N collocations after the treatment

The means and standard deviations for the judgment of V+N collocations across the three groups are given in Table 2.

Table 2: Means and standard deviations for collocation judgment scores

Groups	Mean	SD	N
EI-Group	19	2.6	15
C-Group	18.3	3.2	19
D-Group	16.3	2.5	19
Total	17.8	3.9	53

Table 2 shows that the EI-Group scored the highest and the D-Group the lowest in their judgment of V+N collocations. In order to determine whether there was a significant difference among the three groups, a one-way independent ANOVA was conducted. The normality assumption, checked through a Shapiro-Wilk test, was sustained. The assumption of homogeneity of variance, assessed via Levene's test, was sustained, $F (2, 50) = .857$, $p > .05$, as well.

The ANOVA results indicated that the judgment of V+N collocations across the three groups differed significantly, $F (2, 50) = 4, 455$, $p < .05$, partial $\eta^2 = .151$. Bonferroni post hoc tests revealed a significant difference between the EI-Group and the D-Group, but not between the D-Group and the C-Group, or between the EI-Group and the C-Group.

The opinions of the participants about the use of corpus in learning and teaching collocations were addressed through a self-evaluation questionnaire and semi-structured interviews.

3.3. Self-evaluation questionnaire

Among the ways that corpora can help language learners, the participants indicated that corpora primarily help learners become aware of common uses of collocation in the target language (M=4.21). They also indicated that corpora can help them (Item-2) to become aware of mistakes they make with collocations (M=4.10), to become aware of how collocations are used in the target language (M=4.07), to use frequently occurring collocations in their writing (M=3.59), and to obtain cultural information about the target language (M=3.14). The results also revealed that overall

there were more respondents who were in favor of using corpora for language learning (68.2%).

Since the participants were studying ELT as their major, they would be English language teachers after graduation. The results of the items about whether and how the participants would use corpora in the future revealed that they would use corpora mostly to explore the English language (M=4.14). They would also encourage their colleagues to use corpora for their teaching (M=4.00). Additionally, they would use corpora to develop teaching materials (M=3.90); and make their students use corpora to explore the English language (M=3.76).

The analysis of the qualitative data obtained from the semi-structured interviews presented information on three central themes: the performance of the participants, improvement in English language skills, and learner autonomy. With respect to the performance theme, results revealed that two factors influenced the participants' performance in the study. The first factor was the research task, which in general was found to be motivating because the participants derived a certain pleasure from seeing their collocation mistakes and learning combinations that they had never used before. The other factor was the time of the laboratory sessions, which was the chief complaint of the participants in the D-Group, some of whom found 9 a.m. too early to concentrate on a research task.

The second theme that emerged from the interviews was the role of corpora in language learning, which had three categories. First, the participants commented on the usefulness of the corpus in improving their language skills, mainly for its being a reliable resource, having richness of vocabulary and providing context. Second, the participants emphasized the importance of the collocations in L2 proficiency, particularly in writing. Last, although they stated that they had benefited from the laboratory sessions, they highlighted that it could have been perfect if they had not encountered a problem with retention which was the third category.

The last emerging theme was labelled learner autonomy, because it was specifically about the learning strategies of the participants in the past and the strategies introduced in D-Group. From the explanations of the participants, it was induced that, prior to the experiment, they had all experienced explicit instruction in collocations. While some participants noted that the innovative strategy might have had a negative effect, others found the method more advantageous than explicit instruction. The participants in the C-Group who were exposed to both methods considered DDL more effective.

4. Discussion

For the first research question, the findings showed that at the end of the laboratory sessions, the Explicit Instruction Group (EI-Group) and the Combined Group (C-Group) outperformed the DLL-Group (D-Group) in their recognition accuracy of V+N collocations. With respect to the second research question, the EI-Group outperformed the D-Group in terms of intuition about the acceptability of those

collocations. The fact that the D-Group had the worst performance might be explained by two major issues discussed in the literature. First, as has been strongly stated by several scholars (e.g. Cheng et al, 2008; Howarth, 1998; Granger, 1998; Natinger & DeCarico, 1992; Nesselhauf, 2003, 2005; Siyanova & Schmitt, 2008), there is considerable L1 interference on some collocations, especially the non-congruent ones. Due to non-congruency, L2 learners are likely to make mistakes because they rely on their L1 resources. To solve the problem, Nesselhauf (2003) maintains that it is necessary to teach and learn collocations explicitly. Explicit teaching helps learners enhance their awareness, thereby enabling them to deal with possible L1 interference. Second, the nature of teaching practices focusing on input-based learning and avoiding explicit vocabulary teaching can cause collocation problems (Laufer & Waldman, 2011). Comments by the EI-Group and C-Group suggest that explicit instruction can make the process of learning collocations more memorable (Walker, 2011). The results somewhat correlate with the findings of Sun and Wang (2003) and Chan and Liou (2005) who indicate that explicit collocation instruction is effective in promoting EFL learners' collocation knowledge.

The low performance of the D-Group might also be explained by the novelty of using corpus. As highlighted by Kaur and Hegelhaimer (2005), the novelty of using a corpus can cause learners trouble in fully and productively exploiting the corpus. In addition, DDL, as an innovative strategy, is a method unfamiliar to learners whose educational background is shaped by deductive learning strategies. Therefore, activities designed in light of DDL do not automatically lead to efficient inductive learning in all learners (Sun, 2003; Vannestal & Lindquist, 2007). Also, Webb and Kagimoto (2009) argue that implicit learning of L2 collocations is a slow process requiring a huge deal of input and exposure; and Webb, Newton, and Chang (2013) suggest that the more learners encounter the same collocations, the better they learn.

The results concerning the first two research questions do not support the findings of Steven (1991), Cobb (1997, 1999), Koosha and Jafarpour (2006), Çelik (2011), Huang (2014), Rezaee et al. (2014), or Daskalovska (2015) with regard to the effective use of corpus in teaching vocabulary. However, it should be noted that the present study did not have the same experimental setting as those studies. Steven (1991) and Cobb (1997, 1999) did not examine collocations. Even though Koosha and Jafarpour (2006), Çelik (2011), Huang (2014) and Daskalovska (2015) investigated the effectiveness of corpus consultation in teaching collocations, the focus was on different types of collocations, not on V+N collocations. The length of exposure to concordancing, however, was longer in Rezaee et al. (2014) than the current study in examining verb+noun collocations.

Although DDL based on corpus consultation did not prove to be as effective as explicit instruction in the present study, the results of the self-evaluation questionnaire showed that the students were generally positive about DDL. They agreed that corpus consultation could increase their awareness of the common uses of collocation and of their collocation mistakes. These findings coincide with those of Yoon and Hirvela (2004), Chan and Liou (2005), Aşık et al. (2016). They also echo the

findings of Huang (2014) and Rezaee et al. (2014), who reveal that the use of concordancing made participants realize they needed to pay more attention to collocations. What is more, the participants agreed that they would use corpora in the future to explore the English language for the accurate use of collocations. This finding supports the views of Tsui (2004) and Kaltenböck and Mehlmauer-Larcher (2005), who state that corpora can be an invaluable tool for answering the frequently asked questions of language teachers about collocations. However, as opposed to the findings of Yoon and Hirvela (2004) and Huang (2014), limited number of sentences, unfamiliar vocabulary and cut-off sentences were not reported by the participants of this study as problems encountered during corpus consultation.

The semi-structured interview data validated the self-evaluation questionnaire findings and offered more insights into the participants' assessment of the value of corpus consultation in teaching collocations. The interviews revealed that the participants realized the advantages of corpus consultation mentioned in the literature: being a reliable resource, having richness of vocabulary and giving context. This finding is in line with the views of Tribble and Jones (1990), Kennedy and Micelli (2001), Kaltenböck and Mehlmauer-Larcher (2005) and Huang (2014). The data also indicated that the participants were well aware of the role of collocations in language proficiency, which is also mentioned in the literature (e.g. Bahn & Eldaw, 1993; Granger, 1998; Howarth, 1998; Nesselhauf, 2003, 2005). Furthermore, apart from Sun's (2003) study that suggests prior knowledge, teacher intervention, cognitive skills, and concordancing skills as factors influencing the learning process and strategies in use, the interview data identified prior learning experiences of the participants as a significant factor influencing the effectiveness of corpus consultation.

Even though the results related to the first two research questions of the present study indicated that explicit instruction was more effective than DDL in teaching collocations, the interview data from the participants of the C-Group indicated that they found DDL more effective than explicit instruction. These results pose a striking contradiction, which might be explained by other findings from the interviews with respect to the time of the treatment to the D-Group and a retention problem. The performance of the D-Group might have been different if they had received the treatment later in the day and if an exercise to practice newly learned collocations had been done at the end of each session.

5. Limitations and suggestions for further research

The present study is not without limitations, so the findings should be interpreted with caution. First, in the absence of a list of problematic collocations specific to Turkish learners of English, the researcher chose the nouns whose collocating verbs might be problematic for the participants of the present study to learn, based on his intuition, observations and prior experiences. For future research, English collocations that Turkish learners of English might have difficulty in learning and

producing due to L1 interference should be determined through an analysis of a Turkish learner corpus. Second, the D-Group and the EI-Group lacked compatibility with respect to in-class activities. While the EI-Group was given a gap-fill exercise at the end of each lab session, the D-Group was not given such an exercise due to time restrictions. In a future study, a follow-up exercise should be done at the end of each session across groups in order to establish complete compatibility. Third, the reliability of the collocation judgment test was low. Therefore, using tests with higher reliability in future research would yield more valid results. Fourth, since a pre-test was not administered, the overall collocation gains could not be investigated. In order to see the difference between entry performance and the final level of learners, a pre-test could be used in a future study. Fifth, due to the institutional restrictions, the groups could not be given the same level of treatment at the same time of the day. Therefore, in a future study it would be worth giving treatment to each group at the same time of the day so as to make an accurate assessment of the effectiveness of each method in teaching collocations. Sixth, because of construction in the labs and the fact that the national holidays occurred during the semester, the study had to be conducted over only five class sessions. In a future study, it would be useful to explore the effectiveness of corpus consultation over a longer period of time. In this way, learners would also gain more familiarity with corpora and inductive learning strategies. Last, advanced level language learners from ELT department participated in the current study. A future study with participants with lower level language proficiency and/or from various disciplines would definitely contribute to the field.

Acknowledgements

References

Ang, F. (2006). *Effectiveness of corpus consultation through concordancing on the formulaic academic language use of freshman ELT students.* Unpublished master's thesis, Bogazici University, Istanbul, Turkey.

Aşık, A., Vural, A. Ş., & Akpınar, K. D. (2016). Lexical awareness and development through data driven learning: Attitudes and beliefs of EFL learners. *Journal of Education and Training Studies, 4*(3), 87-96.

Boulton, A. (2008). Evaluating corpus use in language learning: State of play and future directions. Paper presented at the American Association of Corpus Linguistics. Provo, Utah

Chambers, A. & O'Sullivan, Í. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL, 16*(1), 158-172.

Chan, T., & Liou, H. (2005). Effects of web-based concordancing Instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning, 18*(3), 231-250.

Chang, Y., Chang, S. J., Chen, H., & Liou, H. (2008). An automatic collocation writing assistant for Taiwanese EFL learners: A case of corpus-based NLP technology. *Computer Assisted Language Learning, 21*(3), 283-299.

Cheng, W., Warren, M., & Xu-feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, *31*, 173-186.

Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System, 25*(3), 301–315.

Cobb, T. (1999). Applying constructivism: A test for the learner as a scientist. *Educational Technology Research & Development, 47*(30), 15 – 31.

Conrad, S. M. (1999). The importance of corpus-based research for language teachers. *System, 27*(1), 1-18.

Conzett, J. (2000). Integrating collocation into a reading and writing course. In M. Lewis (Eds.), *Teaching collocation: Further developments in the lexical approach* (pp. 70–86). London: Language Teaching Publications.

Cowie, A. P. (1994). Phraseology. In E. A. Ronald (Eds.), *The encyclopedia of language and linguistics* (pp. 3168– 3171). Oxford: Pergamon.

Çelik, S. (2011). Developing collocational competence through web based concordance activities. *Novitas-ROYAL (Research on Youth and Language), 5*(2), 273-286.

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning, 28*(2), 130-144. **http://doi.org/b4th**

Fan, M. (2009). An exploratory study of collocational use by ESL students – A task based approach. *System, 37*, 110-123.

Firth, J. (1957). *Papers in linguistics*. Oxford: Oxford University Press.

Frankenberg-Garcia, A. (2012). Learners' use of corpus examples. *International Journal of Lexicography, 25*(3), 273–296. **http://doi.org/b4tj**

Gardner, D. (2007). Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics*, *28*(2), 241-256.

Gavioli, L., & Aston, G. (2001). Enriching reality: Language corpora and language pedagogy. *ELT Journal, 55*(3), 238-46.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A.P. Cowie (Eds.), *Phraseology: Theory, analysis, and applications* (pp. 79–100). Oxford: Oxford University Press.

Granger, S. (2002). A bird's eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 3–33). Philadelphia, PA: John Benjamins.

Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Eds.), *Phraseology: Theory, analysis, and applications* (pp. 161–186). Oxford: Oxford University Press.

Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexico-grammatical patterns in L2 writing. *ReCALL, 26*(02), 163-183.

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

Johansson, S. (2007). Using corpora: from learning to research. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), Corpora in the foreign language classroom: Selected papers from the sixth international conference on teaching and language corpora (TaLC 6), University of Granada, Spain, 4-7 July, 2004 (pp. 17-28). Amsterdam; New York, NY : Rodopi.

Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning materials. Classroom concordancing. *ELR Journal, 4,* 1–16.

Johns, T. F. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Eds.), *Perspectives on pedagogical grammar* (pp. 293-313). New York: Cambridge University Press.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61*(2), 647-672.

Leech, G. (1997a). Introducing corpus annotation. In R. Garside, G. Leech, & T. McEnery (Eds.), *Corpus annotation: Linguistic information from computer text corpora* (pp. 1-18). Harlow, Essex: Addison-Wesley.

Leech, G. (1997b). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23). New York. Addison Wesley Longman Inc..

Kaltenböck, G., & Mehlmauer-larcher, B. (2005). Computer corpora and the language classroom: On the potential and limitations of computer corpora in language teaching. *ReCALL, 17*(1), 65-84.

Kaur, J., & Hegelheimer, V. (2005). ESL students' use of concordance in the transfer of academic word knowledge: An exploratory study. *Computer Assisted Language Learning, 18*(4), 287-310.

Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.

Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology, 5*(3), 77–90.

Kita, K., & Ogata, H. (1997). Collocations in language learning: Corpus-based automatic compilation of collocations and bilingual collocation concordancer. *CALL, 10*(3), 229–38.

Koosha, M., & Jafarpour, A. A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *The Asian EFL Journal Quarterly, 8*(4), 192-209.

Mishan, F. (2004). Authenticating corpora for language learning: A problem and its resolution. *ELT Journal, 58*(3), 219–227.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nattinger, J. R., & DeCarrico, J. D. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*, 223 – 242.

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

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

Papp, S. (2007). Inductive learning and self-correction with the use of learner and reference corpora. In E. Hidalgo, L. Quereda, & J. Santana (Eds), *Corpora in the foreign language classroom: Selected papers from the sixth international conference on teaching and language corpora (TaLC 6), University of Granada, Spain, 4-7 July, 2004* (pp. 207-220). Amsterdam; New York, NY: Rodopi.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

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

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.

Silverman, D. (2001). Interpreting qualitative data: Methods for analyzing talk, text and interaction. London: Sage.

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

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review, 64*(3), 429-58.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Stevens, V. (1991). Concordance-based vocabulary exercises: A viable alternative to gap-fillers. *English Language Research Journal, 4,* 47-63.

Sun, Y. C., & Wang, L. Y. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning, 16*, 83 –94.

Sun, Y. C. (2003). Learning process, strategies and web-based concordancers: A case study. *British Journal of Educational Technology, 34*(5), 601-613.

Tribble, C., & Jones, G. (1990). *Concordances in the classroom: A resource book for teachers.* Essex: Longman.

Tsui, M. B. A. (2004). What teachers have always wanted to know and how corpora can help. In J. M. Sinclair (Eds.), *How to use corpora in language teaching (Vol. 12)* (pp. 39-61). John Benjamins Publishing.

Vannestal, M. E. & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL, 19*(3), 329-350.

Widdowson, H. (2000). On the limitations of linguistics applied. *Applied Linguistics*, *21*(1), 3-25.

Walker, C. P. (2011). A corpus-based study of the linguistic features and processes which influence the way collocations are formed: Some implications for the learning of collocations. *TESOL Quarterly, 45*(2), 291-312.

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge: Cambridge University Press.

Wray, A. (2008). *Formulaic language: Pushing the boundaries.* Oxford: Oxford University Press.

Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly, 43*, 55–77.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning, 63*, 91–120.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus in L2 writing. *Journal of Second Language Writing, 13,* 257-283.