

Effectiveness of Data-Driven Learning in Reducing Phrasal Verb-related Errors among Intermediate English as a Foreign Language Learners

¹Ali Beikian*

²Saeed Esmailnia

Research Paper

IJEAP-2303-1950

Received: 2023-01-12

Accepted: 2023-03-15

Published: 2023-03-30

Abstract: Phrasal verbs (PVs) are a challenging aspect of English language learning, particularly for English as a Foreign Language (EFL) learners. The current study investigated whether a data-driven learning (DDL) approach could reduce PV-related errors (PVEs) in writing tasks for intermediate EFL learners. In doing so, it employed an experimental design with 30 female EFL learners aged 13 to 20 randomly assigned to an experimental group (n=15) and a control group (n=15). The Nelson Proficiency Test ensured that both groups had similar proficiency levels. The pre-test and the post-test required the participants to write six paragraphs using six separate sets of five related PVs. Each group received six 90-minute instruction sessions, where the Wikipedia Corpus was used to teach PVs to the experimental group with a DDL approach, and explicit instruction, focusing on memorizing and retaining the meanings of each PV, was used for teaching the same to the control group. The study's findings indicated that the experimental group outperformed the control group in reducing PVEs, demonstrating the effectiveness of the DDL method. Therefore, the study recommends using DDL as a tool for teaching and learning PVs in EFL classrooms, as it promotes learners' ability to use PVs accurately and appropriately.

Keywords: Data-Driven Learning (DDL), Intermediate EFL Learners, Phrasal Verb (PV), Phrasal Verb-related Errors (PVEs), Writing Tasks

Introduction

Phrasal verbs (PVs) have long been recognized as a challenging aspect of English for learners of English as a Foreign Language (EFL). They are so ubiquitous that learners may face one PV per 150 words of a typical English text (Celce-Murcia & Larsen-Freeman, 1999; Gardner & Davies, 2007), thereby underscoring the importance of addressing this issue in language teaching. Several studies have documented the difficulty that learners from different linguistic backgrounds encounter when trying to use PVs (Abdul Rahman et al., 2014; Aldukhayel, 2014; El-Dakhs, 2016; Gilquin, 2015; Liao & Fukuya, 2004).

Numerous challenges are associated with acquiring phrasal verbs (PVs) in second language (L2) learning. One fundamental difficulty stems from the vast number of PVs that exist in the English language, which can be overwhelming for L2 learners. McCarthy and O'Dell (2004) have estimated that there are almost 5000 PVs in English, a daunting prospect for learners who may prefer to use single-word synonyms (Siyanova & Schmitt, 2007). Another factor that makes PVs challenging for L2 learners is their language-specific nature. PVs are a unique feature of specific languages, including English, which means that learners whose first language lacks the verb + particle construction may find it difficult to acquire PVs (Hulstijn & Marchena, 1989; Laufer & Eliasson, 1993). PVs are also composed of orthographic words treated as a

¹ Assistant Professor of Translation Studies (Corresponding Author), a_beikian@yahoo.co.uk; English Department, Chabahar Maritime University, Chabahar, Iran.

² MA Graduate of TEFL, saeedesmailnia@gmail.com; English Department, Chabahar Maritime University, Chabahar, Iran.

single unit semantically. This construction may lead learners to process the meaning of the individual components of PVs separately, which can result in a misunderstanding of their overall meaning (Arnon & Christiansen, 2017; Garnier & Schmitt, 2016). In addition, PVs are difficult to interpret due to their unpredictability and non-compositionality, making them challenging for learners to use in context (Celce-Murcia & Larsen-Freeman, 1999; Cornell, 1985; Moon, 1997). Moreover, PVs are both syntactically and semantically complex. Syntactically, they differ in their allowance of intervening words between a verb and particle and, if so, how many they permit (Gardner & Davies, 2018). Semantically, PVs can vary in degrees of idiomaticity, with some having more transparent or opaque meanings. Finally, PVs are highly polysemous, with a large number of PVs carrying multiple meanings, further complicating their use in context (Gardner & Davies, 2007; Garnier & Schmitt, 2014). These various factors contribute to the challenges of acquiring and using PVs in L2 learning.

In the realm of English language learning, PVs pose a significant challenge for EFL learners, who often choose to avoid them (Hulstijn & Marchena, 1989; Laufer & Eliasson, 1993; Liao & Fukuya, 2004). In contrast, native English speakers seem to effortlessly create new PVs (Zareva, 2016). As a result, numerous scholarly articles have been written to discuss the semantic and syntactic complexities of PVs (Dehé, 2002; Jackendoff, 2002). Despite this, their presentation in textbooks often emphasizes their arbitrary nature, suggesting an unsystematic way of teaching them (Cornell, 1985; Darwin & Gray, 1999; Gardner & Davies, 2007; Moon, 1997).

Various approaches have been employed to teach PVs to EFL learners, including explicit instruction, implicit instruction, a combination of both, categorization of PVs by particle, verb, or meaning to aid learners in identifying patterns and generalizing their knowledge, the use of computer-assisted language learning (CALL) tools such as interactive games and quizzes, and data-driven learning (DDL). DDL, in particular, is an approach that has gained popularity in recent years. It is "the attempt to cut out the middleman as far as possible and to give the learner direct access to the data" (Johns, 1994, as cited in Azzaro, 2012, p. 3). The spirit of DDL is to help students engage in direct quality textual analysis, fostering independent learning of collocations and allowing learners to use corpora and tools to search for significant collocations personally. Boulton (2009, p. 38) summarizes DDL's core pedagogical motivation: "It may even be, in some cases, that learning is more effective without a teacher, i.e., when learners discover things for themselves."

With the challenges posed by PVs, intermediate EFL learners often struggle to write accurately. As such, the current study aims to investigate the potential of DDL to reduce PV-related errors (PVEs) resulting from these challenges. Specifically, the study addresses whether DDL can help reduce intermediate EFL learners' PVEs in writing tasks. By exploring the effect of DDL on the reduction of PVEs, this study can contribute to improving the quality of writing among EFL learners. The findings may also have important implications for English language teaching and learning, particularly in the context of phrasal verb usage. If DDL proves effective in reducing errors resulting from the incorrect use of PVs, it could be incorporated into language teaching methodologies and curriculum development.

Although previous research has explored the effectiveness of DDL in various language-learning contexts, a gap in the literature exists regarding its specific application to PVs and the challenges they pose to EFL learners in writing tasks. Furthermore, this study's novelty lies in its focus on a group of intermediate EFL learners in Iran, which has received relatively little attention in the literature on DDL and PVs.

Literature Review

Teaching language to intermediate EFL learners is a considerable challenge for teachers. These learners often have gaps in their knowledge that need to be filled, and their progress can be less noticeable than that of beginners, leading to a lack of motivation to learn. According to Xue (2021), the primary reason is that

many intermediate learners have not been trained to recognize which words go together meaningfully, which hinders their ability to use words in various collocations (Hsu & Chiu, 2008; Wang & Yang, 2020). Phrasal verbs (PVs), which are a subcategory of collocations, are among the most challenging word constructions for learners (Celce-Murcia & Larsen-Freeman, 1999; Cornell, 1985; Moon, 1997; Garnier & Schmitt, 2016; Laufer & Eliasson, 1993; Gardner & Davies, 2018). The traditional method of teaching PVs has been problematic (Ganji, 2011; Talebinezhad & Farhadian, 2014; Taheri & Zarei, 2021). Therefore, developing new and more effective methods for teaching PVs is essential, especially for teaching intermediate EFL learners, to enable them to learn PVs systematically and use them more efficiently with their meanings.

According to Boulton and Cobb's (2017) meta-study, which analyzed 64 empirical studies, DDL is a valuable tool in foreign language contexts for graduate and undergraduate students and individuals at intermediate and advanced proficiency levels. Moreover, DDL can be used effectively for both general and specific/academic purposes and is ideal for examining both small and large corpora, whether learners prefer hands-on concordancing or paper-based exploration. DDL is equally valuable for learning new material and as a reference resource, and it is particularly beneficial for vocabulary and lexicogrammar (Boulton & Cobb, 2017). By analyzing large corpora of authentic language data, educators can identify the most frequently used PVs in different contexts and provide learners with relevant examples of their usage. The field of DDL has been thriving, with a plethora of empirical research, reviews, syntheses, and meta-studies conducted, all indicating the value of DDL for language learning, with meta-studies showing an overall positive effect size (Boulton & Cobb, 2017; Cobb & Boulton, 2015; Lee et al., 2019; Mizumoto & Chujo, 2015). These findings demonstrate an enduring enthusiasm and conviction among researchers about the effectiveness of DDL as a learning aid, as well as an aspiration for it to become more mainstream.

In addition, studies have shown that DDL significantly impacts language learning outcomes. For instance, Mizumoto and Chujo (2015) found that DDL can improve vocabulary acquisition and reading comprehension skills, while Vyatkina and Boulton (2017) demonstrated that DDL could enhance learners' ability to use the target language for specific communicative purposes. Furthermore, DDL is an engaging and motivating approach to language learning, as learners are encouraged to explore language data and discover patterns and regularities independently (Boulton & Pérez-Paredes, 2014). This approach aligns with contemporary views of language learning that emphasize the importance of learner autonomy and the development of language awareness (Lee et al., 2019).

Scholars have extensively investigated the effectiveness of using DDL with the help of corpora in language teaching (Ashkan & Seyyedrezaei, 2016; Barabadi & Khajavi, 2017; Koosha & Jafarpour, 2006; Paker & Ergül-Özcan, 2017; Tekin & Soruç, 2016; Uçar & Yükselir, 2015; Yılmaz & Soruç, 2015). While numerous studies have explored learners' acquisition of English collocations through corpus-based approaches, limited attention has been given to using corpora in teaching PVs (Girgin, 2019). A review of the existing literature on this topic reveals only four previous studies reporting the use of corpora for teaching PVs, as follows.

Azzaro's (2012) study compared dictionary presentation and concordance-based DDL techniques for teaching PVs. The results indicated that DDL could effectively teach new lexical items and reinforce old ones. Similarly, Girgin (2019) investigated the impact of corpus-based activities on teaching PVs to Turkish upper-intermediate teacher candidates. The results showed that students could recognize the form of the verbs and construct correct forms for paraphrasing after the six-hour instruction via corpus-based activities. Troy and Millar (2019) concluded that a DDL approach could aid in teaching PVs, although creating materials could be time-consuming. The study suggested that implementing a DDL course for private language schools might be beneficial, but only if the teacher or the organization had time to create materials that do the methodology justice. Finally, Tsulaia (2021) argued that corpus-based language teaching effectively teaches PVs to EFL students. The study reviewed the advantages of corpus-based language teaching and introduced the valuable tools of the SKELL corpus for teaching English PVs.

To sum up, although corpora are widely recognized as a valuable resource in language teaching, research on the effectiveness of using corpora specifically for teaching PVs remains limited. Therefore, further research is necessary to better understand corpora's efficacy in teaching PVs and to explore ways to maximize their use in language teaching. Therefore, the current research is guided by the following research question:

Research Question One: Is DDL effective in reducing PV-related errors (PVEs) among intermediate EFL learners in writing tasks?

Methodology

Design

The present study utilized an experimental design consisting of two groups: an experimental group and a control group. The study aimed to investigate the effectiveness of the DDL method compared to explicit instruction, which focuses on memorizing and retaining the meanings of each PV. The independent variable was the teaching method, with the experimental group receiving instruction using the DDL method and the control group receiving explicit instruction. The dependent variable was the outcome measure.

Participants

The study population was female intermediate EFL learners from Rahamouz Shokouh Institute in Qaem Shahr. The participants were aged between 13 to 20 years old. Among these intermediate EFL learners, 74 students were selected using voluntary sampling to take a language proficiency test. Based on their test scores, 30 participants were selected and then randomly assigned to the experimental group (n=15) and the control group (n=15).

Instruments

This section describes the instruments used in our study, including their design, administration, and validation procedures.

Nelson Proficiency Test

The Nelson Proficiency Test was given to both groups to ensure that the experimental and control groups' competence levels were uniform. This test is regarded as highly valid since its validity and reliability have been evaluated by various researchers on several occasions (Shahivand & Pazhakh, 2012).

List of 30 PVs

A list of 30 PVs divided into six sets of five PVs was used as the basis for the pre-test and post-test. In choosing PVs for each set, it was essential to consider their relatedness in meaning and context, as well as their intended use and purpose. Five PVs per set provided enough variety and context for writing paragraphs while still being manageable and focused. With this in mind, we used ChatGPT to choose six independent sets of five related PVs for writing a coherent paragraph that flows naturally and communicates ideas effectively. This list, including phrasal verbs used in example paragraphs written by the researchers, has been provided in Appendix A.

ChatGPT

ChatGPT is part of the Generative Pre-trained Transformer (GPT) family of language models, designed to generate human-like text by predicting the next word in a sequence of words.

Pre-test

The pre-test required the participants to write six paragraphs using the six PV sets.

Post-test

The post-test required the participants to write six paragraphs using the same PV sets used in the pre-test.

Wikipedia Corpus

The Wikipedia Corpus, used for teaching PVs with a DDL approach, comprises more than 4 million articles and 1.9 billion words. Searches in this corpus can be conducted based on the word, phrase, part of speech, synonyms, and collocations.

Data Collection and Analysis Procedure

The Nelson Proficiency Test determined the homogeneity of the participants' language proficiency levels. The test was administered to 74 intermediate EFL learners, and their scores were calculated using the SPSS software to compute means and standard deviations in descriptive statistics. From this test, 30 participants were selected based on their scores' closeness to the mean on the minimal curve. Then in the pre-test phase, both groups were given the list of 30 PVs (see 3.3.2) and asked to write six paragraphs using each set of five PVs. It must be noted that the topics of the paragraphs were open to their choices. Then the pre-test writings were scored, and the PVEs were recorded.

After that, six 90-minute sessions of treatment were executed for both groups. The following steps were taken for the experimental group to use the Wikipedia Corpus for teaching PVs with a DDL approach: 1. selection of relevant data: a subset of the Wikipedia Corpus was selected that included texts in English containing examples of PVs included in the list of 30 PVs. The selection criteria were based on the frequency of PVs (see 3.3.2) in the text, as well as the representativeness of the texts for different genres and registers; 2. design of learning materials: based on the selected texts, a set of learning materials was designed that included exercises and activities to help learners identify and use PVs in context. The materials were organized around different aspects of PVs, such as their form, meaning, and use, and included examples and explanations of common patterns and collocations; 3. use of analytical tools: to identify patterns and examples of PVs in context, a range of analytical tools were used, including concordancers, collocation analysis software, and corpus query tools. These tools allowed for identifying common patterns and collocations of PVs and extracting examples of PVs in context; and 4. instructions for using the Wikipedia Corpus: before starting the learning activities, the participants in the experimental group were given instructions on how to use the Wikipedia Corpus for learning PVs. Specifically, they were told to search for and read articles in the selected subset of the corpus that contained examples of the target PVs. They were encouraged to pay attention to how the PVs were used in context and to take notes on any patterns or collocations they noticed. They were also given guidance on how to use the analytical tools (mentioned in step 3) to aid their learning and were given time to practice using them with guidance from the teacher. Finally, they were reminded to use the learning materials provided to reinforce their learning and to practice using the PVs in different contexts.

For the control group, to teach PVs using explicit instruction, the following steps were taken: 1. Selection of materials: the materials used for the control group included the list of PVs and their meanings (see 3.3.2), which were organized by topic or theme, and included examples of PVs in context; 2. instructional procedures: the participants were exposed to the PVs and their meanings through reading and repetition. Specifically, the researcher read the PVs and their meanings aloud to the participants, who then repeated them and were required to memorize them; and 3. practice: to reinforce their learning, the participants were required to practice using the PVs in context. They completed exercises and activities based on the lists, which provided opportunities to use the PVs in sentences and paragraphs. The practice activities were designed to reinforce the participants' knowledge of the PVs and their meanings. Following the six treatment sessions, in the post-test phase, the participants in both groups were asked to write another set of six paragraphs using the same six PV sets used in the pre-test. The post-test writings were then scored, and their PVEs were recorded to be compared to those in the pre-test.

Data analysis

The obtained scores of the participants were calculated using SPSS software to compute their means and standard deviations in descriptive statistics. The normal distributions of the scores were depicted in bell-shaped diagrams. The obtained results were compared using t-tests to investigate the differences in the mean scores of the two groups. The significance level was set at $p < 0.05$. Finally, the PVEs made by the participants in the pre-test and post-test were analyzed and compared to investigate the effectiveness of the DDL method.

Results

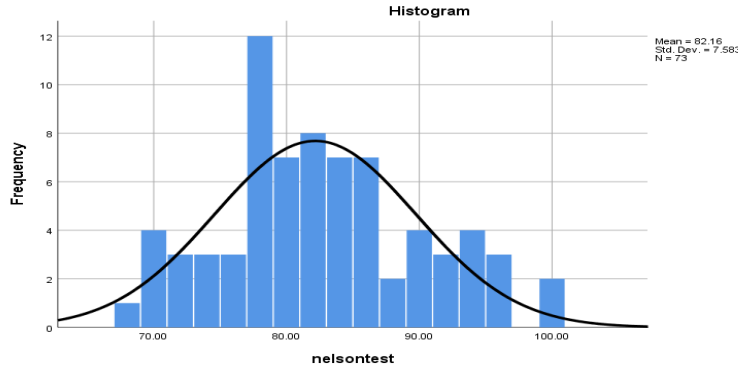
This section presents the study's findings on the effectiveness of the DDL approach in teaching PVs compared to explicit instruction, which focuses on memorizing and retaining the meanings of each PV. The section begins with an overview of the participants' homogeneity, followed by a description of the results of the pre-test and post-test assessments. Finally, the results are presented regarding the participants' improvement in their knowledge and use of PVs.

Participants' Homogeneity

In order to ensure homogeneity among the participants, the Nelson Proficiency Test was administered, and scores were analyzed using SPSS version 25 to calculate the variance and standard deviation. Descriptive statistics were used to calculate the mean, median, mode, standard deviation, and variance of the scores, as illustrated in Table 2. The mean score was 82.1644, with a standard deviation of 7.58 and a variance of 57.500. Participants with scores closer to the mean were selected as homogeneous for the study. In addition, the One-Sample Kolmogorov-Smirnov Test was employed to verify the normality of the data. A significance level of 0.05 was used, with scores exhibiting a significance level above this threshold deemed normal. The normal distribution curve on the histogram (Figure 1) provides further evidence of the normality of the data, with scores between 75 and 85 being the most frequent.

Figure 1

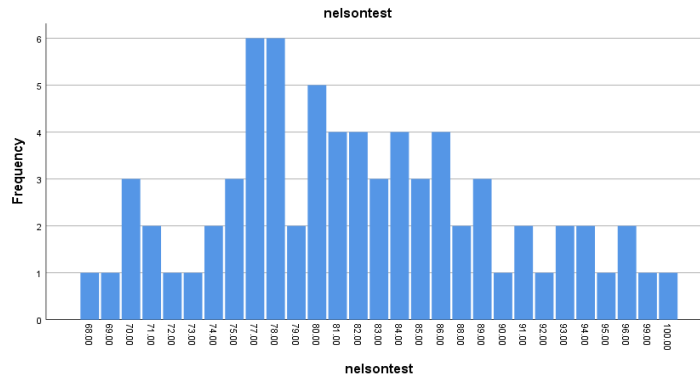
Normality of the Nelson Test Scores by Histogram



The bar chart diagram (Figure. 2) provides detailed information on the scores and their dispersion, further supporting the normality of the data. This diagram also demonstrates the range of scores the researcher used to select homogeneous participants.

Figure 2

Normality of the Nelson Test Scores by Bar-Chart Diagram



Using the Nelson Proficiency Test, we chose 30 homogenous participants, indicating they had comparable characteristics relevant to the study's research question. Once the participants were selected, they were randomly divided into the control group (n=15) and the experimental group (n=15).

Descriptive Statistics

This study aimed to investigate DDL's effectiveness in reducing intermediate EFL learners' PVEs in their writing tasks. To achieve this, a pre-test was given to both the experimental and control groups to assess their initial level of proficiency in using PVs in their writing tasks. After the pre-test, both groups underwent six sessions of treatment. The experimental group was taught using a DDL approach, and the control group was taught using explicit instruction. After the treatment, a post-test was given to both the experimental and control groups to assess their final level of proficiency in using PVs in their writing tasks. Table 1

provides a detailed breakdown of the pre-test and post-test results, including the number of PVEs made by both groups in their writing tasks.

Table 1

Control and Experimental Groups' Pre- and Post-test PVEs

Set	Phrasal Verbs	Control Group		Experimental Group	
		Pre-Test PVEs	Post-Test PVEs	Pre-Test PEVs	Post-Test PVEs
1	jot down	9.00	9.00	7.00	5.00
	cross out	6.00	5.00	4.00	0.00
	fill in	10.00	11.00	8.00	4.00
	look up	8.00	9.00	5.00	3.00
	set aside	4.00	3.00	1.00	0.00
2	cut up	5.00	6.00	5.00	1.00
	mix in	15.00	13.00	10.00	5.00
	boil over	6.00	5.00	3.00	0.00
	simmer down	4.00	4.00	4.00	1.00
3	plate up	2.00	10.00	2.00	3.00
	warm up	6.00	7.00	3.00	1.00
	work out	10.00	9.00	7.00	2.00
	cool down	5.00	6.00	3.00	2.00
	build up	8.00	8.00	7.00	3.00
4	tone up	7.00	6.00	7.00	5.00
	check in	4.00	3.00	4.00	0.00
	take off	1.00	1.00	0.00	0.00
	get around	3.00	3.00	2.00	0.00
	look around	4.00	5.00	2.00	4.00
5	set off	8.00	7.00	8.00	5.00
	follow up	6.00	6.00	5.00	5.00
	hand over	11.00	10.00	9.00	3.00
	take on	5.00	6.00	3.00	1.00
	turn down	5.00	5.00	5.00	3.00
6	work out	6.00	8.00	5.00	3.00
	log in	9.00	7.00	8.00	4.00
	back up	7.00	6.00	6.00	2.00
	transfer over	14.00	11.00	11.00	3.00
	set up	15.00	12.00	14.00	10.00
	bring up to date	7.00	8.00	6.00	3.00

As can be seen from Table 1, the experimental group outperformed the control group. Specifically, the experimental group exhibited fewer PVEs in the post-test compared to the pre-test in all six PV sets. This observation implies that the DDL approach was highly effective in addressing the challenges that intermediate EFL learners face in using PVs in their writing tasks. On the other hand, the control group had fewer PEVs in only four out of the six PV sets in the post-test, which indicates that the explicit instruction approach was less effective in reducing PVEs compared to the DDL approach.

Table 2 presents the descriptive statistics for the control group's pre-and post-tests. Prior to the implementation of the treatment, the mean score in the control group was 7.0000. In contrast, after the treatment, the mean score decreased to 5.4667, indicating reduced errors related to phrasal verb usage. The employment explicit instruction in the control group appears to have contributed to a decrease in PVEs.

Table 2

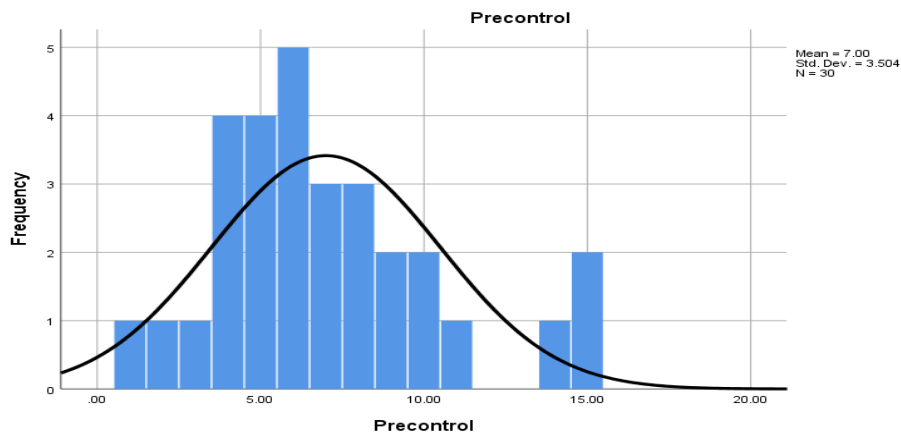
Descriptive Statistics for the Control Group

		Pretest	Posttest
N	Valid	30	30
	Missing	0.00	0.00
Mean		7.0000	5.4667
Std. Error of Mean		.63968	.57081
Median		6.0000	5.0000
Mode		6.00	5.00
Std. Deviation		3.50369	3.12645
Variance		12.276	9.775
Skewness		.819	.658
Std. Error of Skewness		.427	.427
Kurtosis		.495	.592
Std. Error of Kurtosis		.833	.833
Range		14.00	14.00
Minimum		1.00	0.00
Maximum		15.00	14.00
Sum		210.00	164.00

Figures 3 and 4 provide empirical evidence to support the normal data distribution. The figures highlight the normal distribution curves for both pre- and post-control groups. The curves indicate that the data adheres to a normal distribution, essential for conducting various statistical analyses. This observation highlights the accuracy and reliability of the data gathered in the study.

Figure 3

Normal Distribution Curve for the Pre-control Group



Moreover, Figure 4 shows a decline in PVEs in the post-control group, indicating the effectiveness of explicit instruction.

Figure 4

Normal Distribution Curve for the Post-control Group

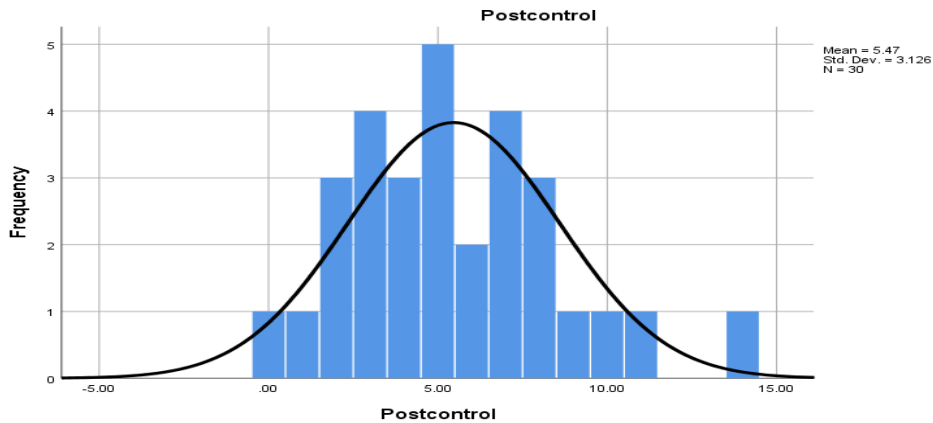


Table 3 provides an overview of the statistical measures obtained for the experimental group. Specifically, the mean value for the pre-test is 6.9667, whereas, for the post-test, it is 2.7000. This significant decrease in the mean value indicates that the use of the Wikipedia Corpus had a remarkable impact on the participants' ability to learn the appropriate usage of the given PVs. In other words, the experimental group significantly reduced the number of PVEs made in the post-test compared to the pre-test. Furthermore, the standard error of the mean, which is .52628 and .40443 for pre and post-tests, respectively, also suggests a substantial reduction in the number of PVEs made by the participants in the post-test. A lower standard error of mean implies that the data points in the sample are closer to the sample mean, which is a positive indicator of the effectiveness of the DDL approach used in this study.

Table 3

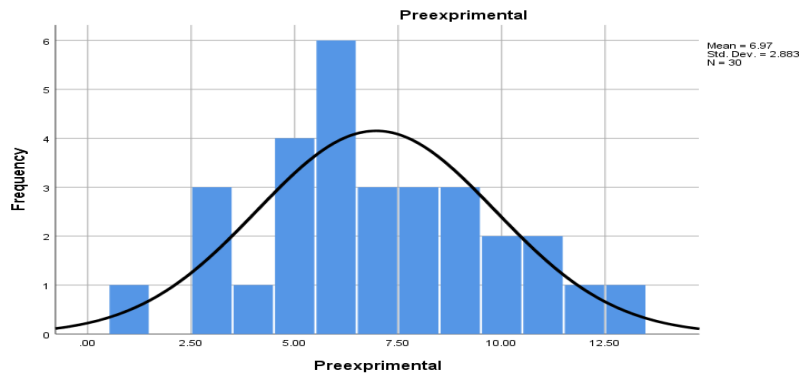
Descriptive Statistics for the Experimental Group

		Pretest	Posttest
N	Valid	30	30
	Missing	0	0
Mean		6.9667	2.7000
Std. Error of Mean		.52628	.40443
Median		6.5000	3.0000
Mode		6.00	3.00
Std. Deviation		2.88257	2.21515
Variance		8.309	4.907
Skewness		.156	1.080
Std. Error of Skewness		.427	.427
Kurtosis		-.343	2.549
Std. Error of Kurtosis		.833	.833
Range		12.00	10.00
Minimum		1.00	.00
Maximum		13.00	10.00
Sum		209.00	81.00

Figures 5 and 6 provide insights into the normality of the data collected and the effectiveness of the treatment employed in the study. These figures offer a graphical representation of the data, allowing for a more in-depth understanding of the changes due to the DDL approach used as a treatment method. The curves in these figures demonstrate a clear shift towards the left in the post-test data for the experimental group, indicating a significant decrease in the number of PVEs made by the participants after the treatment.

Figure 5

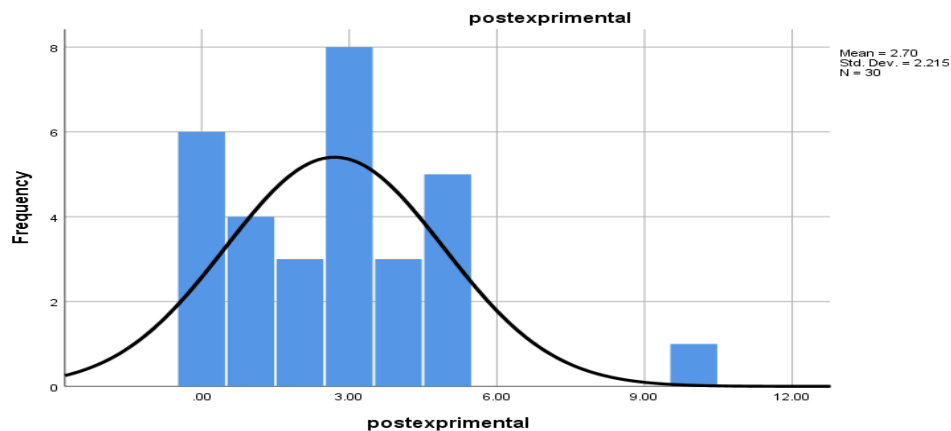
Normal Distribution Curve for the Pre-Experimental Group



Moreover, Figure 6 provides information on the degree of change due to the treatment. The curves indicate a more significant standard deviation in the pre-test data, which decreased in the post-test data, indicating a reduction in the variability of the data. This observation suggests that the DDL approach had a consistent impact on all participants in the experimental group, resulting in a more consistent improvement in their use of PVs.

Figure 6

Normal Distribution Curve for the Post-experimental Group



Inferential Statistics

Table 4 reports on the paired samples statistics conducted to compare the post-control and post-experimental measures in the study, which evaluated the effectiveness of the DDL method in improving participants' use of PVs. The paired samples t-test was utilized to compare the means of the two measures, with the results indicating a statistically significant difference between them. The mean score for the post-control measure was 5.4667, with a standard deviation of 3.12645 and a standard error of the mean of .57081. In contrast, the mean score for the post-experimental measure was 2.7000, with a standard deviation of 2.21515 and a standard error of the mean of .40443. These results suggest that the DDL treatment effectively reduced PVEs, as the mean score for the post-experimental measure was significantly lower than that of the post-control measure.

The difference in means between the two measures was also statistically significant, with a t-value of 7.086 and a p-value of < .001. This further supports the conclusion that the DDL method effectively improved the participants' performance in using PVs. In summary, the paired samples statistics presented in Table 4 provide robust evidence for the effectiveness of the DDL approach in improving the use of PVs by the participants. The statistical analysis suggests a significant difference between the post-control and post-experimental measures, with the latter demonstrating a significant decrease in PVEs.

Table 4

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Post-control	5.4667	30	3.12645	.57081
	Post-experimental	2.7000	30	2.21515	.40443

Table 5 presents the results of the paired samples t-test to compare the mean scores of post-control and post-experimental measures. The table shows that the mean difference between the two measures is 2.76667, with a standard deviation of 2.06253 and a standard error mean of 0.37656. The 95% confidence interval of the difference is between 1.99651 and 3.53683. These statistics suggest that the DDL method treatment approach effectively reduced PVEs among the participants.

Table 5

Paired Samples Test for Post-control and Post-experimental Measures

	Paired Differences							
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Pair 1:	2.76667	2.06253	.37656	1.99651	3.53683	7.347	29	.000

The t-value of 7.347 with 29 degrees of freedom indicates that the difference in means is statistically significant at the 0.05 level of significance. The p-value of 0.000 is less than the alpha level of 0.05, indicating that the results are statistically significant. These findings provide strong evidence that the treatment approach using the DDL method effectively improved participants' performance in using PVs.

In conclusion, the paired sample statistics in Table 5 support the hypothesis that the DDL method is effective for teaching PVs to ESL learners. The results suggest that the DDL approach significantly improved participants' performance using PVs, as evidenced by the significant difference in mean scores between the post-control and post-experimental measures.

Discussion

The present study investigated whether a DDL approach could effectively reduce PVEs in writing tasks for intermediate English as a Foreign Language (EFL) learners. The study found that the experimental group, which received instructions using a DDL approach, outperformed the control group in reducing PVEs. These findings align with previous studies that have examined the effectiveness of DDL in teaching PVs.

Azzaro's (2012) study showed that DDL was more effective than dictionary presentation in teaching new lexical items and reinforcing old ones. Likewise, Troy and Millar (2019) concluded that a DDL approach could aid in teaching PVs, though creating materials could be time-consuming. This suggests DDL is a viable approach for teaching PVs in EFL classrooms. Girgin (2019) also found that using corpus helped students recognize the form of the verbs and construct the correct forms for paraphrasing. Furthermore, Tsulaia (2021) argued that corpus-based language teaching is effective for teaching PVs to EFL learners.

The present study's findings align with previous studies examining DDL's effectiveness in teaching PVs. Nonetheless, the present study contributes to the existing literature by proving that a DDL approach can effectively reduce PVEs in writing tasks for intermediate EFL learners. This highlights the importance of EFL teachers considering using a DDL approach in teaching PVs, particularly for intermediate learners.

Conclusion

This study investigated the effectiveness of DDL in teaching phrasal verbs (PVs) to intermediate EFL learners. The findings suggest that DDL can effectively reduce PVEs in writing tasks. However, this study has some limitations that future research should consider. Firstly, the study only focused on PV use in writing tasks, limiting the findings' generalizability. Secondly, the study only included intermediate EFL learners, which may not reflect the effectiveness of DDL for other proficiency levels. Thirdly, the study was conducted in a controlled environment, which may not reflect the complexity of real-world language use. Finally, the study only used one corpus, which may limit the generalizability of the findings to other corpora. Future research could investigate the long-term impact of DDL, explore other language skills besides writing, examine different types of PVs, test DDL in naturalistic settings, and compare the effectiveness of DDL with other corpora. Despite these limitations, this study recommends DDL as an effective tool for teaching and learning PVs in EFL classrooms.

The practical implications of this study are significant for EFL teachers and learners. Teachers can use DDL to enhance learners' PV proficiency by introducing them to corpora and providing them with opportunities to practice using PVs in different contexts. DDL can also help learners become more independent and effective, as they can use corpora to identify and correct errors in their writing and develop their language skills outside the classroom. Learners who struggle with PVs can benefit from DDL-based instruction as it helps them identify patterns and use them in context, leading to the more accurate and appropriate use of PVs in writing. This study also has theoretical implications, adding to the literature on the effectiveness of DDL in promoting language learning, particularly in the area of PVs. The findings contribute to understanding the role of corpora in language learning, as it demonstrates the potential of using large-scale digital corpora in EFL classrooms. Furthermore, the study provides insight into the

importance of language patterns and their benefits in context, which can enhance learners' understanding and use of PVs.

Acknowledgment

We want to express our heartfelt gratitude to Dr. Hooshang Khoshshima, Professor of TEFL at Chabahar Maritime University, for his invaluable guidance and support throughout this project. Dr. Khoshshima's expertise in TEFL has shaped our research and provided valuable insights.

Declaration of Conflicting Interests

The authors declare that they have no conflicting interests related to the research presented in this manuscript. They also confirm that they have no direct or indirect financial interest in any product, service, or company that could be perceived as influencing the outcome of this research.

Funding Details

The authors declare that they received no funding for this research.

References

- Abdul Rahman, Z. A., & Abid, R. Z. (2014). Rarity or non-existence of phrasal verbs in the written discourse of Omani student-teachers of English. *SAGE Open*, 4(4), 215824401455698.
- Aldukhayel, D. M. (2014). *exposure effect on avoidance of phrasal verbs by Arab ESL learners* [Unpublished doctoral dissertation].
- Arnon, I., & Christiansen, M. H. (2017). The role of Multiword building blocks in explaining L1-L2 differences. *Topics in Cognitive Science*, 9(3), 621-636.
- Ashkan, L., & Seyyedrezaei, S. H. (2016). The effect of corpus-based language teaching on Iranian EFL learners' vocabulary learning and retention. *International Journal of English Linguistics*, 6(4), 190.
- Azzaro, G. (2012). Phrasal Verbs through DDL. *Journal of Theories and Research in Education* 7 (2), 1–23.
- Barabadi, E., & Khajavi, Y. (2017). The effect of data-driven approach to teaching vocabulary on Iranian students' learning of English vocabulary. *Cogent Education*, 4(1), 1283876.
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(1), 37–54.
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393.
- Boulton, A., & Pérez-Paredes, P. (2014). ReCALL special issue: Researching uses of corpora for language teaching and learning editorial researching uses of corpora for language teaching and learning. *ReCALL*, 26(2), 121-127.
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1999). *The grammar book: An ESL/EFL teacher's course*. Newbury House.

- Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching*, 52(4), 460–475.
- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *Cambridge handbook of corpus linguistics* (pp. 478–497). Cambridge, UK: Cambridge University Press.
- Cornell, A. (1985). Realistic goals in teaching and learning phrasal verbs. *IRAL - International Review of Applied Linguistics in Language Teaching*, 23(1–4).
- Darwin, C. M., & Gray, L. S. (1999). Going after the phrasal verb: An alternative approach to classification. *TESOL Quarterly*, 33(1), 65.
- Dehé, N. (2002). *Particle verbs in English: Syntax, information structure and intonation*. John Benjamins Publishing.
- El-Dakhs, D. A. S. (2016). The lexical knowledge and avoidance of phrasal verbs: The case of Egyptian learners of English. *International Journal of Applied Linguistics and English Literature*, 5(1), 132–144.
- Ganji, M. (2011). The best way to teach phrasal verbs: Translation, sentential contextualization or metaphorical conceptualization? *Theory and Practice in Language Studies*, 1(11).
- Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL Quarterly*, 41(2), 339–359.
- Gardner, D., & Davies, M. (2018). Sorting them all out: Exploring the separable phrasal verbs of English. *System*, 76, 197–209.
- Garnier, M., & Schmitt, N. (2014). The PHaVE list: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research*, 19(6), 645–666.
- Garnier, M., & Schmitt, N. (2016). Picking up polysemous phrasal verbs: How many do learners know and what facilitates this knowledge? *System*, 59, 29–44.
- Gilquin, G. (2015). The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. *Corpus Linguistics and Linguistic Theory*, 11(1).
- Girgin, U. (2019). undefined. *Language Teaching and Educational Research*.
- Hsu, J. Y., & Chiu, C. Y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *Asian EFL Journal*, 10(1), 181–204.
- Hulstijn J.H., Marchena E. (1989). Avoidance: Grammatical or semantic causes? *Studies in Second Language Acquisition*, 11, 241–255.
- Jackendoff, R. (2002). English particle constructions, the lexicon, and the autonomy of syntax. In N. Dehé, R. Jackendoff, A. McIntyre & S. Urban (Eds.), *Verb-particle explorations* (pp. 67–94). Berlin, Germany: Mouton de Gruyter.
- Kartal, G., & Uner, S. (2017). The effects of conceptual metaphors on the acquisition of phrasal verbs by Turkish EFL learners. *European Journal of Foreign Language Teaching*, 2(2), 34–51.
- Koosha, M., & Jafarpour, A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal Quarterly*, 8(4), 192–209.
- Laufer B., Eliasson S. (1993). What causes avoidance in L2 learning: L1-L2 difference, L1-L2 similarity, or L2 complexity? *Studies in Second Language Acquisition*, 15, 35–48.

- Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*, 40(5), 721–753.
- Liao, Y., & Fukuya, Y. J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning*, 54, 193–226.
- McCarthy M., O'Dell F. (2004). *English phrasal verbs in use*. Intermediate Level. Cambridge: Cambridge University Press.
- Mizumoto, A., & Chujo, K. (2015). A meta-analysis of data-driven learning approach in the Japanese EFL classroom. *English Corpus Studies*, 22, 1–18.
- Moon, R. (1997). Vocabulary connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 40–63). Cambridge, MA: Cambridge University Press.
- Paker, T., & Ergül-Özcan, Y. (2017). The effectiveness of using corpus-based materials in vocabulary teaching. *International Journal of Language Academy*, 5(1), 62-81.
- Shahivand, Z., & Pazhakh, A. (2012). The effects of test facets on the construct validity of the tests in Iranian EFL students. *Higher Education of Social Science*, 2(1), 16–20.
- Siyanova A., Schmitt N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *International Review of Applied Linguistics*, 45, 109–139.
- Taheri, S., & Zarei, Z. (2021). The Effect of Container Metaphor Awareness on Learning Phrasal Verbs by Iranian Intermediate EFL learners. *Journal of Language, Culture, and Translation*, 4(1), 147-165.
- Talebinezhad, M. R., & Farhadian, N. (2014). A comparative study of two cognitive models in teaching idiomatic phrasal verbs: Tyler and Evan's vs. Lakoff and Johnson's approach. *Theory and Practice in Language Studies*, 4(8), 1621-1627.
- Tekin, B., & Soruç, A. (2016). Using corpus-assisted learning activities to assist vocabulary development in English. *TOJET: The Turkish Online Journal of Educational Technology, Special Issue*, 1270–1283.
- Troy, H. R., & Millar, N. (2019). Evaluating the Use of Data-driven Learning to Teach Phrasal Verbs in a Private Language School in Japan. *Bulletin of Nagoya University of Foreign Studies*, 5, 321–347.
- Tsulaia, N. (2021). Corpus-based teaching of English phrasal verbs. *The International Scientific and Pedagogical Organization of Philologists "East-West."* 7(1), 83–88.
- Uçar, S., & Yükselir, C. (2015). The effect of corpus-based activities on verb-noun collocations in EFL classes. *TOJET: The Turkish Online Journal of Educational Technology*, 14(2), 195–205.
- Vyatkina, N., & Boulton, A. (Eds.). (2017). Corpora in language teaching and learning. *Language Learning & Technology*, 21(3), 1–8.
- Wang, C., & Yang, J. (2020). Revisiting the Explicit Learning of Vocabulary of Chinese EFL Learners. *English Language Teaching*, 13(2), 86–96.
- Xue, L. (2021). Using data-driven learning activities to improve lexical awareness in intermediate EFL learners. *Cogent Education*, 8(1), 1–17.
- Yılmaz, E., & Soruç, A. (2015). The use of concordance for teaching vocabulary: A data-driven learning approach. *Procedia-Social and Behavioral Sciences*, 191, 2626–2630.
- Zareva, A. (2016). Multi-word verbs in student academic presentations. *Journal of English for Academic Purposes*, 23, 83–98.

Appendix A – List of 30 Phrasal Verbs

A list of 30 phrasal verbs, divided into six sets of five phrasal verbs, was used as the basis for the pre-test and post-test. In choosing phrasal verbs for each set, it was important to consider their relatedness in meaning and context, as well as their intended use and purpose. Five phrasal verbs per set provided enough variety and context for writing paragraphs while still being manageable and focused. With this in mind, we used ChatGPT to choose six independent sets of five related phrasal verbs for writing a coherent paragraph that flows naturally and communicates ideas effectively. This list, including phrasal verbs used in example paragraphs written by the researchers, is as follows.

Set 1:

1. Jot down - to quickly write something down
2. Cross out - to draw a line through something to indicate that it should be removed
3. Fill in - to add information to a blank space or a form
4. Look up - to find information in a reference source, such as a dictionary or online search engine
5. Set aside - to save or reserve something for later use

Example paragraph:

To write a great essay, it is essential to start by jotting down your ideas in a brainstorming session. Once you have got your ideas down, you can start organizing them into a rough outline, crossing out any irrelevant or repetitive points. As you start writing your essay, you can fill in any gaps with additional research, taking the time to look up any facts or statistics that will support your argument. Finally, it is essential to set aside some time for revising and editing, making sure that your essay is clear and concise.

Set 2:

1. Cut up - to chop food into smaller pieces
2. Mix in - to combine ingredients together
3. Boil over - to overflow from a container due to boiling
4. Simmer down - to reduce the heat on the stove to a low temperature
5. Plate up - to arrange food on a plate in a visually appealing way

Example paragraph:

Cooking a delicious meal requires attention to detail and careful preparation. First, you need to cut up all of your ingredients into small, uniform pieces. Then, you can mix in your spices and other seasonings, making sure that each ingredient is evenly coated. As you cook your dish, you need to be careful to monitor the heat, making sure that your pot does not boil over. If it does, you can simply simmer down the heat to prevent any spills. Once your dish is ready, you can plate up your creation, arranging the food on the plate in a visually pleasing way.

Set 3:

1. Warm up - to prepare the body for exercise by doing light movements
2. Work out - to exercise, often with a specific goal in mind
3. Cool down - to bring the body back to a resting state after exercise
4. Build up - to increase strength or endurance over time
5. Tone up - to develop and strengthen muscles through exercise

Example paragraph:

A good exercise routine should always begin with a proper warm-up, such as stretching or light cardio, to get the blood flowing and loosen up your muscles. Once you are warmed up, you can work out, focusing on specific muscle groups or overall fitness goals. After your workout, it is important to cool down by doing some light stretching or walking, allowing your body to return to a resting state. Over time, you can build up your endurance and strength through consistent exercise, toning up your muscles and improving your overall health.

Set 4:

1. Check in - to register at a hotel or airport
2. Take off - to leave the ground (for airplanes)
3. Get around - to travel from place to place within a city or country
4. Look around - to explore and see what is nearby
5. Set off - to begin a journey or trip

Example paragraph:

When you are traveling, there are a few key things you will need to do to get started. First, you will need to check in at your hotel or airport, making sure that all of your documents are in order. Then, it is time to take off and start your adventure! Whether you are flying to a far-off destination or exploring a new city on foot, you will need to figure out how to get around. Take some time to look around and explore the sights and sounds of your new surroundings. When you are ready to move on, it is time to set off on your next journey, eager to see what lies ahead.

Set 5:

1. Follow up - to contact someone again for additional information or to check the progress
2. Hand over - to give responsibility or control to someone else
3. Take on - to accept a responsibility or task
4. Turn down - to reject an offer or proposal
5. Work out - to resolve a problem or issue

Example paragraph:

In the world of business, there are many challenges and opportunities to navigate. When you are trying to close a deal or make a sale, it is important to follow up with your clients or customers to make sure that they have all the information they need. Sometimes, you may need to hand over responsibility for a project or task to someone else, trusting in their expertise and experience. Other times, you may need to take on additional responsibilities, pushing yourself to achieve new goals and objectives. When an offer or proposal does not quite meet your needs, you may need to turn it down and keep searching for the right fit. With persistence and determination, you can work out any problems or issues that arise, finding new solutions and building strong relationships along the way.

Set 6:

1. Log in - to enter your username and password to access a computer system
2. Back up - to make a copy of important data or files
3. Transfer over - to move data or files from one device to another
4. Set up - to configure software or hardware on your computer or device
5. Bring up to date - to install the latest version of software or firmware to improve functionality and security

Example paragraph:

Technology has become an integral part of our daily lives, and there are many tasks you will need to accomplish to stay up to date. First, you must log in with your username and password to access your computer or device. To ensure that your data is safe, it is important to back up your files regularly, making a copy of your important data or files. If you need to transfer data or files from one device to another, you can easily do so by transferring it over through a cloud-based service or external drive. Once you have your files, you must set up any new software or hardware on your computer or device, configuring it to suit your needs. Finally, to keep your software and firmware up to date, you must bring them up to date by installing the latest updates, which can improve functionality and security.