Abstract: Since the early twenty-first century, data-driven learning (DDL) approach that is a pedagogical application of corpus linguistics in classroom, has introduced a paradigm shift in EFL instruction. Research output, however, concerning this inductive, discovery-oriented learning is equivocal. This study, thus, explored the application of both native-speaker and local learner corpora, attesting the effect of direct vs. indirect DDL activities on 39 EFL learners’ development in CAF measures of writing. To this end, two experimental groups were taught through corpus consultation, but the control group received the conventional method of using a textbook, teacher explanations, and classroom exercises. Results obtained from three (two experimental and one control) groups of participants’ writing performances pre and post to seven sessions of paragraph writing confirmed the significant role of indirect DDL in writing more accurate and fluent paragraphs; however, no statistical evidence was found as regards syntactic complexity. Moreover, no significant effect of the direct DDL method in improving learners’ writing was observed, which is, thus, interpreted as suggestive that applying indirect DDL could be more effective than the direct DDL approach. It is concluded that classroom-based computers are not necessarily essential tools to implement the DDL pedagogy.

Keywords: data-driven learning, corpus linguistics, local learner corpora, CAF measures
1 Introduction

Writing in English is an essential skill for English as a foreign language (EFL) learners, but it also poses great challenges to them, especially those engaging in academic writing. Despite years of effort and exercises, the majority of learners still struggle with one or more aspects of writing, achieving only a basic or below level of performance (Neumann 2014). They typically produce sentences that are less syntactically complex, contain more grammatical errors, are shorter, have higher percentages of capitalization and spelling errors, and mostly are lower in overall quality than those of skilled native writers (Graham and Harris 2005). Moreover, the problem is intensified as many social, academic, and vocational outcomes in today’s world require proficient written expression (Someya 2000). To communicate socially individuals also increasingly need to use electronic correspondence through email or blogs. Accordingly, the majority of Asian EFL learners’ writing competence lags far behind native-like competence (Jafari and Ansari 2012).

In this digital era, technology has been widely used as a resource for the instruction of foreign language, including writing, as educators increasingly recognize its ability to produce both independent and collaborative learning environments (Kern 2006). Alongside the increasing use of technology in language research, the potential of using corpus (a collection of language texts organized and stored on a computer) as a resource in language pedagogy has been evident to practitioners during the last three decades. Corpus Linguistics, i.e., the study of language as expressed in a corpus, has shown tremendous potential in computer-assisted language learning and teaching (Cobb 2010). These affordances have made some researchers and teachers adopt a data-driven learning (DDL) approach. This approach was first proposed by Tim Johns (1991), to provide inductive, discovery-oriented learning opportunities for learners; that is, second language learners analyze corpus acting like a researcher as they are engaged in an active and autonomous process of learning (Chambers 2010).

The role of DDL in language acquisition, as stated by Faerch and Kasper (1987), is to provide examples for consciousness-raising and hypothesis testing. By this means, learners’ declarative knowledge, in the form of generalized schemata of language use (Aston 1997; Skehan 1998), is formed as hypotheses that are confirmed against the authentic language in corpora. DDL can help students increase their ability to process language input efficiently by providing them with corpus data through concordancer programs.

These computerized programs produce concordance lines that are taken from a corpus and presented in Key-Word-in-Contexts (KWICs) format so that students
can get a large amount of information relevant to the target pattern in a limited time. There are generally two different ways of conducting the DDL approach in EFL classrooms, namely, direct DDL and indirect DDL (Yoon and Jo 2014). Direct DDL (hereafter, DDDL) refers to learners’ direct consultation with corpora through concordancers to learn linguistic features or solve their language problems. In contrast, in indirect use of DDL (hereafter, IDDL), learners receive printouts of concordancer outputs extracted by their teachers or researchers.

One aspect of language pedagogy in which corpus usage has made considerable impact is second language writing pedagogy (Yoon 2011). Since 1980 several researchers (e.g., Conrad 2008; Granath 2009) have hailed DDL as a promising approach that can inform language pedagogy in general and L2 writing in particular. Johns (1991) asserts that introducing DDL activities in a writing classroom permits students to study the target language through a usage-based learning approach. This type of learning can help learners to gain and retain lexico-grammatical patterns, which are helpful to express themselves in writing. Furthermore, many studies have evidenced that the DDL approach is useful in helping EFL learners detect and correct lexico-grammatical errors, increases learners’ language awareness and autonomy, and also provides insights into native speakers language data for second language teachers and learners (Charles 2014; Huang 2011; Lee and Swales 2006).

However, in some studies, a number of potential barriers in the approach have also been detailed (e.g. Chambers 2007; Farr 2008). As Boulton (2010) concludes in his review of research on the direct use of corpus in language teaching, DDL seems to be effective, but the findings were promising rather than conclusive regarding their superiority over conventional or non-DDL (henceforth, NDDL) methods of teaching language. Romer (2011) and Tsui (2004) also believe that although the corpus-based pedagogy has been highly regarded in linguistic research, it does not seem to have been well used in ESL teaching.

2 Literature review

Boulton and Cobb (2017) maintain that during the twenty-first century there has been a paradigm shift from the “process-control” approach of instruction, rooted in the communicative approach, to an “information-resource” model in which learners act as researchers and teachers are merely facilitators who help them research effectively, i.e., data-driven learning (Johns 1991). Gavioli and Aston (2001) provide an even more dialectical view of paradigm change when they integrated DDL with the communicative approach. In their point of view,
in institutional inertia has led to the slow diffusion of DDL, which as Rogers (1983) believes, needs to be explored through a systematic and critical appraisal.

Luo and Zhou (2017) believes that in earlier mostly theoretical studies of DDL, authors tried to discuss the theories, for example, the noticing hypothesis (Schmidt 2001), that underpinned DDL. In contrast, Luo (2016) asserts that recent literature has witnessed an increase in practical studies on DDL. The sheer volume of such studies has confirmed the positive effects of DDL on different aspects of language learning. To name a few, Oghigian and Chujo (2010) explored the use of DDL activities on learning grammar basics in English, Thurstun and Candlin (1998) examined the effect of concordancing on the teaching of English academic vocabularies, and Koosha and Jafarpour (2006) studied the influence of DDL on teaching collocations of prepositions. However, as Yoon (2011) holds, there are a limited number of studies that have directly examined the result of corpus consultation on L2 writing improvement.

In general, DDL studies on L2 writing can be divided into two major categories based on how the corpus is used: Corpora used as research tools or corpora used as reference tools (Yoon 2011). These studies examine the effectiveness of DDL on enhancing different aspects of L2 writing. Some of them have also focused on comparing the use of different corpora. Unlike much of the previous research that has examined using corpora in the classroom, a case study by Chang (2014) examined the use of corpora by learners on their own and also compared the effectiveness of consulting general versus specialized corpora for academic writing. The results of the study indicated that although both types of corpora served the learners well as reference resources, the specialized corpora were more helpful for engineering students who wanted to follow the writing patterns of their own discourse community.

Other researchers have investigated the effect of DDL on error correction in revising written drafts (e.g., Chambers and O’Sullivan 2004; O’Sullivan and Chambers 2006; Yoon 2008). The study done in 2014 by Tono, Satake, and Miura on 93 undergraduate university students in Japan investigated the effect of direct use of corpus on the revision task of L2 compositions. Two sets of timed essay tasks were used. In the first phase, learners were supposed to write a specific type of essay, and in the second phase, they were asked to revise their writing using an online concordancer via their mobiles. Both the treatment and control groups were allowed to use online dictionaries, but before the treatment sessions started only the treatment participants received specific instructions on using concordancers properly. Their results evidenced that corpus consultation positively affected omission and addition errors; however, misinformation errors remained low in the correction.
Another group of DDL studies examined the potential effects of the approach on learning/teaching particular types of linguistic points influential in writing enhancement, with most of them confirming the positive effects of corpus use. These studies include Cresswell (2007) who explored the case for learning connectors in English, Celik and Elkatmis (2013) who did a similar study on EFL learners’ proper use of punctuation marks, and Aluthman (2017) who examined the effects of the method on ESL writers’ citation patterns.

There are other studies (e.g., Luo 2015; Yoon and Jo 2014) that have compared the effectiveness of different types of DDL (i.e., direct and indirect) on improving L2 writing. Amongst these, Vyatkina (2016), in a research similar to the present study, examined the effect of using DDL on learning German lexico-grammatical structures by comparing the effects of computer-based and paper-based DDL tasks. Their findings indicated that both DDL types had a positive impact on learners’ lexico-grammatical features and their overall language proficiency. Another important outcome was the learners desire to continue using concordance output in their future independent learning. However, the findings of these studies were ambiguous on which method produced better results. In other words, as Luo’s (2016) noted, both methods have their own advantages and disadvantages. As Rogers (1983) and Luo (2016) recommend, this slow diffusion of DDL research needs to be evaluated through a systematic and critical evaluation.

Additional studies have found that the use of both native and non-native (learner) corpora (Cotos 2014) can aid learners by providing them a chance to compare their writing deficiencies with those of native-speakers. This method yields even more promising results if DDL activities are conducted across varying levels of proficiency (Chujo and Oghigian 2010).

The current study is an empirical study on EFL learners’ writing pedagogy that will complement existing research and possibly enhance writing pedagogy in second language classrooms. To this end and based on recent EFL research on the micro-level of writing (i.e., syntax, morphology, and lexis) (Brown 2004; Cumming 2001), we analyzed the use of corpus on three characteristics of L2 written performance, namely, complexity, accuracy, and fluency (CAF). These measures were counted as introduced in Cumming’s (2001), Larsen Freeman and Strom’s (1977), and in a review of the writing measurements used in 39 studies of second language development by Wolfe-Quintero et al.’s (1998). CAF measures are considered as dependent variables to evaluate variation with respect to the independent variables like the level of acquisition or task features. Researchers unanimously believe that CAF can be operationalized in different ways, and they do not provide performance description unless specifically determined (Skehan and Foster 1997).
Accordingly, this study was guided by the following research questions:
Can DDL approach compared to the NDDL approach help EFL learners achieve higher improvements in CAF measures of writing?

This can be explained into:
1. Is there a significant difference in the writing complexity scores for the NDDL, DDDL, and IDDL groups while controlling for their pretest complexity scores?
2. Is there a significant difference in the writing accuracy scores for the NDDL, DDDL and IDDL groups while controlling for their pretest accuracy scores?
3. Is there a significant difference in the writing fluency scores for the NDDL, DDDL, and IDDL groups while controlling for their pretest fluency scores?

3 Methodology

3.1 Participants

Three intact groups of intermediate undergraduate EFL learners majoring in English in an English Language Translation program were selected as participants. There were 39 learners, male (19) and female (22), ranging in age from 19 to 26. They were assigned randomly to three classes: one class of 14 students in the NDDL (control) group and two experimental classes with 13 students in the DDDL group and 12 students in the IDDL group.

According to the questionnaire given to all learners at the beginning of the study, most of the subjects were second-year students in a four-year first-degree system and all had passed similar EFL courses at Iranian schools. This information ensured that the participants were comparable and could be used for our research goals.

3.2 Materials

To set the ground for putting the theoretical aspects of the current study into practical practice and to help it run smoothly toward achieving its intended goals, the following materials were utilized:

1. DDL-based materials developed by the researchers of this study (DDL units) were used as teaching materials for both DDL groups (i.e., DDDL and IDDL groups). Each unit included two main sections: concordance lines and corpus-based tasks. The whole unit was devoted to practicing one linguistic element
These elements were micro-level features of the English language that were first determined by the researchers. To develop the units, the following two sources were analyzed by the researchers: (a) The researchers took a critical look at the textbook Paragraph Development by Arnaudet and Barret (1990) which is widely used in academic writing courses in Iran. This was also used as the main teaching material for the traditional method of teaching paragraph writing to the NDDL group. (b) A collection of paragraphs written by Iranian EFL students in a paragraph development course at Urmia and Buali Sina universities in Iran were given a thorough error analysis by the study researchers to detect the most frequent problematic linguistic features. These features provided the main focus of the lessons in the experimental sessions. Rating reliability (Cronbach alpha = 0.84) was established via a double rating of 25% of the written pieces.

2. A mini local learner corpus of 120,000 word tokens (9411 word types) was compiled by the researchers. The necessity of compiling a specific corpus for this study came from the idea that the concordance examples of this corpus would be more appropriate for Iranian learners’ needs and Aston’s (1997) belief that it could facilitate the learner’s task of analyzing concordance extracts. Rohrbach (2006, as cited in Granger 2009) considers this kind of localization of a learner corpus to be one of the most promising avenues in learner corpus pedagogy.

The texts used in the compilation were paragraphs written for classroom assignments by a parallel group of EFL learners from the previous academic year. To avoid any comprehension problems, major grammatical error and structural pattern deteriorations in the texts were corrected.

1. The compiled corpus and two additional corpora were used by the researchers to search for teaching target features and to prepare the activities of the DDL units. The same two corpora plus the compiled corpus were also used by the DDDL group to search for the same target linguistic features via a concordancer to make concordance lines by themselves. Here is a brief explanation of the two corpora: (a) The Corpus of Contemporary American English (COCA) that contains nearly 400 million words taken from texts of spoken conversation from the media and academic journals. Since the focus of this research was academic prose, an academic subsection of COCA taken from 100 academic journals in a wide range of scientific fields from 1990 to 2006 was used as the searching resource in the current study; (b) The Michigan Corpus of Upper-level Student Papers (MICUSP) that contains a collection of 830 papers with around 2.6 million words that have been compiled since 2004.
3.3 Instruments

In an attempt to collect the required quantitative data, the following instruments were used:

1. A slightly modified questionnaire adapted from Sripicharn (2002) was used to elicit personal information (e.g., English learning background, learning strategies, and familiarity with concordancers) about participants. This questionnaire established the extent to which the participants of each group were comparable.

2. A pre and post writing test was given to all groups at the first and last treatment sessions. Learners were asked to write in two of the different genres (descriptive and argumentative) taught during the treatment sessions. The descriptive prompt was ‘describe your university’ in the pretest and ‘describe your city’ in the posttest task. The argumentative prompt required participants to compose a paragraph on the topic of ‘the use of social media by young people’ in the pretest and ‘the use of social media by the elderly’ for the posttest. The two distinct task types used in the current study provided for the collection of diverse written production from the participants.

3. The concordancer program AntConc (version: 3.2.4w), a free downloadable tool developed by Anthony Lawrence in 2012, was used to retrieve concordance lines by the researchers as well as the participants of the DDDL group who were to work with the program on their own. Participants could easily install the program on their computers and run it.

4. Printouts of concordance lines taken from the corpora introduced in 2 and 3 by the authors of this study.

3.4 Procedures

The data collection procedure took place over the academic semester starting from October 2017 to February 2018. The subjects were taking a two credit compulsory “Paragraph Development two” course. Having determined the target linguistic elements for teaching, the researchers prepared a series of seven DDL units. The first group of units included concordance lines in the KWIC format (as described in 4). These units were used as teaching materials for the IDDL group but not for the DDDL group. The DDDL group used the same concordancer, KWIC format, and corpora to individually retrieve their lines. The second group of DDL units consisted of some corpus-driven tasks, each focused on one of the target linguistic patterns taught in the treatment sessions. These tasks were provided to both DDL groups to enhance their learning. In contrast, the NDDL group was taught by the
traditional method of using a textbook, teacher explanations, and classroom exercises. At each session, all groups wrote a specific type of paragraph. A questionnaire and pre-test were administered to the participants at the beginning of the study.

It was the participants’ first experience working with concordancers. Several scholars have stressed the importance of training in corpus consultation before students embark on independent work with concordancer software or its output (e.g., O’Sullivan and Chambers 2006; Yoon and Hirvela 2004). Therefore, it was deemed necessary to hold an introduction to concordancers session prior to beginning the DDL classes. All three groups attended nine weekly classes including two sessions for the pre-test and post-test, but the DDL groups also attended two 1-h sessions: one teaching general paragraph writing information and one introduction to concordancers session.

The training session took place in a computer lab (connected to the Internet) for the experimental participants (DDDL and IDDL) who reported no previous use of a concordancer. A number of objectives were to be achieved in the training session. Firstly, to acquaint participants with the corpus and its use in language learning. Secondly, to provide students with an awareness of the use of AntConc software. Thirdly, to clarify to the participants how to utilize concordance outputs before moving on to the learning sessions.

The first objective was achieved by an introduction to corpus linguistics and types of corpus analyses, a brief overview of their applications in language learning, and distributing printouts of concordance lines of the word commit in both ‘full-sentence’ and ‘KWIC’ to the learners for more clarification. For the second objective, the researchers demonstrated the use of the AntConc concordancer, i.e., how to insert a corpus into the program, where to type keywords or phrases, how to determine the length and width of lines, and the meta-linguistic codes and terminologies (e.g., 1R, 2R, 3R) used on the AntConc. Besides, participants were asked to practice searching for words, phrases, or sentences of their own choice, and generate concordance lines, so as to be able to use the concordancer in a way which was going to be of potential use to them in their further studies.

To fulfill the last and the most essential purpose of this laboratory session, indirect DDL activities were provided to the participants to help them extract and analyze information from paper-based concordance lines. These tasks included fill-in-blanks exercises for missing words in concordance lines followed by questions that asked learners to extract patterns inductively, and some questions following samples of concordance lines that required learners to analyze lines in detail.
The KWIC display was also introduced. This type of concordancer output situates the linguistic terms in context, but often removes the sentence from its paragraph context, and thus, leaves sentence fragments. That is to say, in the KWIC format, concordance output can be confusing and may not provide the appropriate length or span of context (Sripicharn 2002). To avoid it, a new way of reading needs to be applied for the fragmented utterances. That is, the reader needs to focus on the listed keywords both center-out and vertically instead of reading them from left to right. Therefore, through the tasks described above, the researchers also aimed to make the participants practice this untypical reading skill.

4 Data analysis

Participant’s writing performances, in terms of CAF measures, were scored by Larsen-Freeman (2006) and Wolfe-Quintero et al.’s (1998) profiles based on T-units, which are defined as “one main clause with all subordinate clauses attached to it” (Hunt 1965, p. 20). Accordingly, writing accuracy was measured by means of the proportion of the error-free T-units to the total T-units (EFT/T), mean length of error-free T-units (MLEFT) was taken into account as a measure of writing fluency, and, lastly, we calculated writing complexity through the total number of clauses divided by the total number of T-units (C/T).

For the purpose of this study we opted to not teach lexical knowledge, so measures of lexical complexity were removed from the data analyses. Hence, in this study the term complexity means syntactic complexity or “the range of syntactic structures that are produced and the degree of sophistication of such structures” (Lu 2011, p.19).

Following data collection, the handwritten pieces of students’ writing samples were subsequently transcribed into Microsoft Word. In this stage, the written responses underwent some modifications from their original format. Then, they were subdivided into T-units, with each T-unit starting a new line within the writing. In the next step, T-units containing errors were tagged, and the number of clauses (dependent and independent) in each T-unit was identified. The Microsoft files containing only the error-free T-units of each student’s written production were then uploaded to AntConc (version: 3.2.4w) software and identified as a new corpus. The number of words, as ‘word’ is defined by this tool, was calculated for each participant’s group of error-free T-units.

The obtained data for each participant was then imported into a spreadsheet file. The items recorded included all the T-units, number of clauses and errors in
each T-unit, and the total words of the all error-free T-units per writing task; so as in each cell, a T-unit was placed in a cell and in each of the front two cells, respectively, scores of error occurrences and number of clauses for the same T-unit were recorded. In this way, the number of errors in each T-unit was represented in the column of the cells next to the T-units’ column. In the column of number of errors, those rows that contained zero scores were counted and summed up to yield the total number of error-free T-units in each written script.

The columns of T-units and clauses were also summed. The same calculation processes were applied to every piece of writing in the pre and posttest phases.

Sentence boundaries were identified by participant-supplied punctuation marks that were considered as terminating a sentence such as periods, exclamation marks, and question marks. Identification of T-units, however, was not as clear-cut, and in some instances we relied on the researchers’ decision. For instance, if a very long series of words contained none of these terminating marks, it was considered a punctuation error. Thus, the researchers had to base some of their decisions on separating parts of meanings as sentence endings or beginnings. When the participants’ writing included non-standard grammatical structures or sentence fragments that could obscure meaning or clause boundaries, the dataset analysis required some reader interpretation. Therefore, an initial coding was conducted to calibrate and decide on additional guidelines to recognize these units.

Sometimes, learners’ written data was less than clean-cut and contained fragments, incorrect salutations, missing pronouns, etc. These deficiencies required the researchers to make choices that emerged from their data. The following is an example sentence composed by one of the participants in response to the description task, “Many rooms, and all the rooms [are] close to the teachers room”, the researchers identified a fragment, ‘many rooms’, followed by an intact T-unit containing a possessive apostrophe error. Thus, of the three choices: (1) counting the fragment as an independent T-unit, (2) ignoring the fragment, or (3) including it with the following T-unit (Bardovi-Harlig and Bofman 1989; Ishikawa 1995), we agreed on counting such fragments as an independent T-unit.

There were other instances, such as “We can walk to the station and get a taxi to any place”, where the researchers experienced uncertainty on how to treat in the data. Based on the definition of T-units as the minimal terminable unit (Hunt 1965), i.e., the smallest unit containing a grammatical sentence without leaving behind any unaccounted for fragments, this statement could be justifiably viewed as including two T-units.
Besides, the principles outlined by Hunt (1966) were also used as a guideline through the sentence segmentation process of the written paragraphs. Thus:

(1) Each independent clause, plus its subordinate clauses and embeddings counted as one T-unit.
(2) The coordinating conjunctions ‘and’, ‘but’, ‘yet’, and ‘so’ (when it meant ‘and so’) were regarded as markers which separate adjacent T-units (except in cases where they separated two subordinate clauses).
(3) A clause was defined as an expression that contained a subject (or coordinated subjects) and a finite verb (or coordinated finite verbs).
(4) Sentence fragments that resulted from the omission of a single word were counted as T-units, with the missing word supplied.
(5) Interjections, parentheses, unintelligible words, vocatives and speaker tags in conversation were discarded. (p. 95)

Having determined the T-units in the written productions, the researchers assessed the accuracy index (EFT/T) by coding the written pieces for errors. However, to avoid any discrepancies among the researchers in coding (Polio 1997), they first needed to formulate guidelines stating clearly what constitutes an error. To do so, the error coding model based on Bardovi-Harlig and Bofman (1989), and also used by Storch 2007 and Neumann 2014, was applied.

Moreover, to measure accuracy, a decision was made to avoid the thorny issue of error classification altogether by merely identifying whether or not a given T-unit included an error and to just ignore the type of errors. Certain spelling and punctuation errors that proved to be extremely common in the written pieces were also not counted as errors.

To sum up, in the present study the syntactic complexity was calculated as the total number of clauses in each participant’s data divided into the total number of T-units of that participant’s performance. The other ratios, (EFT/T and MLEFT) for which data were recorded on spreadsheets, were calculated in the same manner.

Having access to the controlled pretest scores of CAF measures, a one-way between-groups analysis of covariance (ANCOVA) was run to compare the effectiveness of the three different interventions designed to instruct L2 writing. The independent variable was the type of intervention: direct DDL (DDDL), indirect DDL (IDDL), and non-DDL (NDDD) methods, and the dependent variable consisted of scores on the post-test administered after the intervention was completed. Participants’ writing scores of the three CAF measures on pre-intervention were used as the covariates throughout the analyses.
5 Results

Having checked that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate through Kolmogorov–Smirnov test, we prepared summary statistics for pre and post-treatment tests, presented in Table 1, that was used in all hypotheses testing procedures.

As shown in Table 1, for both accuracy and fluency measures there is an overall increase in the mean scores from pretest to posttests. An opposite trend for the mean scores of complexity measure is indicated. That is, except for the DDDL group that had an increase of 0.15, the mean scores of the other two groups decreased from pretest to posttest. Nevertheless, to make sure that the obtained differences of mean scores are significant, further ANCOVA was run.

Research Question I

As already mentioned, the first question probed whether there is a significant difference in the writing complexity scores for the NDDL, DDDL, and IDDL groups.

As indicated in Table 2, the line relating to complexity as the covariate in the pretest stage, shows that pre-test scores are not a statistical covariate (p > 0.05). Furthermore, as the groups line shows, the three groups of IDDL, DDDL, and NDDL are not significantly different from each other when controlling for the pretest scores \[ F (2, 36) = 0.89, p = 0.41, \text{Eta} = 0.04 \]. Therefore, there was no significant difference in complexity measure across the three groups. Besides, using the guidelines suggested by Cohen (1988) (0.01 = small, 0.06 = moderate, 0.14 = large effect), this finding indicates a rather small effect size.

Research Question II

The second question concerned if there is difference in the writing accuracy scores across the groups when pre-test writing accuracy scores are treated as covariate.

Table 1: Summary statistics of CAF measures in three groups of DDDL, IDDDL, and NDDL.

<table>
<thead>
<tr>
<th>Feature Group</th>
<th>Complexity (C/T)</th>
<th>Accuracy(EFT/T)</th>
<th>Fluency (MLEFT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDDL</td>
<td>1.50</td>
<td>1.00</td>
<td>0.83</td>
</tr>
<tr>
<td>DDDL</td>
<td>1.38</td>
<td>1.04</td>
<td>0.69</td>
</tr>
<tr>
<td>NDDL</td>
<td>1.14</td>
<td>0.94</td>
<td>0.77</td>
</tr>
<tr>
<td>Posttest</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDDL</td>
<td>1.41</td>
<td>0.90</td>
<td>0.87</td>
</tr>
<tr>
<td>DDDL</td>
<td>1.53</td>
<td>1.26</td>
<td>0.79</td>
</tr>
<tr>
<td>NDDL</td>
<td>1.00</td>
<td>0.96</td>
<td>0.80</td>
</tr>
</tbody>
</table>

C/T = clause per T-unit, EFT/T = error free T-unit per T-unit, MLEFT = mean length of error free T-unit.
In Table 3, the groups line indicates that the three groups of IDDL, DDDL, and NDDL are significantly different from each other \([F(2, 36) = 4.25, p = 0.02]\). While accuracy scores in the pretest do not act as a significant covariate \((p > 0.05)\). Therefore, there is a significant, difference with a large Eta squared \((0.19)\), in the writing accuracy scores for the NDDL, DDDL and IDDL groups while controlling for their pretest scores.

Having determined that there was a significant difference among the three groups, at this stage of analysis, we conducted a set of post hoc pair-wise comparisons (Table 4.) to explore the differences across the groups.

In the results presented in Table 4, only IDDL and NDDL groups are significantly different from one another \((p = 0.04)\). That is, the IDDL group and the NDDL group differ significantly in terms of their accuracy scores.

On the other hand, in Table 1, the mean scores of IDDL and NDDL are 0.87 and 0.80, respectively. Therefore, it seems that IDDL has been more effective on the improvement of writing accuracy at posttest than NDDL and DDDL.

---

**Table 2:** Tests of between-subjects effects for complexity scores.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type II sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial eta squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected model</td>
<td>2.806(^a)</td>
<td>3</td>
<td>0.935</td>
<td>0.922</td>
<td>0.440</td>
<td>0.073</td>
</tr>
<tr>
<td>Intercept</td>
<td>16.014</td>
<td>1</td>
<td>16.014</td>
<td>15.787</td>
<td>0.000</td>
<td>0.311</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.646</td>
<td>1</td>
<td>0.646</td>
<td>0.637</td>
<td>0.430</td>
<td>0.018</td>
</tr>
<tr>
<td>Groups</td>
<td>1.824</td>
<td>2</td>
<td>0.912</td>
<td>0.899</td>
<td>0.416</td>
<td>0.049</td>
</tr>
<tr>
<td>Error</td>
<td>35.502</td>
<td>35</td>
<td>1.014</td>
<td>4.254</td>
<td>0.022</td>
<td>0.196</td>
</tr>
<tr>
<td>Total</td>
<td>105.000</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>38.308</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)R Squared = 0.073 (Adjusted R Squared = −0.006).

---

**Table 3:** Tests of between-subjects effects for accuracy scores.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type II sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial eta squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected model</td>
<td>6.838(^a)</td>
<td>3</td>
<td>2.279</td>
<td>3.483</td>
<td>0.026</td>
<td>0.230</td>
</tr>
<tr>
<td>Intercept</td>
<td>30.931</td>
<td>1</td>
<td>30.931</td>
<td>47.263</td>
<td>0.000</td>
<td>0.575</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.540</td>
<td>1</td>
<td>0.540</td>
<td>0.824</td>
<td>0.370</td>
<td>0.023</td>
</tr>
<tr>
<td>Group</td>
<td>5.568</td>
<td>2</td>
<td>2.784</td>
<td>4.254</td>
<td>0.022</td>
<td>0.196</td>
</tr>
<tr>
<td>Error</td>
<td>22.906</td>
<td>35</td>
<td>0.654</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>84.000</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>29.744</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)R Squared = 0.230 (Adjusted R Squared = 0.164).
Research Question III

The last question in this study tested the possible difference in the writing fluency scores across the groups with pretest writing fluency scores considered as covariate.

As Table 5 shows, the same pattern of observed effect for accuracy was obtained for fluency. That is, all groups turned out to be significantly different from each other without a statistical role of the covariate (p > 0.05) and a quite large effect size (Cohen 1988) \[ F(2, 36) = 4.57, p = 0.01, Eta = 0.20 \].

Follow up pairwise comparisons, Table 6, located where exactly the differences existed.

As seen in Table 6, NDDL group performed significantly different from IDDL (p < 0.05) in terms of fluency. This is also observed in Table 1, where the mean score of IDDL group (10.66) is more than that of NDDL (8.85). To sum up, this means that IDDL has been more effective on the improvement of fluency at posttest than NDDL.

---

### Table 4: Pairwise comparisons for accuracy scores.

<table>
<thead>
<tr>
<th>(I) Group</th>
<th>(J) Group</th>
<th>Mean difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence interval for difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDDL</td>
<td>DDDL</td>
<td>0.063</td>
<td>0.325</td>
<td>1.000</td>
<td>-0.754 - 0.880</td>
</tr>
<tr>
<td>IDDL</td>
<td>NDDL</td>
<td>0.829&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.320</td>
<td>0.042</td>
<td>0.024 - 1.633</td>
</tr>
<tr>
<td>DDDL</td>
<td>IDDL</td>
<td>-0.063</td>
<td>0.325</td>
<td>1.000</td>
<td>-0.880 - 0.754</td>
</tr>
<tr>
<td>DDDL</td>
<td>NDDL</td>
<td>0.766</td>
<td>0.317</td>
<td>0.063</td>
<td>-0.031 - 1.563</td>
</tr>
<tr>
<td>NDDL</td>
<td>IDDL</td>
<td>-0.829&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.320</td>
<td>0.042</td>
<td>-1.633 - 0.024</td>
</tr>
<tr>
<td>NDDL</td>
<td>DDDL</td>
<td>-0.766</td>
<td>0.317</td>
<td>0.063</td>
<td>-1.563 - 0.031</td>
</tr>
</tbody>
</table>

Based on estimated marginal means<sup> *</sup> Adjustment for multiple comparisons: Bonferroni.<sup> *</sup> The mean difference is significant at the 0.05 level.

### Table 5: Tests of between-subjects effects for fluency scores.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type II sum of squares</th>
<th>df</th>
<th>Mean square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial eta squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected model</td>
<td>25.348&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3</td>
<td>8.449</td>
<td>3.307</td>
<td>0.031</td>
<td>0.221</td>
</tr>
<tr>
<td>Intercept</td>
<td>142.537</td>
<td>1</td>
<td>142.537</td>
<td>55.790</td>
<td>0.000</td>
<td>0.614</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.036</td>
<td>1</td>
<td>0.036</td>
<td>0.014</td>
<td>0.906</td>
<td>0.000</td>
</tr>
<tr>
<td>Groups</td>
<td>23.374</td>
<td>2</td>
<td>11.687</td>
<td>4.574</td>
<td>0.017</td>
<td>0.207</td>
</tr>
<tr>
<td>Error</td>
<td>89.421</td>
<td>35</td>
<td>2.555</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3955.000</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>114.769</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>R Squared = 0.221 (Adjusted R Squared = 0.154).
6 Discussion

The results of this study revealed that DDL activities in the IDDL experimental group have significantly positive effects on EFL learners’ writing development in terms of fluency and accuracy. In addition, it was generally observed that language learners who studied with the DDDL method did not perform better than those who received instruction in the IDDL or NDDL based method. In contrast, no obvious advantages were found for either direct or non-direct types of DDL instruction as compared with the conventional (NDDL) method in terms of writing complexity.

The findings of the DDDL group are in contrast with the findings of many studies (e.g., Liu and Jiang 2009; Yilmaz 2017) that found direct DDL had positive effects on foreign language writing. However, the insignificant effect of the DDDL method as compared to IDDL on CAF measures of writing, confirmed what Boulton (2010) found in his review of studies on the application of the direct DDL method. His review showed that direct DDL did not seem to have any determined superiority over traditional language learning/teaching methods. In the same vein, Luo’s (2015) study which examined the use of DDL in EFL writing instruction for engineering students also corroborates the current investigation’s findings. Luo reported that the direct DDL method did not have any significant effect on revision tasks as compared with traditional activities. These discrepancies may be due to a lack of computer skills, specifically in the context of the current study, and might act as a barrier in computer-based language learning (Sabti and Chaichan 2014). These results may also be attributed to what Granger and Tribble (1998) suggest about DDL. They believe that if learners are exposed to vast quantities of information there is a considerable risk that DDL activities can become time-consuming and frustrating for learners.

Table 6: Pairwise comparisons for fluency scores.

<table>
<thead>
<tr>
<th>(I) Groups</th>
<th>(J) Groups</th>
<th>Mean difference (I-J)</th>
<th>Std. error</th>
<th>Sig. *</th>
<th>95% Confidence interval for difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDDL</td>
<td>DDDL</td>
<td>0.270</td>
<td>0.648</td>
<td>1.00</td>
<td>−1.359 to 1.899</td>
</tr>
<tr>
<td></td>
<td>NDDL</td>
<td>1.788*</td>
<td>0.653</td>
<td>0.029</td>
<td>0.146 to 3.431</td>
</tr>
<tr>
<td>DDDL</td>
<td>IDDL</td>
<td>−0.270</td>
<td>0.648</td>
<td>1.00</td>
<td>−1.899 to 1.359</td>
</tr>
<tr>
<td></td>
<td>NDDL</td>
<td>1.518</td>
<td>0.620</td>
<td>0.059</td>
<td>−0.041 to 3.078</td>
</tr>
<tr>
<td>NDDL</td>
<td>IDDL</td>
<td>−1.788*</td>
<td>0.653</td>
<td>0.029</td>
<td>−3.431 to −0.146</td>
</tr>
<tr>
<td></td>
<td>DDDL</td>
<td>−1.518</td>
<td>0.620</td>
<td>0.059</td>
<td>−3.078 to −0.041</td>
</tr>
</tbody>
</table>

Based on estimated marginal means *Adjustment for multiple comparisons: Bonferroni. *The mean difference is significant at the 0.05 level.
As observed by the authors of this study, DDDL participants sometimes entered the wrong search word/words during concordancing practices and then got the wrong results; thus, they were only able to obtain what they wanted after having conducted several searches. Moreover, they had to deduce a pattern from a long list of concordance lines, which turned out to be a particular challenge for intermediate EFL learners in this study. This finding is supporting of Huang’s (2011) idea that in computer-based DDL practices students are distracted by the large number of concordance lines, require hardware or software troubleshooting, and are skeptical about whether or not they are extracting the correct data.

The findings in our study also indicated that both groups adopting the DDL (IDDL and DDDL) approach did not make an obvious improvement in their syntactic complexity. This result agrees with Storch’s (2009) study in which learners’ grammatical complexity did not experience any significant development after one semester of corpus consultation. Storch, in line with Ortega (2003), also concluded that an increase in grammatical complexity would require a longer time, possibly a time period up to 12 months, of college-level instruction.

Furthermore, it is possible that topic unfamiliarity may have led to zero gains in learners’ syntactic complexity in our study. Tedick (1990) states that participants produce more complex sentences if the topics are more familiar to them. This is in line with Coto (2014), who brought together both native speaker and learner corpora to investigate DDL pedagogy on the improvement of EFL writers’ knowledge of linking adverbials with the results showing no positive gains in complexity. He attributed this to their incomprehension of the meaning of the input provided by the concordance lines taken from a learner corpus. Coto, in agreement with Chang (2014), concluded that the reason for the incomprehensibility of the data was that the corpus was compiled from a variety of disciplines. As Kennedy and Miceli (2010) and Yoon (2011) point out, a small but focused discipline-specific corpus may provide writing patterns that are directly relevant to students’ needs and interests. Furthermore, the authors of this study think that this finding supports developmental prediction (Cooper 1976) which argues for non-linear complexification as far as subordination is concerned. That is to say, since at upper levels of language proficiency learners should be expected to produce sentences that capitalize on complexification at the phrasal rather than clausal level, such a reduction of C/T at upper proficiency levels is also considered as a result of favoring phrases to clauses (for more details see, Sharma 1980; Wolfe-Quintero et al. 1998). Nonetheless, the findings of the current study contrasts with Yılmaz’s (2017) findings that the experimental group significantly outperformed the control group in the area of syntactic complexity after the use of DDL activities.
With respect to the second research question, the IDDL category of DDL instruction turned out to be more effective on the improvement of writing accuracy at the posttest than the NDDL and DDDL. This is consistent with the study conducted by Tono et al. (2014) which evidenced that the DDL approach, although they did not make a distinction between IDDL and DDDL, significantly improved omission and addition errors. Likewise, in a study done by Luo and Liao (2015), corpus consultation in revision tasks helped learners have an obvious decrease in their writing errors. Similar to the current study, Yoon and Jo (2014) compared the effect of direct and indirect DDL approaches on error correction in L2 writing. They concluded that the indirect use of corpora seemed to have a greater influence than direct DDL in error correction for most participants. The gains in accuracy found in the present study might be due to practice effects. The participants were systematically engaged with authentic data, which could have helped them develop more accurate and academic sentences (DeKeyser 2007). This finding confirms what Skehan (1998) and Aston (1997) hold; they proposed that the DDL approach develops the learners’ generalized schematic of language use because they can test their hypotheses against authentic data in the corpora.

The finding, moreover, provides evidence in favor of the noticing hypothesis (Schmidt 2001), which proposes that noticing is an essential condition for input to be converted into uptake. The abundant concordance lines presented in KWIC format in the DDL activities stimulate students to notice, analyze, and discover correct language patterns. To use Corder’s (1981) terms, this is a particular linguistic pattern that can be crucial in helping learners turn input into uptake.

In response to the third research question, we observed a significant positive effect from IDDL pedagogy on participants’ writing fluency as compared to both DDDL and NDDL approaches. This significant change led us to the conclusion that the use of paper-based DDL activities in a paragraph development course could help learners increase their writing’s quality level. That is to say, we believe that the lengthening of error-free T-units might be interpreted as an aspect of enhanced written expression quality. This finding supports Luo’s (2016) study, which found that there was a significant difference in writing fluency between the pre and posttest compositions of the experimental group in comparison with the control group as a result of consulting the BNCweb corpus.

Chambers and O’Sullivan (2004) and O’Sullivan and Chambers (2006) maintain that the application of corpus data in L2 writing classrooms can help learners notice and learn lexico-grammatical patterns. Accordingly, this may be a reason for the significantly longer essays written by the experimental group in the same length of time.
7 Conclusion

This study explored the implication of both native-speaker and local learner corpora, attesting the effect of the direct, indirect, and non-DDL practices on EFL learners’ writing development in terms of fluency, accuracy, and complexity. The results confirmed the effectiveness of adopting IDDL in writing more accurate and fluent essays, whereas no statistical evidence was found with regard to the improvement of syntactic complexity. Moreover, no significant usefulness of applying the DDDL method to improve learners’ writing CAF measures was found. Nevertheless, this does not mean that direct DDL activities are completely useless in enhancing EFL learners’ written compositions.

The findings imply that learners need to be equipped with sufficient computer skills and enough training to appropriately use a concordancer to consult a corpus. Thus, more concordancer training sessions should be set in future studies. It would also make the picture clearer if teachers provided appropriate guidance when learners get confused in analyzing the concordance lines. Thus acting as active agents, a teacher’s guidance during learner concordancing practices may help them overcome technological and psychological barriers. The positive effects of indirect DDL on accuracy and fluency found in this study have implications for language teachers and materials developers in terms of the advantages and disadvantages of utilizing educational versus authentic materials in boosting or hindering learners’ linguistic awareness, needs, and attractions. These results also suggest that paper-based DDL could be even more effective than computer-based DDL approaches. Accordingly, teachers and students can be assured that computers are not necessary to implement DDL in classrooms. This is particularly good news for EFL contexts that do not have access to classroom-based computers. To sum up, the significant finding that indirect DDL can be effective means that paper-based DDL can be viewed as a teaching method that language teachers can apply along with appropriate tasks to teach the language so that students become active autonomous learners and researchers in the language classroom.

This study was limited in that, like most similar studies (Luo and Khou 2017), it relied on short-term data, i.e., only one semester. However, it usually takes a much longer time for one to improve her/his writing complexity. Therefore, a more longitudinal study is needed to investigate the potential effect of DDL on EFL learners’ writing competence, especially complexity. Additionally, in the present study we did not probe the potential effects of the use of a learner corpus over a native speaker’s corpus; Seidlofer (2002) perceives learner corpora as a strong motivational factor through which corpus-based tasks turn into a personally
relevant learning experience. Thus, in future research the potential effects of a learner corpus in comparison with that of a native-speaker corpus on enhancing CAF measures of L2 writing can be investigated.

References


Cooper, Thomas C. 1976. Measuring written syntactic patterns of second language learners of


skills course. In Encarnación Hidalgo, Luis Quereda & Juan Santana (eds.), Corpora in the

1–23.

Practice in a second language: Perspectives from applied linguistics and cognitive
psychology, 1–18. Cambridge: Cambridge University Press. https://doi.org/10.1017/
CBO9780511667275.002.

Faerch, Claus & Gabriele Kasper. 1987. Procedural knowledge as a component of foreign language
learners’ communicative competence. In Tatiana Slama-Cazacu, H. Bolte & M. Herrlitz (eds.),

Farr, Fiona. 2008. Evaluating the use of corpus-based instruction in a language teacher education
context: Perspectives from the users. Language Awareness 17(1). 25–43.

English Language Teaching Journal 55(3). 238–246.

Graham, Steven & Karen R. Harris. 2005. Writing better: Teaching writing processes and self-
regulation to students with learning problems. Baltimore: Brookes.


Granger, Sylviane. 2009. The contribution of learner corpora to second language acquisition n and
Foreign Language teaching: A critical evaluation. In Karin Aijmer (ed.), Corpora and language

Granger, Sylviane & Chris Tribble. 1998. Learner corpus data in the foreign language classroom:
Form-focused instruction and data-driven learning. In Sylviane Granger (ed.), Learner English


Hunt, Kellogg W. 1965. Grammatical structures written at three grade levels. Champaign: NCTE.

732–139.

Ishikawa, Sandra. 1995. Objective measurement of low-proficiency EFL narrative writing. Journal of

Jafari, Narges & Dariush Nejad Ansari. 2012. The effect of collaboration on Iranian EFL learners’

Johns, Tim. 1991. Should you be persuaded: Two samples of data-driven learning materials. ELR
Journal 4. 1–16.

Kennedy, Claire & Tiziana Miceli. 2010. Corpus-assisted creative writing: Introducing
intermediate Italian learners to a corpus as a reference resource. Language Learning &
kennedymiceli.pdf.

Kern, Richard. 2006. Perspectives on technology in learning and teaching languages. TESOL


**Supplementary Material:** The online version of this article offers supplementary material (https://doi.org/10.1515/iral-2018.0258).
Bionotes

Negin Samoudi
TESOL, Faculty of Department of English and Literature, Urmia University, Urmia, Islamic Republic of Iran
n.samoudi@urmia.ac.ir
https://orcid.org/0000-0002-7868-7020

Negin Samoudi has a Ph.D. in TEFL/TESOL from Urmia University. She is an EFL instructor at several Iranian universities. She is an IELTS teacher trained by IDP. She has been an English teacher since 2002. Her major research areas of interest are learner centered instruction, dynamic assessment, and corpus-based materials development.

Sima Modirkhamene
TESOL, Faculty of Department of English and Literature, Urmia University, Urmia, Islamic Republic of Iran
s.modir@urmia.ac.ir
https://orcid.org/0000-0002-6758-2469

Sima Modirkhamene has a PhD in TEFL/TESOL from the University of Surrey, UK. Since her return to Iran in September 2006, she has been lecturing at undergraduate and postgraduate levels and researching in Urmia University, Iran. Her main research interests include: bilingualism, first and second language acquisition, and cross-linguistic transfer. She has served as the Head of the Department of English Language and Literature and the Deputy Head of the Foreign Languages Center of Urmia University since 2008.