Effects of Data-Driven Learning on College Students of Different Grammar Proficiencies: A Preliminary Empirical Assessment in EFL Classes

Ming Huei Lin

Abstract
This article reports a pre–post comparative study investigating whether the data-driven learning (DDL) approach has different pedagogical effects on grammar students of English as a foreign language (EFL) with different levels of English proficiency. The study entailed a treatment group (TG) of 95 first-year undergraduates who learned grammar using DDL and a control group (CG) of 84 students who received no grammar treatment. Most of the participants were 18 or 19 years old, with only a few outliers, aged 17 or 20. The grammar performance and learning attitudes in both groups (their motivation and self-efficacy) were quantitatively examined through grammar achievement tests and a questionnaire. The data obtained from the groups were then compared at three proficiency levels: high, intermediate, and low. The results of an analysis of covariance show that in grammar performance, the proficiency levels in all the TG students rose significantly and in the posttest they outperformed their counterparts in the CG. However, neither the members of the TG nor those of the CG made any statistically significant improvement in their learning attitudes; no significant differences were found between the groups at any proficiency level. The mixed findings make an important contribution to the field, confirming that DDL is pedagogically suitable for enhancing the linguistic knowledge of university-level grammar learners, regardless of their proficiency, but warning that practitioners who treat the development of learner attitudes (e.g., motivation and self-efficacy) as important should be cautious with this approach.

Keywords
data-driven learning, inductive approach, discovery learning, grammar acquisition

Introduction
The integration of corpus use into language teaching and learning contexts has received considerable attention and interest since Johns (1991) introduced the data-driven learning (DDL) approach. It teaches students no explicit rules but gives them corpus-based material to analyze to generalize contextualized linguistic features/patterns. This rule-inferencing model has hitherto enjoyed support from important learning theories (e.g., the discovery learning and noticing hypothesis) and has capitalized on much empirical evidence. For example, many scholars have found DDL effective for learning collocations (Daskalovska, 2015; Saeedakhtar et al., 2020; Uçar & Yiğiteller, 2015; Wu et al., 2019), writing (M. Chen et al., 2015; Crosthwaite, 2020; Mizumoto et al., 2017; Poole, 2016; Sun & Hu, 2020), and vocabulary (Karras, 2016; H. Lee et al., 2017, 2019, 2020; P. Lee & Lin, 2019; Tsai, 2019). Some researchers have also reported that their students react positively to the DDL approach (M. Chen et al., 2015; Mizumoto et al., 2017; Sun & Hu, 2020). Even so, Kılıçkaya (2015) warned that such a discovering approach may be inappropriate in educational contexts where learners are used to learning from teachers who offer didactic/explanatory instruction. In addition, DDL’s pedagogical suitability seems also limited in the context of adverse learner feedback (Hirata & Hirata, 2013; Kennedy & Miceli, 2001) and its effects cannot be judged superior to those of a traditional deductive approach (TDA) in certain areas of grammar (Smart, 2012). Still, these concerns have been mitigated by the meta-analytical study by Boulton and Cobb (2017), which showed overall consistent evidence in favor of the efficacy of DDL in almost every language learning context. DDL does indeed have strong theoretical and empirical underpinning, but one critical area calling for an empirical investigation is comparing the efficacy of DDL on learners...
with different levels of language proficiency. This issue, which has so far been under-investigated in the field, is particularly important, as proficiency has been deemed by some to have an influential impact on either the affective learning experience of DDL students or their language performance (Hirata et al., 2013; Hirata & Hirata, 2013; H. Lee et al., 2019; Liu & Jiang, 2009). The issue is also worthy of discussion with regard to the current uncertainty about DDL use for lower-level language learners (Aston, 2001; Boulton & Cobb, 2017; Lin, 2016). An educational context that may seem especially suitable for examining this issue is Taiwan, where teacher-centered instruction is still dominant, as Lin and Lee (2019), Meyer (1988), and Wang and Tsai (2012) described it. Investigating whether or not DDL would successfully work in the Taiwanese context should thus shed further light in the field. This being the case, a comparative experiment was conducted on two groups of grammar students from Taiwan where English is mostly studied as a foreign language (EFL). Both groups had students of different proficiencies. One group was learning grammar with the DDL approach (namely, the treatment group, TG) and the other was receiving no grammar treatment (the control group, CG). To judge the pedagogic effects of DDL adequately, the grammar performance, learning motivation, and self-efficacy of both groups were examined and compared before and after the experiment. These two attitude variables are particularly important and should be examined together with grammar performance because motivation and self-efficacy have long been considered critical indications of students’ learning attitudes (Bandura, 1986, 1997; Dörnyei & Ushioda, 2013; Pajares, 2003). They have also been found to correlate strongly to learners’ performance (Zuffianò et al., 2013), in such areas as grammar (Lin, 2016) and writing (Lin, 2014). Considering all three variables should generate much evidence from which to judge DDL’s effects.

Along the line of discussion, the research question that arises to meet the goal of this study is whether or not EFL grammar students of different proficiencies would significantly benefit from the DDL treatment. To answer this question properly, two practical research questions were formulated:

1. After the treatment, to what extent do the TG participants with different grammar proficiency levels improve their grammar performance and learning attitudes?
2. After the treatment, to what extent do the TG participants with different grammar proficiency levels outperform the CG participants in terms of grammar performance and learning attitudes?

Theoretical Support for DDL

The nature of the DDL approach reflects the characteristics of the inductive approach, the discovery learning, and noticing hypothesis. First, an inductive approach has two major components: “1) the students’ attention is focused on the structure being learned; 2) and the students are required to formulate for themselves and then verbalize the underlying pattern” (Shaffer, 1989, p. 396). These operational aspects are consistent with those of DDL-centered treatments, where learners begin by focusing on target linguistic data and then on their own generalize the observed grammatical rules/features. Induction as such is further believed to be capable of reducing students’ cognitive workload, which is often associated with traditional didactic lectures (Sweller et al., 2011), and thus helps learners to concentrate on developing the meaning and structure of language use (Boulton & Cobb, 2017). That learners’ ability to use induction is nurturing by using DDL is generally accepted (L. S. Huang, 2017), but some scholars (H. J. H. Chen, 2011; Kirschner et al., 2006) tend not to recommend the inductive-based approach for low-level language learners, deeming it difficult for them.

Being inductive, DDL also involves discovery learning. The process of discovery learning, as Richards and Schmidt (2002) elaborate, comprises five major steps: observing, inferring, formulating, predicting, and communicating. Essentially, DDL also embraces these constituents. Specifically, it “entails encouraging learners to take the role of language researchers by systematically engaging in discovery learning (Gavioli, 2001) and in learning how to learn through observations, analyses, interpretations, and presentations of language-use patterns in corpus data” (L. S. Huang, 2011, p. 482). The correspondences between discovery learning and DDL have hitherto been widely acknowledged in the field (Boulton & Cobb, 2017; Frankenberg-Garcia, 2016; Liu & Jiang, 2009; Vyatkina, 2013, 2016), although it has also been advised that students participating in discovery learning tasks should have attained a certain level of linguistic knowledge (Johns, 1991; H. Lee et al., 2019).

Furthermore, when involved in DDL-centered activities, learners are focused on observing and analyzing recurring linguistic features of study, so their language awareness is enhanced (Boulton & Cobb, 2017; Flowerdew, 2015; Hadley, 2002; L. S. Huang, 2011; Timmis, 2015), leading to possible gains in language use (L. S. Huang, 2017). This method and its anticipated outcome recall the noticing hypothesis proposed by Schmidt (1990, 2001). As he asserts, increasing language awareness is a key element whereby learners consciously notice the linguistic features of interest; this turns input into effective intake and thus generates successful output. Such an innate quality of noticing in DDL has been added to experimental assessments, and much positive evidence in its favor has been produced (Hadley, 2002; Hong, 2010; Liu & Jiang, 2009; Mizumoto et al., 2017; Moon & Oh, 2018; Sealey & Thompson, 2007).

DDL Effects in Grammar Classes

While ample evidence has been found in favor of DDL treatments for the learning of different linguistic skills (e.g.,
vocabulary, collocation, and writing), this section focuses entirely on addressing DDL’s effects on the learning of grammar, which refers to the principles by which a person puts words together to form phrases, clauses, or sentences. Johns (1994) was one of the earliest scholars to advocate the instructional effects of DDL on grammar students. He submits that allowing learners to analyze a sufficient amount of organized input (i.e., concordance lines) will effectively facilitate their grammar acquisition. Johns’ views are endorsed not only by early scholars (Conrad, 2000) but also by later experimental findings. For example, Hong (2010) empirically examined two groups of grammar students and found that the group taught with corpus-aided instruction developed a stronger consciousness at the level of noticing, that is, the cognitive ability for learning the use of determiners (e.g., zero article and countability). Smart (2014) reported similar effects, writing that inductive DDL treatments resulted in significantly more gains in learning the passive voice than did traditional deductive treatments. Z. Huang’s (2014) comparative study further revealed that referencing a corpus helped students create more accurate lexico-grammatical patterns than those of students who consulted dictionaries alone. In addition, Lin (2016) and Lin and Lee (2019) found that treatments blending DDL and the TDA led to enhanced grammar skills and learning attitudes, although the improvement was mainly nonsignificant when compared with that of a pure TDA. Recently, Moon and Oh (2018) also revealed that secondary students benefited more from DDL grammar than from traditional grammar teaching, specifically in the use of the verb to be.

The above discussion seems to support the use of DDL in EFL grammar classes. However, such approval is mostly for higher-proficiency language learners rather than for lower-level students. In fact, few have empirically focused on verifying whether DDL suits this group. To begin with, Aston (2001) warned that corpus-based activities may not be appropriate for beginning-level students. This is because DDL tasks, being inductive and based on discovery learning, may require a certain level of linguistic knowledge on the part of students to cope with them (H. J. H. Chen, 2011; Johns, 1991; Kirschner et al., 2006; H. Lee et al., 2019). This concern is endorsed by Liu and Jiang (2009), who found empirically counterproductive cases with students who either had low levels of language proficiency or lacked vocabulary knowledge. Some recent works (Boulton & Cobb, 2017; Lin, 2016) also suggest that DDL may benefit intermediate-to-advanced-level learners, but excludes low-level students. Nonetheless, one study by Boulton (2010) showed encouraging results with beginning grammar students: they significantly improved their grammar performance and showed a preference for DDL. In light of the scarce and inconsistent empirical evidence, further empirical experiments are urgently needed to confirm whether or not DDL is pedagogically suitable for EFL students with different proficiencies, particularly those at lower levels.

The Present Study

The general research design and procedure for the current study is presented in Figure 1 and then discussed.

Sample

The study, based at a single university in Taiwan, recruited a convenience sample of 179 first-year non-English-major college students from four general English classes (Classes A, B, C, and D) which aimed at developing students’ general English skills, with a particular focus on reading, vocabulary, and grammar. Class A (52 students) and Class B (43), forming the TG (95), were taught by the present researcher, who had taught English skills using DDL at higher education for several years. Class C (41) and Class D (43) formed the CG (84 students) and were taught by another teacher, who had had several years’ experience of teaching English skills to undergraduates. The gender distribution in the TG was 43 males and 52 females; the CG had 20 males and 64 females. In both groups, most of the participants were 18 or 19 years old, with only a few outliers, aged 17 or 20. Before taking part, all the participants had studied English for approximately 10 years in Taiwan’s educational system. The general English proficiency levels in both groups were mostly at the A2 (TG: 65 students; CG: 63) and B1 levels (TG: 28; CG: 18) in the Common European Framework of Reference for Languages. Only a few reached the B2 level (TG: 2 students; CG: 3). Once enrolled in the experiment, the students’ specific grammar abilities were explicitly examined by means of a grammar achievement pretest (described below in the data collection section) and, for analysis and comparison purposes, set at one of three distinct levels: high, intermediate, and low (described in the “Data Analysis” and “Results” sections).

The Treatment

The TG treatment. The TG had one 90-min DDL grammar lesson per week for 3 weeks. The grammar items for each week comprised the language use of agree and deprive (Lesson 1), adjective clauses introduced by who/whom/whose (Lesson 2), and adjective/noun clauses introduced by that (Lesson 3). These items were purposely selected because, in his years of teaching experience, the current researcher had found that many of his previous EFL grammar students were confused by them, thus deeming them suitable material to test the effects of DDL on grammar learners at different levels of attainment.

Following the instructional model of Lin (2016), the TG was first shown concordance lines containing the node word (the key words in context) for observation (see Figure 2 for illustration). The material, which the researcher compiled and delivered to students in the form of printed handouts, comprised authentic examples from the Corpus of Contemporary
American English (corpus.byu.edu/coca). In reading the concordance, the TG was given several minutes to answer a general question (e.g., Observation in Figure 2); they could either work out on their own or discuss with peers and share their findings. When they were unable to analyze the concordance or offered inaccurate inferences, more guidance or questions were given, such as “Please observe the preposition phrases or patterns after the key words,” “Please judge the part-of-speech of to in the first two sentences and that of to in the following sentences,” or “Can you detect any differences between the functionality of who and whom in these sentences?” After sharing answers, learner understanding was checked by creating sentences, answering multiple-choice questions, and judging correct or incorrect grammar sentences (see Figure 2 for illustration).

The CG treatment. In contrast, over the 3 weeks, the CG was given neither a DDL treatment nor taught the specific grammar items chosen for the TG. Instead, they focused primarily on learning about general reading skills and appreciating textbook articles. For example, reading strategies were explained to them, such as scamming/scanning articles and inferring their main ideas. They were also taught the concepts of topic sentences and concluding sentences, which needed skills that would be advantageous when the students read for main ideas. The teacher also demonstrated how the skills/strategies were acquired by using the actual reading material. Afterwards, the teacher asked the class if they had any questions regarding the knowledge or application of these reading skills/strategies. To all the questions, the teacher would give a full answer. In addition, the CG also learned about appreciating different genres of writing and discussed them. For instance, some articles concentrated more on presenting opinions about social events; some were more intent on reporting scientific findings; while others told interesting historical stories. The classes learned how to interpret or
understand these different genres more effectively. They also discussed the ideas of each article after they had finished reading. The teacher raised questions, invited answers, and gave feedback. Finally, the meaning of some vocabulary items was explained when necessary, but explanations of language use and grammar rules relating to those designated for the TG were avoided.

Data Collection Instruments

Grammar achievement tests. Two sets of self-created grammar achievement tests were administered as pre/posttests for both groups. Specifically, they comprised concordance lines from the COCA. Each set had 30 questions, with every 10 items (six 4-option multiple-choice questions and four correct-incorrect grammar items) focusing on one distinct grammar lesson. Three other senior teachers teaching the same courses at the experimental site were invited to evaluate the quality of the test questions. They agreed with the researcher that the items adequately reflected the target linguistic features. Furthermore, a pilot study involving 117 other undergraduate students was conducted to verify the validity and reliability of the grammar tests. The pilot study results show that the whole test was valid because all its items had good item difficulty (all difficulty values between .40 and .80) (Chase, 1978) and item discrimination (all discrimination values higher than .25) (Noll et al., 1979). In addition, test–retest reliability was also obtained, with Pearson’s $r$ indicating strong correlation ($r = .91, p < .001$) between the test scores that the pilot study participants produced on two different occasions. These results suggest that the tests were suitable test instruments for this study. The sets were then randomly assigned for use in a pre- or posttest, in which each correct answer counted as 1 point, 0 being the minimum and 30 the maximum.
Motivation and self-efficacy questionnaire. To understand learner attitudes, the researcher adapted the 5-point Likert-type-scaled questionnaire designed by Lin (2016), which also examined learner motivation and self-efficacy in regard to grammar learning. The changes made to his version comprised minor wording of the grammar items examined, to make it fit the grammar points of the present study. For example, some original items asked about learner confidence in using grammar passives, relative clauses, or phrases to express purpose. These were changed to noun clauses introduced by that, adjective clauses introduced by that, or agree/deprive and their phrases. The quality of the revision was then examined in a pilot study involving another sample of 173 participants. The results of factor analysis show that the 15-item revision had overall strong validity (67.23% variance explained) and reliability (Cronbach’s α = .918). It also comprised three valid and reliable components: (1) self-efficacy in learning and using grammar (Items 1–7) (27.26% variance explained; Cronbach’s α = .915), (2) self-efficacy in identifying learned grammar (Items 8–11) (22.47% variance explained; Cronbach’s α = .848), and (3) motivation to learn grammar (Items 12–15) (17.51% variance explained; Cronbach’s α = .779).

Data Analysis

The data of the grammar tests and questionnaires were analyzed using several statistical methods. First, descriptive statistics presented the grammar pretest results, which served as criteria dividing all the participants into three levels of grammar proficiency. The top 1/3 of the scorers were assigned to the high level, the bottom 1/3 to the low level, and the remainder to the intermediate. Second, independent t-tests examined whether any significant differences existed between the groups at each level, in terms of both grammar and questionnaire outcomes. Third, paired-sample t-tests certified whether after the experiment each group determined by level improved its grammar performance and learning attitudes. All the t-test results were reported with effect sizes using Cohen’s d (Plosky & Oswald, 2014). Fourth, a set of ANCOVAs (analysis of covariance) compared whether, after the experiment, the groups with different levels differed from each other in their grammar posttest results. Likewise, several MANCOVAs (multivariate analysis of covariance) examined whether the groups with different proficiencies differed from each other in the exit questionnaire results. The effect sizes reported for both ANCOVA and MANCOVA results were partial η².

Results

This section is divided into three parts. The first part presents the results of both groups’ entry behaviors in terms of grammar performance and learning attitudes. The second part shows the results of the participants’ changes (if any) after the experiment, to answer Research Question 1. The last part presents the results of the comparisons between the groups, which answers Research Question 2.

Results of Entry Behaviors

Entry grammar proficiency. Table 1 presents the descriptive statistics of the grammar pretest results of both groups. Each group’s top 1/3 scorers, namely, those who obtained 17 points or above, were labeled high proficiency (TG: 44 students; CG: 22), the bottom 1/3 were labeled low proficiency, with 14 points or below (TG: 17 students; CG 22), and the remaining group, with either 15 or 16 points (TG: 34 students; CG: 40) were labeled intermediate. In addition, independent t-tests further revealed no statistically significant differences (p < .05) among the groups as a whole and between each level of the groups. These results suggest that either as a whole group or at different proficiency levels, the groups had similar grammar proficiency at entry, which made them suitable for purposes of comparison and further analysis.

Entry learning attitudes. A set of independent t-tests further showed that both groups also had similar attitudes (motivation and self-efficacy) to learning grammar whether compared as intact groups or between each level of the groups, since all the examined items evinced no statistically significant t-values (p < .05).

Results of Improvements After the Experiment: Answers to RQ1

Grammar improvements after the experiment. Table 2 presents the pre–post comparisons of both groups’ grammar performance. As the table shows, after the treatments, all the TG participants of different levels significantly improved their grammar performance, Whole: t(94) = −10.51, p ≤ .000; High: t(43) = −4.57, p < .001; Intermediate: t(16) = −6.45,
p < .001; Low: t(33) = −9.47, p < .001, with nearly large to very large effects (d-values between 1.28 and 10.84). In addition, the low level TG had the largest gain (pre–post mean difference: 6.5), the intermediate level had the second largest (5.53), and the high level had the least (2.91). In contrast, however, in the posttest grammar, the CG as a whole had no statistically significant gain, t(83) = 0.43, p > .05, and neither did the high level CG, t(22) = 1.11, p > .05. The intermediate level CG was even found to statistically deteriorate, t(22) = 2.50, p < .05, d = 3.31. The only statistically significant gain for the CG was found in the low-level participants, t(40) = −2.04, p < .05, although its effect size was rather small (d = 0.65). Overall, this shows that the TG’s improvement in grammar was not random but subject to the DDL treatments.

Learning attitudes after the experiment. However, in terms of learning attitudes, there were no statistically significant differences between entry and exit questionnaires of both groups at any level (p > .05) (Table 3), suggesting that the treatment had no effects on the development of the TG’s motivation and self-efficacy with regard to learning grammar.

Differences Between the Groups After the Experiment: Answers to RQ2

Grammar differences. Before examining any differences between the groups on the grammar posttest, all the major ANCOVA assumptions were tested and found tenable. First, all the grammar test scores were entered into SPSS and were checked in terms of normality of distribution via their skewness and kurtosis indices. Both values were found to be within the range of −1 to +1, suggesting that the data were of acceptably normal distribution and thus suitable for developing parametric analysis (Field, 2017; Peng & Woodrow, 2010). Second, as discussed above, the assumption of the independence of the covariate and treatment effect was met, since no significant difference was found between the groups at all levels (p > .05). Third, Table 4 shows nonsignificant F-values at p > .05 at all levels when customizing the ANCOVA model to examine the independent variable and covariate interaction. This indicates no statistically significant interaction between the groups at all levels, justifying the assumption of the homogeneity of the regression slopes. In short, the assumption test results verified the appropriateness of using the grammar pretest results as a covariate in running the main ANCOVA analysis for the posttest grammar performance.

While Table 5 presents the grammar posttest scores before and after adjustment using pretest grammar scores, ANCOVAs for the scores show that the group effects are significant at all levels, Whole: F(1, 176) = 86.716, p < .001; High: F(1, 63) = 13.150, p < .001; Intermediate: F(1, 33) = 45.673, p < .001; Low: F(1, 71) = 37.314, p < .001, with large effects (partial η² = .330, .173, .581, .344). The results indicate that, following grammar treatments, the TG significantly outperformed the CG on the grammar posttest, lending support to the claim that DDL treatment affects grammar students of all levels of proficiency.

Learning attitude differences. Before analyzing the differences between the groups with regard to their learning attitudes, the assumptions for running MANCOVAs on the data were tested. First, the assumption of the independence of the covariate and treatment effect had previously been found to
Table 3. Paired Sample t-Tests for Entry-Exit Questionnaire Results for Both Groups for All Levels.

| Groups/levels | Item   | Test  | N   | M    | SD   | t    | df | p    | d    |
|---------------|--------|-------|-----|------|------|------|----|------|------|
| TG/High       | Dimension 1 | Entry  | 44  | 21.39 | 4.08 | -0.12 | 43 | .909  | 0.01 |
|               |        | Exit   | 44  | 21.43 | 4.37 |       |    |       |      |
|               | Dimension 2 | Entry  | 44  | 12.27 | 2.80 | -0.19 | 43 | .847  | 0.03 |
|               |        | Exit   | 44  | 12.34 | 2.56 |       |    |       |      |
|               | Dimension 3 | Entry  | 44  | 13.82 | 2.37 | -1.35 | 43 | .185  | 0.18 |
|               |        | Exit   | 44  | 14.25 | 2.43 |       |    |       |      |
|               | Overall | Entry  | 44  | 47.48 | 7.74 | -0.75 | 43 | .455  | 0.07 |
|               |        | Exit   | 44  | 48.02 | 7.82 |       |    |       |      |
| TG/Intermediate | Dimension 1 | Entry  | 17  | 21.41 | 3.61 | -1.82 | 16 | .088  | 0.26 |
|                |        | Exit   | 17  | 22.35 | 3.62 |       |    |       |      |
|                | Dimension 2 | Entry  | 17  | 12.41 | 2.00 | -0.66 | 16 | .517  | 0.15 |
|                |        | Exit   | 17  | 12.71 | 2.54 |       |    |       |      |
|                | Dimension 3 | Entry  | 17  | 14.82 | 2.72 | 1.14  | 16 | .270  | 0.21 |
|                |        | Exit   | 17  | 14.24 | 3.44 |       |    |       |      |
|                | Overall | Entry  | 17  | 48.65 | 6.32 | -0.64 | 16 | .533  | 0.10 |
|                |        | Exit   | 17  | 49.29 | 8.29 |       |    |       |      |
| TG/Low         | Dimension 1 | Entry  | 34  | 19.53 | 4.18 | 0.13  | 33 | .897  | 0.02 |
|                |        | Exit   | 34  | 19.44 | 4.23 |       |    |       |      |
|                | Dimension 2 | Entry  | 34  | 11.97 | 2.87 | -1.21 | 33 | .234  | 0.15 |
|                |        | Exit   | 34  | 12.41 | 1.94 |       |    |       |      |
|                | Dimension 3 | Entry  | 34  | 13.50 | 2.95 | -0.59 | 33 | .556  | 0.09 |
|                |        | Exit   | 34  | 13.76 | 2.82 |       |    |       |      |
|                | Overall | Entry  | 34  | 45.00 | 8.80 | -0.60 | 33 | .552  | 0.07 |
|                |        | Exit   | 34  | 45.62 | 7.24 |       |    |       |      |
| TG/Whole group | Dimension 1 | Entry  | 95  | 20.73 | 4.09 | -0.50 | 94 | .619  | 0.04 |
|                |        | Exit   | 95  | 20.88 | 4.31 |       |    |       |      |
|                | Dimension 2 | Entry  | 95  | 12.19 | 2.68 | -1.09 | 94 | .278  | 0.09 |
|                |        | Exit   | 95  | 12.43 | 2.33 |       |    |       |      |
|                | Dimension 3 | Entry  | 95  | 13.88 | 2.66 | -0.80 | 94 | .426  | 0.07 |
|                |        | Exit   | 95  | 14.07 | 2.75 |       |    |       |      |
|                | Overall | Entry  | 95  | 46.80 | 7.96 | -1.25 | 94 | .264  | 0.07 |
|                |        | Exit   | 95  | 47.39 | 7.75 |       |    |       |      |
| CG/High        | Dimension 1 | Entry  | 22  | 23.50 | 4.13 | 0.91  | 21 | .371  | 0.19 |
|                |        | Exit   | 22  | 22.73 | 4.31 |       |    |       |      |
|                | Dimension 2 | Entry  | 22  | 13.41 | 2.13 | 0.49  | 21 | .627  | 0.11 |
|                |        | Exit   | 22  | 13.18 | 2.70 |       |    |       |      |
|                | Dimension 3 | Entry  | 22  | 14.45 | 2.89 | 0.00  | 21 | 1.000 | 0.00 |
|                |        | Exit   | 22  | 14.45 | 2.72 |       |    |       |      |
|                | Overall | Entry  | 22  | 51.36 | 7.42 | 0.70  | 21 | .489  | 0.13 |
|                |        | Exit   | 22  | 50.36 | 8.91 |       |    |       |      |
| CG/Intermediate | Dimension 1 | Entry  | 22  | 21.55 | 4.25 | -2.04 | 21 | .054  | 0.26 |
|                |        | Exit   | 22  | 22.64 | 3.71 |       |    |       |      |
|                | Dimension 2 | Entry  | 22  | 12.55 | 2.36 | 0.33  | 21 | .747  | 0.06 |
|                |        | Exit   | 22  | 12.41 | 2.24 |       |    |       |      |
|                | Dimension 3 | Entry  | 22  | 13.36 | 2.42 | -0.11 | 21 | .916  | 0.02 |
|                |        | Exit   | 22  | 13.41 | 2.77 |       |    |       |      |
|                | Overall | Entry  | 22  | 47.45 | 8.19 | -0.94 | 21 | .360  | 0.12 |
|                |        | Exit   | 22  | 48.45 | 7.54 |       |    |       |      |
| CG/Low         | Dimension 1 | Entry  | 40  | 19.60 | 4.29 | 0.09  | 39 | .926  | 0.01 |
|                |        | Exit   | 40  | 19.55 | 4.16 |       |    |       |      |
|                | Dimension 2 | Entry  | 40  | 12.35 | 2.41 | 1.19  | 39 | .243  | 0.20 |
|                |        | Exit   | 40  | 11.88 | 2.42 |       |    |       |      |

(continued)
be tenable, since nonsignificant independent t-test results emerged from the groups at all levels. Second, the assumption of the homogeneity of regression slopes was also tenable, as shown in Table 6: no statistically significant interactions were found between any level and its corresponding entry questionnaire score (p > .05). Third, Levene’s tests of equality of error variance were all statistically nonsignificant for all the exit questionnaire scores (p > .05), suggesting that the assumption of the homogeneity of variances was met for the scores at all levels. The results endorsed the feasibility of performing main MANCOVA analyses on the data.

Table 7 presents the adjusted mean scores for each exit questionnaire item after controlling for their overall entry scores as the covariates. In Table 8, no significant differences were found between the groups at all levels, Whole: Hotelling’s Trace = .029, F(3, 174) = 1.67, p > .05; High: Hotelling’s Trace = .015, F(3, 61) = 0.303, p > .05;
Table 6. The Interaction Between the Independent Variable Groups (Levels) and the Covariates—Entry Questionnaire Scores When Exit Scores Were Examined as Dependent Variables.

| Group (levels) × covariate | Hotelling’s Trace | F | Hypothesis df | Error df | p  |
|---------------------------|------------------|---|----------------|----------|----|
| High × Entry Questionnaire | .031             | 0.613 | 3 | 60 | .609 |
| Intermediate × Entry Questionnaire | .055 | 0.619 | 3 | 34 | .608 |
| Low × Entry Questionnaire | .052             | 1.197 | 3 | 69 | .317 |
| Whole group × Entry Questionnaire | .009 | 0.496 | 3 | 173 | .686 |

Table 7. Descriptive Statistics for the Exit Questionnaire Results After Controlling for the Covariate.

| Level     | Item | Group | N  | Unadjusted | Adjusted |
|-----------|------|-------|----|-------------|----------|
|           |      |       |    | M       | SE    | M       | SE    |
| High      |      | TG    | 44 | 21.43   | 0.66 | 21.98   | 0.45 |
|           |      | CG    | 22 | 22.73   | 0.92 | 21.64   | 0.65 |
| Dimension 1 |      | TG    | 44 | 12.34   | 0.39 | 12.62   | 0.31 |
|           |      | CG    | 22 | 13.18   | 0.58 | 12.63   | 0.45 |
| Dimension 2 |      | TG    | 44 | 14.25   | 0.37 | 14.49   | 0.32 |
|           |      | CG    | 22 | 14.45   | 0.58 | 13.97   | 0.46 |
| Dimension 3 |      | TG    | 44 | 48.02   | 1.18 | 49.08   | 0.81 |
|           |      | CG    | 22 | 50.36   | 1.90 | 48.24   | 1.16 |
| Overall   |      | TG    | 17 | 22.35   | 0.88 | 22.09   | 0.56 |
|           |      | CG    | 22 | 22.64   | 0.79 | 22.84   | 0.49 |
| Intermediate | Dimension 1 | TG    | 17 | 12.71   | 0.62 | 12.57   | 0.45 |
|           |      | CG    | 22 | 12.41   | 0.47 | 12.52   | 0.40 |
| Dimension 2 |      | TG    | 17 | 14.24   | 0.83 | 14.05   | 0.58 |
|           |      | CG    | 22 | 13.41   | 0.59 | 13.55   | 0.51 |
| Dimension 3 |      | TG    | 17 | 49.29   | 2.01 | 48.71   | 1.12 |
|           |      | CG    | 22 | 48.45   | 1.61 | 48.90   | 0.99 |
| Overall   |      | TG    | 34 | 19.44   | 0.73 | 19.43   | 0.56 |
|           |      | CG    | 40 | 19.55   | 0.66 | 19.56   | 0.52 |
| Low       |      | TG    | 34 | 12.41   | 0.33 | 12.41   | 0.32 |
| Dimension 1 |      | TG    | 34 | 13.76   | 0.48 | 13.76   | 0.39 |
|           |      | CG    | 40 | 13.03   | 0.51 | 13.03   | 0.36 |
| Dimension 2 |      | TG    | 34 | 45.62   | 1.24 | 45.60   | 0.91 |
|           |      | CG    | 40 | 44.45   | 1.33 | 44.47   | 0.84 |
| Dimension 3 |      | TG    | 95 | 20.88   | 0.44 | 20.97   | 0.31 |
|           |      | CG    | 84 | 21.19   | 0.47 | 21.10   | 0.33 |
| Overall   |      | TG    | 95 | 12.43   | 0.24 | 12.47   | 0.20 |
|           |      | CG    | 84 | 12.36   | 0.27 | 12.31   | 0.21 |
| Whole group | Dimension 1 | TG    | 95 | 14.07   | 0.28 | 14.12   | 0.23 |
|           |      | CG    | 84 | 13.50   | 0.33 | 13.44   | 0.24 |
| Dimension 2 |      | TG    | 95 | 47.39   | 0.80 | 47.57   | 0.53 |
|           |      | CG    | 84 | 47.05   | 0.94 | 46.85   | 0.57 |

Note. TG = treatment group; CG = control group.

Table 8. Multivariate Tests for the Groups’ Exit Questionnaire Scores by Hotelling’s Trace.

| Level     | Test          | Value | F   | Hypothesis df | Error df | p     | Partial η² |
|-----------|---------------|-------|-----|----------------|----------|-------|------------|
| High      | Hotelling’s Trace | .015  | 0.303 | 3 | 61 | .823 | .015 |
| Intermediate | Hotelling’s Trace | .055  | 0.619 | 3 | 34 | .608 | .052 |
| Low       | Hotelling’s Trace | .052  | 1.197 | 3 | 69 | .317 | .049 |
| Whole group | Hotelling’s Trace | .029  | 1.670 | 3 | 174 | .176 | .028 |
Interdisciplinary scholars (Johns, 1991; H. Lee et al., 2019) who advised that having students learn inductively, which is the theoretical practice of DDL (L. S. Huang, 2017), does indeed affect students’ receptivity to the DDL approach and/or its effectiveness on their performance. That is, having students learn inductively, which is the theoretical practice of DDL in grammar classes. In addition, the result supports Boulton (2010), whose investigation shows that beginning grammar learners can also benefit from learning with corpus material. Further, the finding has helped allay concerns raised by Lin (2016), H. J. H. Chen (2011), Hirata and Hirata (2013), Hirata et al. (2013), Liu and Jiang (2009), who all cast doubt on low-level students’ abilities to cope with DDL-centered activities, such as analyzing corpus data on their own.

The pedagogical suitability of DDL to grammar learning has been especially validated by the fact that all three different levels of the TG statistically significantly improved their grammar performance and outperformed their counterparts in the CG. To some extent, the positive learning gain may lend support to the theoretical claims for DDL. That is, having students learn inductively, which is the theoretical practice of DDL (L. S. Huang, 2017), does indeed have a beneficial effect on their acquisition (Shaffer, 1989). Similarly, the finding also dissolves the doubts of some scholars (Johns, 1991; H. Lee et al., 2019) who advised against students of different levels learning with a discovery-learning approach such as DDL. Finally, since DDL was operationalized through focusing students’ attention on observing the grammar rules, the significant improvements that the TG participants collectively made here further verifies that DDL does in fact effectively reflect the noticing hypothesis (Schmidt, 1990, 2001), as discussed in the literature review.

The fact that the greatest improvement was found in low-level DDL students is worth special attention. Echoing Saeedakhtar et al.’s (2020) study, this finding goes beyond the belief of most previous researchers on effectiveness of using DDL with high and intermediate proficiency students (Aston, 2001; Boulton & Cobb, 2017; H. Lee et al., 2019; Lin, 2016). The finding goes so far as to recommend it also as beneficial for lower-proficiency learners. It also endorses that of C. Y. Lee and Liu (2003), that low-vocabulary-level students were found to benefit more than students at other levels from concordancing. In addition, the result supports Boulton (2010), whose investigation shows that beginning grammar learners can also benefit from learning with corpus material. Further, the finding has helped allay concerns raised by Lin (2016), H. J. H. Chen (2011), Hirata and Hirata (2013), Hirata et al. (2013), Liu and Jiang (2009), who all cast doubt on low-level students’ abilities to cope with DDL-centered activities, such as analyzing corpus data on their own.

As a matter of fact, many of the student complaints about corpus-aided treatments are mainly about the time and effort needed to understand and interpret the numerous difficult sentences that an electronic corpus randomly throws up (M. Chen & Flowerdew, 2018; Liu & Jiang, 2009; Wu et al., 2018). Likewise, “the cognitive burden of using new technology (the concordancing software tool) during DDL may [have also] inhibit[ed] learning” (Moon & Oh, 2018, p. 51) in the present study. This may even apply to low-proficiency students (H. Lee et al., 2019) who must struggle with linguistic knowledge while learning to cope with an electronic corpus. Likewise, when compiling the printed material, the present researcher not only chose precisely sentences that were complete and stand-alone but also excluded those containing information too difficult to understand, for instance,
carrying too much sophisticated vocabulary or requiring field-specific knowledge. However, the paper-based DDL material used in the study of Lin (2016) clearly did not consider such principles, thus probably increasing the difficulty of the task for his DDL students (Boontam & Phocharoensil, 2018; Moon & Oh, 2018; Yoon & Hirvela, 2004), especially the lower-proficiency learners.

An important reason that low- and intermediate-level students benefited more from DDL grammar than did high-level students should be underlined. First, the advanced DDL students may have been somewhat more confined to their past learning experience. That is, they may have been able to gain greater benefit from the TDA, since the former’s being assigned to the higher-level group at the beginning of the current study resulted from their past learning model (Boulton, 2010). That being said, it does not mean that the DDL approach is inappropriate for higher-level students, because they also made significant gains with DDL. Rather, perhaps more of the higher-level students, such as those in the current study, may benefit more from traditional teaching than can those relatively unsuccessful language learners using a past learning method. Similarly, this reasoning does not suggest that all lower-level students would benefit from DDL treatments. Instead, for those specific low-proficiency students, such as those in the current study, who had no successful learning experience with the TDA, DDL can serve as an alternative and rather more effective pedagogical option.

Possible reasons should also be discussed for the significant improvement in the grammar performance of the low-level students in the CG, but not of the other students of higher-level proficiencies. One possibility is that, through the CG treatment where the main focus was on reading skills and appreciating English articles, the participants might at least have improved their general reading abilities and/or vocabulary. Such enhanced skills, however basic they may seem, may be advantageous for any language learners, especially those who are less successful, since they may be more likely to undergo greater struggles in more than one area in a learning a foreign language than higher-level student are. Therefore, after being empowered to read slightly more effectively and/or to have more vocabulary, the lower-level CG students might have been able to do better than they originally would. This may be more deeply appreciated if one considers that reading was the main skill they had for comprehending the test items; better knowledge of vocabulary probably might also have helped them to better understand the meaning of item statements. Although this may sound like a side issue in this study, the discussion here seems able to suggest that general or broad reading may also be an effective way to help low-level language students to improve. In contrast, to enable higher-level learners to break through their originally strong bottleneck in learning a foreign language, more influential or intensive pedagogical treatments may be needed.

While the discussion tends to acknowledge DDL’s usefulness in the teaching of grammar, the results drawn from the questionnaires somewhat undermine this tendency: the DDL approach did not improve the learning motivation and self-efficacy of the TG students with different proficiencies; nor were differences found between the TG and the CG at all levels. This is in line with Lin (2016), who also found in DDL no superior effect to the traditional approach in terms of the same affective factors examined here. Although the correspondence of the two studies seems to cast some doubt on the effects of DDL on students’ learning attitudes, this may be explained by the fact that both of the studies lasted only 3 weeks, which is perhaps too short a time to nurture such important affective perceptions. This explanation is particularly likely given that, before participating, none of the participants of either study had any experience of learning with the corpus-based inductive approach. Instead, they had long been exposed to learning grammar using the TDA, the mainstream practice even in today’s English grammar classes in Taiwan (H. C. Lee, 2013; Lin & Lee, 2019; Smith, 2011; Tamney & Chiang, 2002). This very experience of the past may have caused those students to take much more time to adopt affirmative attitudes about a wholly new approach such as DDL, rather than picking up linguistic knowledge through it. The observation of M. Chen and Flowerdew (2018, p. 355) lends support to this argument: DDL learners are given more responsibilities in exploring answers from the concordance by themselves, thus "mak[ing] the learning process appear to be more challenging and take longer.”

Finally, it should be acknowledged that the current study has certain limitations that require readers to treat its findings carefully, and it awaits contributions from future researchers. First, the sample size of this study justifies confidence in judging its findings, but it may still fail to represent the whole population of EFL grammar students. Future researchers may consider participants other than college students and explore how such students with different proficiencies react to DDL instruction. This line of inquiry would be more interesting when examining whether paper-based DDL and hands-on DDL with a computer would have different effects on students with different proficiency levels. Likewise, this study determined students’ levels of language proficiency only by means of grammar tests, neglecting other equally important language skills such as vocabulary/collocation. This leaves a gap for future researchers to fill. Furthermore, the current study focused solely on learners’ levels of language proficiency and did not take into account the participants’ learning preferences. This is worth pursuing, given the possibility that before the experiment, those who were deemed high achievers might prefer or be better able to benefit from their past learning model (i.e., TDA) than the low achievers. Understanding whether or not the DDL approach is beneficial for different learning styles can make it easier for future practitioners to tailor its application in class to more or less receptive students. Last but not least, in this
study the TG learned grammar with DDL for a total of only 6 hr, which may have prevented them from sufficiently expe-
rencing its practice, and thus suppressed some of its effects,
positive or negative, on their affective perceptions. Future
studies are thus suggested that they consider providing stu-
dents with more intensive or longer hours of DDL learning
so as to reveal richer and more accurate treatment effects on
this dimension.

Conclusion
This study empirically examines the effects of the DDL
approach on learners with different levels of language pro-
ficiency. The study findings are particularly meaningful in
that they empower the present researcher to verify that
DDL is pedagogically suitable for enhancing the linguistic
knowledge of university-level grammar learners, regardless
of their proficiency. However, it must not be forgotten that
this success seems to capitalize on the camera-ready DDL
material prepared by the teacher. Any inference that most of
the students would continue to benefit from hands-on DDL
by using an electronic corpus by themselves should be
treated with caution. This weighty issue is especially note-
worthy for rigorous educators and practitioners who priori-
tize the importance of nurturing learner attitudes (e.g.,
motivation, self-efficacy, or others), since this study has not
been able to offer robust evidence in favor of or against
DDL on this matter.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect
to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support
for the research, authorship, and/or publication of this article: This
article was written with funding support from Taiwan’s Ministry
of Science and Technology (MOST 108-2410-H-032-027; MOST
109-2410-H-032-063).

ORCID iD
Ming Huei Lin https://orcid.org/0000-0003-4926-7163

References
Aston, G. (2001). Learning with corpora: An overview. In G. Aston
(Ed.), Learning with corpora (pp. 6–45). Athelstan.
Bandura, A. (1986). Social foundations of thought and action: A
social cognitive theory. Prentice Hall.
Bandura, A. (1997). Self-efficacy: The exercise of control. Freeman.
Boontam, P., & Phocharoenril, S. (2018). Effectiveness of English
preposition learning through data-driven learning (DDL). 3L: Lan-
guage, Linguistics, Literature®, 24(3), 125–141. http://doi.
org/10.17576/3L-2018-2403-10

Boulton, A. (2010). Data-driven learning: Taking the computer out
of the equation. Language Learning, 60, 534–572. https://doi.
org/10.1111/j.1467-9922.2010.00566.x

Boulton, A., & Cobb, T. (2017). Corpus use in language learning:
A meta-analysis. Language Learning, 67(23), 48–393. https://
doi.org/10.1111/lang.12224

Chase, C. I. (1978). Measurement for educational evaluation (2nd
ed.). Addison-Wesley.

Chen, H. J. H. (2011). Developing and evaluating a web-based col-
location retrieval tool for EFL students and teachers. Computer
Assisted Language Learning, 24(1), 59–76. https://doi.org/10.
1080/09588221.2010.526945

Chen, M., & Flowerdew, J. (2018). A critical review of research and
practice in data-driven learning (DDL) in the academic writing
classroom. International Journal of Corpus Linguistics, 23(3),
335–369. https://doi.org/10.1075/ijcl.16130.che

Chen, M., Huang, S. T., Chang, J. S., & Liou, H. C. (2015).
Developing a corpus-based paraphrase tool to improve EFL
learners’ writing skills. Computer Assisted Language Learning,
28(1), 22–40. https://doi.org/10.1080/09588221.2013.783873

Conrad, S. (2000). Will corpus linguistics revolutionize grammar
teaching in the 21st century? TESOL Quarterly, 34(3), 548–
560. https://doi.org/10.2307/3587743

Crosthwaite, P. (2020). Taking DDL online: Designing, imple-
menting and evaluating a SPOC on data-driven learning for
tertiary L2 writing. Australian Review of Applied Linguistics,
43(2), 169–195.

Crosthwaite, P., Wong, L. L., & Cheung, J. (2019). Characterizing
postgraduate students’ corpus query and usage patterns for dis-
ciplinary data-driven learning. ReCALL, 31(3), 255–275.

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

Dörnyei, Z., & Ushioda, E. (2013). Teaching and researching:
Motivation. Routledge.

Field, A. (2017). Discovering statistics using IBM SPSS statistics
(5th ed.). SAGE.

Flowerdew, L. (2015). Data-driven learning and language learn-
ing theories. In A. Leikko-Szymańska & A. Boulton (Eds.),
Multiple affordances of language corpora for data-driven
learning (pp. 15–36). John Benjamins.

Frankenberg-Garcia, A. (2016). Corpora in ELT. In G. Hall (Ed.),
The Routledge handbook of English language teaching (pp.
383–398). Routledge.

Gavioli, L. (2001). The learner as researcher: Introducing corpus
concordancing in the classroom. In G. Aston (Ed.), Learning
with corpora (pp. 1000–1130). Athelstan/Bologna.

Hadley, G. (2002). An introduction to data-driven learning. RELC
Journal, 33(2), 99–124. https://doi.org/10.1177/003368820
203300205

Hirata, Y., & Hirata, Y. (2013). Data-driven learning and learner
interviews in a Japanese context. In S. K. S. Cheung, J. Fong,
W. Fong, F. L. Wang, & L. F. Kwok (Eds.), Hybrid learning
and continuing education (pp. 122–133). Springer.

Hirata, Y., Hirata, Y., & Thompson, P. (2013). Two different types
of corpora: Japanese students’ perceptions. In J. Lam, K. C.
Li, S. K. S. Cheung, & F. L. Wang (Eds.), Knowledge sharing
through technology (pp. 1–15). Springer.

13

Lin
Hong, S. C. (2010). EFL learners’ consciousness-raising through a corpus-based approach. *English Teaching*, 65(1), 57–86. https://doi.org/10.1588/entea.65.1.201003.57

Huang, L. S. (2011). Corpus-aided language learning. *ELT Journal*, 65, 481–484. https://doi.org/10.1093/elt/ccr031

Huang, L. S. (2017). Has corpus-based instruction reached a tipping point? Practical applications and pointers for teachers. *TESOL Journal*, 8(2), 295–313. https://doi.org/10.1002/tesj.271

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

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

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

Kılıçkaya, F. (2015). Computer-based grammar instruction in an EFL context: Improving the effectiveness of teaching adverbial clauses. *Computer Assisted Language Learning*, 28(4), 325–340. https://doi.org/10.1080/09588221.2013.818563

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

Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students’ approaches to corpus investigation. *Language Learning & Technology*, 5(3), 77–90. https://doi.org/10.10125/44567

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86. https://doi.org/10.1080/00461520500357524

Lee, C. Y., & Liou, H. C. (2003). A study of using web concordancing for English vocabulary learning in a Taiwanese high school context. *English Teaching & Learning*, 27(3), 35–56.

Lee, H., Warschauer, M., & Lee, J. H. (2017). The effects of concordance-based electronic glosses on L2 vocabulary learning. *Language Learning & Technology*, 21(2), 32–51. https://doi.org/10.10125/44610

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. https://doi.org/10.1093/aplin/amy012

Lee, H., Warschauer, M., & Lee, J. H. (2020). Toward the establishment of a data-driven learning model: Role of learner factors in corpus-based second language vocabulary learning. *The Modern Language Journal*, 104(2), 345–362. https://doi.org/10.1111/modl.12634

Lee, H. C. (2013). Investigating the effects of student learning of English using COL approach based on situational theories. *Computers in Human Behavior*, 29, 2211–2217. https://doi.org/10.1016/j.chb.2013.05.016

Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, 81, 14–25.

Lin, M. H. (2014). Effects of classroom blogging on ESL student writers: An empirical reassessment. *The Asia-Pacific Education Researcher*, 23(3), 577–590.

Lin, M. H. (2016). Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students’ learning attitudes and teachers’ perceptions in Taiwan. *TESOL Quarterly*, 50(4), 871–893. https://doi.org/10.1002/tesq.250

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

Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, 93(1), 61–78. https://doi.org/10.1540/j.mtaj.2009.00828.x

Meyer, J. F. (1988). Moral education in Taiwan. *Comparative Education Review*, 32, 20–38.

Mizumoto, A., Hamatani, S., & Imao, Y. (2017). Applying the bundle-move connection approach to the development of an online writing support tool for research articles. *Language Learning*, 67(4), 885–921. https://doi.org/10.1111/lang.12250

Moon, S., & Oh, S. Y. (2018). Unlearning overgenerated be through data-driven learning in the secondary EFL classroom. *ReCALL*, 30(1), 48–67. https://doi.org/10.1017/S0958344017002246

Noll, V. H., Scannell, D. P., & Craig, R. C. (1979). *Introduction to educational measurement* (4th ed.). Houghton Mifflin.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly*, 19(2), 139–158.

Peng, J. E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning*, 60(4), 834–876.

Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. https://doi.org/10.1111/lang.12079

Poole, R. (2016). A corpus-aided approach for the teaching and learning of rhetoric in an undergraduate composition course for L2 writers. *Journal of English for Academic Purposes*, 21, 99–109. https://doi.org/10.1016/j.jeap.2015.12.003

Richards, J., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics* (3rd ed.). Longman.

Saeeedakhtar, A., Bagerin, M., & Abdi, R. (2020). The effect of hands-on and hands-off data-driven learning on low-intermediate learners’ verb-preposition collocations. *System*, 91, 102268. https://doi.org/10.1016/j.system.2020.102268

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158. https://doi.org/10.1093/aplin/11.2.129

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press.

Sealey, A., & Thompson, P. (2007). Corpus, concordance, classification: Young learners in the L1 classroom. *Language Awareness*, 16(3), 208–223. https://doi.org/10.2167/la453.0

Shaffer, C. (1989). A comparison of inductive and deductive approaches to teaching foreign languages. *The Modern Language Journal*, 73(4), 395–403. https://doi.org/10.1111/j.1540-47 81.1989.tb05319.x

Smart, J. (2012). *Innovative approaches to ESL grammar instruction* [Unpublished doctoral dissertation]. Northern Arizona University.

Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL*, 26, 184–201. https://doi.org/10.1017/S095834401400081

Smith, S. (2011). Learner construction of corpora for general English in Taiwan. *Computer Assisted Language Learning*, 24(4), 291–316. https://doi.org/10.1080/09588221.2011.557024
Sun, X., & Hu, G. (2020). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/1362168820954459

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Tamney, J. B., & Chiang, H. L. (2002). *Modernization, globalization, and confucianism in Chinese societies*. Praegar.

Timmis, I. (2015). *Corpus linguistics for ELT: Research and practice*. Routledge.

Tsai, K. J. (2019). Corpora and dictionaries as learning aids: Inductive versus deductive approaches to constructing vocabulary knowledge. *Computer Assisted Language Learning, 32*(8), 805–826.

Uçar, S., & Yükselir, C. (2015). The effect of corpus-based activities on verb-noun collocations in EFL classes. *Turkish Online Journal of Educational Technology, 14*(2), 195–205.

Vyatkina, N. (2013). Discovery learning and teaching with electronic corpora in an advanced German grammar course. *Die Unterrichtspraxis/Teaching German, 46*(1), 44–61. https://doi.org/10.1111/tger.10128

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

Wang, H.-Y., & Tsai, C.-C. (2012). An exploration of elementary school students’ conceptions of learning: A drawing analysis. *The Asia-Pacific Education Researcher, 21*, 610–617.

Wu, S., Fitzgerald, A., Yu, A., & Witten, I. (2019). Developing and evaluating a learner-friendly collocation system with user query data. *International Journal of Computer-Assisted Language Learning and Teaching, 9*(2), 53–78.

Wu, S., Li, L., Witten, I., & Yu, A. (2018). A systematic review of using discipline-specific corpora for lexico-grammatical pattern learning: A case study for computer science postgraduates. *International Journal of Computer-Assisted Language Learning and Teaching, 8*(1), 31–49. http://doi.org/10.4018/IJCALLT.2018010103

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*, 257–283. https://doi.org/10.1016/j.jslw.2004.06.002

Zuffianò, A., Alessandri, G., Gerbino, M., Kanacri, B. P. L., Di Giunta, L., Milioni, M., & Caprara, G. V. (2013). Academic achievement: The unique contribution of self-efficacy beliefs in self-regulated learning beyond intelligence, personality traits, and self-esteem. *Learning and Individual Differences, 23*, 158–162. https://doi.org/10.1016/j.lindif.2012.07.010