INDONESIAN JOURNAL OF APPLIED LINGUISTICS
Vol. 14 No. 3, January 2025, pp. 626-640

Available online at:
https://ejournal.upi.edu/index.php/IJAL/article/view/64690

https://doi.org/10.17509/ijal.v14i3.64690

The impact of data-driven learning on the improvement of grammatical proficiency in the ESL classroom environment

Sabina Sultana[1], Manvender Kaur Sarjit Singh[2], Rabiul Islam[3*], and Hafizah Hajimia[4]

[1,2]School of Languages, Civilisation & Philosophy, Awang Had Salleh Graduate School of Arts and Sciences, University Utara Malaysia, UUM Sintok, Kedah, Malaysia
[3]School of International Studies, University Utara Malaysia, UUM Sintok, Kedah, Malaysia
[4]Department of English and Linguistics, Academy of Language Studies, Universiti Teknologi MARA, Arau, Perlis, Malaysia

ABSTRACT

Corpus-based data-driven learning (DDL) is an innovative approach that utilises electronic text collections for linguistic analysis, thereby enhancing teaching practices and learning skills for ESL/EFL students. This innovative method surpasses traditional language teaching approaches. The aim of this study was to examine the impact of incorporating corpora in teaching grammatical constructs, specifically subject-verb agreement rules for ESL/EFL students. Furthermore, it examined ESL/EFL students' perceptions of using corpora for grammar instruction. A mixed method research design was employed, collecting both quantitative and qualitative data through triangulation methods. Data collection involved written essays, two timed writing tasks (a pretest and post-test), and an individual semi-structured interviews. Quantitative data were analysed using rubrics and paired samples t-tests, while thematic analyses were applied to the qualitative data. The analysis of the essays revealed that the students made errors in subject-verb agreement. The paired sample t-test revealed a statistically significant *p*-value (.001<0.05), indicating a notable improvement in the students' mastery of subject-verb agreement rules after receiving DDL instruction. In addition, qualitative interview responses indicated that participants held positive opinions about learning through the DDL approach. They described it as enjoyable, fascinating, and challenging, and believed it to be an effective method for acquiring new grammatical skills. The study concluded with recommendations for English Language Teaching (ELT) instructors and curriculum designers.

Keywords: Data-driven learning; EFL; ESL; grammar; subject-verb agreement

Received:	**Revised:**
12 April 2023	19 October 2024
Accepted:	**Published:**
19 December 2024	31 January 2025

How to cite (in APA style):
Sultana, S., Singh, M. K. S., Islam, R., & Hajimia, H. (2025). The impact of data-driven learning on the improvement of grammatical proficiency in the ESL classroom environment. *Indonesian Journal of Applied Linguistics*, *14*(3), 626-640. https://doi.org/10.17509/ijal.v14i3.64690

INTRODUCTION

The use of corpora in language teaching and learning presents a promising prospect for transforming the way languages are taught and learned (Therova & McKay, 2024; Tosun & Sofu, 2023; Zare et al., 2024). In linguistic research, corpus linguistics involves the gathering and analysis of collections of authentic texts to provide evidence for describing the nature, structure, and use of languages (Boulton, 2017; Pérez-Paredes, 2022). Data-driven learning (DDL) is a method for learners to engage with corpora directly or indirectly through materials, representing a more radical approach to language teaching (Johns, 1988). It is defined as "the use of corpus tools and techniques for pedagogical purposes in a foreign/second language" (Boulton & Vyatkina, 2021 p. 68). This approach transforms language learning from a teacher-centred

*Corresponding author
Email: rabiul@uum.edu.my

process to an interactive, student-centred process. By engaging in DDL, learners are empowered to take control of their own learning and develop a deeper understanding of the language they are studying. It has a long tradition of using texts as the empirical basis for linguistic description, and examines all levels of language, including phonology, lexis, grammar, and discourse. Grammar is fundamental in English as a Foreign Language (EFL) or English as a Second Language (ESL) programmes, as it helps students in developing their language skills (Alsehibany & Abdelhalim, 2023). However, grammar mistakes remain a significant challenge for ESL/EFL learners globally, often leading to confusion and misunderstandings (Geluso, 2022; Rahman & Hassan, 2019; Tomas & Dulin, 2021; Tsulaia, 2021; Xu, 2022). Research by Özer and Özbay (2022) highlighted that students frequently overlook grammatical conventions, resulting in writing errors that impede effective communication.

Among these errors, subject-verb agreement (SVA) has been identified as particularly challenging (Goldin et al., 2023). For example, SVA rules are notably complex for Bangladeshi learners due to significant structural differences between Bengali and English grammar (Chawdhury & Kabir, 2014; Seraj et al., 2021; Rezwana et al., 2022; Sultana et al., 2023). In Bengali, the verb form is not influenced by the subject, whereas, in English, the verb must align with their respective subjects. Such differences result in frequent misunderstandings and persistent errors among learners.

Although English grammar is a required subject in Bangladesh from primary through upper secondary levels, many students fail to achieve proficiency in SVA rules (Ali & Hamid, 2020). Despite years of studying English grammar, students continue to make repeated mistakes SVA patterns when writing sentences, often due to a lack of interest in grammar class (Alahmadi, 2019; Sayma, 2020; Rezwana et al., 2022). Numerous studies have explored innovative teaching approaches, including Data-Driven Learning (DDL), to address these gaps effectively (Sun & Hu, 2020, 2023; Oktavianti et al., 2023). This study integrates a compiled corpus of online newspaper articles to provide learners with authentic and contextually relevant grammatical patterns. By employing critical thinking strategies, students engage more deeply with the material, leading to enhanced comprehension and application of grammar rules. DDL is also advantageous because it offers up-to-date knowledge, whereas textbooks often become outdated before reaching students (Crosthwaite, 2020; Fang et al., 2021; Lin, 2021; Salama 2023; Shubha, 2021).

Despite an increasing number of studies on DDL in tertiary education, its application at the secondary level remains underexplored. This study, therefore, seeks to fill this gap by investigating the impact of DDL on improving SVA proficiency among secondary-level learners. Additionally, it examines students' perceptions of using this innovative method for grammar instruction. To address these aims, the following research questions guide this study:

1. What is the effect of incorporating Data-Driven Learning (DDL) techniques in teaching subject-verb agreement rules on the grammatical accuracy and proficiency of English as a Second Language learners?
2. How do students perceive the use of Data-Driven Learning (DDL) in grammar instruction for learning subject-verb agreement rules in English as a Second Language classes?

Subject-Verb Agreement Difficulties in Grammar

Researchers have found that subject-verb agreement (SVA) issues are prevalent in students' writing across all educational levels, including tertiary students (Tafida & Okunade, 2016). Various scholars have focused on SVA errors in their studies, including Stapa and Izahar (2010) and Alahmadi (2019). They concur that SVA rules continue to pose challenges for English language learners. In a study conducted by Nurjanah (2017), it was found that university level students continue to persistently struggle with subject-verb agreement. In the same note, Thiagarajah and Razali (2021) found that most of the students' scores were categorised as poor, reflecting significant weaknesses in their subject-verb agreement skills. Students continue to encounter difficulties when applying subject-verb agreement in their writing. Many students experience issues with the general rules rather than specific ones.

According to Febriyanti (2019), difficulties with subject-verb agreement arise due to the absence of this rule in their first language. Multiple studies have demonstrated that students' errors span both performance and competence, indicating ongoing struggles with subject-verb agreement. The findings suggest that further research on SVA is needed to reduce errors and improve student performance. Goldin et al. (2023) outline that, subject-verb agreement can be segmented into five categories: subject-verb agreement basics, subject-verb agreement with expressions of amount, subject-verb agreement with there + be, and subject-verb agreement involving irregularities in quantity (Febriyanti, 2019; Riadil et al., 2023; Seraj et al., 2021; Tafida & Oikunade, 2016).

One of the common challenges in English grammar is ensuring proper subject-verb agreement, where the verb form must match the number and person of the subject. Students often struggle with this when dealing with collective nouns, such as

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

'team' or 'committee,' which can take singular or plural verbs depending on context. Confusion also arises with indefinite pronouns like 'everyone' or 'nobody', as they are singular and require singular verbs, despite referring to multiple individuals. In complex sentences with phrases or clauses intervening between the subject and verb, maintaining agreement becomes challenging, leading to errors in sentence construction (Goldin et al., 2023). Non-native speakers may encounter challenges when the subject and verb are separated by phrases such as 'along with,' 'as well as,' or 'in addition to,' since the verb must agree with the nearest subject.

Discrepancies in agreement can also occur in sentences with compound subjects joined by 'or' or 'nor,' where the verb agrees with the closer subject, depending on whether the subjects are singular or plural (Febriyanti, 2019). Learners might find subject-verb agreement particularly tricky in sentences with inverted word order, such as questions or sentences beginning with negative adverbs, where careful attention to sentence structure is required (Thiagarajah and Razali, 2021). To address this issue, this study employed Data-Driven Learning (DDL) method to improve subject-verb agreement rules, related to number, person, tense etc. DDL provides learners with real-life language data, such as texts or corpora, where subject-verb agreement occurs naturally.

Data Driven Learning (DDL)

The nature of the DDL approach reflects the characteristics of inductive approach, discovery learning, and the noticing hypothesis (Boulton & Vyatkina, 2021). It is generally accepted that DDL nurtures learners' ability to use inductive reasoning (Huang, 2017). However, Chen (2004) does not recommend the inductive-based approach for low-level language learners, as they considering it challenging. As an inductive method, DDL, also involves the process of discovery learning. According to Richards and Schmidt (2002), discovery learning involves five main steps: observing, inferring, formulating, predicting, and communicating. These components are integral to DDL which motivates learners to act as language researchers by actively participating in discovery process. This includes learning how to learn through observing, analysing, interpreting, and presenting language-use patterns found in corpus data (Huang, 2022).

The connections between discovery learning and DDL have been widely recognised in the field (Boulton & Vyatkina, 2021). However, it has also been suggested that students engaging in discovery learning tasks should possess a certain level of linguistic knowledge (Johns, 1991; Lee et al., 2019). In addition, learners enhance their language awareness when engaged in activities centred around DDL. They focus on observing and analysing recurring linguistic features of study which can potentially lead to improvements in language use. These findings are supported by Boulton & Vyatkina (2021), Flowerdew (2015), and Huang (2017). The method and its expected outcome align with the noticing hypothesis proposed by Schmidt (1990). According to his theory, increasing language awareness is crucial, as it involves learners consciously observing the linguistic features of interest. This process helps convert input into effective intake and ultimately leads to successful language output. Experimental assessments have highlighted the inherent quality of noticing in DDL, and multiple studies have provided positive evidence supporting this approach (Çekiç, 2022; Ma et al., 2023; Shubha, 2021).

Effects of DDL in Grammar Lessons

Many scholars have found DDL effective for learning collocations, writing and vocabulary (Johns, 1991; Shubha, 2021; O'Keeffe, 2021; Wu, 2021; Yu and Shen, 2022). Johns (1991) was one of the earliest scholars to advocate for the instructional effects of DDL on grammar students. He argued that allowing learners to analyse enough organized input, such as, (concordance lines) effectively facilitates their grammar acquisition. Johns' views are endorsed not only by early scholars (Singh, 2014) but also by later experimental findings (Shubha, 2021; Çekiç, 2022; Ma et al., 2023 Therova & McKay, 2024). According to Shubha (2021), a corpus is defined as a large and systematic collection of naturally occurring texts, whether written or spoken, stored electronically. Corpus linguistics refers to language studies based on corpora which first emerged in the 1950s. Researchers and instructors are increasingly recognising the potential and advantages of corpora for language learning. The use of corpora in language training began in the 1980s.

Data-Driven Learning refers to the practise of utilising corpora for language instruction, a term coined by Johns in 1990. In the foreign language classroom, corpus can be used in two different ways. The first way is indirectly, where students use corpus-based resources created by the instructor to identify linguistic patterns. The second way is directly, where students use corpus software to explore the language. There are three types of Data-Driven Learning methods: computer-based, paper-based, and hands-on (Çekiç, 2022; Hajimia et al., 2019; Ma et al., 2023).

In the DDL exercises, students assume the role of researchers. They analyse data related to a specific linguistic feature, classify the data, draw conclusions based on the evidence, and formulate findings. Language pattern analysis can be conducted by utilising corpora in both quantitative and qualitative manners, allowing students to

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

determine frequency of occurrence of the target feature in the corpus.

Yepes and Krishnamurthy (2010) examined the effectiveness of both corpus-based and corpus-driven strategies by teaching Spanish grammar to two groups of participants with varied competence levels using the Aston Corpus Network (ACORN). While the initial learners (GC1) were instructed using a corpus-based approach, while the advanced learners (GC2) were taught in using a corpus-driven manner. The results showed that students appreciated corpus-based grammar training because they believed it would help them complete tasks in the future.

In Özer & Özbay's (2022) study, the corpus was utilised to teach students about adverbs and adjectives. They observe that most students felt positively about the DDL approach nothing it was different from traditional English instruction and effective. The discussion supports the use of DDL in ESL/ EFL grammar classes, particularly for higher-proficiency language learners. In fact, few studies have empirically focused on verifying whether DDL suits this demographic. Given the scarce and inconsistent empirical evidence, further empirical experiments are urgently needed to confirm whether DDL is pedagogically suitable for ESL/EFL students at lower levels especially for improving subject-verb agreement rules.

METHOD

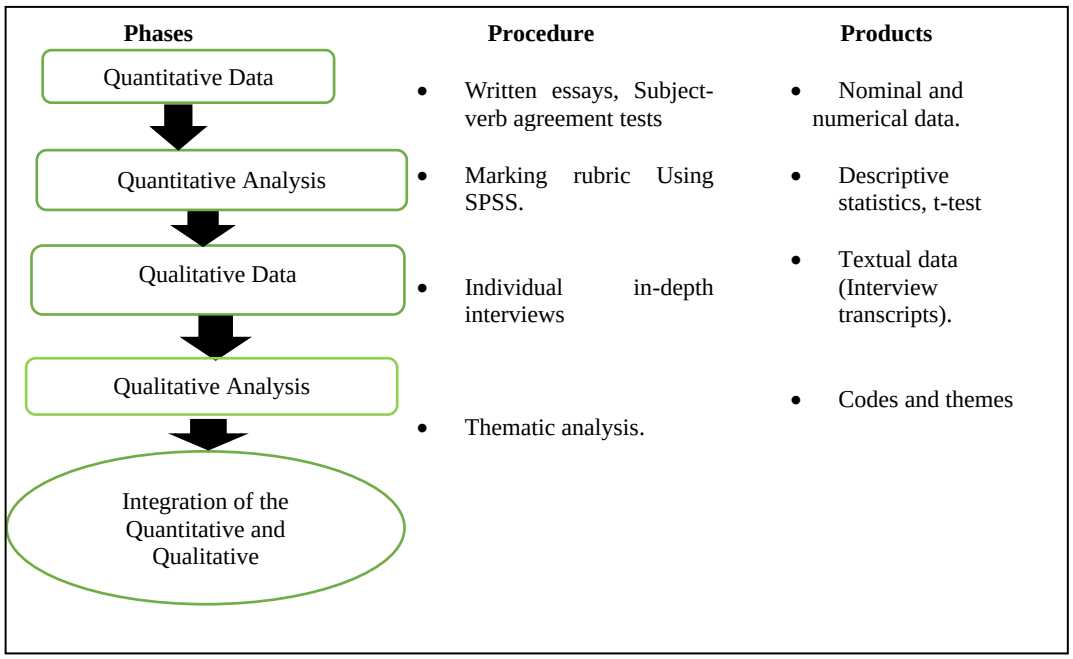
Research Design

The mixed methods of data collection utilised is well suited to dealing with a relatively small number of respondents within a limited time and the funding restrictions on this examination. This dual approach provides a balanced way to address both breadth and depth in understanding the research problem. It is noted that this form of methodology is supported by Bryman (2012) and Bryman and Bell (2015), who consider it advantageous to adopt different data collection methods to offset any perceived weaknesses in the process. Several scholars (Plano Clark & Ivankova 2016; Terrell, 2012; Creswell, 2014) have listed various types of mixed methods research design. The present study employed a mixed-method research design, in particular Creswell's (2014) sequential explanatory mixed methods to produce a more comprehensive understanding of the research problem.

The flowchart developed by Ivankova et al. (2006), reproduced in Figure 1, assists readers with a visual learning style in understanding the operation of the sequential explanatory model. It outlines distinct phases for quantitative and qualitative analysis, thereby clarifying the research process. This dual approach employs a variety of data collection and analysis techniques, offering both a thorough explanatory methodology and a profound intuitive understanding. It aims to advance knowledge of these issues encountered by students in learning grammar, particularly with subject-verb agreement (SVA) problems. By combining quantitative and qualitative studies, this method enables the exploration of the difficulties and underlying causes associated with grammar education among ESL students.

Figure 1

Explanatory Mixed-Methods Design Procedures (Ivankova et al., 2006)



Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

Respondents

The need to generalise the results of the study led the researcher to favour probability sampling over non-probability sampling. The research sample comprised of 41 nine-grade students from one rural high school in Bangladesh, selected from 20 schools in the same area. These participants were typical in age and English proficiency for their grade level, ensuring the study's relevance. The students were chosen based on random sampling recommended by Creswell (2014). To facilitate data analysis, each student was assigned codes, from P1, representing the first student, to P40, representing the last student. For quantitative research, a random sample was used, whereas purposive sampling was employed for qualitative research in explanatory studies. Specifically, ten students from the experimental group at the chosen school were selected for the semi-structured interviews to evaluate their attitudes regarding DDL.

The DDL treatment

After the pretest, the group had one 90-minute DDL grammar lesson per week for 3 weeks. The grammar items for each week comprised the proper use of subject-verb agreement. These topics were selected based on the researcher's teaching experience, which identified common confusion among learners regarding SVA. First, the researcher showed the students concordance lines, which contained the node word and key words in context, for observation. The material, which the researcher compiled and delivered to students in the form of printed handouts, comprised authentic examples from the Corpus of Bangladesh Online Newspaper Articles (CBONA). In reading the concordance, the students were given several minutes to answer a general question regarding SVA; they could either work it out on their own or discuss it with peers and share their findings. When they were unable to analyse the concordance or offered inaccurate inferences, more guidance or questions were given, such as "Please observe the verb after the key words," "Please look at the subject in the first two sentences and that in the following sentences," or "Can you detect any differences between the functionality of I and He in these sentences?" After sharing answers, learner understanding was checked by creating sentences, answering multiple-choice questions, and judging correct or incorrect grammar sentences.

Instruments

To address the research difficulties of the current study, the researcher developed three research instruments: a written essay, a pre-test and post-test experiment, and a semi-structured interview. In the semi-structured interview, only selected students took part, and the written essay, pretest, and post-test would all solely employ student work samples. Using work samples from the same group of students ensures consistency in the data, making it easier to track progress and attribute changes directly to the intervention. According to research in educational assessment, consistent sampling is crucial for reliable measurement of learning outcomes (Khan et al., 2021). Like Khan et al. (2021), Li & Zhang (2022) mentioned that it is more practical to manage and analyse a smaller, selected sample in detail. This is particularly important in educational research, where resources and time are often limited. The essay was structured as a 200-word persuasive task, enabling focused evaluation of grammar usage. The pre-test and post-test for the subject-verb agreement assessment in this study comprised sentences of varying complexity, designed to evaluate participants' grasp of subject-verb agreement rules. To ensure relevance and reliability, questions were adapted from previous studies (Karim et al., 2015; Middleton, 2022; Thiagarajah & Razali, 2021).

Reliability was ensured through consistent formatting and pilot testing. In order to gather qualitative data for the current study and determine students' ideas and impressions of the many components of learning English writing, semi-structured interviews were used to understand their perceptions regarding any challenges. These questions were adapted from various studies conducted by Sah (2015), Crosthwaite (2020), and Çekiç (2022). Then, the questions were refined through expert review and pilot testing. Linguistics experts provided feedback to ensure the questions were clear, relevant, and unbiased. A pilot interview was conducted with a small sample from the study population to test the questions' effectiveness and clarity. Based on this feedback, the questions were further refined to ensure they elicited comprehensive and meaningful responses. Finally, the interview questions were organized into a logical sequence, beginning with general questions to establish rapport, and gradually moving to more specific and detailed inquiries about subject-verb agreement. This structured approach ensured that the interview questions were both reliable and valid for capturing the necessary data.

Data Analysis

The data was analysed quantitatively and qualitatively. An essay marking rubric with unequal points for each area (i.e., content, vocabulary, grammar, and mechanics) was used to evaluate the quality of the texts from the written essays. The scoring scale for each rubric category (poor=1, good=2, and excellent=3) was applied. This rubric is reliable because it is based on a tested high school framework and refined through practical use. Educational experts highlight that such development enhances reliability. It includes clear categories (content, vocabulary, grammar, mechanics) and a

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

specific scoring scale (poor, good, excellent), ensuring consistent assessments. The quantitative scale (1, 2, 3) allows objective measurement, making it more reliable than qualitative assessments alone. It has been adapted to fit the study's context, enhancing relevance and reliability. The unequal weighting of categories (e.g., content over mechanics) ensures a nuanced and accurate evaluation. In order to detect and categorise different subject-verb agreement mistakes produced by school children in writing tasks from the experimental groups, such as tense, number, and so on, the data from the pretest and post-test were analysed. Using SPSS statistical software, descriptive statistical analyses were performed on the pretest and post-test results for the 40 student participants. Paired sample *t*-tests were also conducted to determine if there were any significant differences between the groups at each level regarding the outcomes of the SVA test. The *t*-tests determined if each group, categorised by level, improved their grammar performance and learning attitudes after the experiment. Thematic analysis has been used to examine the current study's qualitative data. The steps of thematic analysis proposed by Braun and Clarke (2006) were used in the current investigation. The process entails a meticulous analysis of the data to identify emerging codes. These codes are subsequently categorised into themes. The data was transcribed and analysed to identify codes. These codes were then categorised into themes based on emerging patterns.

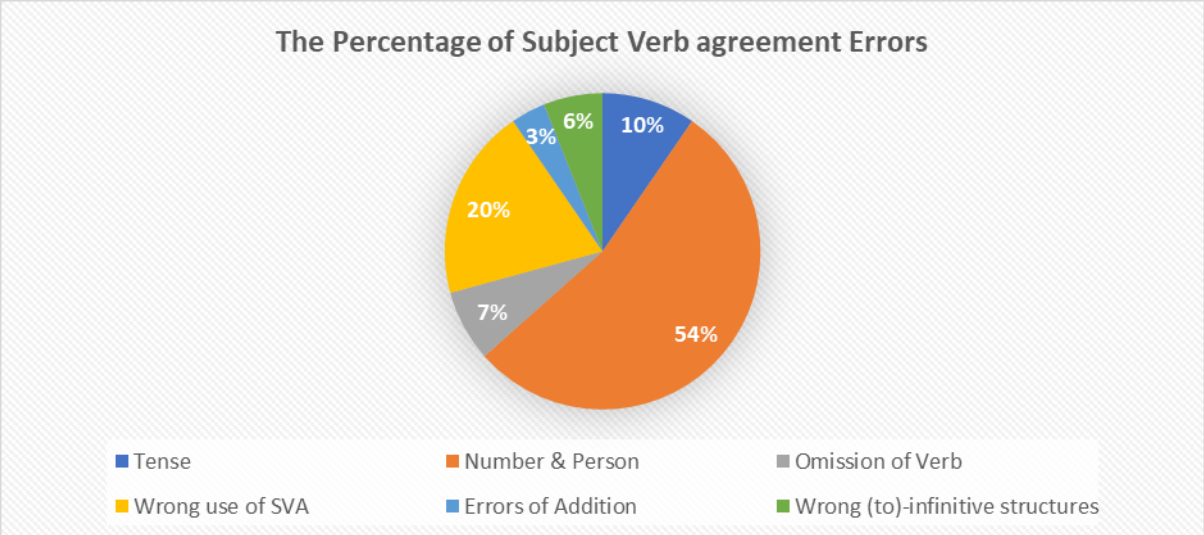
FINDINGS AND DISCUSSIONS

Findings from the Written Essays:

The essay analysis revealed that all participants received lower exam results, underscoring significant challenges in meeting subject-verb agreement requirements. Analysis of writing proficiency indicated that 50% and 25% of student essays, respectively, demonstrated inadequate skills, with 25% of students failing based on established criteria. Most students struggled with verb usage, particularly verb tenses, suggesting a fundamental misunderstanding or lack of knowledge in this area. Detailed error analysis identified six primary types of subject-verb agreement issues: tense, number and person, verb omission, incorrect SVA usage, addition errors, and incorrect infinitive constructions. The researcher recorded 24 out of 40 instances of tense errors, 40 instances of number and person errors, 26 instances of verb omission errors, 40 instances of incorrect SVA usage, 15 addition errors, and 16 incorrect (to)-infinitive structures. Cumulatively, the data revealed 46 tense errors, 258 number and person errors, 35 verb omission errors, 94 incorrect SVA usage errors, 29 incorrect (to)-infinitive constructions, and 17 addition errors, resulting in a total of 479 errors. This comprehensive error analysis highlights the pervasive nature of subject-verb agreement issues among the student population. Figure 2 below visualises the percentage of subject-verb agreement errors.

Figure 2

Percentage of the subject-verb agreement errors



The first findings of this study differ significantly from previous research. This investigation identified that the most common errors in students' writing were related to number and person agreement, which occurred 258 times (53.86%), and incorrect placement, occurring 94 times (19.6%). Tense errors were the second most frequent, with 46 instances (9.6%). Verb omissions were the third most common, occurring 35 times (7.3%), followed by improper (to)-infinitive constructions at 29 instances (6.05%), and addition errors at 17 instances (3.54%). The high prevalence of number and person

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

errors highlights students' challenges with correctly using singular and plural forms.

These findings contrast with Atashian et al. (2018) reported that adverbs, and pronouns were the most common errors among students at the University of Nizwa in Oman. Promsupa (2017) identified grammatical errors as predominantly morphological (81.97%) and syntactic (18.03%). Ibrahim (2020) categorized errors made by Sudanese EFL students into ten types: singular/plural forms (20), verb tenses (7), prepositions (7), subject-verb agreements (27), articles (16), spelling (40), verb forms (7), capitalization (8), misuse of terms (8), and missing words (13). Ibrahim also investigated the causes of these errors, citing uncertainty of basic grammatical structure, carelessness, word-by-word translation, and inadequate vocabulary as key factors. This study's results align with Nguyen et al (2022), Suraprajit (2021), and Sari (2019) who concluded that grammatical errors in essays were due to carelessness, first language interference, subject-verb agreement, tenses, misformation, and translation issues. Similarly, Fang et al. (2021) found that intralingual transfer (35.33%) was the main cause of errors, followed by interlingual transfer (34.50%) and learning context (30.15%). Alahmadi (2019) found that ESL students encounter many difficulties distinguishing between spoken and written words, and they struggle with grammatical issues such as subject-verb agreement and how to correctly combine sentences to write a paragraph. Likewise, Farooq (2020) also asserted that ESL students have difficulty with grammar, which is seen to be the most difficult part of writing. When learners correctly compose sentences, structures, and paragraphs, they often encounter a variety of difficulties. A group of sentences, the usage of different sentence forms, subject-verb agreement, parallel construction, the placement of modifiers, and tense agreement are all examples of grammar skills.

Findings from the pre-test and post-test:

The outcomes of the pre-test and post-test findings are then shown. Table 1 displays the pre- and post-test writing component scores for students reveals that none of the students received good or exceptional marks on the pre-test, with 18 students (or 45%) classified as very poor, 11 students (or 27.5%) as poor, and 11 students (or 27.5%) as average. According to the students' subject-verb agreement on the post-test, there were 5 students (12%) who scored poorly, 13 students (32.5%) who scored averagely, 15 students (37.5%) who scored well, and 7 (17.5%) who scored very well.

Table 1

Frequency and Rate Percentage of the Students' Subject-Verb Agreement Scores

No	Score	Category	Pre-test		Post-test	
			Frequency	Percentage	Frequency	Percentage
1	30-40	Very Good	-	0	7	17.5
2	20-29	Good	-	0	15	37.5
3	15-19	Average	11	27.5	13	32.5
4	10-14	Poor	11	27.5	5	12.5
5	5-9	Very Poor	18	45	-	0
	Total		**40**	**100%**	**40**	**100%**

Moreover, descriptive statistics of the pre- and post-tests on the writing assignments given to ESL students were completed. Table 2 presents the outcomes. Students' scores on the post-tests compared favourably to those on the pretests in terms of structural knowledge. Descriptive data for the Subject-Verb Agreement pre-test and post-test are provided in Table 5. Also, students' scores on the SVA post-test significantly improved as compared to the pretest (Minimum=4, Maximum=19, Mean=10.85, SD=4.458) (Minimum = 9, Maximum =33, Mean = 20.57, SD = 7.196). Due to extreme student performance, the context clues pretest and post-test's standard deviations were high.

Table 2

Descriptive statistics for Subject-Verb Agreement pre-test and post-tests

	N	Minimum	Maximum	Mean	Std. Deviation
Pre-test	40	4	19	10.85	4.458
Post-test	40	9	33	20.57	7.196

To analyse any potential differences in scores with and without the use of DDL, it is required to compare the outcomes of the pre- and post-subject-verb agreement tests. The following findings and results are provided in the following tables because of the analysis of the data using the paired samples test:

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

Table 3
Paired Samples Statistics of the experimental group

		Mean	N	Std. Deviation	Std. Error
Pair 1	Pre-test	10.85	40	4.458	.705
	Post-test	20.58	40	7.196	1.138

The experimental group's mean post-test value was 20.58, the standard deviation was 7.196, and the standard error mean was 1.138. The experimental group's mean pretest value was 10.85, the standard deviation was 4.458, and the standard error mean was.705. The statistics for the experimental group are shown in Table 5 above. The findings demonstrate that there is a substantial (p=.001) difference in the experimental group's mean score between the pre-test and post-test. This suggests that the learners made sufficient progress following DDL therapy. Study of experimental group performance using a paired sample test to compare means. Table 4 further explains the idea/

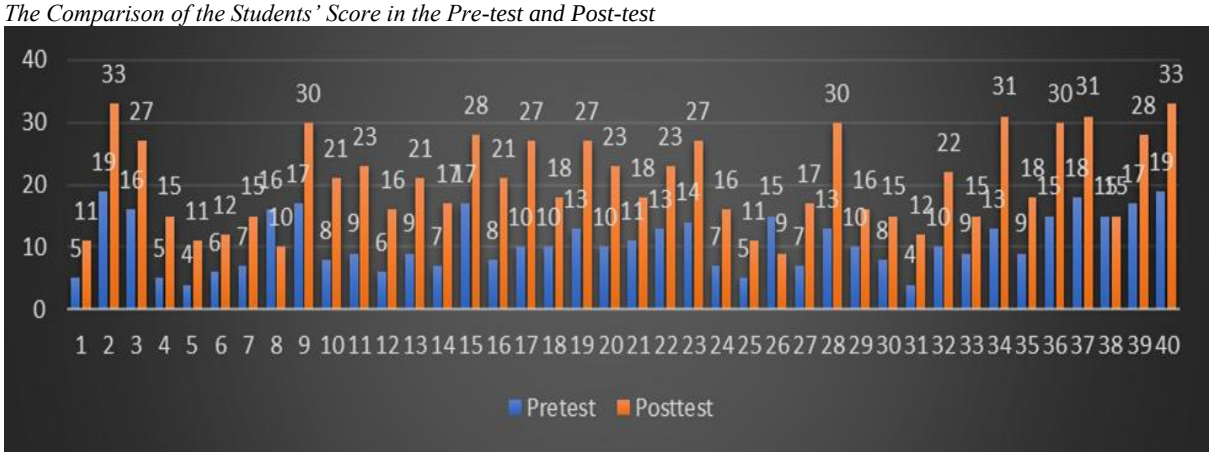
Table 4
Comparison of means using paired sample test.

	Pair 1
	Pre-test and post-test scores
Paired differences mean	9.725
Std. Deviation	5.179
Std Error Mean	.819
95% Confidence Interval of the Lower	-11.381
Difference Upper	-8.069
t	11.877
df	39
Sig.(2-tailed)	<.001

In the experimental group, the paired sample mean difference between the pre-test and the post-test was 9.725 with a standard deviation of 5.179, a standard error of.819, and a t-value of 11.877. Because the Sig. (2-tailed) pair 1 was less than the value of probability 0.05, it was.001.

It indicates that there was a substantial difference between the experimental group's pre-test and post-test results. Data, however, indicated that after getting therapy, pupils in the experimental group improved. The aforementioned table demonstrates a statistically significant change between the pre- and post-test results for all treatment/intervention plan variables for students in EG. After getting therapy, mean treatment plan factors, such as planning (pre-test =10.85, post-test =20.58), were greater. Figure 3 displays how the writing scores changed between the pre-test and post-test.

Figure 3
The Comparison of the Students' Score in the Pre-test and Post-test



The researcher inferred that using the DDL technique in teaching and learning to the ninth-grade students was 100 percent successful based on the research results described above as findings. The difference between the pupils' maximum scores before and after the treatments could be noted. The maximum score for the students on the pretest was 19. On the post-test, it showed signs of progress and was 33. The students' SVA progress is still insufficient in the pre-test. In addition, their concepts on subject-verb agreement in grammar were insufficient to fill in the spaces appropriately.

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

Several of them still committed grammatical errors. However, once the researcher used the DDL approach to treat the students, they were able to reduce their mistakes and write better sentences. Also, the post-average test's score was 20.58. Based on the average ability, it could be inferred that their writing had improved and that they were better able to write appropriately and grasp the SVA regulations. The students after getting treatment exhibit a substantial change between the pre-test and post-test.

These results help answer the research question by showing a statistically significant improvement in mean scores after using Data-Driven Learning (DDL) to teach grammar. This indicates that learners performed better with DDL. The findings support previous studies that highlight DDL's beneficial role in academic English lecture comprehension, task motivation, and learner autonomy (Lin & Lee, 2015; Zare & Aqajani Delavar, 2023). Additionally, the study aligns with research by Huang (2014), Boontam and Phoocharoensil (2018), and O'Keeffe (2021), which found paper based DDL activities effective in helping students independently acquire grammatical knowledge and language patterns. Survey results also mirrored the findings of Yepes and Krishnamurthy (2010), where most learners found DDL helpful for acquiring new grammatical knowledge and expressed a desire to learn other English lessons through DDL activities (Lin & Lee, 2015; Nugraha et al., 2017). Interestingly, this study found that DDL aided long-term memory retention and encouraged active learning, distinguishing it from traditional methods (Lin & Lee, 2015; Nugraha et al., 2017; Boontam & Phoocharoensil, 2018). Their study supports Chen's (2004) notion that transforming learners into language researchers can enhance classroom performance. The research demonstrated that DDL is effective for developing lexico-grammatical skills, particularly in advanced EFL writing. This is supported by literature (Noguera-Díaz & Pérez-Paredes, 2020; Römer et al., 2020; Sun & Hu, 2023), which highlights several advantages of DDL for language learners. For instance, learners actively engage with extensive language material independently, consolidating their autonomy. Exposure to authentic language input increases sensitivity to language variation, and discovery learning is enhanced through DDL's inquiry-based approach, which contrasts with explicit teacher instruction. Using DDL resources such as BNC, COCA, and CANCODE enabled learners to make their own judgments about lexico-grammatical items. The results align with studies by Yoon and Hirvela (2004), Sah (2015), Wu (2021), Yu and Shen (2022), Salama (2023), Emir & Yangın Ekşi (2023), Crosthwaite (2020), and Ma et al. (2023), which reported the effectiveness of DDL in developing lexico-grammatical skills in EFL writing. However, the temporary impact of form-focused DDL on task engagement can be attributed to increased foreign language anxiety (Zare et al., 2022), which negatively affects learner engagement (O'Reilly & García-Castro, 2022). In summary, this study revealed that incorporating DDL form-focused tasks into English language classes can enhance EFL learners' engagement in the short term.

Findings from the Interviews

According to the results of the interview replies, most respondents had a favourable opinion of the DDL method's application in English classrooms in terms of a number of different areas. In addition, 83.33% of the respondents said they are extremely interested and find it to be very simple. When learning to write with DDL exercises, they never became discouraged or irritated. Finally, they discovered that they felt comfortable handling the writing evaluation. While 16.67% of students said they found it extremely difficult, they made an honest attempt. The primary themes and categories in this respect were discovered using thematic analysis. The findings revealed opinions of the DDL-based learning that were both favourable and unfavourable. The topics and categories concerning the positive and negative opinions of the ESL students are then further developed. This is a summary of the key points of these themes:

Theme 1: Importance of ESL teaching strategies

Participants claimed that Data-driven learning techniques enabled them to impart grammatical education. Using visual aids like concordance lines and important phrases in context, they may complete academic activities based on their replies. As S1 put it, "If I'm learning a new vocabulary, I'll normally search one word, several words also appeared to display the vocabulary and sentence patterns to assist pupils remember what the terms imply." The usefulness of DDL techniques for young language learners was also acknowledged by the participants. Thus, it is crucial to add DDL-based exercises to help students become more fluent in their language.

Theme 2: Effective learning

When asked if data-driven learning was successful, two students responded that learning the SVA rules using DDL was more effective. The flexibility of blended learning also aided pupils in their learning. Student 3 stated, "Occasionally, even when one of my classmates is ill, they may still study utilising DDL-based activities from outside the classroom.

Theme 3: Autonomous learning

The DDL system was installed, all interviewees said, making us study more independently. The students were also questioned if they felt more

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

autonomous in their learning using the DDL system. "DDL based learning, students are expected to be more independent in learning since in learning online, for example, needing to download and search terms without being monitored," said student 2. Student 3 stated, however, "I feel they cannot be independent because they can repeat the learning content but cannot ask questions that are not obvious." ". Moreover, student 1 said that "For us as students, using more educational resources simultaneously is sufficient.

Theme 4: Easy to use

All respondents who were asked if they knew how to utilise DDL said that they did it extremely readily. This was brought on by the teacher's frequent use of paper-based DDL in the classroom. Student 1 stated, "It is sufficient since the textbook is real and the teacher uses it frequently. Student 9 said, "Really, I had known how to utilise it," in a similar fashion. Also highlighted by student number eight was the fact that "... yesterday our school had a training, but I didn't go. I know I can learn on my own, and I believe I can even if I study by myself.

Theme 5: DDL motivates students to write

Every participant in the interview was questioned about their thoughts about DDL. All of the participants reported their satisfaction with DDL.

DDL was recognised as being interesting by six participants. "I think it's fascinating. Students were drawn to DDL for a variety of reasons, including the use of computers in the writing process, which greatly aids in the development of their grammatical abilities. Student 7 said, "I don't feel drowsy when I use the computer." "Because I enjoy studying on a computer." I simply had to type. Other from that, one person expressed her excitement about developing DDL.

I feel excited, she said. Another participant said, "I can write more readily because of the offered words," adding that he found it "a bit easier." Glad, happy! was one participant's cheery response because it was novel to her. In addition, when asked if they would use it again, all of the participants gave the same response, i.e., they would.

Theme 6: DDL helps them to write better

It was discovered that several participants claimed that using the DDL's pre-established patterns helped them come up with ideas for writing appropriate phrases. The participants also mentioned that when they worked in groups, their buddies would point out any grammar mistakes. The participants also succeeded in expanding their vocabulary by looking up new words on other websites. "My phrase has gotten lengthier in the narrative writing." I may use a new word, therefore. "I find the new words from the computer based DDL." These results unequivocally demonstrate that participants had a favourable attitude regarding DDL, which they believe will help them become more proficient in grammar.

Theme 7: Improving technological skills

When asked if the computer based DDL might help students develop their technology (ICT) abilities, 100% of the respondents agreed that it could. According to Student 1, "I can run the computer better wherever and whenever I am using this DDL, especially the most important one is establishing the characteristics of DDL programme." Students 3 also noted that "enhancing the skills of my ICT. For instance, I was previously ignorant of how to use technology like Google, Corpus, and others. In my experience, there is an increase since I utilise technology often ". Student 5 added another comment, saying that "since I typically utilise a laptop and network to do homework, I need to improve my ICT abilities. Thus, whether I like it or not, I must be able to utilise DDL tools on a computer to study grammar. I had never learnt it before to applying for DDL.

The results from the interviews both corroborate and challenge previous research in several ways. Firstly, they affirm that Data-Driven Learning (DDL) boosts understanding of English grammar. Participants exhibited notable improvement following the DDL lesson, which aligns with Troy and Millar (2019), who found that DDL significantly enhanced students' learning of English phrasal verbs and improved test scores. Similarly, the findings are in line with Koosha and Jafarpour (2006), who emphasized DDL's effectiveness in teaching prepositional collocations. Additionally, using phrasal verb quizzes as supplementary exercises helped students review their knowledge, thereby improving the effectiveness of the DDL method. Secondly, the study found positive relationship between students' educational levels and their success in learning English grammar through DDL. This indicates that educational background reliably predicts DDL learning outcomes. Therefore, if teachers create materials that match learners' language proficiency, DDL can be effectively used with a wide range of learner groups, from young to adult learners. However, the study also revealed that students preferred learner-centred DDL approach over the traditional teacher-centred methods. Although they acknowledged DDL's utility for learning vocabulary and sentence structures, they did not enjoy the process and found it difficult to understand concordance lines. This finding is consistent with Boulton (2017), where participants found DDL challenging, but it contrasts with the studies of Boontam & Phoocharoensil, (2018) and Zare et al. (2022) which reported positive student attitudes towards DDL and their desire to continue

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

using it. Additionally, the study demonstrated that DDL improved students' knowledge of collocations, confirming earlier research by Vyatkina (2016). This improvement is attributed to the cognitive engagement required by DDL, including activities like noticing, exploring, analyzing, and verifying linguistic patterns, as suggested by Johns (1991) and Flowerdew (2015). Regular practice and exposure to multiple contextual examples of collocations enabled learners to effectively understand and acquire new collocations. Despite its benefits, DDL has some disadvantages. The cognitive load associated with DDL can hinder its effectiveness, as noted by Lin and Lee (2017). Lee et al. (2019) also pointed out that DDL's demanding nature might reduce its effectiveness for some learners. Moreover, participants believed their schools might resist adopting DDL due to its technical challenges and costs, as discussed by Sun and Hu (2020). Resistance to new methods, highlighted by Boulton (2009), also limits DDL's classroom adoption. In conclusion, while DDL has been shown to be effective in improving learners' knowledge of grammar and its technical challenges, and general resistance to new methods restrict its wider use in educational settings. Teachers must consider these factors carefully when incorporating DDL into their teaching practices.

CONCLUSION

The study found that DDL is a useful technique for teaching grammar to ESL/EFL students. The purpose of this study was to enhance comprehension of how DDL activities can be integrated into grammar instruction in an L2 setting to promote autonomy and enhance learning effectiveness for ESL/EFL learners. The written performances of the students were analysed quantitatively in terms of writing development measures. Additionally, the learners' attitudes towards the DDL method used were also examined. The learners were instructed to conduct KWIC searches using a specified corpus database called CBONA. This allowed them to identify patterns and enhance their future language performances. The study's qualitative results indicate that the adoption of DDL-based instruction by students may have facilitated the learning of SVA rules. Patterns from DDL may offer students opportunities to learn the rules of SVA, in comparison to traditional teaching methods. The DDL approach has positively impacted students' overall writing performance. The students' perception of the DDL-based activities also reflects a positive impact. The study has significant pedagogical implications for academicians and researchers. It emphasises the importance of DDL in the course design of ELT methods and approaches Future studies should closely analyse the linguistic, rhetorical, and discourse features of students' joint texts and individual texts produced after DDL activities. This study affirms the significance of discovery learning theory and the noticing hypothesis in small group learning within the ESL/EFL context

.

REFERENCES

Alahmadi, N. (2019). A study of grammatical errors of subject-verb agreement in writing made by Saudi learners. *International Journal of English Language and Linguistics Research, 7*(6), 48–59.

Ali, M. M., & Hamid, M. O. (2020). Teaching English to the test: Why does negative washback exist within secondary education in Bangladesh? *Language Assessment Quarterly, 17*(2), 129–146. https://doi.org/10.1080/15434303.2020.1717495

Alsehibany, R. A., & Abdelhalim, S. M. (2023). Enhancing academic writing vocabulary use through direct corpus consultation: Saudi English majors' perceptions and experiences. *Arab World English Journal, 14*(1), 166–182. https://dx.doi.org/10.24093/awej/vol14no1.11

Atashian, S., & Al-Bahri, K. (2018). Towards Arab students' grammatical errors in academic writing & their perceptions. *Arab World English Journal, 1*, 140–145. https://doi.org/10.24093/awej/mec1.10

Boontam, P., & Phoocharoensil, S. (2018). Effectiveness of English preposition learning through data-driven learning (DDL). *3L: Language, Linguistics, Literature. The Southeast Asian Journal of English Language Studies, 24*(3), 125–141. https://doi.org/10.17576/3L-2018-2403-10

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL, 21*(1), 37–54. https://doi.org/10.1017/S0958344009000068

Boulton, A. (2017). Data-driven learning and language pedagogy. In S. Thorne & S. May (Eds.), *Language, education and technology: Encyclopedia of language and education* (3rd ed., pp. 181–192). Cham: Springer. https://doi.org/10.1007/978-3-319-02237-6_15

Boulton, A., & Vyatkina, N. (2021). Thirty years of data-driven learning: Taking stock and charting new directions over time. *Language Learning & Technology, 25*(3), 66–89.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford University Press.

Bryman, A., & Bell, E. (2015). *Business research methods* (4th ed.). Oxford University Press.

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

Çekiç, A. (2022). Data-driven vs. deductive learning of English L2 collocations using online resources: A convergent mixed-methods study. *Bartın University Journal of Faculty of Education, 11*(1), 101–117. https://doi.org/10.14686/buefad.1045468

Chen, Y. H. (2004). The use of corpora in the vocabulary classroom. *The Internet TESL Journal, 10*(9). http://iteslj.org/Techniques/Chen-Corpora.html

Chawdhury, R., & Kabir, A. H. (2014). Language wars: English education policy and practice in Bangladesh. *Multilingual Education, 4*(1), 1–16. https://doi.org/10.1186/s13616-014-0021-2

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed approaches* (5th ed.). SAGE Publications.

Crosthwaite, P. (2020). Taking DDL online: Designing, implementing, and evaluating a SPOC on data-driven learning for tertiary L2 writing. *Australian Review of Applied Linguistics, 43*(2), 169–195. https://doi.org/10.1075/aral.00031.cro

Emir, G., & Yangın Ekşi, G. (2023). Corpus used as a data-driven learning tool in L2 academic writing: Evidence from Turkish contexts. *TEFLIN Journal, 34*(2), 209–225. https://doi.org/10.15639/teflinjournal.v34i2/209-225

Fang, L., Ma, Q., & Yan, J. (2021). The effectiveness of corpus-based training on collocation uses in L2 writing for Chinese senior secondary school students. *Journal of China Computer-Assisted Language Learning, 1*(1), 80–109. https://doi.org/10.1515/jccall-2021-2004

Farooq, M. S., Uzair-Ul-Hassan, M., & Wahid, S. (2020). Opinion of second language learners about writing difficulties in English language. *South Asian Studies, 27*(1).

Febriyanti, R. H. (2019). Students' ability and factors in using subject-verb agreement: A case study of Indonesian EFL learners in university level. *Journal of Linguistics and Language Teaching, 5*(2), 78–85. https://doi.org/10.30998/scope.v3i2.3334

Flowerdew, L. (2015). Data-driven learning and language learning theories. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15–36). John Benjamins. https://doi.org/10.1075/scl.69.02flo

Geluso, J. (2022). Grammatical and functional characteristics of preposition-based phrase frames in English argumentative essays by L1 English and Spanish speakers. *Journal of English for Academic Purposes, 55*, 101072. https://doi.org/10.1016/j.jeap.2021.101072

Goldin, M., Otero, L., César, J., & Hur, E. (2023). How frequent are these verbs? An exploration of lexical frequency in bilingual children's acquisition of subject-verb agreement morphology. In A. Corr & N. Schifano (Eds.), *Romance grammars: Context and contact*. *Isogloss. Open Journal of Romance Linguistics, 9*(2)/7, 1–25. https://doi.org/10.5565/rev/isogloss.194

Hajimia, H., Singh, M. K., & Rathakrishnan, M. (2019). Analysing frequency and distributional pattern of discourse markers in the Malaysian newspaper reports: A corpus-based study. *Opción, 34*(19), 1692–1714.

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, Y. (2022). Construction of "interactive" English translation teaching model based on data-driven learning. *Wireless Communications and Mobile Computing, 2022*, 5315110. https://doi.org/10.1155/2022/5315110

Huang, Z. (2014). The effects of paper-based DDL on the acquisition of lexicogrammatical patterns.

Ibrahim, M. M. M., & Ibrahim, A. I. (2020). An analysis of grammatical errors in academic writing of EFL students. *Asian Journal of Social Science and Management Technology, 2*(6), 81–90.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods, 18*(1), 3–20. https://doi.org/10.1177/1525822X05282260

Johns, T. (1988). Whence and whither classroom concordancing? In P. Bongaerts, P. de Haan, S. Lobbe, & H. Wekker (Eds.), *Computer applications in language learning* (pp. 9–27). Foris. https://doi.org/10.1515/9783110884876-003

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

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

Karim, S. M., Fathema, F., & Hakim, A. (2015). Analysis of errors in subject-verb agreement among Bangladeshi tertiary level EFL learners. *The International Journal of Social Sciences, 31*(1), 2305–4557.

Khan, M. A., Nabi, M. K., Khojah, M., & Tahir, M. (2021). Students' perception towards e-learning during COVID-19 pandemic in India: An empirical study. *Sustainability, 13*(1), 57. https://doi.org/10.3390/su13010057

Koosha, M., & Jafarpour, A. A. (2006). Data-driven learning and teaching collocation of

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal, 8*(4), 192–209.

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/applin/amy012

Li, H. H., & Zhang, L. J. (2022). Investigating effects of small-group student talk on the quality of argument in Chinese tertiary English as a foreign language learners' argumentative writing. *Frontiers in Psychology, 13*, 868045. https://doi.org/10.3389/fpsyg.2022.868045

Lin, H. C., & Lee, Y. D. (2017). A study of the influence of organizational learning on employees' innovative behavior and work engagement by a cross-level examination. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(7), 3463–3478. https://doi.org/10.12973/eurasia.2017.00738a

Lin, H. M., & Lee, J. Y. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal, 69*(3), 264–274. https://doi.org/10.1093/elt/ccv010

Lin, M. H. (2021). Effects of data-driven learning on college students of different grammar proficiencies: A preliminary empirical assessment in EFL classes. *SAGE Open, 11*(3), 1–15. https://doi.org/10.1177/21582440211029936

Ma, Q., Chiu, M. M., Lin, S., & Mendoza, N. B. (2023). Teachers' perceived corpus literacy and their intention to integrate corpora into classroom teaching: A survey study. *ReCALL, 35*(1), 19–39. https://doi.org/10.1017/S0958344022000180

Middleton, F. (2022). Subject-verb agreement: Examples, rules & use. *Scribbr*. https://www.scribbr.com/language-rules/subject-verb-agreement/

Nguyen, H. M. A., Nguyen, H. Y., Nguyen, T. Y. T., & Luong, M. N. (2022). Grammatical errors in academic writing of English second-year students. *European Journal of English Language Teaching, 7*(6), 70–83. https://doi.org/10.46827/ejel.v7i6.4547

Noguera-Díaz, Y., & Pérez-Paredes, P. (2020). Teaching acronyms to the military: A paper-based DDL approach. *Research in Corpus Linguistics, 8*(2), 1–27. https://doi.org/10.32714/ricl.08.02.01

Nugraha, S., Miftakh, F., & Wachyudi, K. (2017). Teaching grammar through data-driven learning (DDL) approach. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, 9th International Conference on Applied Linguistics (CONAPLIN 9), 300–303. Atlantis Press. https://doi.org/10.2991/conaplin-16.2017.68

Nurjanah, S. (2017). An analysis of subject-verb agreement errors on students' writing. *English Language Teaching-Echo, 2*(1), 13–25. https://doi.org/10.24235/eltecho.v2i1.1590

O'Keeffe, A. (2021). Data-driven learning: A call for a broader research gaze. *Language Teaching, 54*, 259–272. https://doi.org/10.1017/S0261444820000245

O'Reilly, J., & García-Castro, V. (2022). Exploring the relationship between foreign language anxiety and students' online engagement at UK universities during the COVID-19 pandemic. *Language Learning in Higher Education, 12*(2), 409–427. https://doi.org/10.1515/cercles-2022-2054

Oktavianti, I. N., Eriani, E., Rolyna, I., & Prayogi, I. (2023). Investigating the use of corpus-informed grammar materials in Indonesian EFL classrooms. *Indonesian Journal of English Language Teaching and Applied Linguistics, 7*(2), 417–438. https://doi.org/10.21093/ijeltal.v7i2.1415

Özer, M., & Özbay, A. S. (2022). Exploring the data-driven approach to grammar instruction in the ELT context of Turkey. *Australian Journal of Applied Linguistics, 5*(2), 35–63. https://doi.org/10.29140/ajal.v5n2.713

Pérez-Paredes, P. (2022). A systematic review of the uses and spread of corpora and data-driven learning in CALL research during 2011–2015. *Computer Assisted Language Learning, 35*(1–2), 36–61. https://doi.org/10.1080/09588221.2019.1667832

Plano Clark, V. L., & Ivankova, N. V. (2016). *Mixed methods research: A guide to the field*. Sage. https://doi.org/10.4135/9781483398341

Promsupa, P., Varasarin, P., & Brudhiprabha, P. (2017). An analysis of grammatical errors in English writing of Thai university students. *HRD Journal, 8*(1), 93–104.

Rahman, M., & Hasan, K. (2019). Academic writing difficulties of Bangladeshi students at a higher institution in Malaysia. *Journal of Research and Multidisciplinary, 2*(2), 145–171.

Rezwana, M., Hossain, M. M., & Ahmad, S. (2022). The practices and problems of teaching English grammar to young EFL learners. *International Journal of Language and Literary Studies, 4*(3), 139–163. http://doi.org/10.36892/ijlls.v4i2.937

Riadil, I. G., Rahman, A. N. I., & Chonpracha, P. (2023). Investigating EFL learners' difficulties and strategies in academic writing skill: A qualitative study. *ELP (Journal of English Language Pedagogy), 8*(1), 66–76. https://doi.org/10.36665/elp.v8i1.691

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

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

Römer, U., Skalicky, S. C., & Ellis, N. C. (2020). Verb-argument constructions in advanced L2 English learner production: Insights from corpora and verbal fluency tasks. *Corpus Linguistics and Linguistic Theory, 16*(2), 303–331.

Sah, P. K. (2015). An investigation into the relative effectiveness of data-driven learning (DDL) with integration into PPP and III. *The Journal of Teaching English for Specific and Academic Purposes, 3*(2), 347–366.

Salama, A. H. Y. (2023). Unpacking the semantics of coronaviruses' "referent things": A corpus-driven systemic functional linguistic analysis. *Taylor & Francis, 69*(1), 49–90. https://doi.org/10.1080/00437956.2022.2160147

Sari, R. M. (2019). The causes of grammatical errors made by EFL students in writing essays. *Linguists: Journal of Linguistics and Language Teaching, 5*(2), 86–93.

Sayma, R. A. (2020). Fostering non-English major undergraduate students' writing skill: A study in Cumilla region of Bangladesh. *Journal of Literature, Languages and Linguistics, 69*, 45–57. https://doi.org/10.7176/JLLL/69-06

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129

Seraj, P. M. I., Klimova, B., & Habil, H. (2021). Use of mobile phones in teaching English in Bangladesh: A systematic review (2010–2020). *Sustainability, 13*, 5674. https://doi.org/10.3390/su13105674

Shubha, N. S. (2021). Exploration of phrasal verbs in ELT textbooks: A corpus-based analysis of lower secondary level Bangladeshi books. *FOSTER: Journal of English Language Teaching, 2*(4), 456–472. https://doi.org/10.24256/foster-jelt.v2i4.55

Singh, M. K. (2014). *A corpus-based genre analysis of quality, health, safety, and environment work procedures in the Malaysian petroleum industry* (Doctoral dissertation, Universiti Teknologi Malaysia). Universiti Teknologi Malaysia Institutional Repository. https://eprints.utm.my/78163/

Stapa, S., & Izahar, M. (2010). Analysis of errors in subject-verb agreement among Malaysian ESL learners. *Journal of Language Teaching, Linguistics, and Literature, 16*, 56–73.

Sultana, S., Singh, M. S. K., & Islam, R. (2023). The usefulness of digital media implementation in ESL learning among Bangladeshi students. *World Journal of English Language, 13*(6), 471–479. https://doi.org/10.5430/wjel.v13n6p471

Sun, X., & Hu, G. (2020). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research, 1–29*. https://doi.org/10.1177/1362168820954459

Sun, X., & Hu, G. (2023). Direct and indirect data-driven learning: An experimental study of hedging in an EFL writing class. *Language Teaching Research, 27*(3), 660–688. https://doi.org/10.1177/1362168820954459

Suraprajit, P. (2021). An analysis of errors in English essays written by Thai non-English major students. *Theory and Practice in Language Studies, 11*(1), 12–19. https://doi.org/10.17507/tpls.1101.02

Tafida, A. G., & Okunade, S. K. (2016). Subject-verb agreement problems among English as second language learners: A case study of first-year undergraduates of Federal University of Technology, Minna. *International Invention Journal of Education and General Studies, 2*(2), 20–27.

Terrell, S. R. (2012). Mixed-methods research methodologies. *Qualitative Report, 17*, 254–280.

Thiagarajah, A., & Razali, A. B. (2021). Going back to basics: Effectiveness of reading English newspapers on vocabulary knowledge of Malaysian secondary-level ESL/EFL students. *International Journal of Academic Research in Progressive Education and Development, 10*(2), 766–781. https://doi.org/10.6007/IJARPED/v10-i2/10065

Therova, D., & McKay, A. (2024). Addressing discipline specificity in a multidisciplinary EAP classroom through data-driven learning. *Enhancing Teaching and Learning in Higher Education, 1*, 21–40. https://doi.org/10.62512/etlhe.9

Tomas, R. R. S., & Dulin, A. T. (2021). Social media exposure and paragraph writing of pre-service students. *Linguistics and Culture Review, 5*(S3), 144–153. https://doi.org/10.37028/lingcure.v5nS3.1379

Tosun, S., & Sofu, H. (2023). The effectiveness of data-driven vocabulary learning: Hands-on concordancing through a pedagogical corpus. *Journal of Language and Education, 9*(3), 177–191. https://doi.org/10.17323/jle.2023.12426

Troy, H. R., & Millar, N. (2019). Evaluating the use of data-driven learning to teach phrasal verbs in a private language school in Japan. *Bulletin of Nagoya University of Foreign Studies, 5*, 321–347.

Tsulaia, N. (2021). The effect of data-driven learning of grammar on Georgian EFL university students' grammar achievement. *European Scientific Journal, 17*(33), 168. https://doi.org/10.19044/esj.2021.v17n33p168

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468

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

Wu, Y.-J. A. (2021). Discovering collocations via data-driven learning in L2 writing. *Language Learning & Technology, 25*(2), 192–214. http://hdl.handle.net/10125/73440

Xu, J. (2022). Incremental intelligence matters: How L2 writing mindsets impact feedback orientation and self-regulated learning writing strategies. *Assessing Writing, 51*, 100593. https://doi.org/10.1016/j.asw.2021.100593

Yepes, G. R., & Krishnamurthy, R. (2010). Corpus linguistics and second language acquisition – The use of ACORN in the teaching of Spanish grammar. *Lebende Sprachen, 55*(1), 22–32. https://doi.org/10.1515/les.2010.008

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*(4), 257–283. https://doi.org/10.1016/j.jslw.2004.06.002

Yu, L., & Shen, J. (2022). Analysis of the correlation between academic performance and learning motivation in English courses under a corpus-data-driven blended teaching model. *Hindawi Scientific Programming, 2022*, 3407270. https://doi.org/10.1155/2022/3407270

Zare, J., & Aqajani Delavar, K. (2023). A data-driven learning focus-on-form approach to academic English lecture comprehension. *Applied Linguistics, 44*(3), 485–504. https://doi.org/10.1093/applin/amac021

Zare, J., Karimpour, S., & Aqajani Delavar, K. (2022). The impact of concordancing on English learners' foreign language anxiety and enjoyment: An application of data-driven learning. *System, 109,* Article 102891. https://doi.org/10.1016/j.system.2022.102891

Zare, J., Noughabi, M. A., & Al-Issa, A. (2024). The impact of data-driven learning form-focused tasks on learners' task engagement: An intervention study. *ReCALL, 36*, 1–18. https://doi.org/10.1017/S0958344024000120

.

Copyright © 2025, *authors*, e-ISSN: 2502-6747, p-ISSN: 2301-9468