

Research on English Writing Instruction Based on Data-Driven Learning (DDL)

Siying Huang

Guangdong University of Business and Technology, Zhaoqing Guangdong, 526060

ABSTRACT

This study focuses on the application of Data-Driven Learning (DDL) in English writing instruction, exploring its educational effectiveness through theoretical construction and empirical analysis. Based on cognitive learning theory, corpus linguistics, and constructivist learning theory, the study constructs a theoretical framework for DDL, with core elements including corpus utilisation, data-driven autonomous learning, and teacher-guided support. Through a 16-week teaching practice involving 30 non-English major second-year students at a certain university, combined with DDL activities such as vocabulary learning, writing strategy analysis, error correction, and personalised feedback, it was found that students showed significant improvements in vocabulary richness, grammar error rates, frequency of logical transition words, and the depth of argumentation in argumentative essays. Survey results indicated that 85% of students acknowledged the positive impact of DDL on their writing abilities, and 75% reported enhanced self-directed learning skills. The study indicates that DDL effectively enhances students' language accuracy, writing structure rationality, and critical thinking skills through authentic corpus input and self-directed exploration. However, implementation should focus on improving students' proficiency in corpus operations and teachers' instructional design capabilities. This research provides empirical evidence and practical guidance for reforms in English writing instruction.

KEYWORDS

Data-Driven Learning (DDL); English Writing Instruction; Corpus Linguistics; Self-Directed Learning; Instructional Outcomes

1 Introduction

With the acceleration of globalisation, the importance of English as the primary language of international communication has become increasingly evident. English writing skills are not only a key competence for students' academic and professional development but also an essential tool for cross-cultural communication. However, traditional English writing instruction often focuses on grammar and vocabulary instruction while neglecting the cultivation of students' independent learning abilities and writing strategies. This results in students frequently encountering issues such as limited vocabulary, inaccurate expression, and a lack of writing structure during the writing process. Therefore, exploring new teaching methods to enhance students' English writing skills has become a research hotspot in the field of education. Data-Driven Learning (DDL), as an emerging teaching method, has gradually attracted the attention of educational researchers. DDL guides students to analyse and utilise authentic language data from corpora, helping them independently discover language patterns and thereby improve their writing skills. This method not only enhances students' language awareness but also cultivates their critical thinking and independent learning abilities. Corpus linguistics provides DDL with a rich source of language data. Students can gain insights into the actual use of language by analysing large-scale authentic texts, thereby improving their language application skills. This study aims to explore the application effects of DDL in English writing instruction and its implications for writing instruction through theoretical analysis and case studies. Specifically, this study will construct a theoretical framework for DDL in English writing instruction and analyse the impact of DDL on students' English writing abilities through a semester of teaching practice. The study will focus on the effectiveness of DDL in improving students' vocabulary usage, grammatical accuracy, and writing structure, while also exploring the role transformation and teaching reflections of teachers during the implementation of DDL instruction. Through this study, it is hoped to provide valuable references for reforms in English writing instruction and offer specific guidance and suggestions for teachers implementing DDL instruction.

2 Data-Driven Learning (DDL) Theoretical Framework

2.1 Theoretical Foundations

The theoretical foundation of Data-Driven Learning (DDL) primarily draws from cognitive learning theory, corpus linguistics, and constructivist learning theory. These theories provide a solid theoretical basis for the application of DDL in English writing instruction. Cognitive learning theory emphasises that learning is an active cognitive process, where students construct knowledge systems through interaction with their environment. In DDL, students analyse authentic language data in corpora to independently explore language patterns, thereby deepening their understanding and mastery of language. Corpus linguistics provides DDL with abundant language data resources, enabling students to

analyse large-scale authentic texts to understand the actual use of language and enhance their language proficiency. Constructivist learning theory emphasizes that the process of knowledge construction is actively completed by students through interaction with the environment. DDL guides students to discover language patterns in corpora, fostering their autonomous learning abilities and critical thinking skills. These theoretical foundations collectively form the core philosophy of DDL: promoting students' autonomous learning and knowledge construction through authentic data in corpora to enhance language proficiency.

2.2 Core Elements

The core elements of DDL include corpus utilisation, data-driven autonomous learning, and teacher guidance and support. The corpus serves as the foundational resource for DDL, offering powerful and diverse functionalities. Students can observe and analyse the actual usage of language in various contexts through real-world language data within the corpus, thereby gaining a deeper understanding and mastery of language rules. The data within the corpus not only helps students identify commonalities and variations in language but also provides abundant language input, enhancing their language awareness. Data-driven autonomous learning is the core process of DDL. Under the guidance of teachers, students autonomously discover language patterns through the data in the corpus, thereby cultivating their autonomous learning abilities. This process of autonomous learning not only helps students memorise and apply language knowledge but also cultivates their critical thinking abilities. Through the analysis of corpus data, students learn how to pose questions, analyse problems, and solve them, thereby improving their language analysis and problem-solving abilities. Teacher guidance and support are the key to the success of DDL. Teachers need to design effective teaching activities to help students master the use of corpora and provide necessary guidance and feedback during the learning process. The role of teachers shifts from traditional knowledge transmitters to facilitators and promoters, helping students discover language patterns in corpora and cultivate autonomous learning abilities. Teachers also need to adjust teaching strategies in a timely manner based on student feedback and learning progress to ensure the effectiveness and relevance of teaching activities.

2.3 Advantages in English Writing Instruction

Firstly, DDL significantly improves students' language accuracy and richness. Through authentic writing samples in the corpus, students are exposed to a wealth of authentic language expressions and diverse writing structures. This authentic language input helps students observe the diversity and complexity of language in actual use, enabling them to apply language more naturally in writing. For example, by analysing the usage contexts of high-frequency vocabulary and phrases, students can learn richer vocabulary and more accurate grammatical structures, thereby improving the accuracy and richness of their writing. Secondly, DDL emphasises students' independent learning, effectively cultivating their independent learning abilities and critical thinking skills. During the process of analysing corpus data, students need to actively explore language patterns, formulate hypotheses, and verify them. This process of independent learning not only enhances students' motivation but also improves their learning efficiency. Students can become more confident and independent in their writing, no longer relying solely on teacher guidance. DDL fosters students' independent learning abilities and critical thinking by guiding them to discover language patterns on their own, enabling them to apply their knowledge more flexibly in writing. Additionally, DDL provides students with a personalised learning experience. The diverse data in the corpus allows students to select appropriate learning materials based on their own progress and interests. This personalised learning approach meets the needs of different students, helping them improve their writing skills in the most suitable way for them. Teachers can also provide targeted guidance and suggestions based on students' feedback and learning outcomes, helping them overcome difficulties in writing and further improve their writing skills.

3 Case Study of English Writing Instruction Based on DDL

3.1 Case Design and Implementation Process

To validate the effectiveness of data-driven learning (DDL) in English writing instruction, this study conducted a semester-long teaching practice (2 lessons per week, totaling 16 weeks) targeting 30 non-English major second-year undergraduate students at a certain university. The teaching objectives focused on enhancing students' English writing skills through DDL, including language accuracy, vocabulary richness, enhance their ability for independent learning and critical thinking, and master corpus-based writing strategies and techniques. The course revolved around common themes such as descriptive, argumentative, and report writing, incorporating diverse DDL activities: in vocabulary learning, teachers extracted high-frequency vocabulary phrases related to the writing theme from the corpus, such as 'sustainable development' and 'ecological balance' in conjunction with the environmental protection theme, and

designed vocabulary matching, fill-in-the-blank exercises, and introduce corpus keyword search and word frequency statistics functions. Students are organised into groups to analyse example sentence usage, discuss and summarise vocabulary usage, and finally receive teacher feedback. When learning writing strategies and techniques, teachers select authentic excellent writing samples from the corpus, such as academic papers and news reports, to guide students in analysing the logical development of argumentative essay theses and the problem-analysis-suggestion framework of report writing. By presenting samples, students analyse their structure, logic, and linguistic features, then imitate and rewrite them. After group discussions to summarise writing strategies, students engage in writing exercises. Teachers patrol the classroom to answer questions and provide feedback after class. In the error analysis and correction phase, teachers extract common writing errors from the corpus and categorise them, such as grammatical tense errors or inappropriate vocabulary combinations. They guide students to analyse error types and causes, and students discuss in groups and attempt to correct errors. Teachers provide comments and summaries on the discussion results. In the writing task and feedback phase, writing tasks are assigned every two weeks, such as an argumentative essay on 'The Impact of Technology on Life.' Students complete the writing using the vocabulary and writing strategies they have learned and submit their work. Teachers use corpus analysis tools to generate quantitative reports such as vocabulary complexity and grammar error heatmaps, combined with qualitative comments to provide personalised feedback. In class, teachers provide collective explanations and individual guidance, and students revise their work based on the feedback and submit the final draft.

3.2 Student Writing Outcomes and Feedback

To assess the impact of the DDL method on students' English writing abilities, the study collected data through methods such as questionnaires, interviews, and analysis of writing outcomes. The questionnaire survey showed that 85% of students believed the DDL method helped improve their writing abilities, 75% reported enhanced self-directed learning abilities, and 68% stated they were more willing to actively explore language patterns. In interviews, students generally reported that the DDL method enabled them to learn vocabulary and writing techniques more intuitively, improving writing accuracy and richness. For example, learning vocabulary collocations through corpora reduced rote memorisation, and group discussions enhanced critical thinking skills. Teachers believed the DDL method helped students better understand real-world language usage, enhancing the targeting and effectiveness of instruction. Analysis of students' writing outcomes revealed significant improvements in vocabulary usage, grammatical accuracy, and writing structure. Vocabulary richness increased by 22%, grammatical error rates decreased by 35%, the frequency of logical connectives increased by 40%, and the average length of argument development in argumentative essays increased. For example, some students improved vocabulary diversity by searching for synonyms in corpora, while others learned to properly cite data and conduct trend analysis.

3.3 Teacher Roles and Teaching Reflection

In the DDL teaching process, the role of teachers has undergone significant changes. Teachers are no longer merely knowledge transmitters but rather facilitators and promoters. Teachers must possess strong corpus operation skills and instructional design capabilities to effectively guide students in utilizing corpora for self-directed learning. Their primary tasks include helping students understand the functions and usage of corpora, designing DDL activities related to writing themes, providing personalized feedback and guidance, and fostering students' self-directed and collaborative learning. Through this semester's teaching practice, teachers have discovered that the DDL method offers significant advantages in enhancing students' writing abilities but also presents certain challenges. For example, some students are not yet proficient in using the corpus and require more time and guidance to master its functions. Additionally, teachers need to continuously improve their instructional design and classroom management skills to better adapt to the DDL teaching model. In future teaching, teachers should further optimise instructional activity design, provide more diverse corpus resources, and strengthen individualised guidance for students to fully leverage the potential of DDL in English writing instruction.

4 Conclusions

This study confirms through theoretical analysis and teaching practice that data-driven learning (DDL) can significantly improve students' English writing skills. Students show significant improvements in vocabulary richness, grammatical accuracy, and the logical structure of their writing, while their self-directed learning and critical thinking abilities are effectively cultivated. The theoretical framework of DDL centres on corpora as core resources, promoting students' active construction of language knowledge through a model of student-driven analysis of authentic language data and teacher-

guided support. The study found that DDL's personalised learning pathways and quantitative feedback mechanisms have positive implications for writing instruction, but implementation challenges include students' insufficient proficiency in corpus operations and the need for teachers to optimise instructional design. Future research could expand the sample size, explore the application of DDL in different groups, and investigate its integration with new technologies to comprehensively enhance students' overall language proficiency. DDL provides empirical evidence for reforms in English writing instruction, and its 'student-centred' model is worthy of promotion, requiring continuous optimisation in practice to fully realise its potential.

About the Author

Siying Huang, Female, Associate Professor, Doctor, English Education and Teaching.

References

- [1] Geng Yun. An Exploration of University English Vocabulary Teaching Under a Corpus Data-Driven Learning Model [J]. Journal of Jilin Radio and Television University, 2025, (01): 43-45.
- [2] Pan Gaoran. Research on the Optimisation of English Writing Teaching Under Big Data [J]. Modern English, 2024, (22): 68-70.
- [3] Qi Hui. Design and Application Evaluation of Medical English Academic Writing Instruction Based on Data-Driven Learning [J]. Research and Practice in Medical Education, 2024, 32(01): 89-95.
- [4] Qiu Qingwei, Nie Jingjing, Yang Tingjun. An Exploration of University English Vocabulary Teaching Under the Corpus Data-Driven Learning Model [J]. Huazhang, 2023, (08): 55-57.
- [5] Li Wenting. A Discussion on a New Model of Professional English Writing Teaching Based on Data-Driven Learning [J]. Modern English, 2023, (14): 13-16.
- [6] Liang Jinfang, Tan Ying. A Meta-Analysis of Empirical Studies on the Effectiveness of Data-Driven English Teaching [J]. Education Theory and Practice, 2023, 43(21): 48-52.
- [7] Chen Lidan. A Study on Data-Driven Flipped Classroom Teaching in Business English under the New Liberal Arts Paradigm [J]. Journal of Hubei Open Vocational College, 2023, 36(04): 166-168.
- [8] Zhang Qi, Zou Yanli. Application of Data-Driven Learning Concepts Based on Corpora in Foreign Language Teaching [J]. Journal of Changchun University of Science and Technology (Social Sciences Edition), 2022, 35(06): 181-185.