

A Study on the Construction of College English Context Vocabulary Teaching Based on Hands-Off Data-Driven Learning in China

Haojie Li¹ & Tongde Zhang²

¹ School of Foreign Languages, Southwest University of Political Science and Law, Chongqing, China

² School of Foreign Languages, East China University of Technology, Jiangxi, China

Correspondence: Tongde Zhang, School of Foreign Languages, East China University of Technology, Jiangxi, China.

Received: June 7, 2023

Accepted: October 6, 2023

Online Published: December 19, 2023

doi:10.5539/ies.v17n1p1

URL: <https://doi.org/10.5539/ies.v17n1p1>

Abstract

Hands-off data-driven learning is a data-based, student-oriented learning model characterized by inquiry and discovery. English context vocabulary teaching is the key to English teaching in colleges and an important indicator to evaluate the quality and level of college English teaching, which is a language teaching paradigm focusing on the language environment. Combining the two approaches can give students a more realistic, practical, and meaningful language learning experience. This paper analyzes the vocabulary learning level of two non-English major undergraduate classes at Southwest University of Political Science and Law before and after the application of the context experiment. The positive effect of context vocabulary teaching in the control groups is verified by comparing and analyzing the influence of context teaching based on hands-off data-driven learning on their scores and learning results between the experimental and control groups. It shows that the combination of context in English vocabulary teaching with hands-off data-driven learning can help to improve students' ability to understand, absorb, and apply English vocabulary.

Keywords: hands-off data-driven learning, college English, context vocabulary teaching

1. Introduction

The achievements of corpus linguistics development have been applied to many fields of language teaching, and a lot of progress has been made. Data-driven learning (DDL for short) is corpus-based English teaching method that requires learners to explore and summarize language rules, such as collocations of words, semantic characteristics, and so on. It is described as a revolution in foreign language teaching (Hunston, 2002, p. 1) in that it is not only a basis for teachers to make teaching decisions but also an important resource for students to learn.

In light of the use of corpus in teaching, DDL can be classified into hands-on DDL and hands-off DDL. It has increasingly become a hot topic in college English teaching and research. In this paper, it is maintained that a college English context vocabulary teaching model based on hands-off DDL is a more effective teaching method to construct a student-oriented learning model and college English context teaching than traditional vocabulary teaching. This method is to find the problems and needs of students' learning through analysis and excavation of classroom data, and then design and optimize teaching solutions based on data results to improve the effects and quality of English context teaching. The data-driven teaching of college English vocabulary has promoted the renewal of the teaching paradigm structure, enriching the teaching content, diversifying the teaching media, expanding the teaching resources, constructing the classroom environment and the network environment mutually, and playing a crucial role in creating the context of college English teaching.

This paper is organized as follows. Section 2 reviews context teaching. Section 3 examines the study of DDL. Section 4 discusses context teaching based on hands-off DDL. Section 5 presents the research design. Section 6 presents the results and discussion of the experiment. Section 7 concludes the paper.

2. Context Teaching

The concept of context was first put forward by Malinowski, which is the basis of the context theory proposed by Firth. Context teaching is a teaching method based on a language use environment, which aims to help students master the application skills of language. It emphasizes the environment and actual situation of the

language use, paying attention to students' active participation and practical operations, and promoting the development of students' language ability and communication abilities. As a result, context teaching has been widely used and has become one of the important means of English teaching in college English teaching. In recent years, the concept of context teaching has attracted extensive attention from scholars both at home and abroad. Judith and Glisan (2000) argues that "contextualization" includes not only learning and understanding the meaning of language according to the context of the text but also creating a meaningful language teaching environment and placing foreign language teaching in the context of the overall social development needs.

In China, although the research on English context teaching has been paid more and more attention by scholars, there are serious deficiencies. Through a literature survey, Li (2012) points out that the research attention on context teaching has increased from the perspective of research attention, but the understanding of the role of context in the process of English teaching and learning is insufficient, especially in recent years. From the perspective of research methods, Non-empirical research occupies a dominant position, and empirical research is extremely scarce. Zhu and Yu (2010) also point out that although "theoretical researches on context have a long history, these studies mainly focus on the relationship between context and discourse cohesion and coherence, or discuss issues such as English teaching in the network environment. Few studies on comprehensive teaching combine systematic grammar knowledge, learners' foreign language acquisition process factors, and contextual factors to help learners enhance their language and cultural awareness and correctly use vocabulary and grammar."

To solve these problems, some Chinese scholars discuss the creation of a college English classroom context teaching model from the perspective of cultural contextualization (Ke, 2012), from the perspective of optimizing the network teaching environment (An, 2010), and from the perspective of combining multimedia means and multi-model methods (Zhu & Yu, 2010). However, most of these studies only stay at the level of theoretical discussion and lack specific empirical research. The fundamental crux of the problem lies in the lack of effective teaching methods for creating real context.

3. Data-Driven Learning (DDL)

Data-driven learning (DDL) is a corpus-based foreign language learning method proposed by Professor Tim Johns in the early 1990s. This method provides learners with a rich corpus of natural language as input to create a real language environment for them so that they can acquire "accurate and authentic" language (Granger, 1998, p. 199). Johns (1991) maintains that language learners are fundamentally researchers, and their learning needs to be driven by the acquisition of language data. As a result, the term "data-driven learning" is used to describe this method. This language teaching method allows students to search their spoken or written language corpus through language indexing software, providing them with real examples of language use, rather than examples designed in some grammar books. Its main idea is to guide language learners to make full use of network and corpus resources. The rules of language grammar, meaning expressions, and language characteristics will be discovered, and the problems will be actively solved by observing and analyzing a large number of objective and true linguistic data. This "discovery" learning method helps to improve learners' comprehensive grasp of language knowledge, and at the same time stimulates their interest in learning and cultivates their ability to learn independently (Tribble & Jones, 1990, p. 12).

DDL is expected to become the mainstream teaching paradigm gradually in the era of information. It can be classified into two categories: hands-on DDL and hands-off DDL.

3.1 Hands-On DDL

Hands-on DDL is a data-driven learning method in which learners directly use the computer and perform corpus retrieval by themselves. This new method, which relies on computer technology and the autonomy of learners, is in sharp contrast to the traditional foreign language teaching model. A large number of relevant empirical studies have been carried out one after another, but the teaching effect is not significant and has not achieved the expected advantages (Boulton, 2010b). There are three reasons for this phenomenon (Zhang & Zhou, 2012). First, the amount of data provided by the corpus is enormous. For ordinary learners who have no experience in processing corpus data, it is not easy to analyze so many examples (Cheng et al., 2003). Second, the corpus collects data from the language of native speakers, and its difficulty is usually beyond the range of comprehension for foreign language learners. The analysis and induction of these original corpora are also difficult for middle and low-level learners (Keck, 2004). Third, the high dependence on hands-on DDL on computers not only requires learners to have certain computer application skills (Yoon & Hirvela, 2004) but also puts higher requirements on the learning environment.

3.2 Hands-Off DDL

The hands-off DDL introduces corpus resources into conventional classrooms, allowing learners to conduct DDL in traditional classrooms without computers. The retrieval of the corpus is completed by the teacher, who selects and classifies the search results. The corresponding learning materials are compiled and output as paper texts in light of the needs of teaching and the characteristics of learners. For one thing, this hands-off DDL model retains the characteristics of DDL, that is, learners' discovery learning based on corpus resources, thus the requirements for learners are greatly reduced. For another, the corpus carefully selected by teachers is more targeted and accepted easily by students. About hardware, the hands-off DDL uses paper media to get rid of the direct dependence on computers, making the classroom application of corpus resources possible (Boulton, 2010a).

In recent years, with the further deepening of DDL research, hands-off DDL has once again received the attention of researchers (Allan, 2006; Boulton, 2009, 2010a, etc.). In China, researchers are also constantly exploring the best way to use corpus in classroom teaching (Liang, 2009; He, 2010, etc.). Liang (2009) proposes the concept of mini-text, which is composed of classic examples selected from a large number of corpora by sampling the index lines retrieved. It is a learning resource of "what you get is what you need" (Liang, 2009). The mini-file proposed by He (2010) is a different concept. In addition to a small number of examples extracted from the corpus, the mini-file is accompanied by related instructions and reference answers. Generally speaking, it is still in the exploratory stage in theory and practice to use the teaching aids of hands-off DDL to construct the teaching model of college English. Given this, this paper intends to introduce the hands-off DDL into the context teaching of junior college English learners, investigating and analyzing its actual teaching effect.

4. Context Teaching Based on Hands-Off DDL

Traditional context teaching methods are often difficult to meet students' personalized learning needs, and it is easy to cause students' decreased interest in learning and poor learning effects. The key to solving this problem is to create a real context in college English teaching, and DDL has opened up a new way to create it. The combination of hands-off DDL and context teaching can better promote students' learning and construct a student-oriented learning model, optimizing teaching solutions. It is gradually favored by educational institutions in some developed regions and achieved certain results and social influence in educational practice, although the current Chinese data-driven teaching paradigm has just begun to emerge. Teachers must face some practical problems, such as weak data literacy, insufficiently coordinated advancement, few useful big data technology products, and delayed basic theoretical research. Based on this situation, this paper combines the actual needs of China's college English education development to propose three suggestions for constructing a college English context vocabulary teaching model based on hands-off DDL.

First, practical training for data literacy to improve teacher data awareness and data processing capabilities should be carried out. Schools can invite experts, scholars, or industry leaders to carry out data literacy training. Teachers can improve their data awareness and data processing capabilities through curriculum learning and practical operations. To be more exact, schools can organize special training courses to allow experts and scholars or industry leaders to introduce the theoretical foundation, methods, and application cases of data-driven teaching, sharing data processing skills and tools. The training contents can involve data acquisition, processing, analysis, and visualization to help teachers understand the principles and applications of data-driven teaching in depth. For different needs and levels of teachers, training courses can be carried out at different levels, such as primary training, intermediate training, and senior training, so that teachers will gradually grasp the core concepts and skills of data-driven teaching. It is recommended that the education department incorporates teachers' data literacy into the evaluation system of teachers' professional literacy, setting up a teacher data literacy certification mechanism and regular evaluation system to continuously improve the professional processing ability of teachers of various disciplines (Li & Ren, 2016).

Second, big data smart learning platforms should be created. It is suggested that educational big data enterprises and educational informatization engineering technology research and development institutions develop a new generation of intelligent learning platform, focusing on the natural collection of teaching data, seamless integration of multi-source data, deep mining, and early warning analysis of learning behavior data, construction of knowledge map, intelligent diagnosis of subject ability and other aspects (Yang *et al.*, 2017). In terms of platform data acquisition, it is suggested to integrate log analysis, mobile APP, web crawler, dot matrix digital pen, wearable device, and other acquisition technologies to broaden the scope of learning data acquisition, enhancing the granularity of data acquisition to depict students' portraits more accurately (Xing *et al.*, 2016). The school should establish a learning data platform for collecting, storing, and processing students' learning data, including students' test results, classroom performance, homework, etc. to analyze students' learning data,

including visual analysis, trend analysis, association analysis of data, to extract valuable information and insights.

Schools can develop learning support systems, providing students with personalized learning suggestions and guidance through data analysis and excavation, including curriculum recommendation, learning path planning, intelligent counseling, etc. At the same time, schools need to take a series of measures to ensure the security of students' learning data, including encryption storage, access control, data backup, and so on. Schools also need to continue to improve and optimize big data smart learning platforms, including functional optimization, performance improvement, and improvement of user experience to better support the implementation of data-driven teaching.

Third, a data-driven teaching demonstration project should be implemented. Schools can carry out data-driven teaching demonstration projects, exploring new models of data-driven teaching, encouraging teachers to explore and innovate in practice, and forming replicable and promoted experiences to build a data-driven teaching practice community. This demonstration project should include multiple links, such as data acquisition, data analysis, instructional design, instructional implementation, effectiveness evaluation, etc. To be more specific, courses of different disciplines, grades, and difficulties can be selected, and students' learning data can be obtained by various data acquisition methods. Then, through data analysis and mining, students' learning status and problems can be understood, and teaching programs that meet students' needs can be designed and implemented in practice. In the teaching process, students' learning data should be monitored in time, and adjusted and optimized according to the data feedback. Finally, the effect evaluation should be carried out, and the experience should be summarized and extended to other courses. Such demonstration projects can provide a platform for teachers to practice and explore, providing experience and reference for promoting data-driven teaching.

5. Research Design

5.1 Subject

The experimental group and the control group are composed of 46 non-English major students in grade 2022 at Southwest University of Political Science and Law respectively. None of them have had access to corpus before, let alone experience in hands-off DDL.

5.2 Experimental Vocabulary

The experimental words are 20, which are randomly selected from the "College English Reference Glossary" attached to "College English Curriculum Teaching Requirements" (2007).

5.3 Learning Materials

The corpus used in this study is the corpus of New Horizons College English. Before the experiment, the teacher searches the corpus for the target words of the experimental group and extracts the most representative index lines (5 lines for each target word) from the search results for the experiment, making them into mini-texts. The context co-occurrences of the target words in mini-texts are presented in the form of the key word in context (KWIC). The sentence length of the target word, which is bolded in the context, is about 10 to 15 words, and the rest in the context remain unchanged. Finally, the prepared material is output into paper text.

5.4 Experimental Procedure

Before the experiment, a test is conducted to determine the English vocabulary level of the students in the two classes. The difficulty of the test questions is based on the 5000 vocabularies of CET-4.

The experiment is carried out by comparing hands-off DDL with traditional vocabulary learning methods in a regular classroom, which lasts for 40 minutes. The traditional method is adopted to learn the words of the control group. The steps are as follows: the teacher first explains the rules of the use of the target words accurately and clearly and then gives examples. Finally, the subjects are asked to learn the words through practice to strengthen their understanding and consolidation. After learning, an immediate test of the control group of words is performed.

The materials prepared in advance are distributed to the subjects in the experimental group, which uses the hands-off DDL method to learn vocabulary. The subjects are asked to use the materials to find out the parts of speech, meaning, main collocations, and other information of the target words, making sentences with the target words. The whole learning process is completed independently by the subjects without any explanation from the teacher. At the same time, the experiment time is strictly limited to 40 minutes (the same time as the control group). When the time is up, the learning materials are taken back immediately, and an instant test of the words

is conducted in the experimental group.

All the students in the experimental group and the control group participated in two tests before and after the experiment to better understand the students' English learning state before the experiment and better compare the learning level after the experiment.

6. Results and Discussion

The software SPSS20.0 is used to conduct an independent sample T-test on students' scores in two vocabulary tests before and after the experiment. The test results and statistics of the experimental and control groups are shown in Table 1 and Table 2.

Table 1. Test result before experiment

Groups	N	mean	Std. deviation	t	Sig.
Control group	46	63.37	13.10	1.009	0.318
Experimental group	46	65.70	7.94		

Table 2. Test result after experiment

Groups	N	mean	Std. deviation	t	Sig.
Control group	46	66.64	14.46	4.427	P<0.001
Experimental group	46	76.98	9.21		

It can be seen from Table 1, that the significance level of the two groups in the test before the experiment is 0.318, which is above the critical value of 0.05. Thus, there is no significant difference, which indicates that the English vocabulary level of the students in the two groups is similar before the experiment.

It can be shown from Table 2, that the significance level of the two groups in the test after the experiment is less than 0.001, which is lower than the critical value of 0.05. Therefore, there is a significant difference, which indicates that the English vocabulary level of the two groups is different after the experiment. It is obvious from Table 2 that the average score of the experimental group is higher than that of the control group in the test after the experiment. The analysis of the test before and after the experiment shows that the combination of context in English vocabulary teaching can help improve students' ability to understand, absorb, conclude, and apply English vocabulary.

In a nutshell, students can only understand the general meaning of vocabulary, having little or no knowledge of the specific usage of vocabulary through traditional English vocabulary learning methods. However, hands-off DDL can provide students with the real usage of English vocabulary and enable them to understand the specific usage of the target vocabulary. They can also learn vocabulary collocations through the corpus's index lines. Through classroom observation, it is found that hands-off DDL can indeed stimulate students' interest in vocabulary learning and improve students' participation and enthusiasm in class. Students can actively look for useful clues from the context, especially the use of other words and phrases in the context of the target word analysis. In this way, students will be more confident in participating in vocabulary learning, and gradually improve their vocabulary analysis ability and context guessing ability. This ability of "discovery" learning is suitable for the hands-off DDL.

7. Conclusion

Based on the method of comparative analysis, this paper analyzes the vocabulary learning level of two non-English major undergraduate classes at Southwest University of Political Science and Law before and after the application context experiment. The positive effect of this method on vocabulary teaching in the control groups is verified by comparing and analyzing the influence of context teaching based on hands-off DDL on their scores and learning effects between the experimental and control groups. However, this teaching model also has some shortcomings, such as the need for a large amount of data processing and analysis, teachers need to have relevant data literacy and skills, and so on. Therefore, schools need to establish a corresponding support system, including improving teachers' awareness and skills of data, building a big data smart learning platform, and carrying out data-driven teaching demonstration projects to build a data-driven teaching practice community. Through continuous exploration and innovation, combining hands-off DDL with college English context vocabulary teaching can improve the teaching quality and change the teaching model.

References

- Allan, R. (2006). *Data-driven learning and vocabulary: Investigating the use of concordances with advanced learners of English*. Centre for Language and Communication Studies, Occasional Paper, 66. Dublin: Trinity College.
- An, P. (2010). Implementation of Contextualized Teaching in the Cyber Age. *Theory and Practice of Education*, 30(6), 61-63.
- 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. (2010a). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534-572. <https://doi.org/10.1111/j.1467-9922.2010.00566.x>
- Boulton, A. (2010b). Learning outcomes from corpus consultation. In M. Jaén, F. Serrano, & M. Calzada (Eds.), *Exploring New Paths in Language Pedagogy: Lexis and Corpus-Based Language Teaching*. London: Equinox Publishing.
- Cheng, W., Warren, M., & Xu, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31(2), 173-186. [https://doi.org/10.1016/S0346-251X\(03\)00019-8](https://doi.org/10.1016/S0346-251X(03)00019-8)
- Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on Computer*. London & New York: Addison Wesley Longman.
- He, A. (2010). *An Introduction to Corpus-Assisted English Teaching*. Beijing: Foreign Language Teaching and Research Press.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524773>
- Johns, T. (1991). Should you be persuaded: two examples of data-driven learning. In T. Johns, & P. King (Eds.), *Classroom Concordancing*. Birmingham: ELR.
- Judith, L. S., & Glisan, E. (2000). *Teacher's Handbook: Contextualized Language Instruction*. FLTRP.
- Ke, S. (2012). A study on "cultural contextualization" college English teaching model. *Journal of Chifeng University (Natural Science Edition)*, 28(9), 203-205. <https://doi.org/10.13398/j.cnki.issn1673-260x.2012.18.087>
- Keck, C. (2004). Book review: Corpus linguistics and language teaching research: Bridging the gap. *Language Teaching Research*, 8(1), 80-109. <https://doi.org/10.1191/1362168804lr135ra>
- Li, P. (2012). A Review of Research on English Context Teaching in China from 2002 to 2011. *Success*, (7), 48.
- Li, Q., & Ren, Y. (2016). Review of the current research and practice of teacher's data literacy education in foreign countries. *E-education Research*, 37(5), 120-128. <https://doi.org/10.13811/j.cnki.eer.2016.05.018>
- Liang, M. (2009). Mini-text and its application in foreign language teaching. *Technology Enhanced Foreign Language Education*, (3), 8-12.
- Tribble, C., & Jones, G. (1990). *Concordances in the Classroom: A resource guide for teachers*. London: Longman Group Ltd.
- Xing, B., Yang, X., & Li, Q. (2016). Source and acquisition technology of big data in education. *Modern Educational Technology*, 26(8), 14-21. <https://doi.org/10.3969/j.issn.1009-8097.2016.08.002>
- Yang, X., Luo, J., Liu, Y., & Cheng, S. (2017). Data-driven Instruction: A New Trend of Teaching Paradigm in Big Data Era. *E-education Research*, 38(12), 13-20. <https://doi.org/10.13811/j.cnki.eer.2017.12.002>
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2. *Journal of Second Language Writing*, 13(4), 257-283. <https://doi.org/10.1016/j.jslw.2004.06.002>
- Zhang, B., & Zhou, J. (2012). A study on classroom implementation model of data-driven learning. *Foreign Languages and Their Teaching*, (3), 41-45.
- Zhu, X., & Yu, Z. (2010). A study of teaching college English in contexts. *Foreign Language World*, (4), 50-56.

Acknowledgments

We greatly appreciate the valuable contributions of subjects from Southwest University of Political Science and Law. This paper is one of the research achievements of “Research on Constructing College English Context Teaching Model Based on Hands-off Data-Driven Learning” (No. 2022B07) supported by the Key Project of University-level Higher Education Teaching Reform Research of Southwest University of Political Science and Law, and of “Reuven Tsur’s Theory of Cognitive Poetics” (No. 20XWW001) supported by 2020 Western Project of National Social Sciences Fund of China. We would also like to thank every project member who took the time to participate in this study.

Authors contributions

Dr. Haojie Li and Dr. Tongde Zhang were responsible for study design and revising. Dr. Haojie Li, who was responsible for data collection, drafted the manuscript. All authors read and approved the final manuscript. Dr. Tongde Zhang is the corresponding author.

Funding

This paper was supported by the Key Project of University-level Higher Education Teaching Reform Research of Southwest University of Political Science and Law “Research on Constructing College English Context Teaching Model Based on Hands-off Data-Driven Learning” [project number 2022B07], and by 2021 Jiangxi Province Elementary Education Research Project “A Practical Study of Structured English Knowledge Teaching in Senior High School on the Formation of Effective Learning Behavior” [project number SZUDHYY2021-1121].

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Canadian Center of Science and Education.

The journal’s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.