

Research Article

Data-Driven Learning Teaching Model of College English Based on Mega Data Analysis

Jie Zhang 

School of Foreign Languages, Hubei University of Science and Technology, Xianning 437000, China

Correspondence should be addressed to Jie Zhang; zhangjie@hust.edu.cn

Received 3 December 2021; Revised 25 December 2021; Accepted 27 December 2021; Published 17 January 2022

Academic Editor: Tongguang Ni

Copyright © 2022 Jie Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the arrival of the mega data era, CET has developed a new teaching model based on mega data. This teaching requirement is met by a data-driven model based on mega data. By applying it to CET, students will be guided to explore and discover language rules and pragmatic features through quantitative analysis of data-driven technology, which will help compensate for the disadvantages of traditional English teaching and improve students' autonomous learning ability. The concept of data-driven language learning is introduced into teaching in this paper. Corpus is tried to stimulate students' autonomous learning in the teaching process, and the independent learning model of students is further improved in the teaching reform, based on educational theories of corpus linguistics and second language acquisition linguistics. Students' scores have improved, particularly in English listening, according to the findings. The data-driven CET-4 model improves students' learning ability and interest, as well as their ability to think creatively and critically.

1. Introduction

The advancement of computer and network technology has influenced not only people's daily lives but also traditional teaching methods, resulting in fundamental changes in people's methods and means of acquiring knowledge [1]. The deep integration of the Internet and IT has ushered in a new era of mega data, which presents both an opportunity and a challenge for current foreign language instruction. The teaching model has also undergone qualitative changes as a result of the inclusion of scientific and technological elements [2]. English instruction has progressed from a single computer-assisted method to a mixed method based on the Internet and derived mega data [3]. Learning resources, learning objectives, learning contents, and learning tools have all changed dramatically in the college English curriculum as a result of the digital and networked system [4]. People from all walks of life are becoming increasingly aware that whoever can take the lead in realizing mega data and who has a deeper understanding of mega data mining and application will seize the opportunity in the future. In the field of education, the penetration of massive data will

inevitably lead to a high level of technological integration in the classroom. Developing smart classrooms, implementing data-driven learning (DDL) as a teaching model, and providing accurate decision-making will undoubtedly become a new trend in the future development of educational informatization, with a significant impact on profound changes in the field of education [5]. Under the tide of mega data, the DDL teaching mode method model, the DDL teaching mode method framework, and the practice of DDL teaching model are constantly emerging. The DDL teaching model has become an index to evaluate the effectiveness of any given teaching method [6].

The development of mega data technology provides new opportunities for the development of education and endows DDL teaching model with new connotations and missions. The wide application of data mining has brought rapid development to all fields of social life [7]. DDL model came into being in this situation and soon attracted people's attention, which will promote the development of foreign language teaching [8]. Applying it to college English teaching (CET) can create a real linguistic context for students and guide students to explore and discover language

rules and pragmatic features through quantitative analysis of data-driven technology, so as to make up for the disadvantages of traditional English teaching model and improve students' autonomous learning ability [9]. DDL teaching model can improve our teaching efficiency and ensure our teaching quality. For the education industry, mega data is a major opportunity for traditional education research to move towards scientific demonstration. College English courses should use advanced IT to carry out English teaching based on computer and network so as to provide students with a good language learning environment and learning conditions [10]. This paper introduces the concept of data-driven language learning into teaching, based on the educational theory of corpus linguistics and second language acquisition linguistics; tries to use the corpus to stimulate students' autonomous learning in the teaching process; and further improves students' autonomous learning model in the teaching reform.

Through quantitative analysis of data-driven technology, it can be applied to CET to create a real linguistic context for students and guide them to explore and discover language rules and pragmatic features, thereby compensating for the disadvantages of traditional English teaching models and improving students' autonomous learning ability [11]. The corpus' rich and authentic corpus, as well as teachers' active assistance, not only stimulate English learners' interest in learning and cultivate their ability of active exploration, analysis, and induction but also greatly activate learners' innate language intuition and improve their cognitive and metacognitive skills [12]. The DDL method is a corpus-based discovery method for learning a foreign language that transforms the teaching space from closed to open. Learners are independent builders of knowledge, and the learning process is a knowledge construction process of self-creation and self-motivation. The DDL model, which is based on an English-driven decision-making model, has significant value and significance in encouraging empirical English teaching [13].

2. Related Works

Literature [14] points out that the most remarkable feature of DDL is that teachers do not know what rules learners will find in advance. DDL model is to enable English learners to actively explore the rules contained in the English corpus by analyzing the real and abundant data of the English corpus and to produce top-down discovery learning so as to cultivate English learners' autonomous learning ability. Literature [15] divides the research object into the control group and the experimental group. The experimental group adopts the DDL model for teaching, while the control group adopts the traditional way of teaching by teachers. Through pre-test, post-test and qualitative research, it is finally found that using DDL model for vocabulary learning is far better than the traditional teaching model. Corpus is widely recognized in language teaching, especially that language learning should be based on the abundant corpus. Therefore, the corpus-based DDL model becomes the best choice for teachers. Literature [16] indicates that DDL model is still

more effective than the traditional language teaching method for language learners who have not received corpus operation training and primary level. In addition, it is far more beneficial to use keywords in a real and rich linguistic context for teaching than a long context composed of one or more sentences. Literature [17] shows that paper-based materials can eliminate the obstacles encountered in using corpus-based DDL model and make DDL model win more followers. This is because language learners at the primary level and those with lower qualifications do not have enough prior knowledge and patterns to follow, so they cannot learn effectively from a purely inductive teaching pattern. Literature [18, 19] holds that the transformation from DDL to demand-driven corpus should be completed. However, the DDL model cannot fully assume how to grasp the actual needs of learners. Literature [20] holds that the previous corpus-based DDL research ignored the differences of learners in the learning process. That is to say, the corpus-based DDL model only pays attention to the integrity of learners and ignores the particularity of learners.

The abundant corpus provided by the corpus can make up for the deficiency of the existing English textbooks to a certain extent and enrich the process of English teaching in form and meaning. At present, the application of the DDL model in teaching English to speakers of other languages is mostly limited to theoretical discussion but lacks systematic empirical investigation and research. Taking CET as an example, this paper studies the DDL teaching model of college English under the background of educational mega data and analyzes the operation process of this model.

3. Present Situation of CET

Although China's CET has been reformed in many aspects, the traditional classroom teaching model is still being followed, whether it is college English textbooks, teaching contents, or tests at all levels, including CET-4 and CET-6. Students rely more on vocabulary accumulation and grammar knowledge, focusing on various tests, rather than on the practice of communication in the real context. The combination of the corpus with computer and network makes its application in CET show us a new platform. In addition, corpus linguistics also plays an important role in foreign language teaching theory, content, and methods.

Our era is experiencing an information explosion due to the rapid development of the Internet and the emergence of cloud computing. The teaching model has also changed qualitatively as a result of the inclusion of science and technology elements. English teaching has progressed from a single computer-assisted method to a mixed method based on the Internet and mega data derived from it in the past. Corpus linguistics is a science of language analysis and research based on real corpora [21]. It is devoted to text retrieval, sampling, analysis, and statistics. The introduction of corpus linguistics into CET can merely compensate for the aforementioned flaws. Corpus linguistics has become increasingly linked to foreign language teaching as computers have advanced. The human-computer interaction-centered organization of classroom teaching is now the focus of the

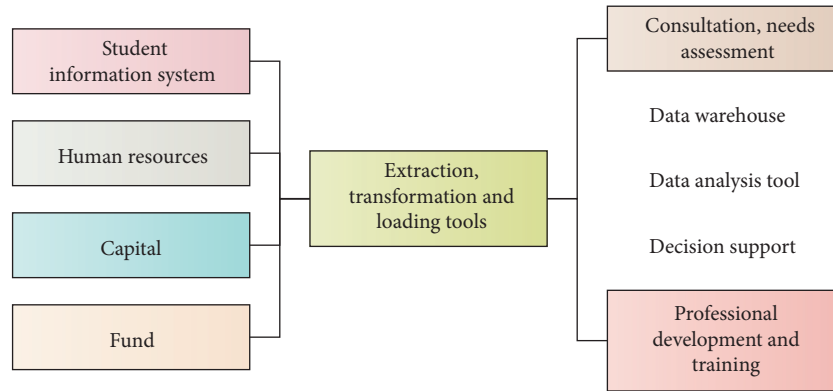


FIGURE 1: Data-driven decision model.

corpus-based data-driven CET model. After the teacher has compiled the language materials on the computer, students can read the materials on the computer and engage in student-centered learning. This model is only appropriate for the core educational viewpoint of today’s college English reform, namely, the autonomous learning education model.

4. Construction of English-Driven Decision Model

4.1. *Generation of Data-Driven Decision Model.* Corpus is a new type of learning resource; it is essentially a warehouse for storing language materials, where many natural language materials of spoken and written languages are gathered. Corpus is a large-scale collection of written and spoken natural language materials that serve as a warehouse for storing language materials, also known as a language database. Because of the advancement and application of computers, the corpora we now use are essentially electronic corpora, which store a large number of language materials in the form of text via computers. The emergence of corpora provides more advanced technical means for linguistic research, as well as significant shifts in language research thought [22]. Knowledge acquisition, according to constructivism, is not a one-time event, and learners’ social and cultural backgrounds and situations play a significant role in constructing the meaning of what they have learned. Learners are cognitive subjects, not passive recipients of knowledge. Learners benefit from learning in real or near-real situations because it allows them to explore the world based on prior experience and to build new knowledge on top of previous knowledge.

With the development of data mining technology, it is possible to make decisions by using various data analyses, which also enables the emergence and development of data-driven systems. The data-driven decision model is shown in Figure 1.

The learning process in DDL model is divided into three stages: pointing out problems, classifying materials, and summarizing. The target language in learning can be summarized as the process of data-driven language learning. Learners classify and search the corpus for context co-occurrence in order to collect a large amount of real

information, which they then input into the materials for induction and summary in order to obtain the law of language. The following aspects of the corpus can be roughly summarized: systematicness, authenticity, flexibility, representativeness, and richness, among others. These characteristics help foreign language learners create complex and realistic task scenarios, which not only help them master more basic language skills but also boost their learning enthusiasm and initiative. The autonomous learning of students dominates the learning process [23]. Teachers typically serve as a guide, organizer, and negotiator in this process, assisting students in determining the overall learning direction. While students are the driving force behind the learning process, they must manage, monitor, and evaluate themselves, as well as further develop their autonomous learning ability, which can have a positive impact on other aspects of students’ lives, forming a virtuous circle and promoting learning.

Language learning under corpus DDL model is the concrete practice of the core content of constructivism theory. With the increasingly mature corpus retrieval tools as the means, it observes and analyzes a large number of real language phenomena by learners themselves or in collaboration with others and sums up their language rules by combining the old information in their minds so as to realize the internalization of knowledge. Sustainability is used to measure the sustainability standard of quality evaluation reform and transformation system and the progress of each subsystem. The development speed of each subsystem is used as a measure as follows:

$$w = (w_{\max} - w_{\min}) \times \frac{I_{\max} - I_i}{I_{\max}} + w_{\min}. \quad (1)$$

Use clustering algorithm function method to find

$$Cr_{(t+1)} = k \times Cr_{(t)} \times (1 - Cr_{(t)}). \quad (2)$$

The calculated trend degree is as follows:

$$f_1(x) = \sum_{i=1}^{D-1} \left[100(x_{i+1} - x_i)^2 + (x_i - 1)^2 \right]. \quad (3)$$

Large corpora have evolved into shared resources, with some corpus retrieval software available for download and

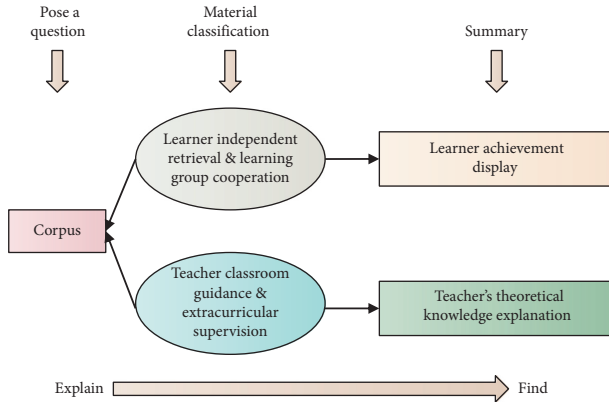


FIGURE 2: College English classroom teaching under the corpus DDL model.

use on computers. All of this provides a research foundation for implementing and developing DDL in foreign language instruction. DDL is a corpus-based language teaching model. Students' exposure to real language is generally limited in traditional teaching models, and the input of a large number of non-authentic languages causes various language errors, which has been fully verified in teaching practice. This is a problem that DDL excels at [24]. This is because all of the language data sources provided by the corpus for students are real communication activities, and they help students improve their language intuition and master a more pure language by creating a real linguistic context for them.

4.2. Construction of English-Driven Decision Model. The education industry has undergone significant changes and innovations as a result of the development of the Internet. The methods used to teach English are varied and innovative. Sources of teaching materials are becoming more diverse. Students and teachers can look up words on the Internet using online dictionaries, as well as use shared teaching resources created by other students. Teachers and students can use the Internet to discuss problems in the classroom and do supplementary learning after class by forming English learning groups or using the teaching platform. On the network platform, teaching is no longer limited to the classroom; it can now take place at any time and in any location. Teachers' lessons to students may be partially or completely uninteresting. This raises an important issue, namely, perception misalignment. As a result, perceived mismatch may be an important factor in causing unsatisfactory teaching results in second language teaching. Because of the traditional teaching model, the DDL model is no exception, and learner analysis is more influenced by test scores. Even though it involves a thorough examination of students' needs, motivations, and preferences, it is primarily based on teachers' intuition and experience. Learners' basic information, learners' learning motivation, learning preferences and needs, learners' vocabulary and grammar learning strategies, perceived mismatch between teachers and students, and

teachers' beliefs are all sources of data in the English-driven decision-making system.

Teachers' beliefs largely determine the success of classroom teaching. Obviously, the DDL model based on the English-driven decision-making model not only takes into account learners' factors but also recognizes the importance of teachers' beliefs and perceived mismatch between teachers and students for data collection and analysis. Corpus DDL model of college English classroom teaching is shown in Figure 2.

Before using the DDL model for teaching English to speakers of other languages, in order to ensure the teaching effect, all factors that may affect the teaching process must be collected from learners' learning motivation, learning preferences, learning needs, teachers' beliefs, and learning strategies. The English-driven decision model is shown in Figure 3.

Students participate in the establishment of key and difficult points in the classroom through study groups, conduct corpus retrieval analysis on the problems discovered or to be solved, observe the retrieval results, discuss in groups, and present the discussion results in the classroom. The English-driven decision-making model incorporates basic learner information, and its foundation is a better understanding of learner characteristics. Differences in thinking and cultural beliefs, for example, caused by differences in nationalization, have a significant impact on learners' learning activities. Students and teachers are confronted with a plethora of data. How can science and technology be used to screen these data and determine their practical value? It's a problem that both teachers and students must work together to solve [25]. The fundamental premise of using the DDL model in English teaching is to analyze the learning needs of English learners. When analyzing the learning needs of English learners, we must take into account the learners' complex expectations and contradictions, as well as the language and cultural needs of the learners. In order to create more targeted teaching materials that meet the needs of students.

Positive indicators, then

$$X'_{ij} = \frac{[\max(X_j) - X_{ij}]}{(\max\{X_j\} - \min\{X_j\})} \quad (4)$$

Inverse indicators, then

$$X'_{ij} = \frac{[X_{ij} - \min(X_j)]}{(\max\{X_j\} - \min\{X_j\})} \quad (5)$$

In order to ensure that the logarithmic calculation is meaningful when calculating the entropy value, all the data is shifted by 0.5, $Y_{ij} = X'_{ij} + 0.5$, $Y_{ij} \in [0.5, 1.5]$, and the shifted value is subjected to a proportional transformation to obtain the normalized value of each evaluation index. The proportional transformation formula is as follows:

$$Z_{ij} = \frac{Y_{ij}}{\sum_{i=1}^m Y_{ij}} \quad (6)$$

Teacher's input means that teachers subconsciously think that this is what learners should learn and master in

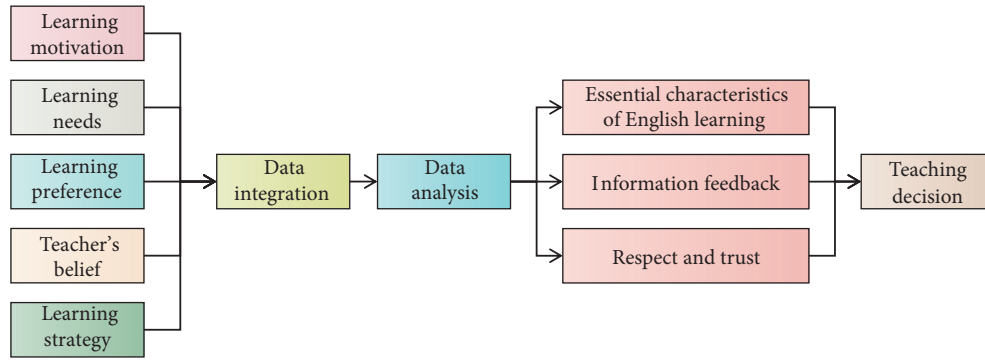


FIGURE 3: English-driven decision-making model.

this class by providing oral and written language materials or other forms of information. Learner acquisition refers to what learners get through personal participation or other learning methods in the course of the class, not just the ready-made information provided by language materials. In the multimedia classroom, teachers and students can experience this classroom teaching model together. However, attention should be paid to the proportion of online classrooms in the whole teaching system and the arrangement of class hours, which not only makes an attempt of classroom reform but also does not affect the normal learning progress of students.

5. Result Analysis and Discussion

It is the main carrier of language, and it is contained in culture, and culture is embodied in language. From the perspective of human beings, it is emphasized that the essence of culture is closely related to the essence of human beings. Culture fundamentally distinguishes human beings from animals, and the essence of culture is creation. Key-word list can be regarded as a common vocabulary or core word group that students often use to express a certain topic. This core word group is associated with other keywords in the thesaurus. Students have reduced semantic differences in the use of words. That is, some words sharing some semantic features are used alternately to express the same concept. This feature makes the students' vocabulary use focus on a few limited core words, and the concepts expressed are monotonous and vague.

Students may not have enough knowledge or confidence in words with more specific meanings, so they use words with broader meanings instead. When students use some fixed or semi-fixed collocations, their choice of collocations is free and changeable, unlike native English speakers, which have certain restrictions [26]. Therefore, to learn and use a foreign language, we must understand the culture closely related to the foreign language, and familiarity with the relevant culture is conducive to using a foreign language appropriately. A learning method that takes words or phrases as a unit and focuses on the sensory memory stage of memory is convenient for students to process vocabulary at a shallow level and memorize vocabulary. Select a number of students to test and compare the memory effect with that of traditional methods, as shown in Figure 4.

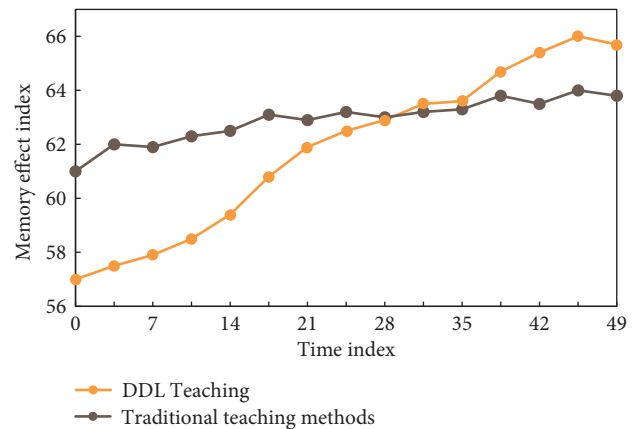


FIGURE 4: Comparison of memory methods.

Students' strong desire to pay attention to society and sense of participation is in sharp contrast with their vague understanding of social modeling. In the expression of students' theme, campus and society are divided into compulsory binary parts so that they can form semantic opposition in the use of words. Teachers should explain the teaching objectives and tasks of this class in class and at the same time conscientiously implement the teaching according to the teaching plan and lesson plan [27].

Set the related parameters of English learning resources distribution, reconstruct the constraint parameters of English learning ability evaluation, and get the time-domain curve of distribution as shown in Figure 5.

Teaching a language is a step-by-step process that progresses from simple to complex. As a result, the teaching of English culture should be done in stages and according to the principle of step by step. The content of culture instruction is determined by the students' language level, acceptance ability, and comprehension ability. Teaching should go from the easy to the difficult, from the simple to the complex, and from the phenomenon to the essence. Teachers should assist poor students in comprehending the cultural information provided by the text itself rather than introducing foreign content [28]. Within the same region, there are linguistic differences. Language has an impact on culture and can also be used to learn how to express culture.

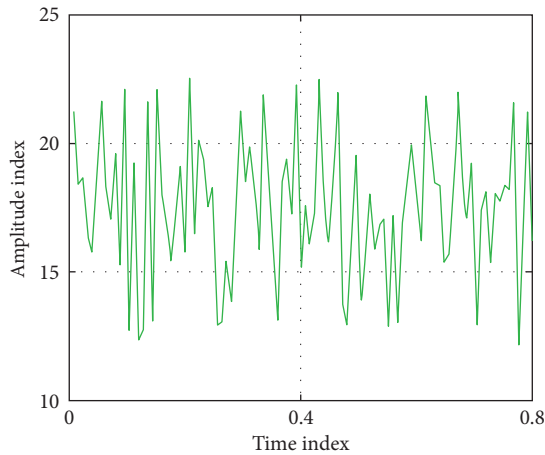


FIGURE 5: Time-domain distribution curve of mega data.

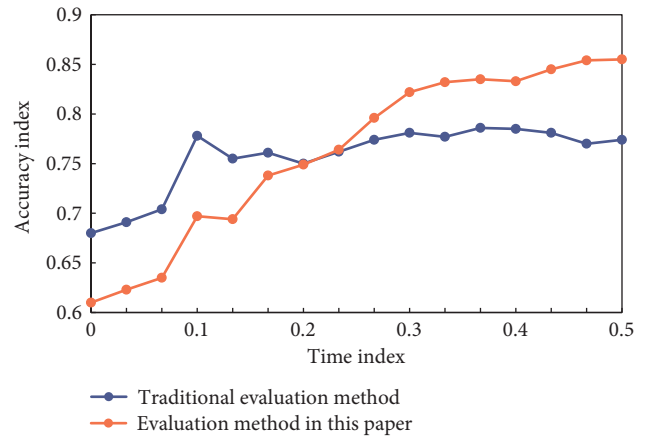


FIGURE 7: Comparison of utilization data of two analysis methods.

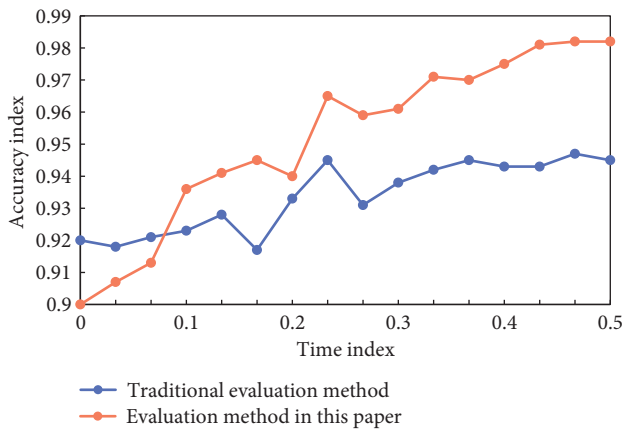


FIGURE 6: Comparison of accuracy data of two analysis methods.

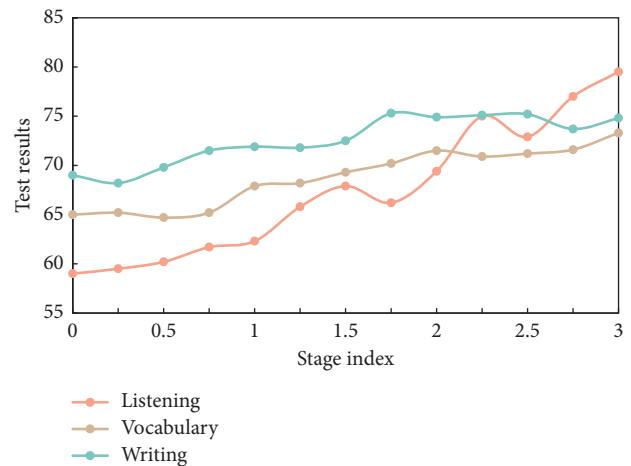


FIGURE 8: Comparison of achievements.

The accuracy of learning ability evaluation with this method is high, and the utilization rate of learning resources is good. A comparison of the two analysis methods is shown in Figures 6 and 7.

Language is not only a way of human communication but also a means of guiding the speaker's views. It provides people with idiomatic patterns for analyzing reality. Cultivating cross-cultural awareness in English teaching can better conform to the trend of globalization, so we must do this link well. The success of students' word use depends on whether they can build a word network around a theme, association, and word collocation. Its significance lies in that foreign language teaching, especially English teaching, centered on a certain topic and semantic association may be more effective than isolated and discrete vocabulary learning. If teachers do not care about the actual situation of students and the content of teaching materials, they will instill cultural knowledge into students, which is unacceptable to students with low English levels. Cultural diversity leads to linguistic diversity. Cultural differences are not only reflected in different regions. Many cultures will change with the passage of time and the influence of other cultures.

Using DDL model in English vocabulary and grammar teaching reflects the complexity of corpus retrieval, the

diversity of corpus content, the dynamics of corpus selection, and the difficulty and time consumption of corpus analysis and collation. It also reflects the dynamics of learners' learning motivation, the complexity of cognitive strategies, the nonlinearity, and self-organization in the learning process. Learners' learning in DDL model, from curiosity and excitement at the beginning to joy at the middle stage to negative emotions at the later stage, fully shows that learners' learning motivation has changed. The test results are shown in Figure 8.

Students' scores have improved, particularly in English listening, according to the findings. There is a pedigree relationship between words when a person uses words to express his understanding of the real world. The semantic space is at the top of this hierarchical relationship. Language learners struggle to match the semantic field of their mother tongue to the semantic field of the target language. The learner's task is to match the mother tongue's semantic field to that of the second language. What they do not realize is that a second language semantic concept is realized as a word [29]. It may, however, be realized in its mother tongue as multiple words, or it may not be realized at all, or it may only be realized as mother tongue through difficult translation

and interpretation. Different cultures are reflected in different languages. Language causes people to think and express themselves in a variety of ways. As a result, people act in different ways. There will be differences, while the globalization of education and the local foundation have an impact. Different thesaurus groups are divided according to semantic functions, and the statistical key thesaurus and thesaurus are grouped and analyzed according to composition topics. By connecting the analysis results with the composition topic and drawing, we can directly observe the basic words used by students to express a certain topic, as well as the relationship between these words.

6. Conclusions

Each teacher has his or her own approach to teaching, and each student has his or her own study habits. This personalized design is made possible by big data. Teachers can select their own teaching materials and develop their own teaching style. Teachers can pay close attention to each student's microscopic performance based on the data and adjust the education plan accordingly through the analysis of relevant student data, resulting in personalized education. With the help of abundant corpus materials and data-driven teaching methods, the corpus-based DDL teaching model of college English can effectively improve teachers' teaching level and students' autonomous learning ability, thus effectively improving the CET effect. The data-driven CET model improves students' learning ability and interest, as well as their ability to think creatively and critically. Students' overall scores have improved as a result of the new learning evaluation system. Students have a strong interest in learning physical education courses as a result of its implementation, and the new learning evaluation system has a positive impact on students' learning. This effect will be further enhanced if teachers can select and build a corpus of appropriate scale and suitable for students in our school from the actual level of students in our school and then carry out teaching activities on this basis.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] S. Wang and S. He, "Research on college English translation teaching design based on data-driven learning," *Journal of Xihua University: Philosophy and Social Sciences Edition*, vol. 36, no. 1, p. 6, 2017.
- [2] X. Xu and C. Yang, "The construction of college English teaching mode under the background of big data," *Science & Technology Information*, vol. 18, no. 2, p. 2, 2020.
- [3] P. Otto, "Choosing specialized vocabulary to teach with data-driven learning: an example from civil engineering," *English for Specific Purposes*, vol. 61, no. 5, pp. 32–46, 2021.
- [4] N. Vyatkina, "Data-driven learning for beginners: the case of German verb-preposition collocations," *ReCALL*, vol. 28, no. 2, pp. 207–226, 2016.
- [5] P. Crosthwaite, L. L. C. Wong, and J. Cheung, "Characterising postgraduate students' corpus query and usage patterns for disciplinary data-driven learning," *ReCALL*, vol. 31, no. 3, pp. 255–275, 2019.
- [6] R. K. Boettger and S. Ishizaki, "Introduction to the special issue: data-driven approaches to research and teaching in professional and technical communication," *IEEE Transactions on Professional Communications*, vol. 61, no. 4, pp. 352–355, 2018.
- [7] K. P. Seul, A. P. Hyeoun, and H. Hee, "[Development and evaluation of electronic health record data-driven predictive models for pressure ulcers]," *Journal of Korean Academy of Nursing*, vol. 49, no. 5, pp. 575–585, 2019.
- [8] G. Cui, "An empirical study of data-driven learning in college English vocabulary teaching," *Chinese Journal of Multimedia and Network Teaching*, vol. 3, no. 3, pp. 27–28, 2018.
- [9] J. Zheng, "The influence of peer support under "output drive" on spoken English," *Journal of Heze University*, vol. 42, no. 1, p. 5, 2020.
- [10] J. Pan and D. Xu, "Research on the project-driven academic oral English teaching model for medical graduate students," *Chinese Journal of Medical Education Research*, vol. 18, no. 2, p. 6, 2019.
- [11] Li Yun, "Data-driven "comprehensive English (1)" Curriculum formative evaluation paradigm construction," *English Teachers*, vol. 20, no. 12, p. 4, 2020.
- [12] X. Xie, "Data-driven practical research on improving junior high school students' English autonomous writing ability," *English Teacher*, vol. 21, no. 12, p. 3, 2021.
- [13] K. Rivers, "Data-driven hint generation in vast solution spaces: a self-improving Python programming tutor," *International Journal of Artificial Intelligence in Education*, vol. 27, no. 1, pp. 37–64, 2017.
- [14] M. Zhao, Q. Liu, A. Jha et al., "VoxelEmbed: 3D instance segmentation and tracking with voxel embedding based deep learning," 2021, <https://arxiv.org/abs/2106.11480>.
- [15] C. Luederitz, M. Meyer, D. J. Abson et al., "Systematic student-driven literature reviews in sustainability science – an effective way to merge research and teaching," *Journal of Cleaner Production*, vol. 119, no. 4, pp. 229–235, 2016.
- [16] L. Wei, "Study on the application of cloud computing and speech recognition technology in English teaching," *Cluster Computing*, vol. 22, no. 4, pp. 9241–9249, 2019.
- [17] X. Liang and Y. Zhang, "Research on English vocabulary teaching from the perspective of data-driven learning," *Contemporary Educational Theory and Practice*, vol. 11, no. 5, p. 6, 2019.
- [18] Y. Yang, "Research on data-driven college English translation teaching model based on bilingual corpus," *Journal of Hubei Second Normal University*, vol. 34, no. 12, p. 4, 2017.
- [19] J. S. Kinnebrew, J. R. Segedy, and G. Biswas, "Integrating model-driven and data-driven techniques for analyzing learning behaviors in open-ended learning environments," *IEEE Transactions on Learning Technologies*, vol. 10, no. 2, pp. 140–153, 2017.
- [20] J. Supovitz and P. Sirinides, "The linking study: an experiment to strengthen teachers' engagement with data on teaching and learning," *American Journal of Education*, vol. 124, no. 2, pp. 161–189, 2018.

- [21] E. Cotos, "Effects of DDL technology on genre learningLanguage," *Learning Technology*, vol. 21, no. 213, pp. 104–130, 2017.
- [22] S. J. Semple, J. K. Stockman, D. Goodman-Meza et al., "Correlates of sexual violence among men who have sex with men in tijuana, Mexico," *Archives of Sexual Behavior*, vol. 46, no. 4, pp. 1011–1023, 2017.
- [23] X. Ni, "The impact of data-driven learning methods on college students' English writing proficiency," *Journal of University of Shanghai for Science and Technology: Social Science Edition*, vol. 38, no. 3, p. 5, 2016.
- [24] H. Guan, "The application of data-driven learning in English article teaching," *Journal of YangLing Vocational and Technical College*, vol. 16, no. 2, p. 5, 2017.
- [25] S. Fu and B. Xu, "An empirical study of corpus methods in the teaching of medical academic English translation and writing," *Continuing Medical Education*, vol. 34, no. 3, p. 3, 2020.
- [26] Y. Xu, "The application of big data analysis in English precision teaching," *English Teacher*, vol. 20, no. 18, p. 4, 2020.
- [27] Y. Feng, "The application of big data analysis in junior high school English teaching [J]," *Beijing Education Teaching Research*, vol. 40, no. 1, pp. 19–22, 2019.
- [28] F. Lin, "Research on the evaluation of business English teaching model in the context of big data," *Microcomputer Applications*, vol. 37, no. 9, p. 4, 2021.
- [29] Y. Liu and W. Wu, "Teaching strategies for postgraduate academic English writing in the context of big data," *Quality Education in West China*, vol. 5, no. 7, p. 2, 2019.