Research Article

Multimodal Teaching of College English Based on Similarity

Jingjing Li

Institute of Humanities and Education, Xijing University, Xi’an 710123, China

Correspondence should be addressed to Jingjing Li; 20140121@xijing.edu.cn

Received 8 March 2022; Revised 25 March 2022; Accepted 28 March 2022; Published 18 April 2022

Academic Editor: Liping Zhang

Copyright © 2022 Jingjing Li. 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.

There are many disadvantages in traditional English teaching. In order to solve these problems, multimodal teaching mode is applied to college English teaching. This paper combines the similarity model with English multimodal teaching and puts forward corresponding teaching methods based on the current situation of students’ English learning. This paper studies the multimodal similarity learning method, defines different relation vectors according to different physical or semantic meanings contained in sample data, and transforms the calculation of single-modal similarity between samples into the calculation of multimodal similarity between samples under different relations. Based on the multimodal discourse analysis theory, this paper explores the influence of multimodal teaching mode on students’ English achievement and attitude by combining quantitative research with qualitative research, so as to verify the effectiveness of this teaching mode. The results show that this method is effective and feasible. The application of multimodal teaching mode in English classroom can have an important influence and has far-reaching research significance. The research in this paper has certain reference value for innovating English teaching mode and improving students’ English achievement.

1. Introduction

With the rapid development of national economy, China’s comprehensive national strength and international competitiveness are constantly improving [1]. Especially since China’s accession to the World Trade Organization, more and more enterprises have gone abroad to carry out global trade, and international exchanges have become more and more frequent [2]. English, as the largest common language for academic and trade exchanges in the world, plays an important role. Proficiency in English has become one of the qualities that talents in the new era must possess, which makes college English learning very important [3]. Although we have adopted many ways to improve our English level, the effect is not satisfactory [4]. There are many disadvantages in the traditional English teaching process, such as, the backward and single teaching mode, the lack of attention to practical teaching, and teachers’ excessive dependence on PPT. As a result, most students’ mastery of English only stays at the stage of word meaning and grammar, but cannot use it skillfully. Driven by the development of the times and the needs of the country, college English teaching in domestic universities has been reformed and innovated, and multimodal teaching of college English has emerged as the times require. Multimodal teaching has become a prominent feature of modern teaching [5]. Multimodal analysis is a new analysis theory initiated in the late twentieth century [6]. Before the multimodal theory was put forward, scholars put forward that using various media to participate in classroom teaching can stimulate students’ interest in learning. After the formal establishment of multimodal discourse analysis theory, domestic scholars often apply multimodal to discourse analysis when studying this theory. Multimodality is a communication mode in which a machine perceives the outside world and obtains information through various senses such as hearing, vision [8, 9], touch, sound, image [10, 11], and smell. It is a term describing the perception of external information.

In fact, multimodal theory can be fully implemented in the classroom, and it can be used by a variety of disciplines. Multimodal teaching, particularly in the field of English education, can stimulate students’ multireading and multimodal communication abilities, allowing them to achieve twice the result with half the effort. All types of symbol systems, along
with the ideographic mode of words, are used to complete the meaning construction of students’ learning in English listening and speaking courses. Textbook typesetting, teaching content setting, and the use of sensory modalities such as audio and video all reflect this. Traditional teaching methods should integrate and coordinate various modal modern teaching methods and promote teaching and learning by mobilising various multimodal teaching resources to form a scientific synergy, according to English multimodal teaching. Multimodal discourse analysis theory, which is based on systemic functional linguistics and sociosemiotics, emphasises the importance of other symbolic systems (such as sound, image, and colour) in the exchange of meaning expression [12]. Multimodal PPT courseware constructs meaning and strengthens language input through words, images, colours, sounds, and other forms as an auxiliary teaching tool [13]. It demonstrates its benefits and uniqueness in capturing students’ attention, increasing information input in class and enlivening the classroom environment. This paper conducts a relatively in-depth investigation into multimodal college English teaching, and its novelty lies in:

1. Based on the research on multimodal teaching of college English, the similarity model is innovatively combined with multimodal teaching of English. This paper studies the multimodal similarity learning method from the data and model perspectives. Based on the analysis of semantic similarity, a semantic synthesis algorithm is proposed, and the best semantic vector weight coefficient is obtained through continuous training. The results show that the performance of this algorithm is better.

2. Teachers can present and explain language knowledge in a variety of ways using modal resources such as language, sound, pictures, and videos, in order to activate students’ senses and help them better understand language knowledge. And address the issue of traditional teaching’s low quality, allowing students to actively learn and practise independently in a multimodal teaching mode to the greatest extent possible and enhancing college students’ comprehensive English literacy and overall English learning ability.

The specific chapters of this paper are arranged as follows.

This paper consists of five sections. The first section is the introduction of the article, which introduces the research background of this article and the corresponding challenges and states the significance of this study and the innovations of this article. The second section is the literature review part of this paper. This section mainly introduces the related content of multimodal and English teaching, and summarizes the related research at home and abroad. The third section is the research and design part of the article. Section 3.1 mainly analyzes the application of multimodal teaching mode in English teaching and summarizes multimodal discourse analysis theory, multimodal teaching mode, and other related contents. Section 3.2 explores the multimodal teaching mode of college English based on similarity model. The research method and process of this paper are given. In Section 4, we conduct the experiment of this study and analyze and discuss the research results in detail. The fifth section summarizes the research on multimodal teaching of college English based on similarity model and looks forward to the future.

2. Related Work

With the deepening of multimodal discourse research, scholars at home and abroad have shown great interest in the research on English multimodal teaching. Maluleke proposed that the development of a multimodal English teaching model is an important part of educational reform. The multimodal teaching mode incorporates a variety of communication methods, emphasising not only the use of multimodal “teaching” by teachers but also the use of multimodal “learning” and multimodal interactive assessment by students [14]. In their study, Lin proposed a multimodal teaching method, believing that communicative activities in classroom teaching and learning should be conducted in multimodal communication. Multimodality should be present in all aspects of teaching, including methods, content, and evaluation [15]. According to Liu and Jiang, in the traditional model, teachers typically dominate English classrooms, limiting students’ subjective initiative, resulting in a lack of interest in English and slow improvement in English scores [16]. The multimodal teaching method, according to Song, can be applied to courses such as English reading, writing, listening, and speaking [17]. P Wang conducted extensive research on the use of multimodality in English instruction. They believe that multimodal education is now an unstoppable trend [18]. Pan and Zhang distinguished multimedia symbols from multimodal symbols in their study, pointing out that multimodal semiotics is a new direction for future research and that people should cultivate multimodal and diverse comprehensive abilities [19]. Multiple modes are not used at will in teaching, according to Wei, and their selection and collocation must adhere to scientific laws in order to achieve the goal of promoting teaching [20]. Zhao investigated the application of multimodal teaching in the New Horizons College English classroom, including the principles of modality selection and specific applications [21]. Ma differentiated multimedia and multimodality, pointing out that multimedia and multimodality are two different learning modes that should be applied scientifically [22]. Teachers should combine students’ English learning level, English learning needs, teaching conditions and environment of colleges and universities, as well as the goals and requirements of innovative teaching system reform, etc., in the actual teaching process, according to Liu et al., and actively integrate multimodal teaching. The model is used to teach English in the classroom [23]. Multimodal discourse analysis and systemic functional linguistics theory are closely linked, according to Yu and Li [24], and systemic functional linguistics theory provides theoretical guidance for multimodal discourse analysis. The multimodal teaching mode was applied to English writing instruction by Liu, who
believed that it could improve students’ self-efficacy, intrinsic motivation, and extrinsic motivation for writing [25]. According to Lidar et al., using a multimodal teaching mode in an English classroom can transform the language from boring to active and from students from passive to active. It can be beneficial without causing harm if teachers use it properly, and it can improve students’ multiple reading abilities as well as their listening, speaking, and communication abilities [26].

Based on the previous studies on multimodal and English teaching modes, this paper combines the similarity model with multimodal English teaching and puts forward corresponding teaching methods based on the current situation of students’ English learning. Based on the similarity model, the difficulty judgment of English texts is studied, and a training and testing corpus is constructed. The effective features that can judge the difficulty of English text are extracted, and the inverse process of clustering algorithm is introduced to ensure that individuals with small similarity do not belong to the same class. Based on various classifiers, the text difficulty judgment algorithm based on single model, single feature, multifeature of single model, and multifeature fusion of multiple models is studied and implemented. In order to understand the feasibility of the method proposed in this paper, an interview survey was conducted before and after the experiment, and a questionnaire survey was conducted after the experiment. The test data are processed by algorithm analysis. The experimental results show that the multimodal teaching method in this paper is feasible and effective.

3. Methodology

3.1. Application of Multimodal Teaching Mode in English Teaching. Traditional language teaching only explains language from a single angle, using blackboard and books as the main teaching materials. This kind of learning method is boring, students cannot put more enthusiasm into learning, and the teaching effect is not good. If the teaching materials are enriched and the teaching methods are diversified, the classroom will change from single to diverse, from boring to active, so that students can be more actively involved in practice and promote the development of English teaching and students’ English communication ability. Multimodal refers to the communicative practice that transcends the language as a social symbol and uses two or more symbolic resources (languages, images, etc.) to complete meaning construction [27]. Multimodality provides a new perspective and direction for English teaching. Multimodal English teaching is to reasonably and appropriately apply various modal symbols and resources to English teaching practice and to stimulate learners’ interest in reading and arouse their enthusiasm and initiative in learning with the help of various symbols and resources. Through the cooperation of various symbols and resources, the meaning of discourse can be strengthened, supplemented, or expanded, so as to help learners better understand and master the teaching content.

In the process of multimodal teaching, teachers can make full use of multimedia and network platforms to provide sound, images, and so on as the real learning environment. Provide auxiliary information from various aspects, so that students can perceive, understand, and process the visual, listening, watching, and other modal information of contact through their brains, so as to strengthen the memory of the received information and digest knowledge from the actual perception, so as to realize the multimodal complementarity and transformation. Under the multimodal English teaching mode, there are many innovations in teaching objectives, contents, activities, and methods [28]. Teachers have changed from a single oral blackboard writing and a single PPT teaching input to rich teaching auxiliary resources such as video, audio, animation, pictures, and text materials. The classroom activities changed from empty and abstract virtual speech exercises to students’ learning “real environment speech examples” before speaking exercises. Multimedia participation can mobilize students’ hearing, sight, touch, and even smell, and stimulate students’ multireading ability and language application ability.

Multimodal teaching requires teachers to integrate various communication modes in teaching and adopt multimodal “teaching.” Therefore, English teachers should apply theoretical knowledge to practical teaching to train students’ practical ability. At the same time, teachers should change backward teaching ideas and teaching models. Teaching is regarded as an important process to cultivate students’ skills and strengthen students’ communicative experience. The process of learning English is also a process in which students acquire information. Teachers can use scientific teaching methods to enable students to efficiently acquire what teachers teach, so as to learn English efficiently. In the process of designing multimodal teaching classroom, teachers should integrate teaching conditions, teaching contents, teaching objectives, and other factors to design related teaching activities such as interactive teaching and scene simulation teaching, so that students can become the main body of multimodal teaching classroom and guide students to actively study and train in listening, speaking, reading, and writing. At the same time, teachers can use body language to assist teaching. For example, posture, movement, facial expression, and eye contact can convey different visual signals. Appropriate body language can also stimulate students’ visual senses and deepen students’ understanding of English teaching content. Figure 1 shows the selection of modal symbols and the classification of interests in English teaching.

Teachers can use new media technology to practise constantly, stimulate students’ vision and hearing by using words, pictures, animations, videos, and other forms, and actively carry out innovative teaching around the points that students are interested in, thereby stimulating students’ learning interest and self-learning awareness. Teachers, for example, can use PPT to organise pictures and small animations. And make the dead teaching materials come alive by designing the picture effect, loading the background music, and making the dead teaching materials come alive. Teachers should, however, pay attention to a reasonable layout, a clear theme, properly displaying the teaching content,
and avoiding fancy when creating PPTs. Multimodal teaching can use various modes, scientific collocation, scientific choice, and scientific design in a scientific way so that students can perceive language more stereoscopically and fully experience English. At the same time, in college English classrooms, teachers can divide students into several groups based on the knowledge points in microclass videos and textbooks, and then ask students to prepare corresponding group learning activities based on the principle of voluntariness or the grouping principle of homogeneity between groups and heterogeneity within groups.

Teaching evaluation should be multimodal in today’s world, given the popularity of multimedia technology. To improve the rationality, scientificity, and efficiency of teaching evaluation in multimodal teaching, teachers should pay attention to the multimodal transformation of teaching evaluation and adopt multimodal teaching evaluation methods. The following aspects of teaching are evaluated by teachers: teaching attitude, teaching preparation, homework correction, and so on. It is also important to look into whether teachers’ teaching methods are using multimodal multimedia teaching methods in a scientific and reasonable way, whether the goal is to cultivate students’ multi-English application ability and cross-cultural communication ability, and so on. Furthermore, the evaluation data is analysed and fed back, and the ability of teachers and students to use, identify, and process multimodal information is summarised in real time, allowing the advantages to be maintained while the disadvantages are corrected.

Multimodal English teaching is mainly divided into three stages: preclass preparation stage, course progress stage, and course ending stage. All teaching activities are carried out under these three stages. Multimodal teaching mode aims at cultivating students’ multi-English application ability and scientifically using modern science and technology to apply multimodal senses to English teaching, which makes students’ English learning more efficient and scientific. The application of multimodal teaching mode can change English from boring to active, from decadent to magical, from boring to enjoyable in class, and from passive to active. Teachers’ application of multimodal teaching mode in English classroom teaching has more positive significance, mainly because multimodal teaching mode can give full play to students’ self-awareness and exploration enthusiasm, promote students to explore and master more English knowledge in English classroom, and strengthen students’ comprehensive English literacy.

3.2. Multimodal Teaching of College English Based on Multimodal Similarity Learning Model. Similarity is a simple and complicated word that is used frequently in philosophy, information theory, and linguistics. The simple reason is that this word is frequently used, but the more complicated reason is that there is currently no unified concept. Semantic similarity refers to judging and scoring the semantic similarity of two texts or sentences. The degree of substitutability between two sentences, as well as the degree of word meaning agreement, is reflected in sentence similarity. The minimum number of editing times required to convert one string into another is referred to as the editing distance. Character replacement, insertion, and deletion are among the editing operations available. The smaller the similarity between two strings, the greater the editing distance. Speech recognition, image processing, automatic summarization, and other fields all use editing distance. The editing distance of strings is used in English sentence similarity. The total number of documents in the corpus is divided by the total number of documents containing the word, and the result
is logarithmic. At the same time, the mathematical expression is meaningless in order to prevent the total number of documents containing the word from being zero. Because relationships in practical tasks are frequently hierarchical, the single-layer multimodal similarity calculation must be expanded to hierarchical multimodal similarity calculation. The shared function determines how similar two people are. The similarity between two individuals is high if the shared function value of the two individuals is large; otherwise, the similarity between the two individuals is low. As a result, the shared function can be used to calculate the shared function values of each individual and all other individuals, and the individual sharing degree is the sum of all shared function values for this individual. Figure 2 is the flow chart of this algorithm to solve multimodal problems.

The calculation of sentence similarity is based on word units, taking into account all kinds of components in sentences and their corresponding weight values determined by preliminary experiments. The sentence similarity calculation scheme proposed in this paper takes the word order, sentence length, keywords, nonkeywords, and other factors into consideration. For many practical problems, there must be useful local structures in weak features and useless local structures in strong features. Different weights are given according to the effectiveness of structures in different features. If more features have a similar local structure, this local structure is considered to be more effective. On the contrary, if fewer features have a similar local structure, this local structure is considered less effective. Semantic dependency-related sentence similarity calculation methods start with the syntactic structure of sentences and analyze sentences into dependency tree structures. These methods need to introduce a large number of grammar rules. Text similarity should not only consider the similarity of the sentences that make up the text but also consider the structure of the text. The method of quantifying and learning the correlation between input samples and samples is called similarity learning.

In this paper, the multimodal similarity learning method defines different relation vectors to represent different topics in the input data and calculates multimodal similarity between samples under different topics to reflect similarity between samples under different relationships. Each particle finds the particle with the least similarity to it among them, ensuring that the two particles do not share the same evolutionary niche. In the beginning, all of the particles in the population are assigned to the same niche. The population is automatically divided into several subpopulations and searched in parallel as the population evolves. Word segmentation can be performed manually or automatically by a computer using a preprogrammed algorithm. Manual word segmentation has a high workload, but it also has a high accuracy. Although computer word segmentation is quick, it is less accurate than manual word segmentation. In this paper, a combination of computer automatic word segmentation and manual word segmentation will be used to improve the speed and accuracy of word segmentation when dealing with small-scale corpora.

By designing a similarity calculation equation, the correlation can be quantified for the three elements (such as formula (1)) composed of sample  \( i \), sample  \( j \), and relation  \( t \). The higher the value, the higher the similarity between  \( i \) and  \( j \) under  \( t \), and vice versa, the lower the similarity. Inspired by the bilinear similarity function, the similarity calculation function is shown in Equation (2).

\[
    s_{ij}^{(t)} = \Phi_1(e_i^T e_j)\Phi_2 = (\Phi_1 e_i^T) (\Phi_2 e_j^T).
\]

On the basis of the above similarity calculation function, a projection vector  \( p_t \) is added. The samples can be calculated by  \( \Phi_1 p_t^T p_t \). Therefore, the modified similarity calculation function can be written as:

\[
    s_{ij}^{(t)} = (\Phi_1 p_t^T p_t e_i^T) (\Phi_2 p_t^T p_t e_j^T).
\]

The similarity between  \( i \) and  \( j \) can be calculated as:

\[
    s_{ij} = \sum_{t=1}^{c} \ln \left( 1 + \exp \left( s_{ij}^{(t)} \right) \right).
\]

The optimization goal of the model is based on the principle of empirical risk minimization, and its objective function is:

\[
    J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \text{Cost} \left( h_\theta \left( x^{(i)} , y^{(i)} \right) \right).
\]

When  \( y = 0 \):

\[
    \text{Cost} \left( h_\theta \left( x^{(i)} , y \right) \right) = - \log \left( 1 - h_\theta \left( x \right) \right).
\]

When  \( y = 1 \):

\[
    \text{Cost} \left( h_\theta \left( x^{(i)} , y \right) \right) = - \log h_\theta \left( x \right).
\]

From this, we can get:

\[
    \text{Cost} \left( h_\theta \left( x^{(i)} , y \right) \right) = - \left( y \log h_\theta \left( x \right) + (1 - y) \log \left( 1 - h_\theta \left( x \right) \right) \right).
\]

Then, the gradient descent method is used to solve the parameter  \( \theta \), and the final decision function is obtained.

In addition to being taken into account from the standpoint of the model, the quality of the input data has an impact on the similarity learning results. As a result, it must be designed from the standpoint of data input. As a method of data optimization, representation learning not only denoises and reduces the dimension of the data, but it also highlights the data's nature. Individuals with I/CF were
chosen at random from the population to form a temporary subpopulation, and the newly generated individuals were compared to each other in the temporary subpopulation. The number of matched alleles was used to determine similarity, and the individuals with the highest similarity were replaced with new individuals. The goal of multimodal similarity learning is to improve the interpretability of traditional methods by measuring the similarity of input samples under different relation vectors. The algorithm will save the found extreme points until all the extreme points have been found, then reduce the fitness value of the extreme points in a small range in a specific way to prevent the optimization from searching the extreme points again. To extract different features from the same data, different feature extraction methods are used. After defining some features that have a significant impact on similarity, this paper provides several definitions related to English sentence similarity calculation and then proposes a similarity calculation scheme based on these definitions, as well as concrete implementation steps for sentence similarity calculation.

Accuracy is determined by calculating the average of the errors between all extreme points found by the search and those that already exist, as shown in

\[
\text{Accuracy} = \frac{1}{|Np|} \sum_{i=1}^{Np} f(p_i) - f(\text{global-optima}_i). \tag{9}
\]

Among them, \(Np\) is the number of global extreme points, and \(\text{global-optima}_i\) are a pair of known extreme points and extreme points found by searching; Accuracy represents the accuracy of the multimodal optimization algorithm. The average number of iterations can be calculated using Equation (10):

\[
\text{ave} = \frac{1}{N_p} \sum_{i=1}^{Np} \text{iter}, \tag{10}
\]

where \(Np\) is the number of global extreme points; iter, represents the number of iterations needed to find the \(i\)th extreme point. The smaller the value of \(\text{ave}\), the better the performance of the multimodal optimization algorithm.

The obtained hierarchical structure is used to control the number of neurons in each layer and the link state between them. The similarity between samples and samples on each neuron can be regarded as whether there is such a linking pattern in a sentence and the degree of such a linking pattern. At the same time, the cluster center of the first layer is regarded as a relation vector, which is used to calculate the similarity model of the single layer. Similar individuals are likely to belong to the same population. Based on this idea, a module similarity model is proposed to determine whether individuals belong to the same niche. Therefore, if two particles have closer fitness value and Euler distance, then they probably belong to the same niche. In order to solve the influence of various elements in complex samples on similarity learning, attention mechanism is introduced on the basis of multimodal similarity learning. Similar to the human perception system, attention mechanism can give different attention weights to different blocks in the sample according to the specific needs of training, so that blocks needing attention get more attention, whereas blocks that do not need attention are ignored accordingly.
4. Result Analysis and Discussion

In order to verify the effectiveness of the similarity model proposed in this paper, this model is applied to the field of text retrieval, and every part of the design of the model is evaluated. After that, it is compared with the latest existing model. In semisupervised learning, some experimental samples already know the corresponding labels. Sample pairs with the same label are called positive example pairs, and sample pairs with different labels are called negative example pairs. Different from the traditional keyword extraction...
method, the experiment extracts keywords from English sentences. The traditional keyword extraction is generally word segmentation and part-of-speech tagging of Chinese sentences. The advantage of this method is that it can extract the keywords of sentences quickly and accurately, and it is suitable for large-scale corpus. Its disadvantage is that some words play an important role in the sentence, but the part of speech does not meet the requirements and is ignored. Keyword extraction according to the role of words in sentences in this chapter is not necessarily related to the part of speech of words. This keyword extraction method is suitable for small-scale datasets, and the accuracy of keyword extraction is higher than that of traditional methods. We use four methods LSA (implicit semantic analysis), LSA+N-Gram, Word2Vec, and this method to do the experiment of screening candidate words. The experimental results are shown in Table 1.

The above table shows that the correct rate of candidate word screening experiment with this method is 96.31%, which is higher than that of LSA, LSA+N-Gram and Word2Vec. Therefore, the effect of this method is the best. The algorithm runs for a fixed number of iterations, and the accuracy is used to measure the error between the global extremum found by the algorithm and the actual global extremum. Each population needs to compare the global extremum points found by its own search with the existing global extremum points. The smaller the error between the extremum points found by searching and the existing extremum points, the better the performance of the algorithm. In order to verify the performance of this algorithm, we compare this algorithm with traditional algorithm and word vector algorithm. The errors of different algorithms are shown in Figure 3.

By providing an index text, the text retrieval algorithm searches the database for texts that are similar to the index text. As a result, the similarity calculation between the indexed text and the retrieved text has a direct bearing on the quality of text retrieval results. Five different datasets are used to evaluate the effectiveness of text retrieval. We use the Word2Vec tool to convert word representation into a real-valued vector and perform vector operations on it to calculate the similarity of text semantics. The supervised multimodal similarity learning method uses the previously expressed learning results as input and defines and optimises different relationship vectors and projection vectors to learn the similarity of different samples under different relationships. The test dataset includes long and short sentences as well as punctuated sentences, which are then processed into vector representations containing only keywords, all words, and nonkeywords. Finally, the similarity calculation method discussed in this paper is used to compare the results of these word vectors’ similarity calculations. The recall index is used to compare different algorithms, and the results are shown in Figure 4.

For the training of hierarchical similarity, first, all the neurons are fully linked for training as a whole pretraining. Then, according to the results of hierarchical clustering, the weights between neurons that should not be linked are set to 0. A simple and effective way to find English collocation of search words in corpus is counting. If two words appear many times, it is more likely that they are collocations. However, it is not ideal to select only the most frequently occurring doublet. We have counted the words in the corpus and listed the most frequent binary groups in the corpus and their frequency of occurrence, as shown in Table 2.

In actual tasks, the relationship between samples follows the distribution of semantic relationships, and there is hierarchical subordination. Therefore, the single-layer relationship between samples is extended to multilayers, and
hierarchical multimodal similarity learning method is used. Replace the synonyms in the relevant English texts in the test set, insert the replaced test set into the mysql database, and use the c3p0 database connection pool and the dbutil database connection tool to connect to the database, the programming tool is Eclipse, and the programming language is Java. In the text retrieval process, firstly, the query word should be entered in the input box. After analyzing and matching the word, it is judged whether the matching result is empty or not. If the match is successful, it means that the matched example sentences are retrieved and displayed. If the matching fails, it means that there is no matching example sentence, and further spelling check is needed, relevant words are recommended, and then, retrieval is conducted again. Success rate is the most important performance index to evaluate multimodal optimization algorithm. The success
rate is mainly through running the optimization algorithm for $N$ times in a row. If the optimization algorithm finds all the global extremum points in the feasible solution space of the test function after a certain run, then it is considered that this run is successful; otherwise, it is considered that this run is failed. We have conducted many experiments to verify the success rate of this method. Draw the result as a line chart, as shown below.

From the results in Figure 5, the advantages of this algorithm in success rate are verified. The selection of external sample memory describes the different characteristics of a given sample by looking for different neighboring samples of the given sample and using different neighboring samples. And choose the samples by giving the initial memory between them, which will give higher weight to the sample pairs that are more in line with the actual conditions. In order to improve the accuracy of the obtained collocation, we use a simple heuristic rule to filter the candidate phrases. A stands for adjective, P stands for preposition, and N stands for noun. In order to better reflect the hierarchical relationship between samples, all neurons in the hidden layer are not fully linked here, but the number of neurons in the hidden layer and their links are linked by hierarchical clustering. Considering nonkeywords, in order to verify that the similarity calculation method proposed in this paper is superior to other models, some English sentences are randomly selected, and then, these sentences are processed. For each sentence, two similar English sentences are given. The accuracy of the algorithm is shown in Figure 6.

From Figure 6, it can be seen that compared with other algorithms, the accuracy of this algorithm is higher. In this algorithm, not only the characteristics of the input data as a whole are considered but also the significant blocks that need attention inside the input data are considered, similar to the human perception system. In order to apply the attention mechanism, the input data is divided, and the attention weight of each part in the data is automatically learned by the algorithm. The keyword query module mainly extracts the keywords of the returned example sentences, and accurately expresses the main information of the example sentences through the keywords, so that users can select and learn a large number of returned example sentences according to their own interests. Text labels and thesaurus included in the dataset, etc. The above information can be subject clustered by a hierarchical clustering method, and the hierarchical structure obtained by clustering can be regarded as the hierarchical network link structure. The success rate can be used to measure the stability of the algorithm. If the optimization algorithm can successfully search all the extreme points at one time, but it cannot run successfully at other times, then the performance of the algorithm is unstable. The stability of the optimization algorithm is tested, and the results shown in Figure 7 are obtained.

It can be seen from Figure 7 that the stability of the optimization algorithm in this paper is high. Divide the input samples into different blocks, and use the attention mechanism to update the attention weights of different blocks in the samples in real time according to the actual situation. It ensures the interpretability of the training results, improves the learning effect of similarity, and verifies the effectiveness of the method through experiments. By analyzing the wrong experimental data and calculation formula, it is found that when calculating the similarity between two sentences, the similarity calculation method based on relation vector model mainly considers the influencing factors of shorter sentences, ignores the influence of longer sentences on the calculation of sentence similarity, and considers the similarity of sentence length unreasonable. However, the similarity in this paper can accurately measure the influencing factors of two sentences, so the accuracy is also high.

5. Conclusions

With the development of teaching reform, multimodal teaching model has been paid more and more attention and studied in college English teaching in domestic universities. Multimodal teaching relies on the digital environment, providing students with a platform to use language. Using this resource, teachers should make their language learning ability active through existing theories and specific organizational methods and urge them to choose language expressions in changing situational contexts, so as to promote the correspondence between language and context. Combine the article similarity model with English multimodal teaching. From the model point of view, aiming at the lack of interpretability of parameters and calculation process of traditional similarity learning methods, several relationship vectors with different meanings are proposed and defined to quantify the similarity between samples under different relationships. And provide some methods for college English multimodal teaching. This method enables teachers, learners, subjects, and objects to interact in many ways, fully arouses learners’ enthusiasm and initiative, cultivates foreign language learners’ innovative consciousness, and makes language learning a more natural behavior pattern. At the same time, compared with the existing algorithms, the effectiveness of this method is verified by experiments. The experimental results show that this method is interpretable.

In short, in the process of college English classroom teaching, teachers should actively apply multimodal teaching mode to cultivate students’ learning autonomy and English comprehensive quality and further improve the effectiveness of English classroom teaching. With the construction and application of multimodal teaching model, it is believed that it will be further promoted and popularized in the future, and it will further promote the development of English teaching and promote the reform of English teaching.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author does not have any possible conflicts of interest.
References


