

Retraction

Retracted: Optimization and Evaluation of Oral English CAF Based on Artificial Intelligence and Corpus

Security and Communication Networks

Received 26 December 2023; Accepted 26 December 2023; Published 29 December 2023

Copyright © 2023 Security and Communication Networks. 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.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] Z. Wang, "Optimization and Evaluation of Oral English CAF Based on Artificial Intelligence and Corpus," *Security and Communication Networks*, vol. 2022, Article ID 4649643, 10 pages, 2022.

Research Article

Optimization and Evaluation of Oral English CAF Based on Artificial Intelligence and Corpus

Zhan Wang 

Foreign Language Department, Qinhuangdao Vocational and Technical College, Qinhuangdao 066100, China

Correspondence should be addressed to Zhan Wang; wangz@qvc.edu.cn

Received 1 July 2022; Revised 26 July 2022; Accepted 2 August 2022; Published 29 August 2022

Academic Editor: Hangjun Che

Copyright © 2022 Zhan Wang. 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 development of the times, the exchanges between countries are increasing. China is becoming a superpower, and the number of international cities is increasing. This requires the communication level of Chinese people to be improved. English, as the second largest communication language in China, should be better understood and studied. This paper makes an in-depth discussion on the optimization and evaluation of oral English CAF based on artificial intelligence and corpus and makes an experimental analysis. The results are as follows: (1) introducing oral English teaching based on artificial intelligence and oral English based on a corpus, so as to deepen the public's cognition of both and make oral English more deeply rooted in people's hearts and get attention. (2) Analyzing the algorithms of phoneme errors in spoken English. Errors in spoken English are very common. Algorithms can be used to identify them better. When evaluating spoken English, algorithms are needed to evaluate them more accurately. (3) There are many examples of the benefits of artificial intelligence to oral English teaching. By comparison, it is found that the method of evaluating using artificial intelligence is more accurate, a corpus can improve oral English, and CAF optimization is also of great help to oral English.

1. Introduction

The work of artificial intelligence is explained by defining intelligent agents and their functions in production systems, reaction agents, real-time conditional schedulers, neural networks, and theoretical decision systems. Proxy learning is interpreted as extending programmers' reach to unfamiliar environments and shows how this role restricts their design and promotes knowledge representation and clear thinking. Robotics and vision are only defined as elements to achieve goals [1]. The second volume of artificial intelligence manual focuses on improving artificial intelligence (AI) and its growing applications, including programming languages, CAI intelligent systems, and the application of AI in medicine, science and technology, and science and education. First of all, this book develops a programming language for artificial intelligence research and application-oriented artificial intelligence research. The discussion focuses on scientific applications, chemical applications, dependencies and assumptions, artificial intelligence, and the functionality of the

LISP programming language [2]. The work of artificial intelligence is explained by defining intelligent agents and their functions in production systems, reaction agents, real-time conditional planners, neural networks, and theoretical decision systems. Proxy learning is interpreted as extending the programmer's scope to unfamiliar environments and shows how the role limits his design to facilitate knowledge representation and clear thinking. Robotics and vision are only defined as elements to achieve goals. This book emphasizes the importance of the task environment, which is the decisive factor in correctly designing agents [3]. Almost all the literature on artificial intelligence is expressed in computer terms, full of complex matrix algebra and differential equations. Unlike many other books on computer intelligence, this shows that most ideas about intelligent systems are simple and clear. It is designed for lectures and for students with little computer skills, and readers do not need any prior knowledge related to programming language knowledge. The methods used in the book have been thoroughly tested in several courses led by the author. This paper introduces the field of

computational intelligence, including rule-based expert system, fuzzy expert system, framework expert system, artificial neural network, evolutionary computation, hybrid intelligent system, knowledge engineering, and data mining [4]. A new eyeliner/ancient phase method for measuring terms and concepts has been proposed. It combines the classification structure of words with the information of statistical database, so that it can analyze the computational evidence obtained from the data distribution of database, thus eliminating the old distance between nodes in the old space. Specifically, the proposed calculation is combined with the edge method in the edge calculation scheme and further expands from the node information calculation method. The keywords of public data collection similar to other computer models are tested [5]. Microblog has become a popular communication tool nowadays. Millions of users exchange views on different aspects of life every day. Therefore, Weibo website is a rich data source for opinion polls and emotional analysis. Since Weibo has only recently appeared, there are research articles specifically aimed at this topic. In our article, we will focus on using Twitter, the most popular microblogging platform, for emotional analysis tasks. We show how to automatically collect corpus for emotion analysis and investigation. We analyze the collected corpus and explain the phenomena found. Using a corpus, we construct an emotion classifier that can identify positive, negative, and neutral emotions in a document. Experimental evaluation shows that our proposed technology is more effective and has better performance than previous methods. In our study, we use English, but the proposed technique can be used in any other language [6]. As a medium of communication, English plays an important role in the world, which makes it more necessary to learn better spoken English. However, for many learners who learn English as a second language, speaking English seems to be more difficult than other English skills such as writing and reading. At the same time, independent college students in China have their own challenges in learning English in the first five minutes of each class [7]. Perplexed by uncertainties in learners' emotions, a survey was conducted to find out what other factors besides attitude, motivation, and language will affect language production. Causing a heated debate, the survey results lead to language barriers and lack of self-confidence. The author suggests that these can be enhanced by fuzzy learning, programmed learning, and personality matching. Finally, a flexible task system is proposed to maximize the impact of oral English teaching [8]. The purpose of this study is to find out the related factors that affect the implementation of oral English teaching assessment in rural middle schools. First, the purpose of this study is to determine the language evaluation level of schools according to the selected demographic factors. Secondly, the purpose of this study is to find out whether English teachers are familiar with the content, function, and application of oral English assessment in schools. Then, the purpose of this study is to determine the cognitive level of English teachers in schools. Then, this study further explores the relationship between oral assessment in the school environment and factors affecting

teachers' content, function, and consciousness [9]. In the current college teaching of English majors in China, it is common for many students to learn a lot of grammar knowledge and vocabulary, but they still cannot speak English accurately. Teaching methods have changed from paying attention to the grammatical structure of language to task-based teaching or communicative teaching. However, in these two approaches, fluency is the main goal of language teaching, while accuracy is often neglected. In recent years, corrective feedback has become a hot research topic in second language acquisition abroad because of its potential role in promoting oral English development [10]. This paper probes into the basic rules of oral English teaching from the perspectives of the connotation of communicative competence, the characteristics of oral communication and learners' willingness to communicate, and puts forward six suggestions on the reform of oral English teaching based on Chinese oral English and the modern teaching reality [11]. A skill is an activity aimed at improving oral fluency. This study investigates the effect of this skill on oral fluency and accuracy of 10 non-English majors. The main findings include the following: (1) compared with the former, the subjects can produce more fluent and accurate speech in the latter's conversation; (2) from the comparison of every two conversations, that is, the comparison of 4-minute conversation with 3-minute conversation, the comparison of 3-minute conversation with 2-minute conversation, and the comparison of 4-minute conversation with 2-minute conversation, the subjects showed the greatest progress from 4-minute conversation to 2-minute conversation, which indicated that the more opportunities for repetition, the more fluent and accurate the speech [12]. In the field of SLA, there are many research studies on the quality and condition of input and output, but there are few research studies on the influence of the input mode on oral output. In this experiment, the microgenetic method is used to study the influence of input methods on oral English. The results show that input patterns have different effects on oral production, and input and output frequencies play an important role in oral production. The research results are of great significance to the oral English teaching method and the measurement of oral English production [13]. In traditional English language teaching, summative assessment is widely adopted and accepted, but it is neither scientific nor reasonable. Language teaching and learning is a process, the result of which cannot be evaluated by one or two tests, especially formative assessment is a form of test that requires students to complete a task. This paper studies the significance, purpose, and principle of formative assessment and puts forward the mode and method of applying formative assessment in universities [14]. The study of corpus linguistics shows that, in actual communication, there are a large number of ready-made lexical chunks that constitute the core of the language structure. Based on the lexical approach, the present study aims to explore the effectiveness of lexical approach in improving students' oral English. The results show that teaching lexical chunks can improve students' communicative competence [15].

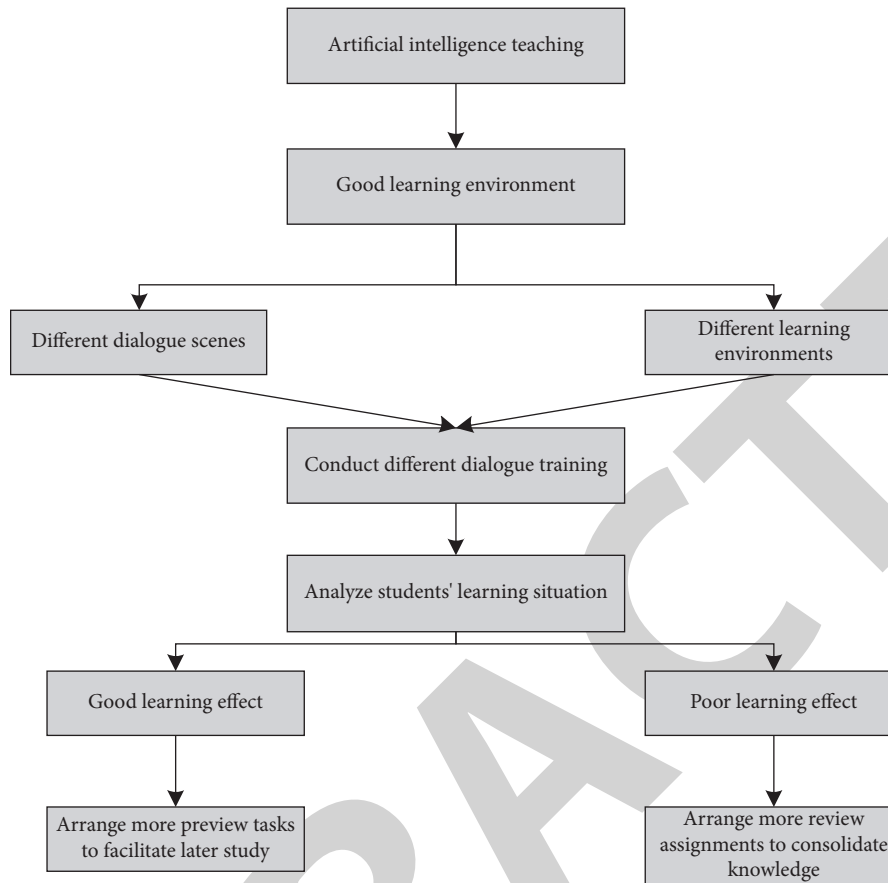


FIGURE 1: Teaching steps of learning oral English by artificial intelligence.

2. Artificial Intelligence and Spoken English in the Corpus

2.1. Artificial Intelligence Robots Carry Out Oral English Teaching Activities. As the most important part of learning English, oral English is an opportunity for students to use communicative tools skillfully. The level of oral expression will not only affect the learning effect cross-language communication, but also objectively reflect the development of students' ability to use English. In the traditional college English classroom teaching mode, teachers neglect the creation of context, students lack the opportunity to communicate, and students are ashamed to express their ambiguity. Using a language and a good environment, a robot starts a dialogue with students and creates a permanent language practice environment for them and has in-depth communication with students. At the same time, according to different dialogue scenes and different learning conditions, we should carry out different dialogue scenes, guide dialogue exercises, and analyze students' oral English communication and some hints on pronunciation and vocabulary to ensure the correct speech expression. After completing the dialogue exercise, the robot can simulate the teacher's role, summarize and evaluate the students' oral practice, and give targeted suggestions to the students. In teaching, the students' oral English level declines due to lack of practice. As an assistant in the communication activities between teachers and students,

teachers can understand students' oral ability more comprehensively and lay a good foundation for the future study. Set up practice tasks as the object of students' oral communication and provide students with various sentence patterns of oral exercises and suggestions for sentence adaptation. It is shown in Figure 1.

2.2. Using the Artificial Intelligence Corpus to Innovate Teaching. English teaching is the focus of teaching English and the beginning of students' learning English. In the era of rapid rise of artificial intelligence, various language platforms provide more targeted help for learners. First of all, in the teaching process, teachers will help students integrate educational resources. When students have a large amount of data, artificial intelligence can change many behaviors about learning. In addition, teachers can also sum up common words according to the vocabulary habits of each student in the system and add corresponding explanations in sentences to help students better understand the meaning of words and reduce their workload. Secondly, teachers will use intelligent technology to provide students with a better learning environment. Students are free to choose different passages and play English, which makes oral English more vivid and profound. After class, students can communicate in English. Students in daily life can scan English with their mobile phones.

2.3. *Classification of Corpus.* There are many types of corpus, and the main basis for determining the types is its research purpose and use, which can often be reflected in the principles and methods of corpus collection. Some classify corpus into four types: (1) heterogeneous: there is no specific principle for collecting corpus, and different corpus is collected and stored as it is; (2) homogeneity: only corpus with the same content type is collected; (3) systematicness: combining corpus according to predetermined principles and proportions, so that corpus is balanced and systematic and can present language facts in a given field; (4) specific: collect corpora for specific purposes only. It is shown in Figure 2.

3. Corpus-Based Spoken English Algorithms

3.1. *Error Determination Threshold.* Assuming that the difference between the score of the i th phoneme of the 40 phonemes of the selected TIMIT corpus language in sentence j of the TIMIT corpus language and the average pronunciation level \bar{p}_i of the phoneme is d_{ij} , the average difference between the phoneme score and the intermediate phoneme pronunciation in sentence j is

$$\bar{d}_{ij} = \frac{1}{c} \sum_{k=1}^c |p_{ij} - \bar{p}_i|, \quad (1)$$

where C is the ranking number of phoneme i in the j th sentence, thus obtaining the average difference D_i between phoneme i and the standard average level. The conclusion is that students make mistakes in pronunciation, so we can get

$$D_i = \frac{1}{M} \sum_{j=1}^M \bar{d}_{ij}, \quad (2)$$

where M is the number of TIMIT voice texts selected in actual application or experiment of the system. Formula (2) indicates that each phoneme has a corresponding threshold, which will change with the number of standard sounds provided by the scoring system. For this, we can calculate the error threshold to observe whether the pronunciation is standard or not.

$$th_i = D_i, \quad (3)$$

$$t\hat{h} = \frac{1}{k} \sum_{i=1}^k D_i. \quad (4)$$

Equation (4) represents the default average overall threshold, which will be used in the following experiments, where k represents the number of phonemes.

3.2. *Phoneme Level Errors.* When detecting and judging phoneme level errors, we should first compare the difference between the learner's phoneme score p_i and the corresponding TIMIT standard speech phoneme score. Feedback judgment is given, and the formula is as follows:

$$d_i = |SP_i - ST_i|, \\ P = \begin{cases} \text{encourage,} & d_i \leq 50000, \\ \text{no suggestion,} & 50000 < d_i \leq th_i, \\ \text{suggestion,} & d_i > th_i. \end{cases} \quad (5)$$

3.3. *Continuous Evaluation.* Evaluate according to the pronunciation of spoken English. Let N_r represent the number of parts of the training set that are labeled as linked according to the uninterrupted rule and belong to different linked groups. N_m represents the number of speakers in the training set who belong to different continuous reading lengths after being labeled artificially. R stands for the linking pronunciation rate and is defined as

$$r = \frac{N_m}{N_r}. \quad (6)$$

In order to improve the reliability of the evaluation, some measures are introduced: correct recognition rate c , wrong recognition rate e , and missing rate γ , which are defined as follows:

$$c = \frac{N_{\text{right}}}{(N_{\text{right}} + N_{\text{error}})}, \\ e = \frac{N_{\text{error}}}{(N_{\text{right}} + N_{\text{error}})}, \quad (7) \\ \gamma = 1 - \left(\frac{N_{\text{missed}}}{(N_{\text{right}} + N_{\text{error}} + N_{\text{missed}})} \right).$$

The meaning of letters in the formula: N_{right} is the number of links marked by both manual and automatic linking systems in the training corpus; N_{error} is the number of training corpus labeled as linking by automatic linking labeling system but not manually labeled as linking; N_{missed} is manually labeled as linked, but the automatic linked labeling system does not recognize the number of linked. Through these three formulas, we can also calculate the measures in the corpus.

According to the hypothesis of linking category dependence, we use the same training corpus to count the three measures mentioned above.

When testing and evaluating spoken English, the practical significance of ambiguity measure is that pronunciation examples belong to or are better than different scoring levels (excellent or good), and the formula is as follows:

Fuzzy measure of "belonging to or better than good."

$$\mu(A) = \begin{cases} 0.0, & |A| = 0, \\ \left[1 + \left(\frac{|A| - \alpha}{\beta} \right)^{-2} \right]^{-1}, & 1 \leq |A| \leq L. \end{cases} \quad (8)$$

Fuzzy measure of "degree of superiority."

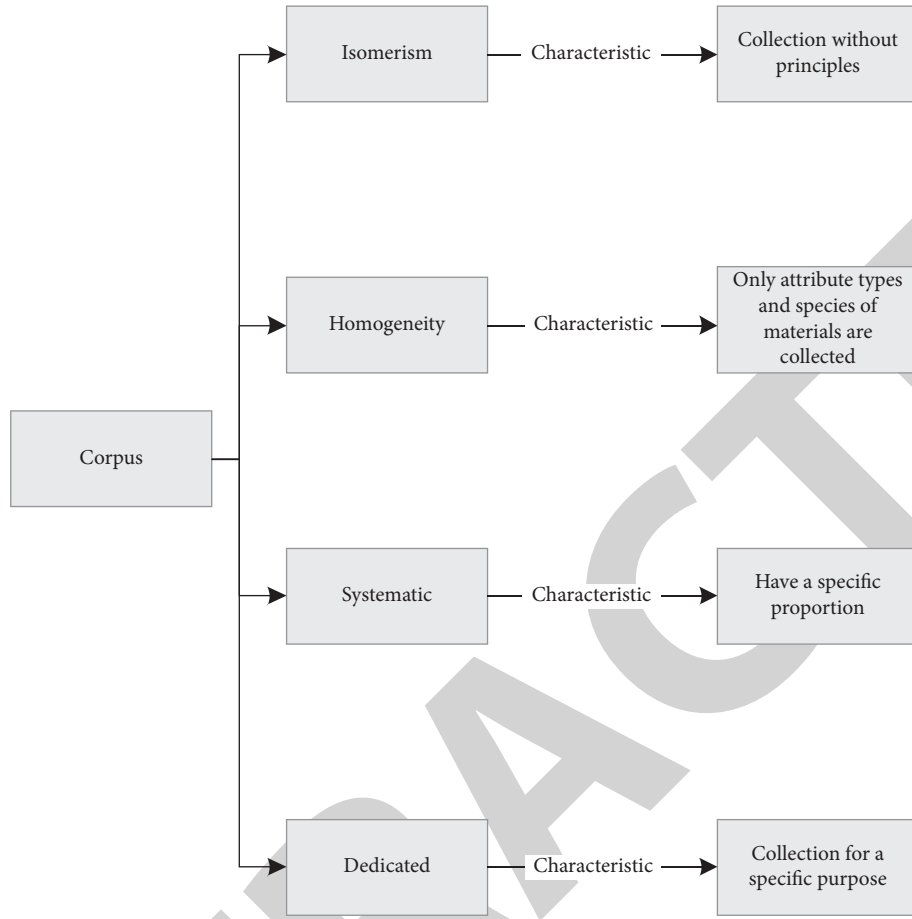


FIGURE 2: Corpus classification and characteristics.

$$\mu(A) = \begin{cases} e^{-((|A|-c)^2/2\lambda^2)}, & 1 \leq |A| \leq c, \\ 1.0, & c \leq |A| \leq L \text{ or } L < |A|. \end{cases} \quad (9)$$

For continuous evaluation, we also need to pay attention to two independent situations:

- (1) When a link is recognized by the system, the reliability of the link is defined as

$$f(x) = c[\text{Cat}(x)]. \quad (10)$$

- (2) When a link is missed by the system, the probability that the link is actually pronounced by the practitioner depends on $\sim\gamma$. Therefore, in this case, its reliability is defined as

$$f(x) = \sim\gamma[\text{Cat}(x)] = 1 - \gamma[\text{Cat}(x)]. \quad (11)$$

3.4. Corpus Analysis. Corpus is a multilanguage database, which contains a variety of information. When we need to find this information accurately in order to obtain the required data, in this paper, we study the algorithm of finding data in corpus.

Assuming that A and B are randomly assigned words in the corpus, the total oscillation amplitude of the corpus is W , S is the period, and the actual observation frequencies are $F(A)$ and $F(B)$; the mutual information value is calculated as follows:

$$I(a, b) = \log_2 \frac{P(a, b)}{P(a) \cdot P(b) \cdot 2S} = \log_2 \frac{F(a, b) \cdot W}{F(a) \cdot F(b) \cdot 2S}. \quad (12)$$

$F(A)$ is the observation frequency of word-controlled structure, $F(B)$ is the observation frequency of word-controlled structure, and $F(A, B)$ is the speech dual frequency of two parts, and its value can be calculated in the following formula:

$$I(a, b) = \log_2 \frac{W \cdot F(a, b)}{F(a) \cdot F(b)}. \quad (13)$$

If the total size of the corpus is W and the frequency of observation of a certain collocation word in the corpus is $C1$, the average frequency of occurrence of the collocation word in each word position is calculated as $C1/W$. 1 is the lexical position occupied by node words, but this paper can calculate for lexical chunks and similar sentence patterns, so the lexical position may not be 1 in the design. However, when considering the co-occurrence probability of a node word with the observation frequency of N , the probability P is calculated as follows:

TABLE 1: The influence of CAF triplet on English learning.

CAF triplet	First-class index	Secondary index
Complicated vocabulary and sentence patterns	Complicated vocabulary Syntactic complexity	Form-symbol ratio, vocabulary frequency change rate Sentence length, subordinate sentence ratio
Oral accuracy	Error ratio Speed Pause	Pronunciation error ratio, word error ratio Speech speed, pronunciation speed, and phonation time ratio Pause frequency, pause duration
Oral fluency	Repetition Self-modification Hesitate	Repeated superposition, paragraph chunk Modification times, modification bands Frequency and length of hesitation

$$P = \frac{C_1 \cdot (2S + 1)}{W} \cdot \frac{N}{W}. \quad (14)$$

The expected frequency of co-occurrence of collocation words is as follows:

$$SD = \sqrt{(2S + 1) \cdot N \cdot \left(1 - \frac{C_1}{W}\right) \cdot \frac{C_1}{W}}. \quad (15)$$

The standard deviation of collocation distribution is as follows:

$$E = \frac{C_1 \cdot (2S + 1) \cdot N}{W}. \quad (16)$$

The Z value is as follows:

$$Z = \frac{C_2 - E}{SD}. \quad (17)$$

4. Research on Oral English Optimization and Evaluation

4.1. The Influence of CAF on Oral English. At present, CAF is an important standard for testing foreign language output. When developing CAF, we have grasped the test level and error analysis. CAF is an important and interesting research topic in oral English output and English language development and has many influences on oral English and English learning. According to our investigation, the influence of CAF on English learning is shown in Table 1.

The effects of CAF on English learning and oral English are as follows: complexity and accuracy (writing, dictation, reading, shorthand, etc.); complexity and fluency (speaking, reading aloud, pause, dictation, etc.); and accuracy and fluency (speaking, interpreting, reading, visual translation, etc.). Its influence on it is shown in the table. From the first index, it has influence on lexical complexity, syntactic complexity, error ratio, speed, pause, and repetition of oral English. From the second index, each first index corresponds to the corresponding second index, which is also the key to the influence of CAF on oral English. In this regard, we have made a survey on the influence of CAF on oral English. By comparing the oral English level before applying CAF optimization with the scores before and after using CAF (out of one point in all aspects), we can get the comparison as shown in Figure 3.

Figure 3 introduces the comparison of scores in various aspects of oral English before and after using CAF. Whether

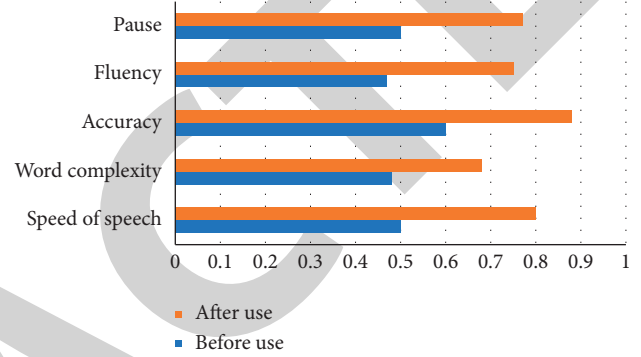


FIGURE 3: Comparison of scores in various aspects of oral English before and after using CAF.

CAF is helpful to oral English learning needs experimental comparison. We mainly compare five aspects: speed, word complexity, accuracy, fluency, and pause. The scores in five aspects are higher after using CAF than before using CAF. I have a better understanding of oral English in the next study of this paper.

From the figure, we can see that, after using CAF optimization, the oral English score has obviously increased a lot compared with that before using CAF, and the accuracy has increased to 0.88, which is the highest score. The score before using CAF is also the highest before using all the data, and the lowest score after using CAF optimization is also 0.68 of the word accuracy. From the figure, we can see that CAF optimization is positive for the adjustment of oral English, and it is of great help to oral English.

4.2. Research on the Influence of Artificial Intelligence on Oral English Learning Evaluation. The general trend is to introduce artificial intelligence into English teaching. Artificial intelligence is actually a science that simulates human intelligence by computer. With the rapid development of the information age, the development of computer technology also affects the progress of society. The use of artificial intelligence is actually what they are most interested in. Intellectual property rights have changed people's production and living habits and played an irreplaceable role in education and teaching. There is no doubt about the importance of language as a basic tool for human communication. Artificial intelligence technology, speech processing, machine translation, and speech recognition are inseparable

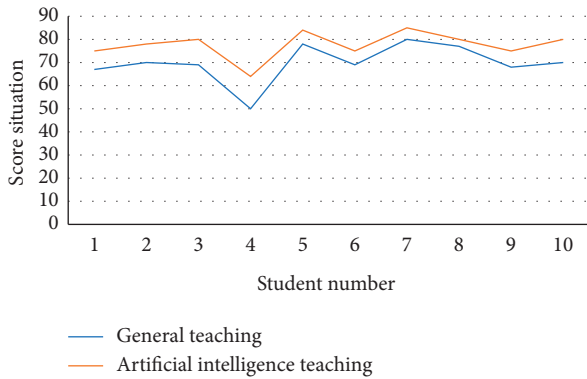


FIGURE 4: Comparison of students' scores in two teaching methods.

from human language learning. Therefore, it is meaningful to use artificial intelligence technology to learn and develop human language. English is the second language in China, and lack of English communication ability is a difficult problem in personnel training in every school stage, and the influence of English teaching in primary schools is ignored. It is an academic topic discussed by experts and scholars, and it is also a difficult problem for teachers, parents, and students. To sum up, the practical application of artificial intelligence technology should be increased in intermediate English teaching, and the revolutionary power provided by artificial intelligence technology will affect the practice of intermediate English teaching. In oral English learning, artificial intelligence products have rich language resources and many functions such as reading samples, student guidance, student repetition, reading aloud, and pronunciation correction. The purpose of developing artificial intelligence products is to enable students to master correct pronunciation in a short time, so as to achieve a higher English level. The software can classify and arrange the existing oral resources, set different levels, and then start testing students and provide corresponding learning resources according to the students' oral test results. In this regard, we put a group of students in different learning environments to study for a period of time. The teaching content and teaching time are the same, but the teaching methods are different. After studying for a period of time, we compare the learning situation and the score distribution as shown in the figure.

Figure 4 shows the results of ten students in two learning environments. In a normal learning environment, their academic performance is not as good as that in the AI learning environment. The highest score for AI is 85. The lowest score is only 80 points in the normal learning environment, and the lowest score is only 50 points. At present, the oral English test mainly includes self-introduction, reading articles, situational questions, listening to essays, and answering questions. It can be seen that the main problems we encounter when implementing smart assessment are speech recognition and content understanding. In order to be evaluated automatically, it must be completed from two angles: oral presentation and reviewer's content. When evaluating expressions, we should first capture the paragraphs of the speech tester and then analyze the

captured contents to obtain features and comprehensively evaluate them from the aspects of sound quality, color, and pitch and finally summarize a reasonable single result to get the result. For example, the content of self-statement must be verified, and the scoring principle is more complex. It is necessary for the scoring system to use natural language processing, collect multiple sets of data for averaging processing, and obtain the comparison of the scoring accuracy between artificial intelligence evaluation and ordinary evaluation in timbre, tone, sound quality, and self-introduction as shown in Figure 5.

Figure 5 is mainly through the comparison of two evaluation methods. In the evaluation and comparison, we should pay attention to accuracy. In the general evaluation, human emotional factors account for a larger impact. At this time, we need to use the evaluation method of artificial intelligence to make a fair evaluation. The evaluation of artificial intelligence is based on the set scoring standard, which is more scientific. Through Figure 3, we can also get that the evaluation accuracy of artificial intelligence is higher than that of ordinary evaluation methods. This paper can better study the evaluation of oral English.

According to Figure 5, we know that the accuracy of artificial intelligence evaluation is higher than that of ordinary evaluation in the most important aspects of evaluation. The accuracy of ordinary evaluation in timbre is 0.6, and that of artificial intelligence evaluation is 0.86; the general evaluation of sound quality is 0.67, and the artificial intelligence is 0.89; the general evaluation of tone is 0.78, and the artificial intelligence evaluation is 0.95; in self-introduction, the accuracy of general evaluation score is 0.56, and artificial intelligence is 0.85. From the above data, it can be shown that the artificial intelligence evaluation method is better than the ordinary evaluation in oral English.

4.3. A Corpus-Based Study of Oral English. There are a lot of data in the corpus, and spoken English is a very common language, which needs a lot of vocabulary and sentence patterns. At this time, the knowledge support in corpus is very needed, but it is inconvenient to find too much data. Therefore, according to the word list, Nation (2001) divides words into four groups. The first two groups are 2,000 high-frequency words. The third group is academic vocabulary, which is mainly used in a written form. The fourth group is low-frequency words commonly found in written and spoken styles. Table 2 is obtained by investigating the vocabulary in the corpus.

The figures in Table 2 show that the first two types of high-frequency words account for the majority of Chinese students, among which the coverage rate of the first type is 82.14%, and the minimum incidence rate of academic words is 4.89%. With the investigation, it is found that the distribution of high-frequency words in different corpora is as follows.

Table 3 shows that there are many high-frequency words distributed in the corpus in different environments. In order to find the required information and words quickly, we conducted a survey on the vocabulary used by various

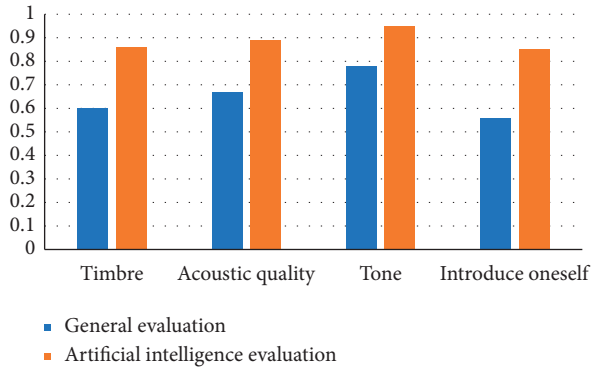


FIGURE 5: Comparison of scoring accuracy between two evaluation methods.

TABLE 2: Distribution of high-frequency words in HVC.

Vocabulary	Quantity
The first category	19,926
Category II	1543
Academic category	1186
Low frequency use	1603
All	24258

TABLE 3: Distribution of high-frequency words in different environments.

Vocabulary	HVC (%)	Local spoken English (%)	Local English writing (%)
High-frequency words	88.5	85	80
Academic category	4.89	5	10
Low-frequency words	6.61	10	10

groups of people in oral English. In the HVC corpus, high-frequency words account for 88.5%, academic words account for 4.89%, and low-frequency words account for 6.61%. In the local spoken language, high-frequency words account for 85%, academic words account for 5%, and low-frequency words account for 10%. In spoken English, locals use lower high-frequency words than Chinese people, and academic and low-frequency words are higher than Chinese people. Locals are more familiar with oral English.

The distribution of high-frequency words in the corpus is shown in Table 3. In foreign environments, the proportion of high-frequency words in spoken language is lower than that in China, but higher in the other two items, which also shows that local people are relatively more accustomed to using low-frequency and academic words. We have also investigated the benefits of corpus for oral English, such as providing words for oral English better and making oral English more fluent. The public's percentage of the benefits of corpus is as follows:

According to Figure 6, it can be seen that, in the eyes of the public, the benefits of the corpus for oral English mainly

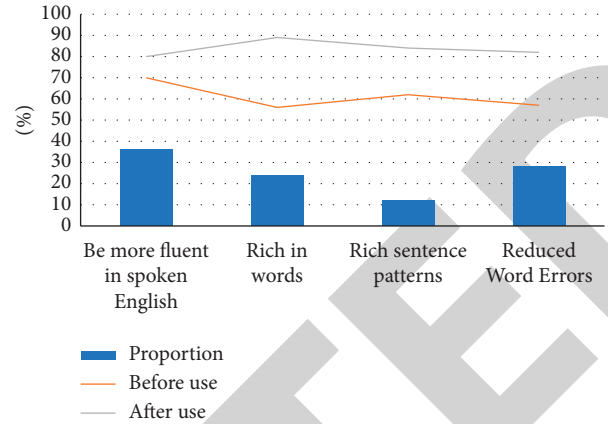


FIGURE 6: Comparison of the benefits of corpus to English compared with the situation after using corpus.

includes the following four aspects: more fluent oral English accounts for 36% at most, rich words account for 24%, rich sentence patterns account for 12%, and reduced word errors account for 28%. The comparison between before and after the use of the corpus also shows that the oral English scores have risen greatly after the use of the corpus.

4.4. Influence of External and Self on Oral English. Besides oral English, external influences are also important. In education, students have less time to practice speaking English. Most students have no way to communicate with foreigners, and of course, there is no environment to communicate in English. The only English communication takes place is in class, and there is no atmosphere conducive to communication after returning home. In today's quality education, testing is still the main means to measure and evaluate students. To pass the exam, you need to master grammar, vocabulary, reading, and writing skills, but students miss the opportunity of oral communication training in class. Many teachers insist that mastering grammar, vocabulary, and sentence structure are the only way to get good test scores, but ignore listening and speaking. Teachers have different language abilities. Senior high school exams have relatively low requirements for senior high school students' oral English, which leads to the decline of many middle school teachers' oral English ability. Daily oral communication mainly takes place in the classroom, which is limited to the classroom, that is to say, teachers' oral games also have certain limitations. Finally, teachers' oral level affects the development of students' oral ability. Students themselves have many emotional changes when learning oral English because oral English is not a rigid study but to communicate with others, and personality has a great influence on it. We conducted a survey on people with poor oral English and found out their differences in personality and their influence on oral English, as shown in Figure 7.

Figure 7 shows several states that affect spoken English, such as shyness, indifference, conformity, and inferiority complex. For the study of oral English, these psychology will affect its exertion. At present, the most serious influence on

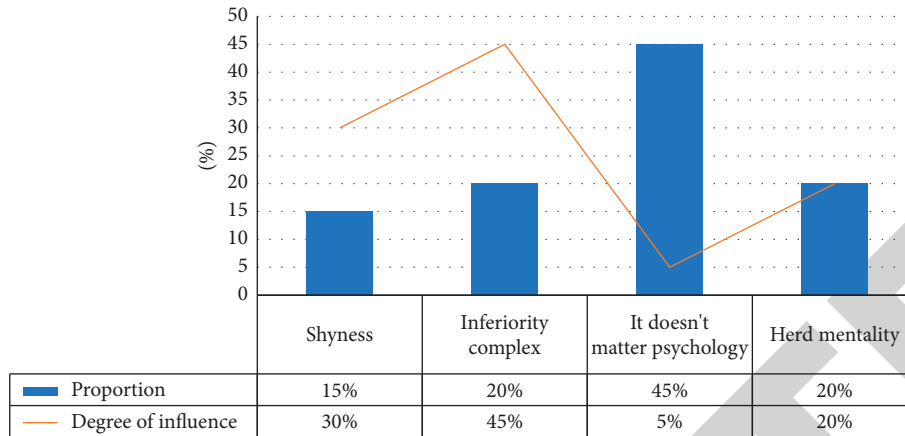


FIGURE 7: Psychological proportion and influence degree of losers in oral English learning.

spoken English is inferiority complex. For inferiority complex, their self-confidence is frustrated and they dare not to speak and communicate. For oral English, the most important thing is to use it frequently, as practice makes a person perfect. Inferiority complex cannot be changed for a while and needs long-term treatment, so it has the greatest impact on oral English.

For psychological differences as shown in Figure 7, the unsuccessful psychology of learning oral English mainly includes the following four aspects: shyness exists in many beginners, such as shy girls and introverted boys. These colleagues are nervous, stuttering, quiet, and vague when communicating. They are always afraid of laughing at themselves when they say the wrong thing. Sometimes they blush and bow their heads, especially some students from rural families. Accounting for 15%, the reason for inferiority complex is weak cognitive ability. Some students have this inferiority complex even if they fail for the first time, while other students stutter. Ambiguous voice, clumsy pronunciation, and answers full of sick sentences have long been unmatched by others. These people account for 20%, and their psychology does not matter. Students with this psychology often think that their grades are not applicable to themselves, and whether they can say it does not matter. It is enough to know words, write and spell, and whether they can say it does not matter. Such people account for 45% at most. Herd mentality: this kind of psychology thinks that many people cannot speak English, everyone does not learn it, and they do not have to learn it themselves, which accounts for 20%.

4.5. Comparison of Comprehensive Oral English Fluency. We have studied above the influence of CAF optimization, artificial intelligence, and corpus on oral English, all of which have positive effects on oral English. We have conducted below mixed experiments to observe the comparison of these items on oral English fluency and the comparison when they are mixed.

According to Figure 8, we can see that the fluency scores of oral English are different under various methods. The

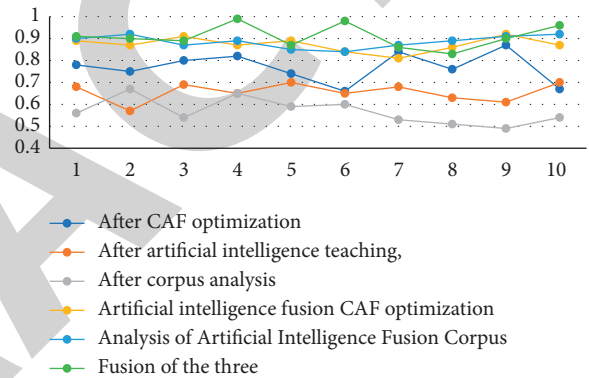


FIGURE 8: Scoring of oral English fluency by various methods.

single method is far worse than the hybrid method. When using corpus analysis and artificial intelligence alone, the average fluency score is only about 0.65, and when using the combination of the three, the average fluency score reaches 0.94.

5. Conclusion

This paper makes an in-depth analysis of the optimization and evaluation of oral English CAF based on artificial intelligence and corpus. As the society moves towards internationalization, oral English becomes more and more important. It is necessary to study oral English by combining artificial intelligence, corpus, and CAF, but their roles in oral English are not exactly the same. This paper introduces this, and the oral English combined by several is fully explained. This paper introduces the algorithm of evaluating and determining accuracy in spoken English and introduces its definition, so that readers can fully understand the importance of spoken English and the research in artificial intelligence corpus.

Data Availability

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

Conflicts of Interest

The author declares that they have no conflicts of interest regarding this work.

References

- [1] N. J. Nilsson, "Artificial intelligence: a modern approach," *Applied Mechanics and Materials*, vol. 263, no. 5, pp. 2829–2833, 2013.
- [2] A. Robinson, J. K. Eaton, and L. Eaton, "Handbook of logic in artificial intelligence and logic programming," *Oxford University Press*, vol. 90, no. 1, pp. 90–100, 2010.
- [3] P. Norvig and S. Russell, "Artificial intelligence: a modern approach (all inclusive), 3/E," *Applied Mechanics and Materials*, vol. 263, no. 5, pp. 2829–2833, 2021.
- [4] M. Negnevitsky, "Artificial intelligence: a guide to intelligent systems," *Information & Computing Sciences*, vol. 48, no. 48, pp. 284–300, 2015.
- [5] J. J. Jiang and D. W. Conrath, "Semantic similarity based on corpus statistics and lexical taxonomy," *Rocling*, vol. 12, Article ID 11512, 2017.
- [6] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *Proceedings of the International Conference on Language Resources & Evaluation*, New York, NY, USA, May 2010.
- [7] Z. Ying, X. Shao, and S. Pan, "An approach on oral English practice in the first five minutes of each class of the independent colleges in China," *Value Engineering*, vol. 115, no. 47, pp. 23501–23511, 2011.
- [8] M. A. Lu and W. Ding, "A different interpretation of "affect" in the teaching of oral English," *Foreign Language World*, vol. 90, no. 1, pp. 10–20, 2015.
- [9] Md. Hamzah, Omar, and S. Paramasivan, "Between the ideal and reality: teachers' perception of the implementation of school-based oral," *English Assessment*, vol. 45, no. 1, pp. 34–90, 2019.
- [10] R. Chu, "Effects of teacher's corrective feedback on accuracy in the oral English of English-majors college students," *Theory and Practice in Language Studies*, vol. 1, no. 5, pp. 20–21, 2021.
- [11] H. A. Liu and W. Yu, "Communicative competence and oral English teaching," *Foreign Languages And Their Teaching*, vol. 90, no. 1, pp. 12–19, 2018.
- [12] A. Zhou, "Fluency and accuracy improvement in oral English," *Foreign Language Education*, vol. 45, no. 3, pp. 12–14, 2012.
- [13] W. J. Zhou, "Effects of input modes on oral English production," *Journal of PLA University of Foreign Languages*, vol. 90, no. 2, pp. 34–35, 2015.
- [14] N. Xiao, "Application of formative assessment to teaching of oral English class," *Journal of Yunnan Agricultural University*, vol. 34, no. 2, pp. 455–456, 2010.
- [15] Y. Chen and Q. Q. Zhao, "A relative study of chunk acquisition and students' oral English," *Shandong Foreign Language Teaching Journal*, vol. 78, no. 1, pp. 13–15, 2014.