Computer Fraud and Security
ISSN (online): 1873-7056

Construction of English Oral Information Security System Based on Reinforcement Learning Algorithm

Xiaodan Chang*
Zhengzhou University of Economics and Business, Zhengzhou, Henan, China
**Corresponding author*

Abstract:

English oral teaching is one of the important ways to improve English oral ability. Traditional English oral teaching often overlooks the needs of learners for oral learning, adopts a single and unified teaching method, emphasizes theory over practice, and cannot effectively improve learners' oral ability. This article constructs an English oral teaching system based on reinforcement learning, and combines deep learning to construct an oral evaluation module and improve the reward mechanism on the basis of recognizing English oral communication. The experimental results show that the system model proposed in this paper has better oral recognition and evaluation performance, is more in line with practical situations, has higher accuracy, and stronger stability. It can effectively adjust reward situations and stimulate learners' learning motivation based on their status. The application experiment results show that the system in this article can significantly improve the oral performance of learners within the same teaching time, and provide feedback to learners on the existing oral problems from multiple perspectives. Most learners indicate that the system can propose targeted learning strategies, enhance their enthusiasm for oral learning through reward mechanisms, and improve learning efficiency and quality.

Keywords: reinforcement learning; English speaking; English teaching; Effect evaluation; Information Security System

1. INTRODUCTION

With the acceleration of globalization and the rapid development of technology, English, as one of the universal languages for international communication, is an important communication bridge connecting different countries [1]. Mastering English speaking skills is crucial for personal career development, international communication, and cultural understanding. Especially with the increasing trend of multinational corporations and international cooperation, having good English speaking skills can not only help individuals better understand the thinking patterns, values, and behavioral habits of different cultural backgrounds, promote mutual understanding and friendly cooperation among nations, but also enhance their professional competitiveness [2]. In daily life, good English oral teaching can also enhance personal cultural understanding and cross-cultural communication skills, help people deeply understand the culture, history, and social customs of English speaking countries, enhance their cultural literacy and cross-cultural communication skills, and better adapt to the globalized social environment [3]. With the acceleration of teaching reform, many problems in traditional English oral teaching have gradually been exposed, as shown in Figure 1, which is the result of a survey questionnaire on the current situation of traditional English oral teaching models.

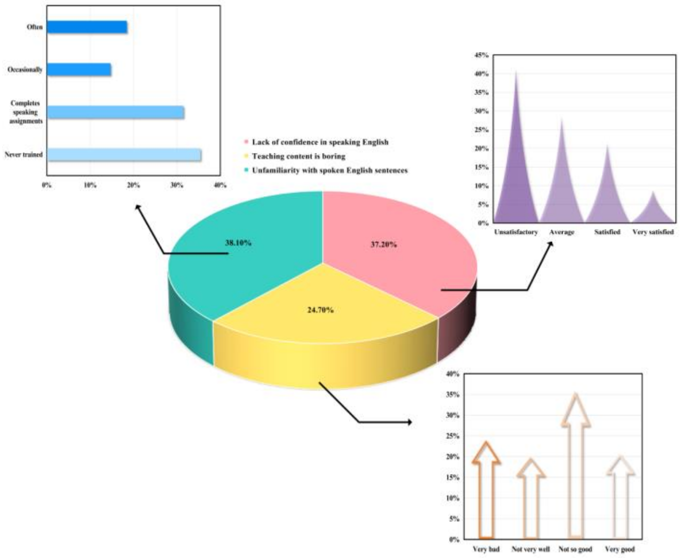


Fig.1 Survey questionnaire results on the current situation of traditional English oral teaching mode

According to the results in the figure, it can be found that there are four main problems in traditional English oral teaching. In terms of teaching content, traditional English oral teaching often focuses on explaining grammar and vocabulary, while neglecting the importance of oral practice. By explaining grammar rules and vocabulary usage for a large amount of time, learners lack the opportunity to actually use the language, making it difficult for them to truly master oral skills and effectively improve their oral expression ability. In terms of teaching methods, traditional English oral teaching lacks sufficient interaction and participation. Teachers are usually the leaders in the classroom, while students are in a passive state of receiving knowledge. This one-way teaching method leads to a lack of opportunities for students to actively participate in classroom discussions and interactions [4]. In terms of teaching environment, traditional English oral teaching often lacks authentic contextual simulation. Learners usually only engage in simple dialogue exercises in the classroom, with little opportunity to apply their learned knowledge in real-life contexts [5]. This teaching method, which lacks practical context, makes it difficult for students to adapt to the real language environment and effectively improve their oral fluency and accuracy. In terms of teaching mode, traditional English oral teaching often overlooks individual differences among students. There are differences in the oral proficiency and learning ability of different learners, but traditional teaching methods often adopt a one size fits all approach, which cannot meet the needs of different learners, thereby increasing their stress and sense of impact [6].

Artificial intelligence technology can effectively solve the problems in traditional English oral teaching mentioned above. Reinforcement learning, as an advanced machine learning method, is based on a trial and error mechanism and aims to maximize expected returns by continuously trying and adjusting strategies [7]. In English oral teaching, reinforcement learning can simulate real language environments, provide personalized learning experiences for learners, and effectively improve their oral abilities. By analyzing the learning data and behavioral characteristics of learners, reinforcement learning algorithms can automatically adjust teaching content and difficulty to meet the needs of different learners [8]. During the learning process, reinforcement learning algorithms can also provide corresponding rewards or punishments based on the performance of learners, thereby helping them identify their strengths and weaknesses and adjust learning strategies [9]. Therefore, this article introduces reinforcement learning algorithms to construct a spoken English teaching system, and combines other algorithms to improve the performance of the teaching system.

The main innovative points of this article are as follows:

Firstly, an English oral teaching system based on reinforcement learning algorithms can provide learners with oral practice scenarios that are close to real environments.

Secondly, this article combines deep learning and reinforcement learning algorithms to extract English oral feature data from learners, and improves the quality of English oral evaluation by adding first language oral phonemes, providing effective suggestions for improving learners' oral problems.

Finally, this article adds a reward mechanism to the English oral teaching system based on reinforcement learning algorithms and verifies through comparative experiments that the reward mechanism has a positive effect on learners' English oral learning.

2. RELATED THEORETICAL FOUNDATIONS

The core idea of reinforcement learning is to enable agents to learn how to take the best action to maximize rewards through continuous trial and error during their interaction with the environment. Its core concepts include agents, environment, states, actions, and rewards [10]. An intelligent agent is a learning agent that perceives states through interaction with the environment and selects and executes actions based on the current state. The environment is the external world in which the intelligent agent is located, and it will respond to the agent's actions and provide corresponding reward signals. Rewards are the core driving force of reinforcement learning, used to evaluate the quality of actions taken by intelligent agents. The basic models of reinforcement learning mainly include Markov Decision Process (MDP). MDP is a framework used to describe reinforcement learning environments, based on the assumption that the future state of the environment depends only on the current state and actions, and is not directly related to previous states and actions [11]. This assumption simplifies the complexity of the problem and makes the solving process more convenient. The value function is a key concept in reinforcement learning, which includes the following two parts:

The Markov process includes the environmental state, the agent's execution of actions, the probability of the state transitioning to another state after executing actions, and the reward value. In order, its description is represented as (S, A, P, R). When the intelligent agent is in n step, its state value function is shown in formula (1):

$$V_{\pi}(s) = \mathrm{E}\left[H_n \middle| S_n = s\right] \quad (1)$$

Among them, the strategy is denoted as π, the current state is represented as s, and the reward return value is represented as H_n. Reinforcement learning requires the best strategy to reach the peak expected return value.

The action value function, also known as the Q function, is shown in formula (2):

$$q_{\pi}(s,d) = \mathrm{E}\left[H_n \middle| S_n = s, D_t = d\right] \quad (2)$$

Among them, the execution action is represented as D, and the current execution action is recorded as d. The value of the Q function is recorded as the expected return on executing an action in the current state and its maximum value is obtained.

According to the Bellman equation, the decomposition of the state value function and action value function can be obtained, as shown in formulas (3) and (4):

$$V_{\pi}(s) = \sum_{d} \pi(d|s) \sum_{h'} \sum_{r} p(s', r|s, d)[r + \alpha v_{\pi}(s')] \quad (3)$$

$$q_{\pi}(s,d) = \sum_{s', r} p(s', r|s, d)[r + \alpha v_{\pi}(s')] \quad (4)$$

Among them, α represents the discount factor.

Through temporal differential learning, policy updates can be achieved after the completion of actions, as shown in (5) and (6):

$$V(S_t) \leftarrow V(S_t) + \beta\left[R_{t+1} + \alpha V(S_{t+1}) - V(S_t)\right] \quad (5)$$

$$Q(s,d) \leftarrow Q(s,d) + \beta\left[r + \alpha Q(s',d') - Q(s,d)\right] \quad (6)$$

The teaching and learning of spoken English is a dynamic process, and curriculum learning methods can simulate this process by imitating the process of learners gradually deepening their understanding and application from basic knowledge [12]. Traditional machine learning methods often use random input data for training, without fully considering the hierarchical and ordered nature of knowledge [13]. However, when learning new knowledge, humans always start with simple concepts and gradually expand to more complex fields [14]. This learning approach, which goes from shallow to deep, helps us better grasp and apply knowledge. The schematic diagram of the course learning method process is shown in Figure 2.

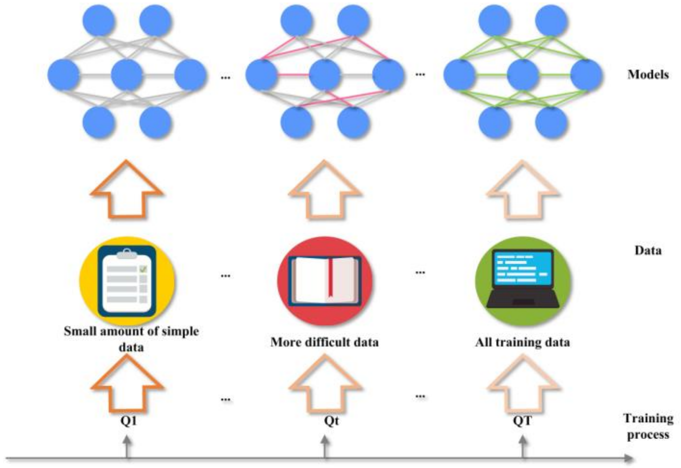


Fig.2 Schematic diagram of the process of course learning methods

Assuming that training can be decomposed into N steps, the corresponding standardized optimization sequence is represented as $C = (O_1, O_2, ..., O_N)$. For the target training distribution $P(z)$, each sequence is weighted accordingly, as shown in formula 7:

$$O_t(z) \propto W_t(z)P(z) \quad (7)$$

Among them, the training samples are denoted as z and $\forall z \in A$, and the training data is denoted as A.

The above formula also needs to meet three conditions, as shown in (8):

$$\begin{cases} G(O_t) < G(O_{t+1}) \\ W_t(z) < W_{t+1}(z) \\ O_t(z) = P(z) \end{cases} \quad (8)$$

3. DESIGN OF AN ENGLISH ORAL TEACHING SYSTEM BASED ON REINFORCEMENT LEARNING

3.1 Dialogue System Design

The dialogue environment in English oral teaching plays an important role in the oral expression ability, language application ability, learning interest and motivation, cross-cultural communication ability, and cooperation and communication ability of learners [15]. Previous dialogue models have focused more on training dialogue strategies, lacking traceability of dialogue states, and improving dialogue performance based on easily generated erroneous outputs [16]. Therefore, this article introduces the decoding of dialogue history in the dialogue system, that is, the generation basis of the current dialogue state is the previous round and historical dialogue state, combined with database retrieval information to generate system actions, and finally complete the system dialogue through natural language generation templates [17]. The schematic diagram of the model structure of an English oral teaching dialogue system based on reinforcement learning is shown in Figure 3.

Fig.3 Schematic diagram of the dialogue system model for English oral teaching based on reinforcement learning

In the dialogue tracing module of the system, the encoder is divided into three types: dialogue history, dialogue state, and implicit dialogue state. The generator is the dialogue state generator [18]. There is a strong correlation between the content before and after English oral conversations, so using a dialogue history encoder can classify content based on bidirectional GRU encoded vectors to understand the learner's intention in English oral practice. The calculation formula is shown in (9):

$$\begin{cases} \vec{b}_{t,l}^{ctx} = GRU_{fwd}(\vec{b}_{t,l-1}^{ctx}, \omega_{t,n}) \\ \overleftarrow{b}_{t,l}^{ctx} = GRU_{bwd}(\overleftarrow{b}_{t,l-1}^{ctx}, \omega_{t,n}) \\ y_t^l = sof\max(W^l \left[\vec{b}_{t,l}^{ctx}; \overleftarrow{b}_{t,l}^{ctx} \right] \end{cases} \qquad (9)$$

In the formula, the probability distribution of learner intentions is described as $y_t^l \in R^{|M_u|}$, the number of intentions is M_u, and the training parameter is denoted as W^l.

The dialogue state encoder can obtain the current state situation based on the previous round of spoken dialogue state, thereby determining the slot value sequence relationship. Through one-way GRU, it can be obtained as shown in (10):

$$b_{t,j}^{st} = GRU_1(b_{t,j-1}^{st}, e_{t,j}^{c} + e_{t,j}^{b}) \qquad (10)$$

Among them, $e_{t,j}^{b}$ is obtained after index encoding, and $e_{t,j}^{c}$ is obtained from the previous round of dialogue state encoding.

Generate system dialogue states through a dialogue state generator, as shown in (11):

$$\begin{cases} P_{t,g}^{gen} = soft\max(E \cdot (b_{t,j}^{dec})^{T}) \\ P_{t,g}^{ctx} = soft\max(b_{t,j}^{ctx} \cdot (b_{t,j}^{dec})^{T}) \\ P_{t,g}^{st} = soft\max(b_{t,j}^{st} \cdot (b_{t,j}^{dec})^{T}) \\ P_{t,g}^{final} = P_{t,g}^{ptr_1} \times P_{t,g}^{gen} + P_{t,g}^{ptr_2} \times P_{t,g}^{ctx} + P_{t,g}^{ptr_3} \times P_{t,g}^{st} \\ P_{t,g}^{ptr_1} = soft\max(W_1 \cdot \left[b_{t,j}^{dec}; w_{t,g}; b_{t,j}^{ctx}; b_{t,j}^{st} \right] \end{cases} \quad (11)$$

Among them, the probability output of GRU after being mapped by the trainable matrix is represented as $P_{t,g}^{gen}$, and the two attention points are represented as $P_{t,g}^{ctx}$ and $P_{t,g}^{st}$ respectively. The final decoding probability is denoted as $P_{t,g}^{final}$, the trainable matrix of $\left[P_{t,g}^{ptr_1}, P_{t,g}^{ptr_2}, P_{t,g}^{ptr_3} \right] = P_{t,g}^{ptr} \in R^3$ is denoted as W, and the input of the decoding model is represented as $w_{t,g}$.

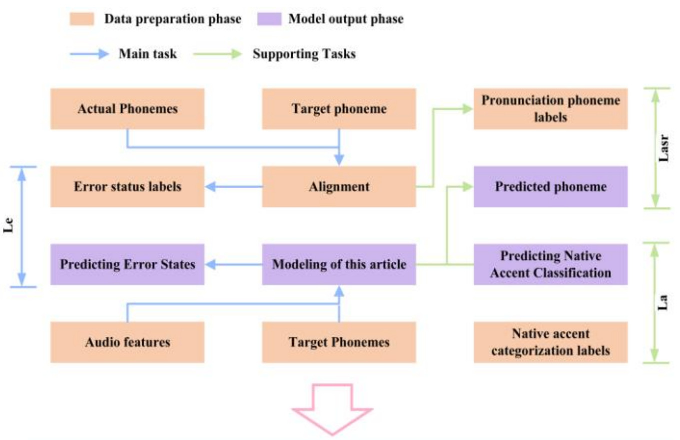
Due to the temporal nature of spoken English conversations, all dialogue history is maintained by using an implicit dialogue state encoder, as shown in (12):

$$\begin{cases} b_t^{state} = GRU_3(b_{t-1}^{state}, v_t) \\ u_t^s = \left[d_{t-1}^s; m_t; b_t^{state}; b_t^{kd} \right] \\ \pi^s(d_t^s, slot_t^s \middle| u_t^s) = sotf\max(MLP[u_t^s]) \end{cases} \quad (12)$$

Among them, after encoding the dialogue, m_t is obtained, and the corresponding probability distribution of the dialogue strategy in the action and strategy spaces is represented as $\pi^s(d_t^s, slot_t^s | u_t^s)$. The binary classification result corresponding to each slot is represented as $slot_t^s$.

3.2 English Speaking Evaluation Module Combined with Deep Learning

For English learners, English is not inferior to language, so they may be influenced by their first language during the learning process. The incorrect pronunciation that occurs during oral pronunciation may be partly based on the annotated pronunciation of the English target pronunciation, or partly based on the annotated pronunciation that cannot be recognized by English pronunciation. In order to improve the performance of reinforcement learning models in evaluating learners' oral performance, this paper introduces Transformer in deep learning [19]. Meanwhile, considering the situation where alignment cannot be optimized during the oral evaluation process, this model will obtain error status labels that align the actual English pronunciation with the target text pronunciation [20]. In addition, the model should add a task of classifying first language spoken phonemes, based on relevant features of first language spoken language, to assist in improving English speaking evaluation results. The English speaking evaluation process combined with Transformer and an example of evaluation alignment annotation are shown in Figure 4.



Target Text	You		know		how		I	thank				you	
Target phoneme	Y	UW	N	OW	HH	AW	AY	TH	AE	NG	K	Y	UW
Actual Pronunciation													
Original labeling	Y	UW	N	AO	HH	AW	AY	TH	AE	<DEL>	K(+S)	Y	UW
Expanded Phonemes	Y	UW	N	OW*	HH	AW	AY	TH	AE	<DEL>	K(+S)	Y	UW
Expanded phoneme	Y	UW	N	AO#	HH	AW	AY	TH	AE	<DEL>	K(+S)	Y	UW
Error status	0	0	0	1	0	0	0	0	0	1	1	0	0

Fig.4 English Speaking Recognition and Evaluation Process Combined with Transformer and Example of Evaluation Alignment Annotation

According to the process in Figure 4, it can be seen that the decoder input of the model is not the learner's actual pronunciation, but a prior target text factor, in order to obtain direct prediction results of actual pronunciation and incorrect labels. The high-level abstract audio features are denoted as G, the actual pronunciation is denoted as $\hat{p}$, and the labels and error labels are denoted as $E=\{e_1,e_2,...,e_n\}$ and $\hat{E}=\{\hat{e}_1,\hat{e}_2,...,\hat{e}_n\}$, respectively. The training objectives between the two are achieved through a binary cross entropy loss function, as shown in (13):

$$L_e^{BCE}=BCE(\hat{E},\ E) \quad (13)$$

In the first language assisted classification task, the mother tongue accent description label is marked as m, and the corresponding prediction is marked as $\hat{m}$. The training of both is shown in (14):

$$L_m = CrossEntropy(\hat{m},m) \quad (14)$$

Among them, $\hat{m} = StatisticsPooling(G)$.

The complete loss function of the model is shown in (15):

$$L = L_e + \alpha L_m + \beta L_{asr} \quad (15)$$

Among them, α and β represent the weights of the first language recognition auxiliary task and the loss function weights of the speech recognition auxiliary task, respectively.

In the process of oral evaluation, not only first language phonetic features are required, but also English phonetic features need to be added. However, the increase in English phonetic features may lead to problems such as mismatched inference data and weakened model performance. Therefore, this article adopts Wav2vec2 self supervised learning, as shown in Figure 5.

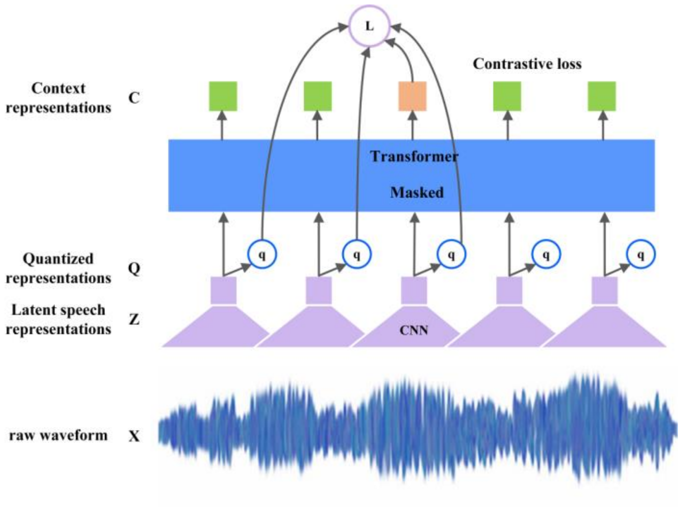


Fig.5Wav2vec2 Schematic diagram of self supervised learning training

Through Wav2vec2 self supervised learning, the model can complete data meaning feature extraction without manual text annotation, improving the efficiency of model task training.

After the pre training of the model is completed, it is necessary to verify its corresponding performance. In order to better test the model performance, this article chooses the L2 Arctic dataset for English speaking evaluation experiments. Firstly, this article first statistically analyzes the common oral errors of English learners who use Chinese as their first language in this dataset, as shown in Figure 6. From the results in the graph, it can be seen that common oral errors among English learners who use Chinese as their first language have a high degree of consistency, that is, verifying the cognitive impact of the first language on English speaking. The addition of auxiliary tasks for first language phoneme classification in the model has a positive impact on English speaking assessment.

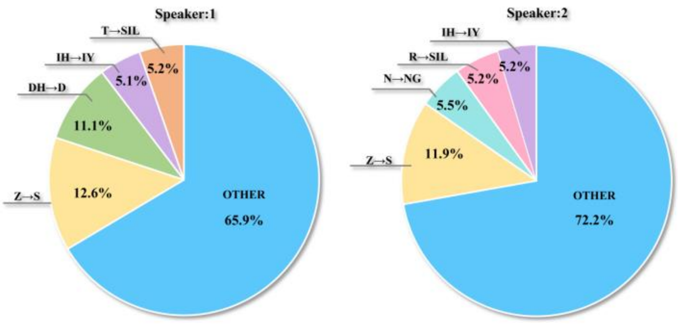


Fig.6 Statistical results of common oral errors among English learners whose first language is Chinese

In order to further verify the influence of first language oral information, this study selected two English oral evaluation models and our model for F1 performance comparison. The results are shown in Figure 7. From the results in the graph, it can be seen that without adding oral information in the first language, the F1 value of our model is the highest under three different factor types. However, the improved ASR model has a significantly higher F1 value than the ASR model under the other two phoneme types, except for the F1 value in the original phoneme that is similar to the F1 value in ASR. After adding first language oral information, both auxiliary classification tasks and first language oral labels can significantly improve the F1 values of the three models. In addition, the F1 value of the model in this article is still at its highest state in different situations. This indicates that the increase in phonemes in first language spoken language can indeed effectively improve the effectiveness of English oral evaluation, and the model in this paper performs the best in English oral evaluation among the three models.

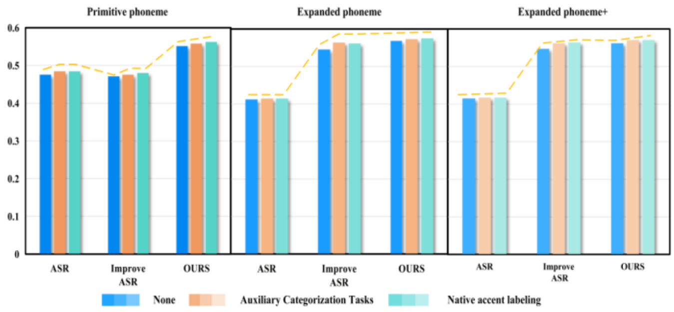


Fig.7 Experimental results of three English oral evaluation models F1

In order to verify the impact of the Wav2vec2 self supervised learning module on the performance of the English oral evaluation model, this paper conducted comparative experiments on the multi accent dataset L2 Abstract and the single accent dataset SpeedOcean, respectively, with this model and the other two models. As shown in Figure 8, the acoustic unit results of similar pronunciations determined by the model in this article based on semantic vectors without linguistic knowledge. The results in the figure show that the model in this article can determine adjacent acoustic units based on the distance distribution characteristics of acoustic units, and the visual distribution of acoustic units is uniform, avoiding the negative impact caused by uneven distribution. This indicates that compared with traditional methods, the model proposed in this paper may obtain more detailed and controllable misreading data at the granularity, and implement training based on the distance labels of front and back semantic vectors.

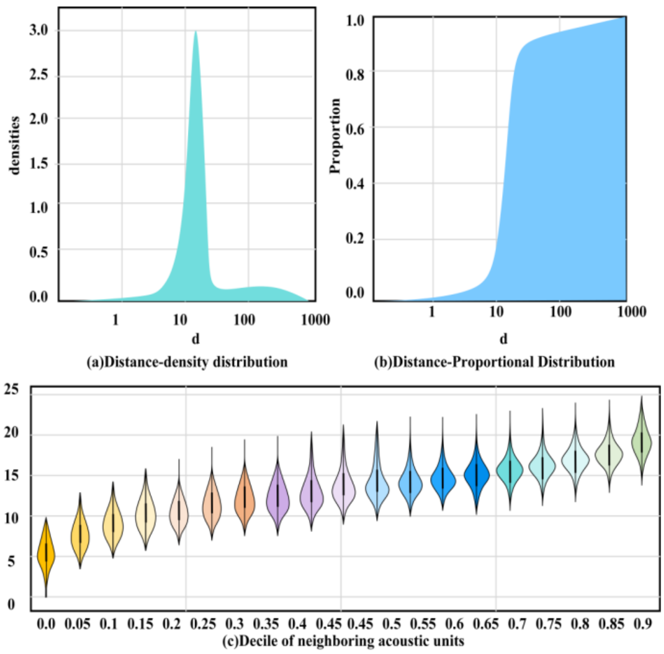


Fig.8 The acoustic unit results of similar pronunciations determined by the model in this article based on semantic vectors without linguistic knowledge

As shown in Figure 9, the comparison results of F1 values for English speaking evaluation using three models in different datasets are presented. The results in Figure (a) show that when the proportion of the training dataset is 20%, the F1 improvement of the model without increasing self supervised learning is not significant compared to the ASR model, but there is a significant improvement in the F1 value of the model with increasing self supervised learning. As the proportion of training datasets increases, the F1 values of all three models continue to improve, and the model with increased supervised learning performs best in F1 values. The results in Figure (b) show that in the single tone dataset, the F1 value of the model with added self supervised learning still performs the best, while the ASR model performs the worst. This indicates that adding the Wav2vec2 self supervised learning module can significantly improve the effectiveness of comment oral evaluation, providing more reliable data information for future application experiments.

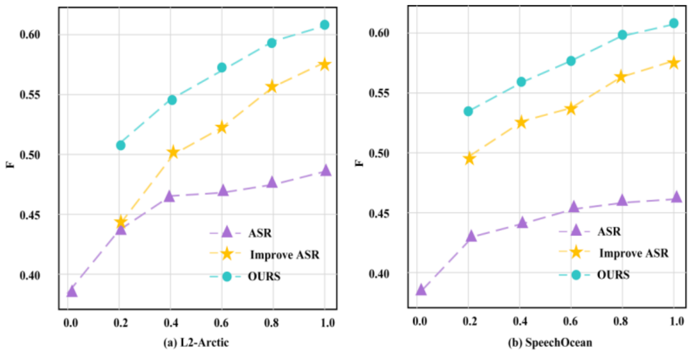


Fig.9 Comparison of F1 values in English oral evaluation using three models in different datasets

3.3 Improving reward mechanisms

Non first language English learners may be affected by the interference of their first language during oral practice, which means they need to control their first language and reduce the cost of switching between two languages. At the same time, English learners need to choose appropriate English words to express themselves based on the current situation in the oral process, so that their oral expression can be fluent in vocabulary, which involves their ability to control their reactivity. Therefore, English learners may be in a state of language system contradiction during the learning process, requiring further interference or stimulation from external factors to promote their English learning process. In English oral teaching, different rewards are important stimuli that push English learners out of the contradictory state of the language system, and corresponding punishments are likely to cause learners to stop learning English oral behavior. Therefore, this article adds a reward mechanism to the English oral teaching system based on reinforcement learning, which provides learners with corresponding rewards based on their English oral evaluation and analysis results.

According to existing research, the emergence of rewards itself does not have a significant motivating effect on learners' learning behavior, but rather the differences in reward prediction play a motivating role in learning behavior. The difference in reward prediction refers to the difference between the rewards received by learners and their own predicted reward results. When the rewards and predicted results are consistent, it can better stimulate the learning motivation mechanism of learners, and vice versa, it may inhibit the learning motivation mechanism of learners. There are significant differences in the prediction results of rewards among different learners, so it is necessary to continuously adjust the reward mechanism for different learners based on their oral assessment and feedback on different rewards in the reinforcement learning module. At the same time, learners will also re recognize their own errors in predicting rewards based on the results of English oral evaluation at different stages, thereby further adjusting their behavior and improving their motivation to learn English oral skills. The control expected value theory based on reinforcement learning is shown in Figure 10.

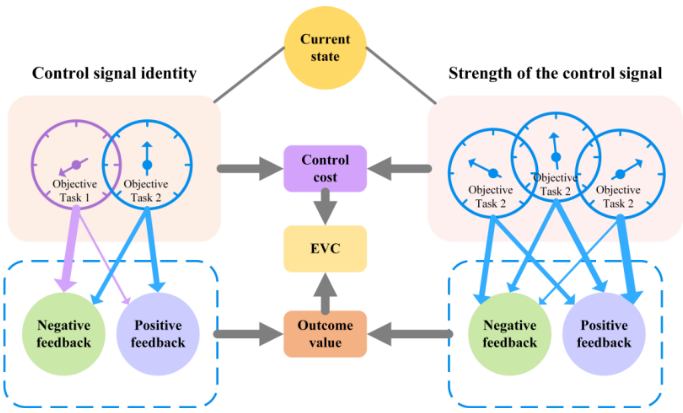


Fig.10 Control Expectation Value Theory Based on Reinforcement Learning

Learners can choose their future English oral learning behavior based on the expected value of control, that is, the reward mechanism will have a decisive impact on their oral learning behavior. The value updates and expected reward errors calculated by learners based on reward errors are shown in (16) and (17):

$$J_t = J_{t-1} + \gamma RPE_{t-1} \quad (16)$$

$$RPE_{t-1} = R_{t-1} - J_{t-1} \quad (17)$$

Among them, the current reward sequence number is t, the corresponding value is denoted as J_t, the prediction error of the previous reward is denoted as RPE_{t-1}, the actual reward result is denoted as R_{t-1}, and the learning rate is denoted as $\gamma, 0 < \gamma < 1$.

In order to verify the motivating effect of the reward mechanism on learners' English oral learning, this study selected two groups of learners for comparative experiments. One group is the experimental group, which introduces the reward mechanism in the learning system, and the other group is the control group, which adopts the general reward method in the learning system. In order to better understand the impact of the reward mechanism on the motivation of learners to learn English speaking, based on passing as the baseline, passing above is divided into two grades, namely 0.5 and 1, while passing below is also divided into two grades, namely -0.5 and 1. The results are shown in Figure 11. From the results in the figure, it can be seen that in the experimental group, the first two experimental results can be seen as the training period, that is, the reward mechanism in the model extracts more feature information through interaction with learners and continuously adjusts it. At this time, the learners' oral performance and expected reward error results are in a relatively disordered state. In the last two experiments, it can be seen that the learners' oral evaluation scores gradually stabilized and showed a significant improvement, and the expected reward error also gradually stabilized and increased. In the comparison group, the performance of the learners in the four experiments showed poor stability, and the corresponding expected error disorder of rewards was strong. This indicates that the reward mechanism in this article can effectively adjust rewards based on the results of learners' oral evaluation, thereby better motivating learners to learn oral English. Learners can also adjust their learning behavior and reward expectations based on different reward situations. The interaction between the two can significantly improve learners' oral learning, and the reward mechanism can play a positive role in English oral teaching.

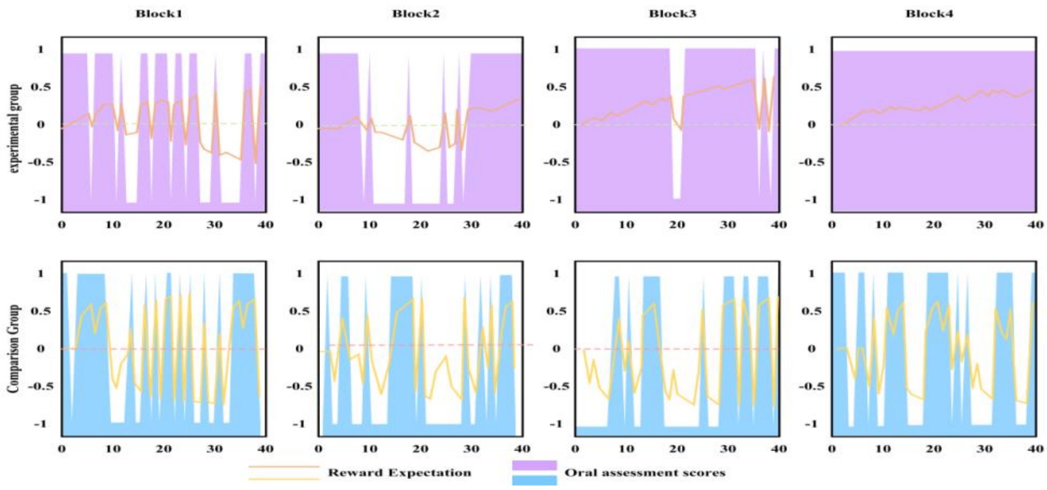


Fig.11 Error analysis of English oral evaluation and reward prediction for learners under different reward mechanisms

3. Evaluation of the application effect of reinforcement learning in English oral teaching

To test the application of reinforcement learning based English oral teaching system in practical teaching, this article selected two English learning classes from a certain school for comparative experiments. Among them, Class A is the control class, using a general English oral teaching model, while Class B is the experimental class, using an English oral teaching system based on reinforcement learning. Before conducting the comparative experiment, this article first conducted a performance test on the oral evaluation of learners in Class B. Ten learners from Class B were randomly selected to evaluate their English oral pronunciation level, and the system evaluation performance of this article was analyzed. The results are shown in Figures 12 and 13. From the results in the graph, it can be seen that the model in this article can score learners from five aspects: intonation, speed, intonation, rhythm, and emotion, and based on this, obtain the final result. This indicates that the model can effectively and targetedly point out the problems that learners have in English speaking, and based on this result, propose effective learning strategies.

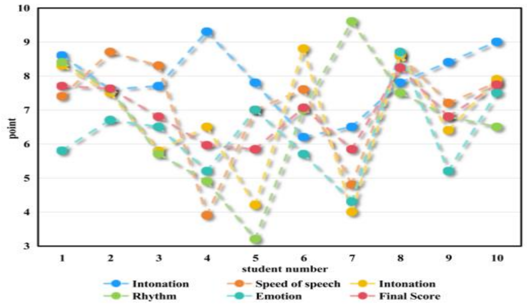


Fig12. Evaluation Results of English Spoken Pronunciation among 10 Random Learners in Class 12B

The performance analysis results of the English oral teaching system based on reinforcement learning are shown in Figure 13. From the results in the graph, it can be seen that in terms of precise consistency, adjacency consistency, and Pearson correlation coefficient results, the system results in this paper are significantly higher than those of big data analysis systems and multi-source feedback systems. This indicates that the English oral evaluation performance of the system in this article has higher accuracy. In terms of misjudgment rate, the misjudgment rate of the system in this article is significantly lower than the other two systems, showing lower error and higher stability. Based on the above analysis, it can be concluded that the English oral teaching system based on reinforcement learning can be effectively applied to English oral teaching.

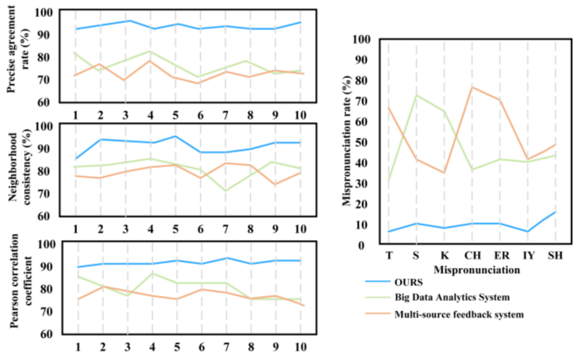


Fig.13 Performance analysis results of English oral teaching system based on reinforcement learning

In the comparative experiment, the English speaking scores of the two classes are tested first. There are 35 learners in each of the two classes, and the comparison of English speaking scores between the two classes before the experiment is shown in the figure. From the results, it can be seen that there is no significant difference in English speaking scores between the two classes of learners, indicating that there is no significant difference between the two. A comparative experiment can be conducted.

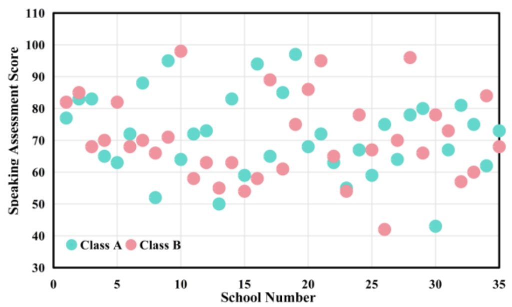


Fig.14 Comparison of English speaking test results between the two classes before the experiment

As shown in Figure 15, the results of the English speaking test after the experiment for two classes show that overall, the grades of Class B are generally significantly improved compared to Class A. In terms of the improvement rate of grades at different levels, the number of failed grades in Class B has significantly decreased, and the proportion of excellent grades has increased by a higher rate than that of Class A. Although there has been some improvement in the grades of Class A, the number of failed grades has hardly changed. This indicates that although general English oral teaching methods can improve learners' English oral performance to a certain extent, their ability to stimulate their learning initiative and enthusiasm is relatively weak. And the model in this article can not only target the oral problems that different learners have. Moreover, it can also provide corresponding rewards based on the learning situation of different learners, stimulate their enthusiasm and motivation for learning English speaking, and significantly improve learning efficiency.

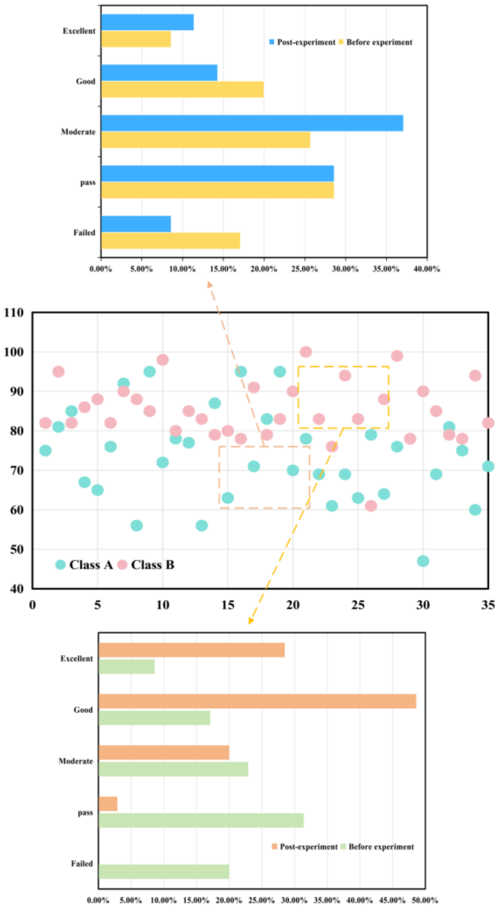


Fig.15 The results of the English speaking test after two class experiments

The evaluation results of two classes on the English oral teaching mode are shown in Figure 16. The results in the figure show that most learners in Class A believe that the effectiveness of English oral teaching in this class is not very good or very poor. The main reason is that learners are unable to identify their own oral problems, which can easily lead to learning bottlenecks. Secondly, it is due to insufficient motivation for oral learning. More than 69% of learners in Class B believe that the English oral teaching system in this class is very effective, while only a small number of learners believe that the teaching effect is not very good. The main reason why the learners in this class recognize the teaching system is that it can effectively conduct oral evaluation, allowing learners to have a more intuitive view of their own oral learning status. And propose targeted questions and learning strategies, followed by reward mechanisms that can effectively meet the expectations of learners and stimulate their motivation.

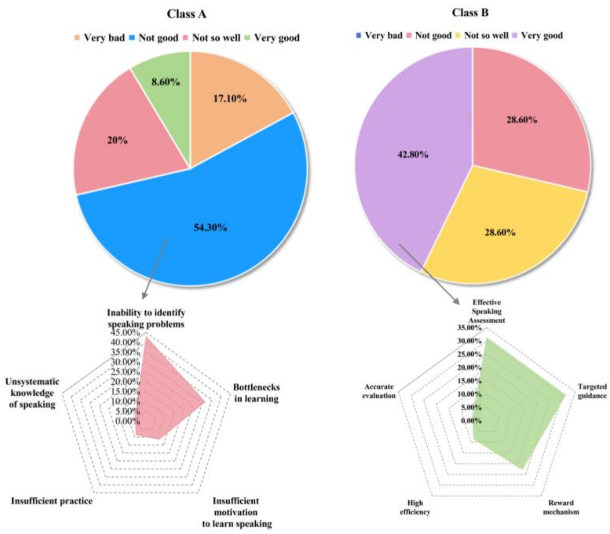


Fig.16 The evaluation results of two classes on the English oral teaching mode

5. CONCLUSION

English oral teaching is one of the important ways to improve learners' English oral skills, but the traditional English oral teaching model is single, one-way, and lacks effective oral practice, which cannot meet the diverse needs of different learners in English oral practice. Therefore, this article constructs English oral teaching based on reinforcement learning, and combines deep learning to construct an English oral evaluation module, and improves the reward mechanism. The performance test results show that the English oral evaluation module combined with deep learning can determine the acoustic units of spoken language without manual annotation. Compared with other evaluation models, this evaluation model has higher performance stability, lower error rate, and better evaluation effect. Furthermore, the improved reward mechanism can effectively enhance the interest of learners in learning English speaking. The experimental results of teaching application show that the model proposed in this article can accurately identify learners' spoken English, reflect the problems existing in learners' spoken English from multiple perspectives, and propose targeted improvement strategies to help learners further improve their speaking ability. Meanwhile, most learners recognize the English oral teaching system based on reinforcement learning, believing that its reward mechanism can effectively stimulate learning motivation and improve oral learning enthusiasm. Future research can further expand the experimental scale and extend the experimental period to comprehensively evaluate the application effect of reinforcement learning in English oral teaching. At the same time, more advanced reinforcement learning algorithms and technologies can be explored to further optimize the teaching framework and improve teaching effectiveness.

REFERENCE

[1] Geden M, Emerson A, Rowe J, et al. Predictive student modeling in educational games with multi-task learning . Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 654-661. DOI:10. 1609/aaai.v34i01. 5406.

[2] Wu Z Y, He T, Mao C J, et al. Exam paper generation based on performance prediction of student group. Information Sciences, 2020, 532: 72-90. DOI: 10. 1016/j. ins.2020. 04. 043.

[3] Hussain M, Zhu W H, Zhang W, et al. Using machine learning to predict student difficulties from learning session data. Artificial Intelligence Review, 2019, 52(1): 381-407. DOI:10. 1007/s10462-018-9620-8.

[4] El Fazazi H, Elgarej M, Qbadou M, et al. Design of an adaptive E-learning system based on multi-agent approach and reinforcement learning.Engineering, Technology &Applied Science Research, 2021, 11(1): 6637-6644.

[5] W. Intayoad, C. Kamyod, P. Temdee. Reinforcement learning based on contextual bandits for personalized online learning recommendation systems.Wireless Personal Communications,2020, 115(4): 2917-2932

[6] Abe K, Xu Z, Sato I, et al. Solving NP-hard problems on graphs by reinforcement learning without domain knowledge[J]. arxiv preprint arxiv:1905.11623, 2019: 1-24. DOI:10.48550/arXiv.1905.11623.

[7] Quang P, Hadjadj-Aoul Y, Outtagarts A. A deep reinforcement learning approach for VNF Forwarding Graph Embedding. IEEE Transactions on Network and Service Management. 2019, 16 (4): 1318–1331.

[8] Saito T, Watanobe Y. Learning path recommendation system for programming education based on neural networks. International Journal of Distance Education Technologies,2020, 18(1): 36-64. DOI:10. 4018/ijdet. 2020010103.

[9] Riordan M A, Abo-Zebiba Z, Marsh J. A Comparison of Multimedia and Traditional Paper Assignments in an Introductory Psychology Course.Teaching of Psychology,2022 (2):130-135.

[10] Ran D, Yingli W, Haoxin Q. Artificial intelligence speech recognition model for correcting spoken English teaching. Journal of Intelligent & Fuzzy Systems, 2021, 40(2): 3513-3524.

[11] He H, Yi S, Liu W. Intelligent Learning Model of Financial Spoken English Teaching based on BPTT algorithm and LSTM Network Model. Journal of Intelligent & Fuzzy Systems, 2020, 39(4): 4835-4846.

[12] Antony Rosewelt L, Arokia Renjit J. A content recommendation system for effective E-learning using embedded feature selection and fuzzy DT based CNN.Journal of Intelligent & Fuzzy Systems, 2020, 39(1): 795-808.

[13] Solozabal R, Ceberio J, Sanchoyerto A, et al. Virtual network function placement optimization with deep reinforcement learning. IEEE Journal on Selected Areas in Communications, 2019, 38(2): 292-303.

[14] Nabizadeh A H, Gonçalves D, Gama S, et al. Adaptive learning path recommender approach using auxiliary learning objects. Computers & Education, 2020, 147: 103777.

[15] Vinyals O, Babuschkin I, Czarnecki W M, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature, 2019, 575(7782): 350-354.

[16] Van Huynh N, Hoang D T, Nguyen D N, et al. Optimal and fast real-time resource slicing with deep dueling neural networks. IEEE Journal on Selected Areas in Communications, 2019, 37(6): 1455-1470.

[17] Valladares W, Galindo M, Gutiérrez J. Energy optimization associated with thermal comfort and indoor air control via a deep reinforcement learning algorithm .Building and Environment. 2019, 155: 105–117.

[18] David K. Dickinson, Kimberly T. Nesbitt, Molly F. Collins, Elizabeth B. Hadley, Katherine Newman, Bretta L, Kathy Hirsh-Paske. Teaching for breadth an depth of vocabulary knowledge: Learning from explicit and implicit instruction and the storybook texts.Early Childhood Research Quarterly, 2019 (47):341-356.

[19] M. Takaso, M. Niitani, T. Imai. What educator could do to facilitate students' use of a deep approach to learning: A multisite cross-sectional design. Nurse Education Today.2020 (89):122-124.

[20] Susanto D A, Oktavia M, Sembiring L T A B. Students' Understanding on Spoken Discourse in The Context of English Language Teaching (ELT) at University. Eralingua: Jurnal Pendidikan Bahasa Asing dan Sastra, 2021, 5(1): 265.