

Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

University English Translation Teaching in the Context of Artificial Intelligence: Problems and Countermeasures

Yanan Sun^{1,†}

1. School of Foreign Languages, Zhengzhou University of Science and Technology, Zhengzhou, Henan, 450064, China.

Submission Info

Communicated by Z. Sabir
 Received January 31, 2024
 Accepted February 8, 2024
 Available online April 1, 2024

Abstract

With the rapid development of artificial intelligence, university English translation teaching faces many challenges and opportunities. This article studies the problems in teaching in this context and proposes effective countermeasures. Diagnostic analysis and factor analysis are used in the article to conduct an in-depth study of the three levels of the teaching process: knowledge skills, autonomy, and emotional development. The students achieved a 70% mastery rate in knowledge and skills, while their independence and emotional development performance was 60% and 50%, respectively. This paper proposes an English speech translation system based on artificial intelligence to enhance teaching effectiveness. After a 14-week teaching experiment, the average grade of the students in the experimental class improved by about 6.82 points compared to the traditional teaching class, proving the system's effectiveness. The study diagnosed the current problems in university English translation teaching, verified the potential of AI technology in improving teaching quality through empirical research, and combined with AI technology can effectively improve the quality and efficiency of university English translation teaching.

Keywords: Artificial intelligence; English translation teaching; Knowledge skills; English speech translation system.
AMS 2010 codes: 65Y04

†Corresponding author.

Email address: clover890601@163.com

ISSN 2444-8656



<https://doi.org/10.2478/amns-2024-0791>



© 2023 Yanan Sun, published by Sciendo.



This work is licensed under the Creative Commons Attribution alone 4.0 License.

1 Introduction

University is a major turning point in the student's learning career, which is an important period of the student's values and outlook on life. In this period, the main goal of the teacher should not be to help students cope with the exam but to make the students better adapt to society and become good people who can wave economic value and personal value in society [1-3]. For the English language as a subject, students must have the ability to apply flexibly, and teachers should pay more attention to the improvement of the student's English language ability and translation abilities so as to lay a solid foundation for the student's life and work in the society and to provide an important safeguard [4-5]. In the context of intelligent technology, university teachers should use the power of science and technology, and in the teaching of English translation, they are using the intelligence of people and the maximum degree of learning is stimulated, and the students' thinking can be helped by the breakthrough and innovation [6-7].

At the same time, with the development of science and technology and education informatization, modernization of the construction of the depth of the advancement of universities to carry out English translation activities in the process of teaching activities began to actively explore the application of artificial intelligence technology. Strive to play the supportive role of artificial intelligence technology, comprehensively and systematically promote the teaching reform of English translation activities, build a platform for students to learn English translation knowledge, and optimize and innovate the teaching effect so as to implement positive and effective translation teaching for students in order to implement positive and effective translation teaching guidance for students, promote the good cultivation of English translation ability and the comprehensive quality of the application of English knowledge of students in colleges and universities, and create favorable teaching reform conditions for the cultivation of students' professionalism in English translation [8-10].

Social development promotes the innovation of the course, and the ability of students to develop their ability in English is more likely to be a comprehensive application of English ability so that they can actually use knowledge to solve practical problems [11]. Literature [12] proposes the Robotic Socio-Cultural Educational Interaction (RI-SCE) method, which aims to solve the problem of students' fear of using English to express their speech and opinions in the classroom, and the simulation analysis confirms that the proposed RI-SCE method can achieve 95.2% translation recognition accuracy. Literature [13] examines the consistency of the conceptualization of teaching classroom practices between Chinese and American English teachers and points out that there are many similarities between the two, with significant differences in grammatical preferences and grammatical translation preferences. There are many similarities between the two, with significant differences in grammatical preferences and grammatical translation preferences. Literature [14] investigated how computer-assisted prosody training (CAPT) and instructor-based prosody teaching (IBPT) affect the development of participants' speaking skills, and the positive effects of CAPT and IBPT on language learning and teaching were confirmed through teaching experiments. Literature [15] extracted brain activation patterns related to Chinese/English visual translation based on the functional near-infrared spectroscopy (fNIRS) technique, and the neural changes related to translation in this study varied according to the direction of translation. Literature [16] builds a Japanese translation teaching corpus based on the bilingual non-parallel data model and conducts a large number of training for the translation teaching machine Moses, and the results of the simulation experiment show that the proposed model has a good translation retrieval function, which can improve the effect of Japanese translation teaching. Literature [17] combines the constructivist learning theory, blended learning theory, and the theory of instructional design, and probes into the status quo of the translation teaching in colleges and universities based on the PACTE translation competence model, and points out that the teachers need to combine with the theory of teaching and supplementing with the teaching of English translation, which can improve the English ability of the students and the awareness of the

English language effectively. Literature [18] proposed a semi-automatic evaluation method of machine translation systems based on fuzzy mathematics and an adapted translation teaching program, which is of positive significance for the cultivation of applied English translators. Literature [19] constructed an English composition scoring model to realize the evaluation of English translation in colleges and universities and confirmed the feasibility of the proposed model through experimental performance testing. It is believed that the future learning mode of human-computer interaction is the development trend of education.

We analyze the three levels of the university English translation classroom: the knowledge and skills level, the autonomy and competence level, and the affective development level through the diagnostic analysis method to determine the current status of students' learning at these three levels. Factor analysis is utilized to explore the intrinsic factors that affect students' learning effects, including teacher level, learning environment, and teaching situation. This paper proposes an English speech translation system that uses artificial intelligence technology to enhance the quality and efficiency of teaching. A 14-week comparative teaching experiment at S University is used to empirically analyze the effectiveness of the system. The system was able to significantly improve students' academic performance, which confirmed the research hypothesis.

2 Diagnostic Methods of Problems in Teaching English Translation at University

2.1 Classroom Diagnostic System for University English Translation Teaching

The diagnostic system of the English translation course is shown in Fig. 1, and the diagnosis of English classroom teaching is mainly divided into "three levels," which are "knowledge and skill level," "autonomy level," and "emotion level." First of all, the diagnosis of these three levels is based on diagnostic activities for the purpose of teaching and training. The diagnosis of the "Knowledge and Skills Level," "Autonomy Level," and "Affective Development Level" are the diagnosticians' analyses and judgments of the problems in the classroom from the perspective of the objectives of teaching and learning. Secondly, the three levels are based on the development of the three types of students with differences in their learning bases and abilities. The "Knowledge and Skills Layer" is aimed at the basic teaching requirements of all students, especially those who are weak in knowledge acquisition. The "Autonomous Ability Layer" is aimed at the teaching requirements of students with medium and strong learning abilities, and the "Emotional Nurture Layer" is aimed at the non-intellectual factors of all students, especially those who have strong learning abilities. Finally, the characteristics of the diagnostic methods of these three levels vary from bottom to top because, from knowledge and skills to affective development, these three levels are a process of tilting from relative objective content to subjective affective content, so the higher the level, the more complex and flexible the design of the diagnostic activities, and the more subjective they are, and vice versa, the more standardized and objective the design of the diagnostic activities is.

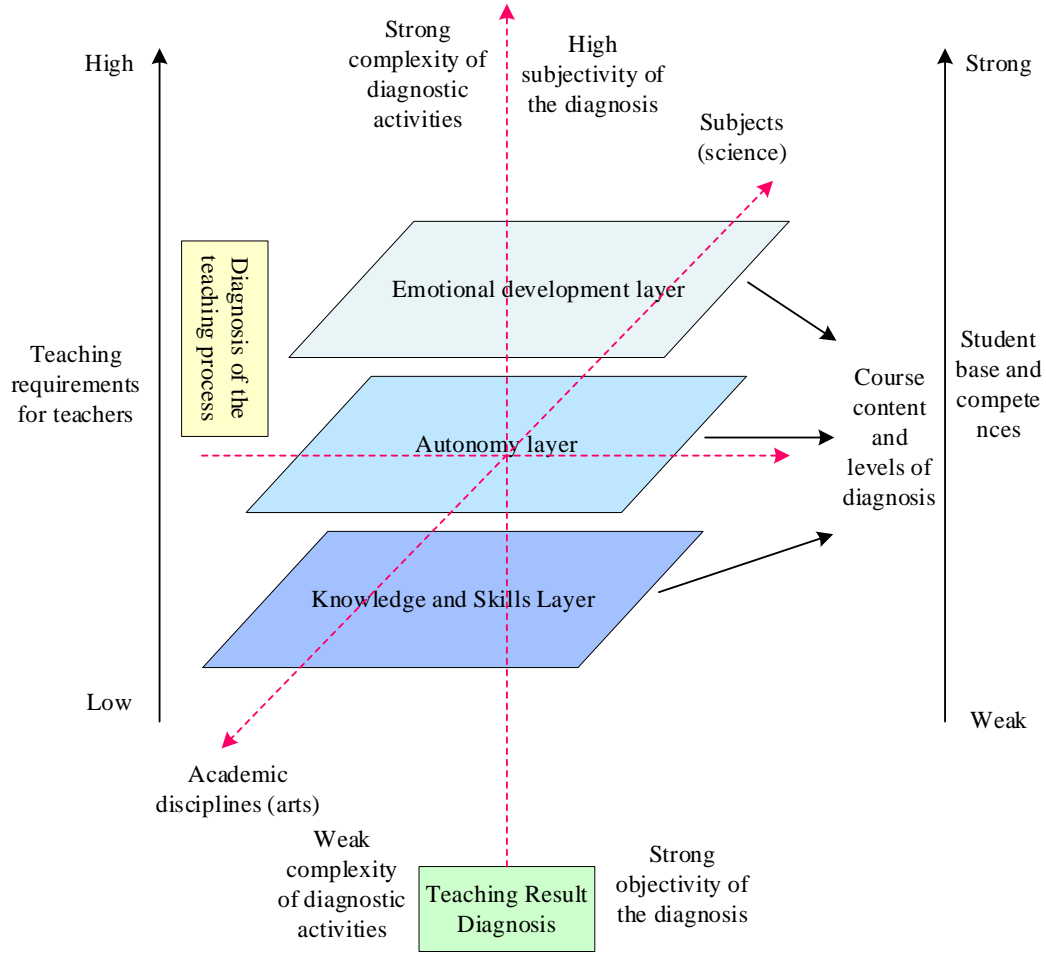


Figure 1. A schematic diagram of the classroom teaching diagnosis system

2.2 A Factor Analysis of Problems Facing the Teaching of English Translation

Factor analysis involves using unobservable potential variables to explain the correlation or covariance between the original variables.

Given a p -dimensional observable random vector $X = (x_1, x_2, \dots, x_p)'$, whose covariance matrix is Σ , decomposed into a k -dimensional ($k < p$) common factor and a p -dimensional special factor, the basic model can be written as follows.

$$X_{p \times 1} = Af_{k \times 1} + \varepsilon_{p \times 1} \tag{1}$$

Where, $X = (x_1, x_2, \dots, x_p)'$ is an observable random vector, $f = (f_1, f_2, \dots, f_k)'$ is a vector of common factors, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)'$ is a vector of special factors, and A is called the factor loading matrix.

It is usually assumed that.

$$E(f) = 0, V(f) = I \tag{2}$$

$$E(\varepsilon) = 0, V(\varepsilon) = D = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2) \quad (3)$$

$$\text{Cov}(f, \varepsilon) = 0 \quad (4)$$

Given the observed data array X of p dimensional correlated variables x_1, x_2, \dots, x_p .

The principal component method is to set the eigenvalues of the covariance matrix Σ of the sample X as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, and the corresponding orthogonal unit eigenvector as $\gamma_1, \gamma_2, \dots, \gamma_p$, and to choose a relatively small number of factors and make the cumulative contribution

$\sum_{i=1}^k \lambda_i / \sum_{i=1}^p \lambda_i$ reach a high percentage, then Σ can be decomposed in the following approximation.

$$\begin{aligned} \Sigma &= \lambda_1 \gamma_1 \gamma_1' + \lambda_2 \gamma_2 \gamma_2' + \dots + \lambda_k \gamma_k \gamma_k' \\ &\quad + \lambda_{k+1} \gamma_{k+1} \gamma_{k+1}' + \dots + \lambda_p \gamma_p \gamma_p' \\ &\approx \lambda_1 \gamma_1 \gamma_1' + \dots + \lambda_k \gamma_k \gamma_k' + D \\ &= AA' + D \end{aligned} \quad (5)$$

In which:

$$\begin{aligned} A &= (\sqrt{\lambda_1} \gamma_1, \sqrt{\lambda_2} \gamma_2, \dots, \sqrt{\lambda_k} \gamma_k) \\ &= (a_{ij})_{p \times k} \\ D &= \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2) \end{aligned} \quad (6)$$

For maximum likelihood estimation, given a common factor $f \sim N_k(0, I)$ and a special factor $\varepsilon \sim N_p(0, D)$, which are independent of each other, there must be an original variable $X \sim N_p(\mu, \Sigma)$.

The likelihood function is calculated from the sample x_1, x_2, \dots, x_n as the $L(\mu, \Sigma)$ function of μ and Σ , Since $\Sigma = AA' + D$, the likelihood function can be expressed as $L(\mu, A, D)$, and the great likelihood estimation of (μ, A, D) is recorded as $(\hat{\mu}, \hat{A}, \hat{D})$, i.e.

$$L(\hat{\mu}, \hat{A}, \hat{D}) = \max L(\mu, A, D) \quad (7)$$

It can be shown that $\hat{\mu} = \bar{x}$, then \hat{A} and \hat{D} satisfy the following equations.

$$\begin{cases} \hat{\Sigma} \hat{D}^{-1} \hat{A} = \hat{A} (I_k + \hat{A} \hat{D}^{-1} \hat{A}) \\ \hat{D} = \text{diag}(\hat{\Sigma} - \hat{A} \hat{A}') \end{cases} \quad (8)$$

Where $\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})'$, because the solution of A is not unique, can be attached to make

computationally convenient uniqueness condition: $A'D^{-1}A$ is a diagonal array, (*) \hat{A} and \hat{D} in the formula can generally be solved by iterative methods.

Least squares is the method whereby if given a matrix $X_{n \times m}$, consider approximating it by a matrix of rank $k \leq \min(m, n)$, where the best approximation is the one that minimizes the distance between vectors when the matrix is viewed as a vector. That is, given $X_{n \times m}$, a matrix $Y_{n \times m}$ of rank k is required such that $\text{tr}(X - Y)'(X - Y) = \min$. This requires a UDV decomposition of the matrix $X_{n \times m}$. The decomposition is.

$$X_{n \times m} = U_{n \times n} \begin{bmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_r & \\ & 0 & & 0 \end{bmatrix} V_{m \times m}' \quad (9)$$

After singular value decomposition (i.e., UDV decomposition) of $X_{n \times m}$, there is then the Equation

$$X = \sum_{i=1}^r \lambda_i u_i v_i', \text{ noting that } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r, \text{ i.e.}$$

$$\begin{aligned} \left\| \dot{X} - \sum_{i=1}^k \lambda u_i v_i' \right\|^2 &= \text{tr} \left(\sum_{i=k+1}^r \lambda u_i v_i' \right)' \left(\sum_{i=k+1}^r \lambda u_i v_i' \right) \\ &= \text{tr} \sum_{i,j=k+1}^r \lambda_i \lambda_j v_i' u_i u_j v_j' \\ &= \sum_{t,j=k+1}^r \lambda_t \lambda_j v_j' v_t u_t' u_j \\ &= \sum_{i=k+1}^r \lambda_i^2 \end{aligned} \quad (10)$$

It can be shown that the right end of the above Equation is indeed a minimum.

3 Artificial Intelligence English Speech Translation System

3.1 English Speech Recognition Model

3.1.1 Mathematical Modeling of Speech Signals

The speech signal generation model can be seen in Fig. 2. The main reason for speech generation is the resonance of the vocal tract excitation, and the vocal tract is in constant motion during vocalization, which can be simulated by a time-varying linear system. Therefore, the vocal tract model can be set as a linear time-invariant model when intercepting speech signals in a small interval of time.

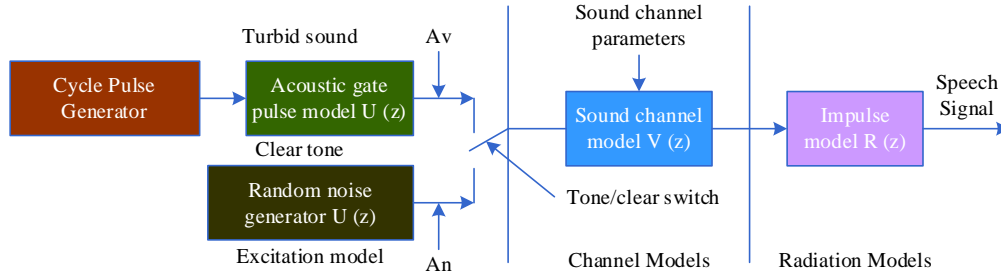


Figure 2. Speech signal model

The complete speech model can be obtained, as shown in Eq. (11), using a tandem representation of the three models: excitation model, vocal tract model, and radiation model.

$$H(z) = A * U(z)V(z)R(z) \quad (11)$$

$U(z)$ is the excitation signal, the z transform of the glottal pulse for turbid tones, the z transform of the random noise for clear tones, and the period pulse is the time it takes for the vocal folds to open and close once.

The acoustic channel model adopts the resonance peak model, and the resonance peak frequency can be obtained from the acoustic channel transfer function $V(z)$, which is an all-poles model. Each pole corresponds to a resonance peak frequency, and its model equation is.

$$V(z) = \frac{G}{1 - \sum_{k=1}^N a_k z^{-k}} \quad (12)$$

The model equation for the radiation model $R(z)$ is.

$$R(z) = R_0 (1 - z^{-1}) \quad (13)$$

3.1.2 English Speech Recognition Model

Linear predictive analysis is the use of human voice principle, the vocal tract model using all-pole digital filter in the form of a moment of the signal as a number of previous moments of the signal as a linear combination of the basic idea is that: there is a correlation between the voice signal samples, can be made by the actual voice of the sample value and linear prediction of the sample value of the average variance of the LMS is minimized, to obtain a unique set of LPC coefficients.

The linear prediction coefficient is the basic parameter of linear prediction, and some other related parameters can be obtained by deforming the parameter. According to the model of speech generation, the speech signal $S(z)$ is a linear non-shift causally stabilized system $V(z)$ stimulated by the signal $E(z)$ to produce the output if the system function of the vocal tract model is.

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 + \sum_{j=1}^p a_j z^{-j}} \quad (14)$$

Where, p is the order of the LPC linear predictor. Let the shock response be $h(n)$, calculate its cepstrum $h'(n)$, according to the definition of cepstrum.

$$\ln H(z) = H'(z) = \sum_{n=1}^{\infty} h'(n)z^{-n} \quad (15)$$

Bringing Equation (15) into (16) and taking partial derivatives with respect to $z-1$ in two passes, we obtain.

$$\frac{\partial}{\partial z^{-1}} \ln \left[\frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} \right] = \frac{\partial}{\partial z^{-1}} \sum_{n=1}^{\infty} h'(n)z^{-n} \quad (16)$$

Which is:

$$\sum_{n=1}^{\infty} nh'(n)z^{-n+1} = \frac{\sum_{k=1}^p ka_k z^{-k+1}}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (17)$$

Thus, we have $\left(1 - \sum_{k=1}^p a_k z^{-k}\right) \sum_{n=1}^{\infty} nh'(n)z^{-n+1} = \sum_{k=1}^p ka_k z^{-k+1}$, whose constant terms and coefficients of each power of z^{-1} are equal in the left and right times, respectively, to obtain the recurrence relation between $h'(n)$ and a_k .

$$\begin{cases} h'(0) = 0 \\ h'(1) = a_1 \\ h'(n) = a_n + \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) a_k h'(n-k), 1 \leq n \leq p \\ h'(n) = \sum_{k=1}^p \left(1 - \frac{k}{n}\right) a_k h'(n-k), n > p \end{cases} \quad (18)$$

LPC order p is generally selected 12, less than a frame of speech sampling points N_s . Therefore, LPCC only represents the first $(n = 1, 2, \dots, N_s)$ values of $h'(n)$. if the cepstrum coefficient is greater than p , then according to Equation (18) can be derived.

The larger the order of the cepstrum coefficient is, the more information it contains; however, as the order of the coefficient increases, the $h'(n)$ will become very small, which, in fact, does not play a big role but increases the amount of computation; therefore, in practice, it is usually based on the current LPCC for first-order and second-order differentials.

The calculation process of MFCC is as follows.

- 1) The input speech signal is preprocessed by frame splitting and pre-emphasis, etc. The spectrum is obtained by a short-time Fourier transform, and the time-domain signal is converted into a frequency-domain signal.
- 2) Find the energy spectrum and the frequency domain of the energy spectrum for bandpass filtering. The bandpass filter $H_m(k)$ center $f(m)$ is divided by the Mel frequency scale, the frequency of the two bottom points is equal to the center frequency of the two neighboring filters, which is defined as follows: the center of the bandpass filter HMK is divided by Mel frequency scale, the frequency of the two bottom points is equal to the center frequency of the two filters, which is defined as follows.

$$f(m) = \frac{N}{F_s} B^{-1} \left[B(f_1) + m \frac{B(f_n) - B(f_1)}{M + 1} \right] \quad (19)$$

Where, f_1 , f_m are the minimum and maximum frequencies applied by the filter frequency, F_s is the sampling frequency, N is the number of points of the FFT transform, $B^{-1}(b) = 700(e^{b/125} - 1)$. Let the number of filters be M , and the output result after filtering is: $X(k)$, $k = 1, 2, \dots, M$.

- 3) Calculate the logarithmic power spectrum of the output result of each filter, and then do the $2M$ -point Fourier inverse transform to obtain the MFCC coefficients. According to the symmetry, this transform can be simplified as follows.

$$C_n = \sum_{k=1}^M \log X(k) \cos[\pi(k - 0.5)n / M], n = 1, 2, \dots, L \quad (20)$$

Here, the number of MFCCs L is usually taken as 12~16, which is due to the fact that the first several dimensions and the last several dimensions of MFCCs are more distinctive for speech.

- 4) The MFCC features obtained after the previous three steps reflect the static features of speech, and the dynamic features of speech can be obtained after first-order and second-order differencing of the MFCC coefficients.

The basic DTW algorithm is to find an optimal path so that the distance measure of each point on the path is minimized, assuming that the reference template is expressed as follows: $R = \{R(1), R(2), \dots, R(m), \dots, R(M)\}$, m is the label of the speech frame, $m = 1$ stands for the starting frame, $m = M$ stands for the termination of the speech, and $R(m)$ is the speech feature vector of the first m frame, and the test template is expressed as follows: $T = \{T(1), T(2), \dots, T(n), \dots, T(N)\}$, $n = 1$ stands for the starting frame of the speech to be tested and $n = N$ stands for the termination of the frame.

In order to test the similarity of the two templates, the distance $D[T, R]$ between them is computed. Let n and m be the arbitrarily chosen frame numbers in T and R , respectively, then $d[T(n), R(m)]$ denotes the distance between two frames of feature vectors. The Euclidean distance is utilized to represent the degree of similarity, and the smaller the distance, the higher the degree of similarity. The temporal axis i of the test vectors is nonlinearly mapped to the temporal axis j of the reference templates, and the distances are computed as follows.

$$D = \min_{\phi(i_n)} \sum_{i_n=1}^N d(T(i_n), R(\phi(i_n))) \quad (21)$$

If $N = M$ can be calculated directly, otherwise a linear expansion method should be used to align $T(n)$, $R(m)$, T in the linear expansion into a sequence of M frames does not take into account the fact that the duration of each segment of the speech signal is also different in different cases, resulting in poorer recognition, therefore, the dynamic programming (DP) method is used in more cases.

The speech signal itself is an observable time-varying sequence of parameter streams of phonemes emitted by the brain based on grammatical knowledge and speech needs. Hidden Markov Models (HMMs) mimic this dual stochastic process as an implicit stochastic process that models the variation of the statistical properties of the speech signal using a Markov chain with a finite number of states and as a stochastic process that has a sequence of observations associated with each state of the Markov chain. According to the different representations of the probability of the output observations, HMM can be divided into discrete HMM (DHMM) and continuous HMM (CHMM), which are only different in the probability density function used in the calculation of state probabilities.

1) Hidden Markov, for speech recognition, must solve three problems.

- (1) The observed sequence $O = \{O_1, O_2, \dots, O_T\}$ and the model $\lambda = (A, B, \pi)$ are known, how to make a finite number of computations to produce the probability $P(O | \lambda)$ of the observed sequence O under the given model conditions.
- (2) The observation sequence $O = \{O_1, O_2, \dots, O_T\}$ and the model $\lambda = (A, B, \pi)$ are known, and how to choose the corresponding sequence of states that is optimal in some sense.
- (3) Given a sequence of observations, how to adjust the parameter (A, B, π) to maximize the conditional probability $P(O | \lambda)$.

2) HMM algorithm

Forward-backward algorithm, this algorithm solves the first problem. The way to solve $P(O | \lambda)$ is: Given a sequence of stationary notions: $Q = q_1, q_2, \dots, q_T$, there are.

$$P(O | Q, \lambda) = \prod_{t=1}^T P(o_t | q_t, \lambda) = b_{q_1}(o_1) b_{q_2}(o_2) \dots b_{q_T}(o_T) \quad (22)$$

Where, $b_{q_t}(o_t) = b_{jk} |_{q_t=\theta_j, o_t=v_k}$, $1 \leq t \leq T$, for a given λ , the probability of generating Q , $P(Q | \lambda) = \pi_{q_1} a_{q_1 q_2} \dots a_{q_{T-1} q_T}$, and hence the probability of O and Q occurring simultaneously under λ conditions is.

$$\begin{aligned} P(O | \lambda) &= \sum_Q P(O | Q, \lambda) P(Q | \lambda) \\ &= \sum_{q_1, q_2, \dots, q_T} \pi_{q_1} b_{q_1}(o_1) a_{q_1 q_2} b_{q_2}(o_2) \dots a_{q_{T-1} q_T} b_{q_T}(o_T) \end{aligned} \quad (23)$$

Viterbi's algorithm, this algorithm solves the second problem. The optimal state sequence Q is the sequence determined to maximize the $P(Q, O | \lambda)$. First define the variable $\gamma_t(i)$.

$$\gamma_t(i) = P(q_t = S_i | O, \lambda) \quad (24)$$

At the moment of t , the maximum probability of arriving at the S_i -like idea along a path that generates the observation sequence $O = \{O_1, O_2, \dots, O_r\}$ is:

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P[q_1, q_2, \dots, q_t = S_i, O_1, O_2, \dots, O_t | \lambda] \quad (25)$$

Then, the step to find the optimal state sequence Q is to initialize:

$$\delta_1(i) = \pi_i b_i(O_1), 1 \leq i \leq N, \varphi_1(i) = 0 \quad (26)$$

Iterative calculation:

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(O_t), 2 \leq t \leq T, 1 \leq j \leq N \quad (27)$$

$$\varphi_t(j) = \arg \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}], 2 \leq t \leq T, 1 \leq j \leq N \quad (28)$$

Final calculations:

$$p^* = \max_{1 \leq i \leq N} [\delta_T(i)] \quad (29)$$

$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i)] \quad (30)$$

Determine the optimal sequence of states:

$$q_t^* = \varphi_{t+1}(q_{t+1}^*), t = T-1, T-2, \dots, 1 \quad (31)$$

The Baum-Welch algorithm, which is used to solve the third problem, adjusts the model parameter (A, B, π) to maximize the probability of occurrence of an observed sequence under a given model and uses an iterative process (the Baum-Welch method, or the expected value correction method) to select the (A, B, π) that maximizes the $P(O | \lambda)$, i.e., the problem of parameter reassessment.

The basic idea of the Baum-Welch algorithm is to use the observation sequence $O = \{O_1, O_2, \dots, O_r\}$ and the model $\lambda = (A, B, \pi)$, to find a new model $\bar{\lambda} = (\bar{A}, \bar{B}, \bar{\pi})$, prove the $P(O | \bar{\lambda}) > P(O | \lambda)$, and repeat the process until the $P(O | \bar{\lambda})$ converges, at which time the $\bar{\lambda}$ is the desired model.

First, define the variables $\xi_t(i, j)$:

$$\xi_t(i, j) = P(q_i = S_i, q_{t+1} = S_j | O, \lambda) \quad (32)$$

Based on the definition of backward and forward variables, it can be concluded that $\xi_t(i, j)$:

$$\begin{aligned}\xi_t(i, j) &= \frac{\alpha_t(i, j)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O|\lambda)} \\ &= \frac{\alpha_t(i, j)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i, j)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}\end{aligned}\quad (33)$$

3.2 Phrase-based Statistical English Machine Translation Modeling

3.2.1 English Machine Translation Model

The IBM model is a word-based statistical machine translation, which takes words as the basic translation unit, but due to the small modeling unit, the model generated from training is inaccurate, the translation performance is greatly limited, and the adaptability of the model is poor. In order to improve the translation performance, researchers began to study the phrase-based translation method based on the characteristics of phrases that are semantically related to the context, and the method has been used up to now.

The basic idea of this model is to decompose the translation process into a sequence of phrases and derive various features associated with the phrases to estimate the translation probability. The basic process is as follows: first, a continuous Chinese sentence is cut into phrases, then the translated phrases are translated into English, and finally, the English phrases are intoned.

Generally, the translation model and language model are included. Meanwhile, in order to make the translation result closer to the English grammatical specification and have better readability, a tuning model Ω is also needed. In view of the modeling problem, the translation system in this paper adopts the logarithmic, linear direct translation model proposed by F.J. Och, and the translation scores are computed by the logarithmic, linear equations. The model is as follows:

$$P_r(e_1^I | f_1^J) = P_{\lambda, \Omega}(e_1^I | f_1^J) = \frac{\exp\left[\sum_{m=1}^M \lambda_m h_m(e_1^I, f_1^J)\right]}{\sum_{e_1^I} \exp\left[\sum_{m=1}^M \lambda_m h_m(e_1^I, f_1^J)\right]}\quad (34)$$

Where $f_1^J = f_1 \dots f_j \dots f_J$ denotes the source language sentence string and $e_1^I = e_1 \dots e_i \dots e_I$ denotes the target language sentence string, $h_m(e_1^I | f_1^J)$ is the feature function ($m=1, \dots, M$) and λ_m is the feature function weight, in the decoding process, the denominator of Eq. The translation is constant, so the optimal one can be obtained as follows.

$$\hat{e}_1^I = \arg \max_{e_1^I} \left\{ P_r(e_1^I | f_1^J) \right\} = \arg \max_{f_1^I} \left\{ \sum_{m=1}^M \lambda_m h_m(e_1^I, f_1^J) \right\}\quad (35)$$

The ordering model Ω adopts the maximum entropy-based lexicalization model in Deyi Xiong et al. First, define two consecutive blocks A^1 and A^2 and their order o , the order has only two states, forward and inverted, such as the following relation.

$$\Omega = f(o, A^1, A^2) \quad (36)$$

Integrating block features to predict block reordering probabilities using a maximum entropy model:

$$\Omega = P_\theta(o | A^1, A^2) = \frac{\exp\left(\sum_i \theta h_i(o, A^1, A^2)\right)}{\sum_o \exp\left(\sum_i \theta h_i(o, A^1, A^2)\right)} \quad (37)$$

Where feature function $h_i \in \{0,1\}$, θ_i is the feature weight, which can be obtained by different algorithms.

The first field is Chinese phrase pairs, the second field is English phrase pairs, the third field is a set of elements associated with the entries, and there are six features in the field. The first feature in the third field is the probability of phrase translation from Chinese to English $P(\bar{t} | \bar{s})$ (where \bar{s} is the Chinese text and \bar{t} is the English text), and the second feature is the probability of lexicalized translation, $P_{lex}(\bar{t} | \bar{s})$, which expresses the degree of alignment of the phrases in the \bar{s} and the phrases in the \bar{t} . Assuming that $\bar{s} = s_1 \dots s_J$, $\bar{t} = t_1 \dots t_I$ and a are the phrases that are aligned between the two, the calculation of the $P_{lex}(\bar{t} | \bar{s})$ is as follows.

$$P_{lex}(\bar{t} | \bar{s}) = \prod_{j=1}^J \frac{1}{|\{j | a(j,i) = 1\}|_{\forall (j,i) a(j,i)=1}} \sum \omega(t_i | s_j) \quad (38)$$

Where $\omega(t_i, s_j)$ is the weight of (s_j, t_i) . The third feature is the phrase translation probability $P(\bar{s} | \bar{t})$ for English to Chinese translation, the fourth feature is the probability of the degree of alignment of the phrase in \bar{t} and the phrase in \bar{s} , the fifth feature denotes the attached phrase, and the sixth is an undefined feature that can be defined by the user.

3.2.2 English Machine Translation Training Process

The flowchart of phrase translation model training can be seen in Fig. 3. After preparing the bilingual corpus with Chinese and English, the data are preprocessed, and the phrase translation probability table is obtained through the training of the phrase translation model.

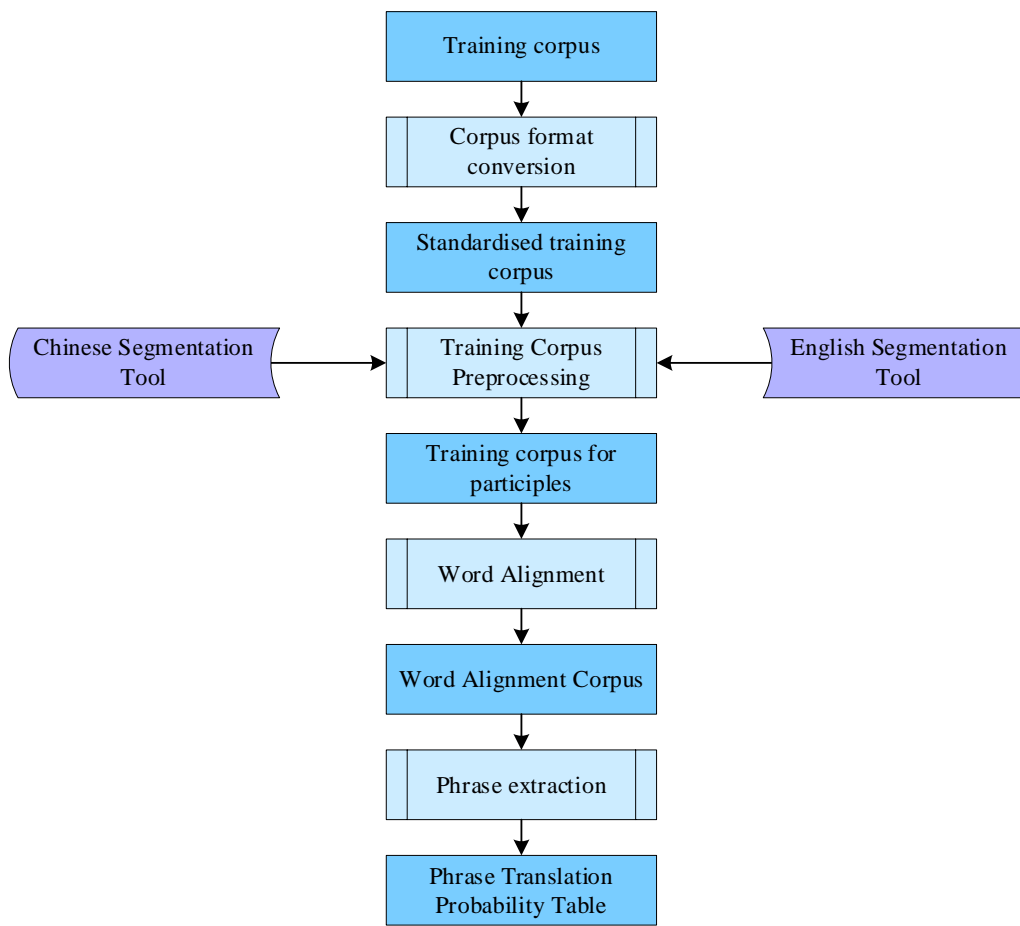


Figure 3. Phrase translation model training process

4 The Effectiveness of University English Translation Teaching Practices

4.1 Analysis of the Effectiveness of English Speech Translation

4.1.1 Performance Analysis of English Speech Translation

The bilingual corpus used for the experiments is DECC, which mainly consists of daily life conversation speech, including 15000 pairs of aligned bilingual sentences. There is no standardized method for calculating the correctness of English speech translation values. In the corpus, a correct Chinese translation that matches an English word is considered correct, and the opposite is also true. In order to differentiate, the dictionary synthesized from level 1 to 4, i.e., the dictionary with a trust value greater than 1, is called a “statistical dictionary,” the dictionary filtered from the interpretation dictionary is called a “filtered dictionary,” the dictionary synthesized from all the generated dictionaries in the algorithm is called “synthesized dictionary,” and the correctness rate, the recall rate and the F1 value are taken as the evaluation criteria, and the results of each criterion of the translated dictionaries are shown in Table 1. All the English translation dictionaries have a correct rate of over 50% and an F1 value of over 0.65. The first and second-level dictionaries have a relatively low correct rate, while the filtering dictionary has the highest correct rate of 98.489% and an F1 of 0.656. Comprehensively, in the comprehensive dictionary obtained by synthesizing all the generated dictionaries, its correct rate, recall rate, and F1 value are 90.142%, 99.98%, and 0.968, respectively,

which indicates that the English speech translation system in this paper has good performance and effect, and can accurately recognize and translate English speech.

Table 1. The standard results of the dictionary

Dictionary	Accuracy (%)	Recall rate (%)	F1
Primary dictionary	50.754	64.15	0.675
Secondary dictionary	69.014	50.17	0.697
Tertiary dictionary	72.848	89.45	0.758
Four dictionary	87.745	79.24	0.823
Statistical dictionary	89.452	98.87	0.834
Filter dictionary	98.489	40.54	0.656
Filter dictionary + 4 dictionary	95.452	93.48	0.957
Integrated dictionary	90.142	99.98	0.968

4.1.2 An Empirical Analysis of English Speech Translation

In order to further study the performance of the English speech translation system proposed in this paper, 80 entries were randomly extracted from various dictionaries, including a source language lexical word and its extracted target language translation items and their correct rates, recall rates, and F1 values were counted, and the statistical results of the dictionaries at all levels are shown in Table 2, and it can be observed that, basically, the correct rates of the translations of the English dictionary increase with the level of the dictionary, but the correct rates of those entries which satisfy one condition, i.e., There are many entries in the first-level dictionary, but they are very low, with only 43.554% and 53.844%, respectively. It can be observed that the correct rate of English dictionary translation basically increases with the increase of the dictionary level, but the correct rate of the words that satisfy one condition, i.e., there are a lot of words in the first level dictionary, is very low, only 43.554% and 53.844%, respectively, which suggests that it is very difficult to extract the bilingual dictionaries with a high correct rate by relying on a single condition alone, and not every source language word has a target phrase corresponding to it. In the level 4 dictionary, the correct rate is 97.730% and 95.907%, respectively, which indicates that the correct rate of English translation can be effectively improved under multiple conditions, but the recall rate is around 0.9, which indicates that the source language vocabulary corresponds to the target language word strings in a relatively limited way. On the other hand, because the corpus is too small, it should be enlarged appropriately to achieve better translation results. Overall, the English translation system constructed in this paper has a better performance, which can provide effective help in teaching English translation.

Table 2. Statistical results of dictionaries at all levels

Dictionary	Accuracy (%)	Recall rate (%)	F1
English-Chinese 0 dictionary	94.845	89.241	0.561
English-Chinese 1 dictionary	43.554	48.280	0.234
English-Chinese 2 dictionary	95.725	88.345	0.508
English-Chinese 3 dictionary	96.845	90.243	0.631
English-Chinese 4 dictionary	97.730	90.541	0.516
English-Chinese 0 dictionary	92.043	85.047	0.519
English-Chinese 1 dictionary	53.844	41.650	0.402
English-Chinese 2 dictionary	95.845	88.346	0.583
English-Chinese 3 dictionary	95.725	89.948	0.526
English-Chinese 4 dictionary	95.907	93.343	0.572

4.2 Analysis of the Problems of University English Translation Teaching

4.2.1 Diagnosis of Problems in University English Translation Teaching

Through the analysis, the factors of the influencing factors are put into the regression analysis using the predictive multiple step-by-step regression. Firstly, the quality of English translation teaching is put into the multiple regression analysis as the dependent variable, and the teacher's level, the learning environment, the translation limitation, the teaching situation, and the teaching support are put into the multiple regression analysis as the independent variables, and the effective influencing factors are obtained, and then, the effective factors are put into the step-by-step regression analysis as the dependent variable, and the others are put into the step-by-step regression analysis as the independent variables. The coefficients of the factors influencing the teaching of English translation are shown in Table 3. The teacher's level, learning environment, teaching situation, teaching support, and translation limitations have a direct influence on the quality of teaching English translation. The standardized coefficients of influence, i.e., the direct effect values are 0.241, 0.310, 0.545, 0.150 and 0.430 respectively, among which the effect influence values of teaching situation and translation restriction are higher, which indicates that the grasp of the teaching situation and translation restriction are the important problems of university English teaching, and the translation restriction is the delay caused by students' classroom's inability to understand the English language translated by the teacher or the course in time, and the interaction between the teaching situation and translation restriction, thus deepening the influence on the quality of English translation courses. The total effect value of 0.975 indicates that it is necessary to reduce the restriction on English translation and improve the teaching situation.

Table 3. The coefficient of the influence factors of English translation teaching

Model		Nonnormalized coefficient		Standard coefficient	T	Sig.
Dependent variable	Independent variable	B	Standard error	Beta		
English translation teaching quality	Teacher level	0.367	0.074	0.241	6.870	0.000
	Learning environment	0.411	0.045	0.310	9.701	0.000
	Teaching situation	0.417	0.042	0.545	11.797	0.000
	Teaching support	0.175	0.059	0.150	3.179	0.003
	Translation restriction	0.685	0.051	0.430	12.501	0.002

4.2.2 Analysis of University English Translation Teaching Practice

In order to verify whether the English speech translation system in this paper can effectively improve the quality of English translation teaching when it is put into the practice of university English translation teaching and whether the quality of English translation teaching is more obvious when the time of putting the English speech translation system into the practice is prolonged, the practice of university English translation teaching is designed. Two English translation classes of University S with comparable levels of performance were subjected to a 14-week comparative teaching experiment, in which Class A was an experimental class put into the English auxiliary system for English translation, and Class B was a traditional class for traditional English translation teaching, and an achievement test was conducted once every two weeks on a biweekly basis, and each time the score of the English test was set at 100, and a total of seven English achievement tests were conducted, and the changes in the achievement of the English translation of the two classes are shown in Fig. 4. The gap between the performance of the experimental class and the traditional class in the English translation course gradually appears after 3 tests, in the 5th, 6th and 7th English translation tests, the

average performance of the experimental class compared with the traditional class has improved by 7.23, 5.83 and 7.40 respectively, the average performance gap in the 1st, 2nd and 3rd is poorer. In the first 3 English tests, the effect of the English speech translation system is in the range of 20% to 28%. The effect has not yet appeared. In all the first 3 tests, the average grade difference between the two classes is not big. With the change of time, the English speech translation system continues to put in, the effect of its use in the 5th achievement test reached more than 60%, and the experimental class of the average grade of the English translation course improved significantly. An analysis shows that the English speech translation system in this paper and the improvement of the average grades of the experimental class of the English translation course also improved. The analysis shows that with the input and effect of this paper's English speech translation system, the average grade of the experimental class in the English translation course also improves, and compared with the average grade of the traditional class, it has a significant advantage, which demonstrates that the English speech translation system proposed in this paper can effectively solve the learning problems caused by the delay in translation in the English translation course, so as to improve the quality of the teaching of English translation.

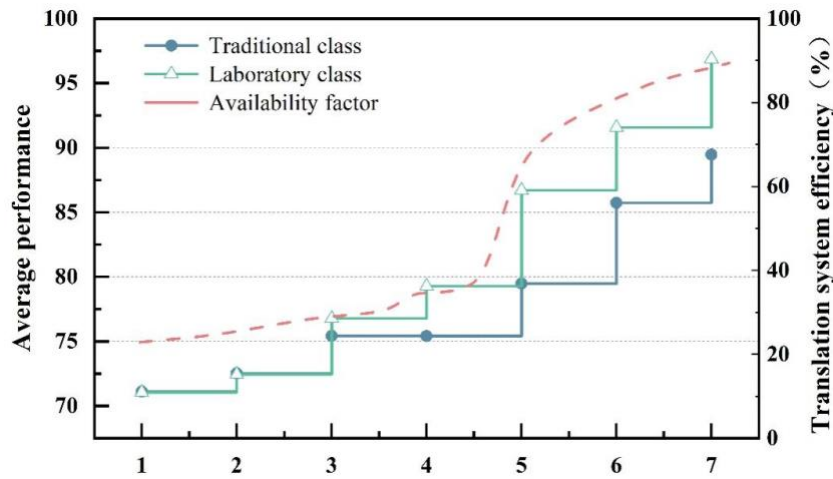


Figure 4. The English translation of the two classes changed

5 Conclusion

This study addresses the challenges faced in university-level English translation teaching, pinpointing key issues such as students' lack of knowledge, autonomy, and emotional development, alongside factors like teacher expertise and the learning environment. We introduce and implement an artificial intelligence (AI)-based system for English speech translation, which notably enhanced student performance during a 14-week experiment at University S. Results showed an average increase of 6.82 points in student grades compared to those in traditional teaching settings, highlighting the effective role of AI in augmenting English translation teaching. This paper underscores the importance of integrating AI into educational practices, offering insightful prospects for future teaching reforms.

References

- [1] Harris, K., English, T., Harms, P. D., Gross, J. J., & Jackso, J. J. (2017). Why are extraverts more satisfied? personality, social experiences, and subjective well-being in college. *European Journal of Personality*.
- [2] Solomon, G. T., Alabduljader, N., & Ramani, R. S. (2019). Knowledge management and social entrepreneurship education: lessons learned from an exploratory two-country study. *Journal of Knowledge Management*, 23(10), 1984-2006.

- [3] Lihua, D. (2022). An extended model of the theory of planned behavior: an empirical study of entrepreneurial intention and entrepreneurial behavior in college students. *Frontiers in Psychology*.
- [4] Liu, M. (2019). Design of intelligent english writing self-evaluation auxiliary system. *Informatika*(2).
- [5] Zheng, X. (2023). Robot translation based on computer vision for cultural psychology of english culture education. *International journal of humanoid robotics*(2/3), 20.
- [6] Yang, X. (2020). Research on integration method of ai teaching resources based on learning behavior data analysis. *International Journal of Continuing Engineering Education and Life-Long Learning*, 30(1), 1.
- [7] Li, N. (2020). An improved machine learning algorithm for text-voice conversion. *Journal of Intelligent and Fuzzy Systems*, 40(2), 1-11.
- [8] Ma, J. (2021). Intelligent decision system of higher educational resource data under artificial intelligence technology. *International Journal of Emerging Technologies in Learning (iJET)*, 16(5), 130.
- [9] Wei, X. (2021). Simulation of english intelligent system based on ca-iafsa algorithm and artificial intelligence. *Journal of Intelligent and Fuzzy Systems*, 1-11.
- [10] Li, X. (2017). The construction of intelligent english teaching model based on artificial intelligence. *International Journal of Emerging Technologies in Learning (iJET)*, 12(12), 35-44.
- [11] Liu, S. (2021). Development of english teaching theory and practice research from the cross-cultural perspective. *International Journal of Electrical Engineering Education*, 002072092110077.
- [12] Zheng, X. (2023). Robot translation based on computer vision for cultural psychology of english culture education. *International Journal of Humanoid Robotics*, 20(02n03).
- [13] Clark-Gareca, B. G. M. (2019). Chinese and american efl teachers' beliefs about curricular and pedagogical practices: cross-cultural similarities and differences. *Nature reviews Cancer*, 19(2).
- [14] Yenkimaleki, M., & Van Heuven, V. J. (2019). The relative contribution of computer assisted prosody training vs. instructor based prosody teaching in developing speaking skills by interpreter trainees: an experimental study. *Speech Communication*, 107, 48-57.
- [15] He, Y., Wang, M. Y., Li, D., & Yuan, Z. (2017). Optical mapping of brain activation during the english to chinese and chinese to english sight translation. *Biomedical Optics Express*.
- [16] Guo, Z., & Jifeng, Z. (2020). Japanese translation teaching corpus based on bilingual non parallel data model. *Journal of Intelligent and Fuzzy Systems*, 40(5), 1-11.
- [17] Liu, Y., & Bai, H. (2021). Teaching research on college english translation in the era of big data. *International Journal of Electrical Engineering Education*, 002072092098431.
- [18] Wu, H. (2021). Multimedia interaction-based computer-aided translation technology in applied english teaching. *Mobile Information Systems*.
- [19] Yuan, Z. (2021). Interactive intelligent teaching and automatic composition scoring system based on linear regression machine learning algorithm. *Journal of Intelligent and Fuzzy Systems*, 40(2), 2069-2081.