

Publication Date: 31 July 2024

Archs Sci. (2024) Volume 74, Issue 4 Pages 185-194, Paper ID 2024424.  
<https://doi.org/10.62227/as/74424>

# Research on Precision Teaching Mode of Language Translation in Colleges and Universities under the Perspective of Intercultural Communication

Yanan Wu<sup>1</sup>, Xiaohui Zhang<sup>2</sup> and Duo Zhang<sup>3,\*</sup>

<sup>1</sup>Russian Center, University of Sanya, Sanya, Hainan, 572022, China.

<sup>2</sup>School of Foreign Languages, University of Sanya, Sanya, Hainan, 572022, China.

<sup>3</sup>School of Economics, Management and Law of Jilin Normal University, Siping, Jilin, 136000, China.

Corresponding authors: (e-mail: jlspszd@jlnu.edu.cn).

**Abstract** Within the scope of cross-cultural communication perspective, this paper proposes to use big data to investigate the language and cultural background, combined with personalized language translation learning resources recommendation model to carry out targeted language translation teaching training. With the help of machine translation technology to assist language translation, it establishes a precise teaching mode of language translation in colleges and universities under big data technology. Delineate the domain of precise language translation teaching and the corresponding organization of language teaching. Explore the relationship between language translation vocabulary factors and the implementation of precise teaching using ANOVA. Using the clustering method to classify language translation learning resources, carry out a two-month practice of precision teaching of language translation in colleges and universities, and evaluate the effect of precision teaching. The number of access times and the mean of the video learning period were 7.689 and 16.030, 2830.05, respectively. The growth in video learning is particularly significant. After the post-t test, sig= 0.000 of the practice and the number of assignments submitted. This paper confirms the validity of the language translation precision teaching mode proposed in this paper.

**Index Terms** learning resource recommendation; clustering method; machine translation; precision teaching; language translation in universities

## I. Introduction

The main goal of language translation teaching in colleges and universities is to cultivate high-quality language translation professionals, so as to better realize the national culture and economic exchanges, which requires language translators not only to have good listening, reading and writing skills, but also need to understand the relevant cultures of other countries, in order to better realize the communication and exchanges with other countries [1]–[3]. Language among culture is not only a tool, but also a kind of cultural symbols, only with the culture of each other to play a role, before learning the language of other countries is the most important thing is to understand the culture of other regions, in order to realize the improvement of the language level [4]–[6].

In the process of teaching language translation in colleges and universities, teachers do not conduct in-depth research on cross-region cultural content because it does not take up a very heavy proportion in teaching, and therefore they are more inclined to language symbols as well as language lecturing work [7]–[10]. The teaching method is also mainly based on classroom teaching and repeated training for students, which leads to the lack of vividness in the classroom, which is not

only unfavorable to students' mastery of cross-regional culture, but also leads to the loss of students' interest in language translation, and fails to effectively improve the teaching effect.

Teachers are the most important guides in language translation teaching work in colleges and universities, in which they play a key role, and teachers' teaching philosophy determines whether cross-culture can be integrated into language translation teaching [11]–[13]. In the cross-cultural perspective, the reform of language translation teaching work is a long-term and arduous task, and the effect can not be seen in a short period of time, which requires teachers to reform the language translation teaching in a long-term unremitting efforts, and gradually integrate the cross-cultural content to realize the cross-cultural teaching, so as to let the students realize the deep meaning of cross-cultural teaching and correct the students' attitude to learning, so that the students' enthusiasm for learning, initiative, and ultimately improve the level of language translation teaching [14]–[17]. Measures should be taken to gradually infiltrate cross-cultural contents into all aspects of language translation teaching, so as to enrich the contents of language translation teaching and cultivate students' cross-cultural thinking. In the teaching of language

translation classroom, teachers can increase the communication with students, so that students can more deeply grasp the cross-cultural knowledge content and language and cultural background, and strive to improve the students' comprehensive cultural competence level.

This paper constructs a precise teaching mode of language translation in colleges and universities under the perspective of cross-cultural communication. It analyzes the language and cultural background using statistical data methods, recommends personalized language translation learning resources using dual clustering algorithms, generates the top-N recommendation list of learning resource objects using CF recommendation algorithms, clarifies the process of recommending language translation learning resources, and carries out personalized language translation teaching training. It also assists the precise teaching of language translation with the advantage of machine translation. Divide the organization form of precise teaching according to the role domain of precise teaching of language translation in colleges and universities. Explore the relationship between the vocabulary factor of language translation and the readability of precise language teaching, cluster analyze the language translation learning resources, and carry out the actual comparison and analysis of the participation, interaction and attention before and after the precise language translation teaching mode.

## II. Language Translation Teaching in Colleges and Universities from the Perspective of Intercultural Communication

### A. Scope of intercultural communication perspectives

Intercultural communication is a multifaceted, complex and dynamic field that involves interaction and communication between different cultures and languages. Definitions of the field usually include the process of transferring information and understanding across cultural boundaries. Intercultural communication is not only about words, but also includes non-verbal communication and symbolic interaction between cultures. Its scope is not limited to communication between individuals, but also involves large-scale interactions between cultures in various fields, including business, politics, and society.

The scope of intercultural communication also includes a wide range of aspects such as information transfer between cultures, cultural understanding, cultural integration, and cultural conflict. This includes not only the process of intercultural communication, but also its results and effects. Therefore, the study of intercultural communication covers multiple dimensions such as cultural differences, cultural interaction, and cultural adaptation in order to fully understand the definition and scope of intercultural communication.

### B. Technical Support for Accurate Teaching of Language Translation in Tertiary Institutions

1) Statistical analysis of the linguistic and cultural context of data

#### 1) Utilizing Big Data Technology for Cultural Research and Background Understanding

Cultural research and background understanding are key steps in addressing cross-cultural challenges in language translation. Before translating, translators should gain an in-depth understanding of the target culture to ensure that they can accurately convey the message and avoid misunderstandings caused by cultural differences. This includes studying the language, customs, values, beliefs, and social background of the target culture. For example, if translators want to translate advertisements or slogans from a certain culture into the language, they need to understand the cultural meanings and historical background of these advertisements or slogans in the target culture. Only with a deep understanding of the target culture can translators carry out accurate language conversion and ensure that the message is conveyed in its original flavor.

#### 2) Application of Statistical Classification Techniques in Linguistic and Cultural Contexts

Through data mining, the hidden intrinsic connection between language translation and various factors is comprehensively analyzed.

At present, many institutions usually remain in the method of querying the database when dealing with a large amount of data information in the teaching process.

In this project, it is proposed to adopt the classification algorithm in data mining to mine the linguistic and cultural background, so as to transform a large amount of data into classification rules to better analyze these data for precise language translation teaching.

- STEP 1: Determine the mining object and target
- STEP 2: Data Acquisition
- STEP 3: Data Preprocessing
- STEP 4: Data Classification Mining
- STEP 5: Analyze the results of classification rules
- STEP 6: Application of knowledge

2) Personalized Language Translation Teaching Training

#### 1) Targeted language translation training

Translation training is another key element in language translation. Training helps to improve the professionalism of translators, especially when working in a cross-cultural context. Training can include the enhancement of language skills, including grammar, vocabulary and language expression. In addition, translation training can cover intercultural communication skills to help translators better understand and cope with differences between cultures. Targeted training can also be adapted to the needs of different fields, for example, translators in the medical field, the legal field or the business field require different specialized knowledge and skills.

2) Recommended Resources for Language Translation Learning

Collaborative filtering imputation method is widely used in many fields, and its recommendation accuracy mainly lies in the selection of similarity, the following is the similarity metric in the near-neighbor based collaborative filtering.

In the case of a two-dimensional space, the cosine similarity of its pinch angle is computed utilizing the calculation by assuming that the vector  $A = (A_1, A_2), B = (B_1, B_2)$ . The cosine value is:

$$\cos \theta = \frac{A_1 \times B_1 + A_2 \times B_2}{\sqrt{A_1^2 + A_2^2} + \sqrt{B_1^2 + B_2^2}}. \quad (1)$$

The calculation in the case of  $N$ -dimensional space assumes vector  $A = (A_1, A_2, \dots, A_n), B = (B_1, B_2, \dots, B_n)$ . The cosine value is:

$$\cos \theta = \frac{\sum_1^n A_i \times B_i}{\sqrt{\sum_1^n A_i^2} \times \sqrt{\sum_1^n B_i^2}}. \quad (2)$$

The value domain of the easily provable cosine value is  $[-1,1]$ , and the closer the cosine value is to 1, the smaller the angle between the two vectors is, and the more similar they are, which means that their directions are probably the same.

Clustering algorithm, as a processing tool for data classification, plays a vital role in educational data mining, and the personalized learning resource recommendation algorithm studied in this paper is designed on the basis of clustering analysis.

From the personalized learning resources recommendation model, it is known that the focus is on the learner model and learning resources model, so through the double clustering of learners and learning resources that is the double clustering algorithm, after the analysis of the learner clustering, learning resources clustering, and the evaluation matrix of the learner on the resources.

In this paper, we denote the set of learners by  $L$ , i.e.,  $L = \{l_1, l_2, \dots, l_n\}$ . We denote the generated clustering of  $N$  learners by  $LC$ , i.e.,  $LC = \{lc_1, lc_2, \dots, lc_n\}$ . Then learners belonging to the same learner clusters in a single learner cluster rate each resource as similar as possible in Pinto, and learners belonging to different learner clusters in different learner clusters rate each resource as differently as possible, while conforming to the following requirements:

$$\begin{cases} \cup lc_2 \cup \dots \cup lc_n = L, \\ lc_i \cap lc_j = \emptyset (i \neq j, 1 \leq i \leq n, 1 \leq j \leq n). \end{cases} \quad (3)$$

From the characteristics of the clustering algorithm, it can be seen that generating  $n$  learner cluster also generates  $n$  learner clustering centers, denoted by set  $LCC = \{lcc_1, lcc_2, \dots, lcc_n\}$ , and the average ratings of the learners in the clusters on the learning resources are denoted by  $LCC_i$ . In order to improve the accuracy of the recommendation, the recorded distances should

be selected for the first  $m$  learning clusters with similar users, denoted by  $NL$ .

Let the set of learning resource objects be  $I = \{i_1, i_2, \dots, i_n\}$ , and denote the generated  $m$  learning resource clusters by  $IC$ , i.e.,  $IC = \{ic_1, ic_2, \dots, ic_n\}$ . Then learners belonging to the same learning resource cluster rate each item as similarly as possible, and learners belonging to different learning resource clusters rate each item as differently as possible, while meeting the following requirements:

$$\begin{cases} ic_1 \cup ic_2 \cup \dots \cup ic_n = I, \\ ic_i \cap ic_j = \emptyset (i \neq j, 1 \leq i \leq m, 1 \leq j \leq m). \end{cases} \quad (4)$$

From the characteristics of the clustering algorithm, it can be seen that generating  $n$  learning resource cluster also generates  $n$  learning resource clustering centers, denoted by set  $ICC = \{icc_1, icc_2, \dots, icc_m\}$ , and the similarity of the whole virtual seat resource clusters is denoted by learning resource clustering center  $icc_j$ . In order to improve the accuracy of the recommendation, the recorded distances should select the first  $m$  learning clusters with similar learning resources, denoted by  $NL$ . Based on the above generated  $n$  learner  $m$  learning resource pair clustering evaluation matrix  $R[n, m]$ , the  $n$  rows of the matrix represent the  $n$  learners, the  $m$  rows represent the  $m$  learning resources, and the  $i$  rows and  $j$  columns of the cross section of the matrix are the ratings of the learners on the learning resources. That is:

$$R_{nm} = \begin{bmatrix} r_{11} & \dots & r_{1k} & \dots & r_{1m} \\ \vdots & & \vdots & & \vdots \\ r_{21} & \dots & r_{2k} & \dots & r_{2m} \\ \vdots & & \vdots & & \vdots \\ r_{n1} & \dots & r_{nk} & \dots & r_{nm} \end{bmatrix}. \quad (5)$$

By clustering learners and learning resources separately, learners in the same type of learner clusters have strong and consistent attention to the corresponding learning resources, and there is similarity between learning resources in the same type of learning resource clusters (similarity between the attributes of each knowledge point). In the process of resource recommendation, it is necessary to select the learners in the  $n$  learner clusters that are the nearest neighbors of the target learners as the new learner space, and select the learning resources in the top- $n$  learning resource clusters that have the highest ratings of these learner clusters as the new resource space, and after that the collaborative filtering algorithm is applied to the new dual space to generate the list of Top- $n$  learning resource recommendations. That is, based on the target learners' predictive scores of learning resources, learner clustering and learning resource clustering, the CF recommendation algorithm is utilized to generate the top- $N$  recommendation list of learning resource objects.

The steps to generate Top-N Learning Resource Recommendation Algorithm are shown below in the algorithm LCCR (LR (Learner Recommendation Form), LID (Student ID), NeedC (Number of Recommended Resources Needed), LR (Learner Rating), LC (Learner Clustering), NL (Nearest Neighbor  $m$  Learner Clustering)), IC (Learning Resource Clustering).

Inputs: LR, LID, NeedC, LR, LC, NL, IC.

Output: the RID (Recommended List of Learning Resources for Predictive Scoring and Recommendations).

Methods:

- Step 1: Filter the nearest neighbor  $m$  learning clusters  $LC = \{lc_1, lc_2, \dots, lc_m\}$  of the target learner LID in NL and the similarity between LID and the center of these learner clusters is  $sim(LID, lc_i) (i = 1, 2, \dots, m)$ .
- Step 2: Filter the top  $s$  learning credentials with the highest  $lc_i (i = 1, 2, \dots, m)$  ratings in the learner rating matrix LR to generate  $IC = \{lc_1, lc_2, \dots, lc_m\}$ .
- Step 3: The learners in the learner clustering LC are used as the new learner space, and the resources in the learning resource clustering IC are used as the new resource space, and the corresponding small module regions are intercepted in the learner scoring matrix. Synthesize the new learner-resource scoring matrix NR.
- Step 4: Use collaborative filtering algorithm for recommendation in the new scoring matrix NR, CF (NR, LID, NC), and return the recommendation result.

### 3) Personalized Language Translation Learning Resources Recommendation Process

The process of personalized language translation learning resources recommendation is shown in Figure 1.

Firstly, we obtain the basic learning data of university students, and determine the group to which the students belong through clustering after preprocessing the student behavior record data. Then the basic information of students' registration, online practice test scores, and learning behavior records are combined with the comment information, rating information and teacher guidance information in the interactive management subsystem to construct the transaction set. And normalized data processing and clustering algorithms for data regularization and data mining, to discover the rules describing the user learning characteristics with given confidence and support, as well as mining the "technology roadmap" of student learning, and then carry out user matching and information recommendation, and recommend possible "interesting" information to the users of the currently visiting students. We can recommend knowledge points and teaching resources that may be of "interest" to the currently visiting student users, and finally realize personalized learning recom-

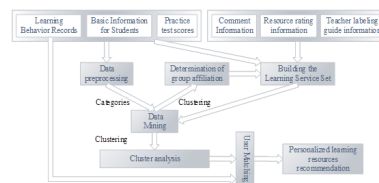


Figure 1: Language translation resources recommendation implementation process

mendation.

### 3) Machine Translation Aids Precision Teaching and Learning

#### 1) Advantages of using technological tools (machine translation)

Technological tools play an increasingly important role in language translation. Tools such as translation memory software, machine translation and online dictionaries can greatly improve translation efficiency and accuracy. Translation memory software ensures consistency by storing previously translated text for reuse in future translations. Machine translation can be used to quickly translate large volumes of text, which can then be post-edited and proofread by translators. Online dictionaries and translation platforms can help translators by providing timely terminology and translation advice.

Taken together, cultural research, targeted translation training and the use of technological tools are key strategies and approaches for addressing cross-cultural challenges in language translation. These approaches help to improve translation quality, reduce misunderstandings, and increase translation efficiency to better meet the needs of intercultural communication. In the evolving globalized environment, these strategies and approaches will continue to play a key role in supporting successful intercultural communication.

#### 2) Machine Translation Technology

Regarding the mathematical definition of the neural machine translation model, the input source language word sequence is denoted by  $x = \{x_1, x_2, \dots, x_m\}$  and the output target language word sequence is denoted by  $y = \{y_1, y_2, \dots, y_n\}$ , where  $m$  denotes the length of the source language word sequence and  $n$  denotes the length of the target language word sequence. The target language translation  $\hat{y}$  with the highest probability of translation result is found given the source language sentence  $x$ , as shown in Eq. (6):

$$\hat{y} = \arg \max_y P(y|x). \quad (6)$$

The neural machine translation model generates the target translation word by word from left to right. When translating the current word, the result of the previous translation is used as the current input. Conditional

probability modeling yields  $P(y|x)$  as shown in Eq. (7):

$$P(y|x) = \prod_{t=1}^n P(y_t|y_{0\sim t-1}, x). \quad (7)$$

When  $t = 0$ ,  $y_0$  represents the beginning of the target sequence and  $y_{n+1}$  represents the end of the target sequence.  $P(y_t|y_{0\sim t-1}, x)$  represents the probability of generating the  $t$ th target language word  $y_t$  based on the source language sentence  $x$  and the generated target translation  $y = \{y_1, y_2, \dots, y_{t-1}\}$ . It is known that the source language sentence is  $x = \{“?”, “?”, “?”\}$ , then the target language translation  $y = \{“I”, “love”, “you”\}$  probability is as in Eq. (8):

$$\begin{aligned} &P(\{“I”, “love”, “you”\} | \{“I”, “love”, “you”\}) \\ &= P(“I” | \{“I”, “love”, “you”\}) \\ &\cdot P(“love” | “I”, \{“I”, “love”, “you”\}) \\ &\cdot P(“you” | \{“I”, “love”, “you”\}) \end{aligned} \quad (8)$$

Theoretically, when obtaining the target language translation  $\hat{y}$ , all  $y$  can be exhaustively enumerated, and then each  $y$  can be evaluated using Eq. (7) to find the one with the highest probability  $y$ . This method is called full search. Although full search can guarantee the global optimum, it cannot exhaust all word sequences due to its excessive time complexity. Therefore, in the decoding stage, the two commonly used decoding methods are greedy search and cluster search.

For each target language translation position  $t$ , the greedy search selects the word with the highest conditional probability as the translation for that position  $\hat{y}_t$ , as shown in Eq. (9):

$$\hat{y}_t = \arg \max_{y_t} P(y_t|\hat{y}_{0\sim t-1}, x), \quad (9)$$

where  $\hat{y}_{0\sim t-1}$  denotes the generated target translation sequence. The target translation sequence is formed by combining the words with the highest conditional probability at all positions. However, this method is prone to fall into local optimization without obtaining the global optimum.

Cluster search, as a compromise between greedy search and full search, alleviates this problem to some extent. When determining each target language translation position  $t$ , cluster search does not select the word with the highest probability as the translation  $\hat{y}_t$  for that position, but selects  $k$  words as candidate values according to the probability in descending order,  $k$  which is called the cluster width, as shown in Eq. (10):

$$\{\hat{y}_{t1}, \dots, \hat{y}_{tk}\} = \arg \max_{\{y_{t1}, \dots, y_{tk}\}} P(y_t | \{\hat{y}_{0\sim t-1}\}, x), \quad (10)$$

where  $\{\hat{y}_{t1}, \dots, \hat{y}_{tk}\}$  denotes the top  $k$  words with the highest probability of translation at position  $t$  of the target language translation.  $\{\hat{y}_{0\sim t-1}\}$  denotes the set of all histories consisting of the first  $k$  words in the first  $t - 1$  positions. Each element in the set represents

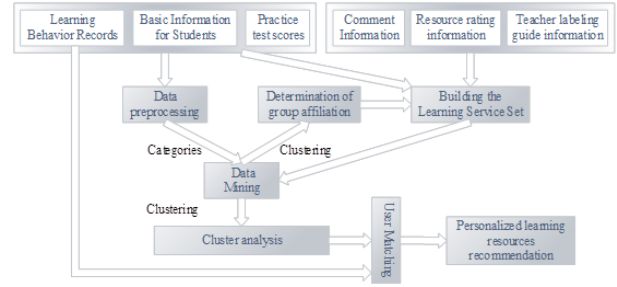


Figure 2: Attention score calculation

a sequence of words translated in the target language,  $\{“I”, “love”\}$  is a sequence of words at time  $t = 2$ .  $P(y_t | \{\hat{y}_{0\sim t-1}\}, x)$  represents the probability of generating  $y_t$  based on a certain path of  $\{\hat{y}_{0\sim t-1}\}$ .

- a. The attentional mechanism will first scan the entire image area and focus attention on the target area that contains key information. The attention score is computed as shown in Figure 2 by considering the sequence of words in the source language sentence  $Source$  as dictionary data for  $\langle Key, Value \rangle$ . Given a word  $Q$  in  $Target$ , the similarity or correlation between  $Q$  and each  $Key$  can be calculated to obtain the weight coefficients of each  $Key$  corresponding to  $Value$ . Then a weighted summation of  $Value$  is performed to obtain the attention score. The whole process can be represented by Eq. (11):

$$Attention(Q, Source) = \sum_{i=1}^L a_i * Value_i, \quad (11)$$

where  $L$  denotes the length of the source language sentence  $Source$ .  $a_i$  denotes the similarity  $Similarity$  between the normalized  $Q$  and a certain  $Key$ ,  $a_i$  as in Eq. (12):

$$a_i = \text{Softmax}(Similarity_i) = \frac{e^{Similarity_i}}{\sum_{j=1}^L e^{Similarity_j}}. \quad (12)$$

There are various ways to compute the similarity  $Similarity$  between  $Q$  and some  $Key$ , which can be obtained by solving for the dot product in Eq. (13):

$$Similarity(Q, Key_i) = Q \cdot Key_i. \quad (13)$$

It can also be obtained by finding the cosine similarity between the two through Eq. (14):

$$Similarity(Q, Key_i) = \frac{Q \cdot Key_i}{\|Q\| \cdot \|Key_i\|}. \quad (14)$$

It can also be obtained by introducing other neural networks through Eq. (15):

$$Similarity(Q, Key_i) = MLP(Q \cdot Key_i). \quad (15)$$

- b. Transformer cleverly introduces a multi-head attention mechanism. Multihead means that the original query vector  $Q$ , key vector  $K$ , and value vector Value are divided equally into multiple parts. Assuming that the original query vector is divided into 8 parts, the result is  $Q = \{q_1, q_2, \dots, q_8\}$ ,  $K = \{k_1, k_2, \dots, k_8\}$ ,  $V = \{v_1, v_2, \dots, v_8\}$ . The multi-head attention mechanism uses each of the  $q, k, v$  divided parts to compute the attention separately. The result of the  $i$ th head attention calculation is shown in Eq. (16):

$$\text{head}_i = \text{Attention}(q_i, k_i, v_i). \quad (16)$$

The advantage of the multi-head attention mechanism is that it allows the model to learn in different subspaces. Each attention head can learn different information and can obtain more comprehensive features. For example, some heads capture syntactic information total and some heads capture lexical information total, and can reduce the error brought by single head. The formula for the multi-head attention mechanism is shown in Eq. (17):

$$\begin{aligned} &\text{Multi Head}(Q, K, V) \\ &= \text{Concat}(\text{head}_1, \dots, \text{head}_8) W^o \text{head}_i. \end{aligned} \quad (17)$$

In RNN network, the latter unit directly depends on the output information of the former unit, which has a certain temporal order. This is in line with the textual language characteristics. Transformer uses the self-attention mechanism, but the units in the self-attention mechanism are independent of each other, and there is no positional information. In order to solve this problem, Transformer introduces positional encoding based on the original word vector input, which is used to represent the positional relationship between units. Transformer uses sine-cosine functions with different frequencies for positional encoding, as shown in Eqs (18) and (19):

$$PE(\text{pos}, 2i) = \sin\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right) \quad (18)$$

$$PE(\text{pos}, 2i + 1) = \cos\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right) \quad (19)$$

where  $E$  denotes the function of position encoding,  $\text{pos}$  denotes the position of the word,  $i$  represents the dimension in the position encoding vector, and  $d_{\text{model}}$  denotes the size of the hidden layer for each position.

### C. Precision Teaching Mode of Language Translation in Tertiary Institutions

- 1) Accurate teaching model of language translation based on big data

The "big data" language translation accurate teaching model includes three dimensions, using big data to establish accurate teaching goals, designing a procedural teaching process

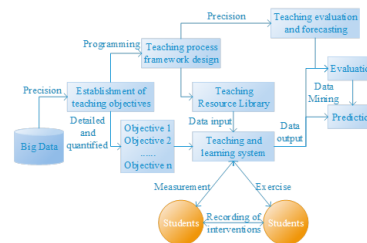


Figure 3: Based on large number of language translation precision teaching mode

framework based on the goals, and realizing the accuracy of teaching evaluation and prediction. The precise teaching model of language translation based on big data is shown in Figure 3.

- 1) Implementation Process

The most important feature of this precise teaching model is to deeply explore the educational value of "big data" and design teaching activities based on "big data".

- 2) Evaluation

The teaching model takes "big data" related technology as the key driver, breaks through the operational difficulties of the previous precision teaching, has clear and concise implementation procedures, and has a reliable and effective operation mechanism, which provides a strong guarantee and support for the further popularization and application of precision teaching.

- 3) "Precision teaching" mode based on "technology platform".

As the degree of education informatization continues to improve, the construction of teaching technology platforms is becoming more and more mature, gradually becoming an important support for the practical application of "Precision Teaching".

- 2) Precise pedagogical scopes of action and their pedagogical organization

The first step in the information-supported "precision teaching" model is to realize the precise determination of teaching objectives. The model uses recursive thinking to build a personalized and precise goal tree to determine the goals accurately. The root node of the target tree points to the overall goal of teaching, and the sub-nodes point to the sub-targets that vary from student to student, so as to break down the knowledge and skills layer by layer, forming a knowledge and skill tree and designing the corresponding test bank accordingly. On this basis, students' shortcomings in knowledge or skills are progressively localized, and the precise focus on sub-targets is realized through regression.

The second is to develop materials and teaching processes that address knowledge and skill gaps. Learning materials refer to interactive digital textbooks that can make students "excited" and break through the concept of traditional textbooks, which are composed of three aspects: "learning materials", "learning materials" and "creative materials". The develop-

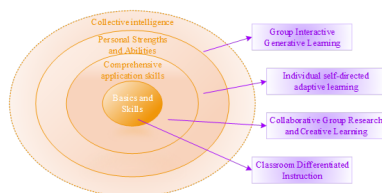


Figure 4: The precision teaching scope and its corresponding teaching group weave form

ment of the teaching process refers to the creation of efficient and intelligent classrooms with the help of information technology, and classroom teaching can be manifested in various forms of teaching organization due to different scopes. The scope of precision teaching and its corresponding teaching organization form are shown in Figure 4, and the four scopes of basic knowledge and skills, comprehensive application ability, individual strengths, and collective wisdom correspond to class differentiated teaching, group cooperative research and creative learning, individual autonomous adaptive learning, and group interactive generative learning, respectively.

### III. Practical Exploration of Precision Teaching Mode of Language Translation in Tertiary Institutions

#### A. Teaching Language Translation Precision

##### 1) ANOVA for language translation difficulty

This section explores the relationship between language translation vocabulary and the translation of college language translation precision, and translates the vocabulary of 14 languages as self-variables, and the numerical model is the primary, intermediate, quasi-advanced and advanced four difficulty levels of the university language translation materials.

The language translation lexical factors include high-frequency words, commonly used words, new words, literary words, colloquial words, false words, pronouns, associated words, abbreviated words, idiomatic and familiar words, polysemous words, monosyllabic words, disyllabic words, lexical diversity, and the average number of words in a sentence.

The results of the four-level ANOVA for language translation are shown in Table 1, where 0.05 is generally set as the threshold of determination, and a p-value of less than 0.05 for a factor means that the factor has an effect on the teaching and learning of language translation precision. Less than 0.01, it has a significant effect on language translation precision teaching, and less than 0.001, it has an extremely significant effect on language translation precision teaching.

The p-values of new words and polysemous words are  $5.56 \times 10^{-5}***$  and  $0.000382***$ , respectively, which show a significant effect relationship with the teaching of language translation precision at the 0.001 level. This may be related to the subject matter of the translated texts, where the use of new words is mainly concentrated in the fields of science and technology, economy and other areas of more rapid development and change. The number of new words in language

Language translation vocabulary factor	F statistic	P value
High frequency word	4.628	0.00452**
Common word	5.0121	0.00289**
Neologism	10.0352	$5.56 \times 10^{-5}***$
Prose	0.361	0.899
Colloquial	3.652	0.0912
Imaginary word	3.025	0.03201*
Pronoun	0.596	0.734
Related words	1.805	0.183
Abbreviations	0.378	0.881
Idioms	5.023	0.00203**
Polysemy	7.318	0.000382***
Monodin	0.489	0.814
Duonic	3.256	0.0329**
Vocabulary diversity	5.512	0.00335**
Average number of words	2.336	0.0890
p<0.1 "•" p<0.05 "***" p<0.01 "****" p<0.0001 "*****"		

Table 1: The analysis of the four-level variance analysis of the language translation

translation textbooks is closely related to the subject matter of their internal chapters, the more network and technology-related subject chapters, the higher the proportion of new words in the textbooks, and the more important it is to utilize the advantages of machine translation to assist language translation precision teaching in colleges and universities.

##### 2) Clustering-based Analysis of Language Translation Learning Resources

The purpose of the cluster analysis is to investigate into which categories the design of learning resources for accurate teaching of language translation in higher education oriented to the results of teaching practice can be categorized into, so as to determine how different levels of difficulty in language translation can be practiced in the same way in the process of designing learning resources for language translation. If the categorization meets the expectations, it can also be verified that the design of the different categories of learning resources for language translation achieves the expected results.

In order to analyze the clustering of the teaching practice of language translation learning resources, this paper adopts the method of using SPSS to conduct cluster analysis.

According to the initial clustering center of each item, it can be seen that the result of clustering is 3 categories, and the final clustering center of language translation learning resources in higher education is shown in Table 2. This shows the clustering process of university language translation learning resources in terms of scientificity, richness, difficulty, reasonableness, completeness, curriculum arrangement, practicality and programming ability.

If clustered into 3 categories, the final clustering result of the 30 language translation learning resources is shown in Figure 5. From the figure, it can be seen that words and basic knowledge are categorized into the first category, while PPT, translation tools, VISIO, speech translation, automatic translation, intelligent classification, and hierarchical analysis are categorized into the second category. Language data, machine translation, language teaching, translation teaching, artificial

	Cluster 1	Cluster 2	Cluster 3
Scientificity	54.5652654212	35.2515690056	38.2400001512
Richness	1.10239900142	10.0236500000	8.1250000000
Difficulty	25.0500000000	8.3125500000	15.0000005393
Reasonableness	19.5000000000	9.5210000000	5.1252100000
Integrity	25.9622000000	36.9825420000	53.8965680220
Course arrangement	35.6582659541	45.7589351200	15.2032200019
Practicability	24.8796525348	55.0635210000	5.1238000000
Programming ability	15.93518522221	7.05000000000	3.59362512000

Table 2: The final clustering center of the university language translation resources

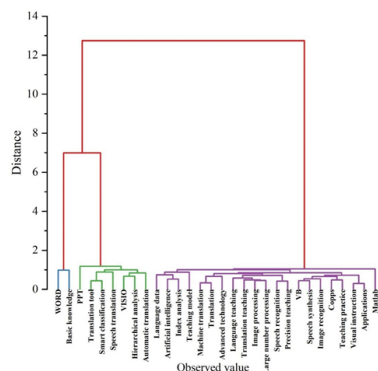


Figure 5: The final clustering results of the language translation resources

intelligence, Matlab, VB, science class, visualization teaching, speech recognition, image recognition, image processing, translation science, speech synthesis, big data processing, indicator analysis, application programs, advanced technology, precision teaching, teaching practice, teaching mode can be classified into the third category. The classification results show that the 30 different teaching schoolbooks are classified into three categories if the results of teaching practice are oriented, and the classification results are consistent with the attributes of interesting or practical school-based learning resources corresponding to their themes, which indicates that the designed school-based learning resources for high school information technology have achieved the expected results.

The distance from the cluster center of each classification and the number of clustered cases are shown in Table 3, which shows the distance from the cluster center of each classification and the number of cases in each cluster. The distances from cluster center 1 to cluster centers 2 and 3 are 6.94943 and 18.11313, respectively. The most representative observations for clusters 1, 2, and 3 are words, translation tools, and teaching translation, respectively. It can be understood that the significance of the results of this clustering is better and can represent the classification results of 30 language translation learning resources oriented on the results of teaching practice.

**B. Experimental results of the precise teaching mode of language translation**

In order to test the effectiveness and implementation effect of the precision teaching model of language translation in colleges and universities, this chapter carries out a two-month

practice study of precision teaching in an authentic teaching environment, implements the precision teaching model, and employs a quasi-experimental study to verify the effectiveness of the precision teaching model.

1) Experimental design of the refined teaching model

The study adopts the idea and method of quasi-experimental research to design a comparative experiment of the precision teaching model, with the aim of applying the designed big-data-based precision teaching model for language translation in colleges and universities in the actual teaching process, and examining the effect of the precision teaching strategy on the enhancement of students' English learning.

This study analyzes the effect of the university language translation teaching model in depth from two aspects: quantitative evaluation and qualitative evaluation. Pre-post behavioral data and outcome data are collected, and after pre-post comparison, the effectiveness of the precision teaching model is analyzed based on the amount of change and significance test.

The object of the study is to select 30 third-year language translation majors in a university, and the precise teaching of language translation is carried out for a period of two months, and the pre- and post-student behavioral data are statistically analyzed respectively.

2) Data collection and analysis of results

The data collection period before and after teaching was two months, and 30 learning resources were released in the late stage of teaching, including micro-teaching resources, text support materials, extended reading, etc. The length of each micro-teaching video was controlled to be around 3 minutes, with a total length of 5,680 seconds.

Compared with the early stage, the later stage after the teaching mode intervention put forward specific requirements for students' online interaction, and timely statistics and reminders. There were a total of 25 practice sessions, counting the students' performance on the classroom exercises or the Wisdom Learning Network quizzes. There were 50 homework assignments, which were submitted to the homework book or photographed and uploaded, and were reviewed by the teacher to record the results of homework quality evaluation. Comparison of student behavioral data in the pre and post periods is shown in Table 4, which collects the behavioral data of each student in the post period and compares it with the results of the pre period data.

From the table, it can be seen that overall there is a different degree of growth in each indicator. In terms of engagement, every student accessed the learning resources and accessed almost 90% of the progress of each video, and most of the students accessed all of the videos, which also included re-watching, so the level of engagement increased significantly. The number of accesses to resources and the length of video learning were 7.68 and 893s respectively in the first stage of the teaching intervention for language translation accuracy,



Clustering	The distance between the cluster centers		
	1	2	3
1		6.94943	18.11313
2	6.94943		11.17332
3	18.11313	11.17332	
Clustering	The most representative observations		Minimum representative observation
1	WORD		WORD
2	Translation tool		PPT
3	Translation teaching		Language data

Table 3: The distance of the cluster center and the number of clustering cases

and 16.03 and 2830.05s respectively in the second stage of the teaching intervention.

The increase in duration was particularly significant for video learning. In terms of interaction, both online and classroom interactions showed multiplicative growth, and the growth of classroom interactions was much larger than that of online interactions. In terms of concentration, students were basically able to submit all the assignments, and those who did not submit the assignments in time due to leave of absence also made up for it. The average score of the exercises reached 88.49, and the average quality of the assignments was 89.76.

In order to compare whether there is a significant difference between the before and after items, the data of the later items are harmonized according to the proportion of the earlier indicators. The before and after data were imported into SPSS for paired-sample t-test, and the paired-sample t-test of the before and after behavioral data is shown in Table 5, and there is a significant difference between the indicators of behavioral characteristics at the level of 0.05.

Through comparison, it is found that students' learning behavior performance after the intervention of teaching mode has obvious improvement, students' participation, interaction performance and task submission have been improved to different degrees, and they still have the best performance in practice and homework. The sample t-test data of the number of practice submissions before and after were 4.88, standard deviation 2.69, standard error mean 0.87, t-value 15.002, sig 0.000. The paired differences of the number of homework submissions before and after were 18.04, standard deviation 1.25, standard error mean 0.21, t-value 6.795, sig 0.000. 0.000. From the before and after t-tests, the data of the six indicators were significantly different before and after the intervention of the teaching model, which likewise indicates the effectiveness of the intervention of the precise teaching model.

#### IV. Conclusion

This paper combines the clustering algorithm to divide learning resources for personalized language translation training, aids language translation teaching in colleges and universities with the advantage of machine translation, and constructs a college language translation accurate teaching mode based on big data.

- 1) Divide the language translation teaching materials of colleges and universities into four levels, and analyze the relationship between the vocabulary factors of language translation and the readability of language trans-

lation precision teaching in colleges and universities. Among them, the p-value of new words and polysemous words is  $5.56 \times 10^{-5}***$  and  $0.000382***$  respectively, which shows a significant influence relationship with language translation precision teaching at the 0.001 level.

- 2) Using cluster analysis to divide the learning resources of language translation in colleges and universities, forming three clustering centers, with the most representative observations in clusters 1, 2, and 3 being words, translation tools, and translation teaching, respectively.
- 3) Statistical data before and after the precise teaching of language translation in colleges and universities, comparing the degree of participation, interaction and concentration before and after the teaching. The length of video learning was 893s in the pre-intervention period of language translation precision teaching and 2830.05s in the post-intervention period, which is a significant increase in the length of video learning. Similarly the assessment of assignment quality in concentration increased from 85.66 to 89.76 points.

In this paper, the proposed big data technology-based language translation precision teaching in colleges and universities improves the accuracy of language translation from the perspective of language and cultural background. The use of clustering methods to classify learning resources and select more representative learning resources, as well as the use of machine translation advantages to assist language translation teaching and enhance the accuracy of language translation teaching, so as to realize the precision and specialization of language translation teaching in colleges and universities from the perspective of cross-cultural perspective.

#### References

- [1] Robertson, E. (2017). Writing in tongues: multilingual poetry and self-translation in france from dada to the present. *Nottingham French Studies*, 56(2), 119-138.
- [2] He, H. (2019). James joyce, language and translation: an interview with fritz senn. *Foreign Literature Studies*, 41(1), 1-12.
- [3] Gould, R. R. (2018, April). Hard translation: Persian poetry and post-national literary form. In *Forum for Modern Language Studies* (Vol. 54, No. 2, pp. 191-206). Oxford University Press.
- [4] Pi, Q. (2023). The information theory of translation. *Information & Culture*, 58(2), 166-179.
- [5] Underhill, J. W. (2017, January). Humboldt in Translation Theory: Pearls of Wisdom, or Splashes in the Ocean?. In *Forum for Modern Language Studies* (Vol. 53, No. 1, pp. 58-70). Oxford University Press.
- [6] Mellinger, C. D., & Jiménez, L. G. (2019). Challenges and opportunities for

Preliminary data	Participation		Interaction degree		Concentration			
	Resource access times	Videoduration(s)	Online interaction times	Class interaction number	Practice submission times	Number of assignments	Practice average score	Quality evaluation
1	12	1250	26	35	15	29	89.63	89.22
2	13	1310	32	26	16	32	83.05	80.12
3	5	986	28	63	19	30	89.11	91.08
.....	.....	.....	.....	.....	.....	.....	.....	.....
28	8	1047	21	22	21	28	92.18	65.02
29	12	2125	22	20	22	42	93.85	69.54
30	10	1152	56	37	25	22	65.79	78.52
Mean value	7.68	893	28.33	25.08	18.04	29.07	75.48	85.66
Late data	Participation		Interaction degree		Concentration			
	Resource access times	Videoduration(s)	Online interaction times	Class interaction number	Practice submission times	Number of assignments	Practice average score	Quality evaluation
1	16	1532	45	50	22	46	86.2	90.11
2	20	2405	50	52	25	48	91.02	89.56
3	10	3010	62	70	18	48	89.15	92.05
.....	.....	.....	.....	.....	.....	.....	.....	.....
28	14	3255	35	42	23	49	93.36	75.43
29	18	2380	48	38	21	50	92.78	73.48
30	23	3360	72	40	22	42	75.43	80.68
Mean value	16.03	2830.05	42.38	48.0	22.92	47.11	88.49	89.76

Table 4: The data comparison of student behavior in and after period

Behavioral data(front - back)	Pair difference					t	Freedom	Sig.
	Mean value	Standard deviation	Standard error mean	The difference is 95% confidence interval				
				Lower limit	Upper limit			
Resource access number	8.35	7.96	0.96	12.56	18.23	18.052	31	0.000
Video duration	1937.05	1698.20	236.69	4201.99	5783.23	31.025	31	0.000
Class interaction number	14.05	9.08	1.65	18.63	23.04	17.968	31	0.000
Online interaction times	22.92	7.03	0.92	6.98	11.03	9.405	31	0.000
Practice submission times	4.88	2.69	0.87	5.96	7.17	15.002	31	0.000
Number of assignments	18.04	1.25	0.21	0.83	1.68	6.795	31	0.000

Table 5: The behavior data matching sample t test

heritage language learners in interpreting courses in the US context. *Revista Signos. Estudios de Lingüística*, 52(101).

[7] Deygers, B., & Malone, M. E. (2019). Language assessment literacy in university admission policies, or the dialogue that isn't. *Language Testing*, 36(3), 347-368.

[8] Brune, K. (2018). The necessities and dangers of translation: brazilian literature on a global stage. *Comparative Critical Studies*, 15(1), 5-24.

[9] Hosington, B. M. (2022, October). On Mediated Translation: A Fruitful Dialogue between Specialists in Translation Studies and Early Modern Translation. In *Forum for Modern Language Studies* (Vol. 58, No. 4, pp. 478-487). Oxford University Press.

[10] Belle, M. A. (2022, October). Introduction: 'Delivered at second hand'? Mediated Translations at the Interface between Translation Studies and Early Modern Studies. In *Forum for Modern Language Studies* (Vol. 58, No. 4, pp. 469-477). Oxford University Press.

[11] Whitehead, P. (2017). Ritchie Robertson and Michael White, eds, Fontane and Cultural Mediation: Translation and Reception in Nineteenth-Century German Literature.

[12] Mole, C. (2024, January). CODE-Switching and polylingual reality: an aesthetics of translation in Nancy Huston's *Trois fois septembre* and *Danse noire*. In *Forum for Modern Language Studies* (Vol. 60, No. 1, pp. 41-58). Oxford University Press.

[13] Sales, M. J. (2019). Tagalog missionary grammars as a translation resource: translation, book history and the production of linguistic knowledge in the spanish philippines. *Comparative Critical Studies*, 16(2-3), 301-322.

[14] Lijewska, A., & Błazkowska, H. (2021). Non-identical cognates yield facilitation in translation—does the way foreign vocabulary is learned affect its processing?. *Poznan Studies in Contemporary Linguistics*, 57(3), 329-358.

[15] Cui, H. (2017). Chinese and foreign film and tv drama translation from the perspective of cross cultural communication. *International Journal of Technology, Management*, 000(001), P.1-3.

[16] Pavlovi, V. (2019). Massive corpora and models of crosscultural communication styles in cognitive linguistics:the case of the n1 v (for) n2to-infinitive construction in english. *Review of Cognitive Linguistics*, 17(1), 29-52.

[17] Yang, X., Hou, J., & Arth, Z. W. (2021). Communicating in a proper way: how people from high-/low-context culture choose their media for communication:. *International Communication Gazette*, 83(3), 238-259.

...