International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-12

https://doi.org/10.70917/ijcisim-2025-0257
Article

A Hidden Markov Model-Based Analysis of the Interaction Patterns between Cognitive Styles and Lexical Productivity in Second Language English Learners

Junling Wang [1,2] and Suthagar Narasuman [2,*]

[1] School of Foreign Languages, Ningxia Medical University, Yinchuan, Ningxia, 750100, China.
[2] Faculty of Education, Universiti Teknologi MARA, Kampus Puncak Alam, Selonger, 42300，Malaysia; suthagar@uitm.edu.my
Junling Wang：sarahwang1234@163.com

Abstract: Vocabulary productiveness refers to learners' ability to use vocabulary correctly and actively in oral or written expressions, and its learning pattern is often influenced by learners' cognitive styles. In this paper, we integrate the feature indicators related to learning effect in learners' behavioral data, predict learners' learning effect through machine learning algorithms, and construct a model of learning effect of learners' cognitive style based on multimodal data. Meanwhile, based on the learner behavioral feature data, the decision tree of learner cognitive styles is established, and the Hidden Markov Algorithm is used to establish a Hidden Markov Model for the preference of each cognitive style, thus proposing the recognition method of learner cognitive styles. With the support of this method, the average score of learners' vocabulary output test in the experimental group increased by 23.08 points compared to the preexperimental period, and showed a statistically significant difference with the average score of the students in the control class (P=0.000). The proposed method of identifying learners' cognitive styles provides a new design idea and empirical basis for teaching vocabulary productiveness and interaction model for second language English learners.

Keywords: lexical productiveness; cognitive style identification; Hidden Markov Model; multimodal data; decision tree; second language English learning

1. Introduction

In English language learning, vocabulary is an independent constituent that expresses meaning, words form sentences that express particular meanings, and any change in any word in a sentence affects the meaning of the sentence. However, compared with vocabulary and phonemic, research on language teaching and acquisition is more inclined to the field of grammar [1]. Influenced by Chomsky's transformational generative grammar in the 20th century, second language acquisition researchers believe that the study of grammar is a shortcut to the study of language [2]. At this stage, grammar still occupies a very important position in classroom teaching, but with the development of cognitive linguistics and systems theory, etc., vocabulary is gradually turning into a research favorite [3-4]. It is generally believed that vocabulary knowledge includes two aspects of vocabulary comprehension (reception/input) and use (output/output). In the past practice of foreign language teaching, language input was highly valued and regarded as an important element of language acquisition [5-6]. At the theoretical level, the input assumption theory even provides theoretical support for this view [7]. From the perspective of cognitive theory, the importance of output for language acquisition cannot be ignored either; through language output, learners are able to learn the language actively and more responsibly, and process the language deeply [8-9]. Students pay attention to output, and conscious control of output can continuously internalize their own language knowledge, overcome language barriers, and improve their language expression [10-11]. In the process of output vocabulary knowledge development, students' conscious attention to language problems can enhance the depth and breadth of language output [12-13].

Since the 1960s, with the focus of language teaching shifting from how teachers teach to how students learn, learners' individual differences and their roles in foreign language teaching have gradually gained the attention of foreign language researchers [14-15]. Among the studies on individual differences, cognitive style is considered to be an important factor affecting the effect of second language acquisition [16]. Many studies have shown that different cognitive styles (field-dependent (FD) versus field-independent (FI), contemplative and impulsive, etc.) have different effects on second language learning. In terms of vocabulary, Ujang [17] (2022) emphasized that vocabulary acquisition is inextricably linked to learners' cognitive styles and multiple regression analysis found that the two work together to promote learners' reading comprehension. Motallebzadeh and Samadi [18] (2017) demonstrated that, in an online collaborative task for English as a second language learning, the contrast between impulsive learners, contemplative learners were more effective for unintentional vocabulary acquisition. Wang [19] (2019) found that both learners, FD and FI, were better than verbal output effects in output vocabulary acquisition in different collaborative output tasks, but FD had weak effects in vocabulary memorization. Xiaoxiao and Sijia [20] (2020) pointed out that there was a significant difference in output task unintentional vocabulary acquisition effect, FI learners were better than FD learners, while the former preferred verbal output and the latter preferred written output. Heidari [21] (2022) stated that under the psycholinguistic point of view, output vocabulary acquisition effect was better for FI learners than for FD learners, whereas the output vocabulary knowledge of FD learners was better than the receptive vocabulary knowledge trial. Mehdipour-Kolour and Ali [22] (2024) tested two cognitive styles of second language learners, FD and FI, both of which contribute to short-term vocabulary memory in a mobile-assisted vocabulary acquisition environment, but FD was more effective. Izmalkova and Blinnikova [23] (2024) explored learners' cognitive strategies during second language vocabulary shuffling by utilizing eye-tracking methods to track learners' gaze times for target vocabulary words and their sentences in reading. Lu et al. [24] (2023) explored the relationship between reading style and cognitive style in language learning with the help of eye-tracking and cognitive style tests, and pointed out that such eye-tracking analysis could potentially help in foreign language material design. It can be seen that with the development of computer technology, the study of cognitive styles has shifted from static to dynamic, exploring the role played by cognitive styles in language learning with the help of eye-tracking and other methods. In addition, Chuk et al [25] (2020) conducted an analysis of eye-movement patterns in a cognitive task with changing cognitive states by switching Hidden Markov Models (HMMs), and identified quantitative criteria for individual differences based on the cognitive behavior/style side of the equation.

HMM is a statistical model commonly used to describe a random sequence containing implicitly unknown states. The model can effectively handle time-series data and has a wide range of applications in various fields such as signal processing, time-series analysis, and natural language processing, which supports the research of language development modeling and cognitive processing processes, enabling new explorations of the interaction patterns of cognitive styles and vocabulary [26-28]. Wang et al. [29] (2018) integrated the HMM and cognitive diagnostic models and embedded them in a in a family of learning models, which can effectively track the evolution of students' skills, be used to combine the assessment of learning intervention models, and make learning adjustments in a timely manner with more assessment feedback. Chen [30] (2022) used HMM for English speech recognition and introduced feed-forward neural networks to optimize the recognition efficiency and accuracy, and this combined algorithm for speech recognition outperforms the single algorithm's recognition effect. Almutiri and Nadeem [31] (2022) pointed out that HMM can model dynamic systems, and in the field of natural language, it can be used to predict hidden labels from the observed words, and complete tasks such as natural language generation and lexical annotation.

This paper firstly explains the construction principle and framework composition of the learner cognitive style learning effect model, especially explains its extraction method for the multimodal data features of learner behavior. Then, it discusses the operation and realization of decision tree algorithm and hidden Markov algorithm in turn, and establishes the cognitive style recognition model of learners. Subsequently, the process of mining and analyzing the cognitive characteristics of students' self-directed learning is demonstrated based on learners' research behaviors. Determine the convergence accuracy of the model and its performance in classifying data of different cognitive styles, and complete the preparation for the application of the proposed cognitive style recognition model. Finally, we set up a teaching application experiment of the proposed model, statistically sorting out and comparing the vocabulary output performance data of the experimental subjects to verify its feasibility.

2. Hidden Markov-Based Recognition of Learners' Cognitive Styles

2.1. Learner Cognitive Style Model of Learning Effectiveness

2.1.1. Modeling

In this study, learning effect modeling of different cognitive styles is carried out through a system dynamics approach. Learning effect is a key research content in the field of education, which is the most intuitive evaluation index of learners' learning effectiveness, and can also indirectly reflect the teaching level of teachers, the advantages and disadvantages of course design, and whether the function of the learning platform is perfect or not. However, the current evaluation of learning effectiveness is mostly carried out by means of post-test questionnaires or behavioral sequence assessment of learners. Although the post-test paper has a certain degree of objectivity, it does not reflect the learning effect in real time, and due to the existence of the forgetting curve, this kind of data has a certain degree of delay. The process of assessing the learner's behavioral sequence is more cumbersome, and it also cannot reflect the learning effect in real time. With the development of technology, biometric technology has been gradually applied to teaching assessment. Physiological data can not only eliminate all kinds of influencing factors and have absolute objectivity, but also reflect the current learning and cognitive state of learners in real time, so as to achieve real-time prediction of learning effects. Therefore, the assessment of learning effect has gradually developed from the previous test papers and behavioral sequence analysis to physiological data analysis. At the same time, previous studies have proved that eye movement, EEG, expression and other modalities of physiological data have a certain function of predicting learning effect, so the combination of multi-modal physiological data to predict learning effect can have more accurate prediction performance. With the continuous development of artificial intelligence and machine learning methods, machine learning algorithms have high feasibility and accuracy in processing and predicting complex data, and are an effective method for processing physiological data and predicting learning effects. At the same time, it is clear that the physiological performance of learners with holistic and analytical cognitive styles will be different during the learning process, and this data difference will affect the prediction performance of learning effects. Therefore, modeling the data of learners with two cognitive styles separately can help to improve the accuracy of learning effect prediction.

Based on the above analysis, this study is dedicated to extracting the feature indicators related to learning effect from a large amount of complex eye movement, EEG and expression data, fusing the related data, and selecting appropriate machine learning algorithms to realize the prediction of learning effect. The model framework is shown in Fig. 1, which is divided into three modules: data acquisition module, data processing module and algorithm application module, which will be introduced in turn. Among them, the feature extraction and data fusion in the data processing module and the algorithm selection in the algorithm application module are the difficult points when constructing the model.

2.1.2. Multimodal data feature extraction

(1) Expression feature extraction

This study utilizes the video-based facial tracking system penFace 2.0 and the Microsoft ONNX framework to analyze and encode the emotional state of the captured facial behavior data. This includes:

1) Extracting relevant features of video frames of students' facial expressions and categorizing different facial action units (AUs) through OpenFace 2.0.

2) Combining the emotional characteristics of the learners in the learning environment, the learners' emotional responses were categorized into six categories: happy, sad, fear, anger, disgust and surprise.

3) Recognize and encode the emotional responses in the video using the ONNX framework.

(4) Calculate the mean and standard deviation of each facial action unit, as well as the frequency of occurrence of different emotional responses.

(2) Eye movement feature extraction

In this study, eye movement features are calculated for each region of interest, including the total number of gaze points, gaze duration, first gaze time, average pupil diameter and its standard deviation, etc., in order to provide data support for analyzing subjects' cognitive processing of different visual elements.

(3) EEG feature extraction

In this study, the time and frequency domain features of 10 bands of data were extracted to refine the key features reflecting brain activity from the raw signals, in order to provide data resources for exploring and understanding EEG signals.

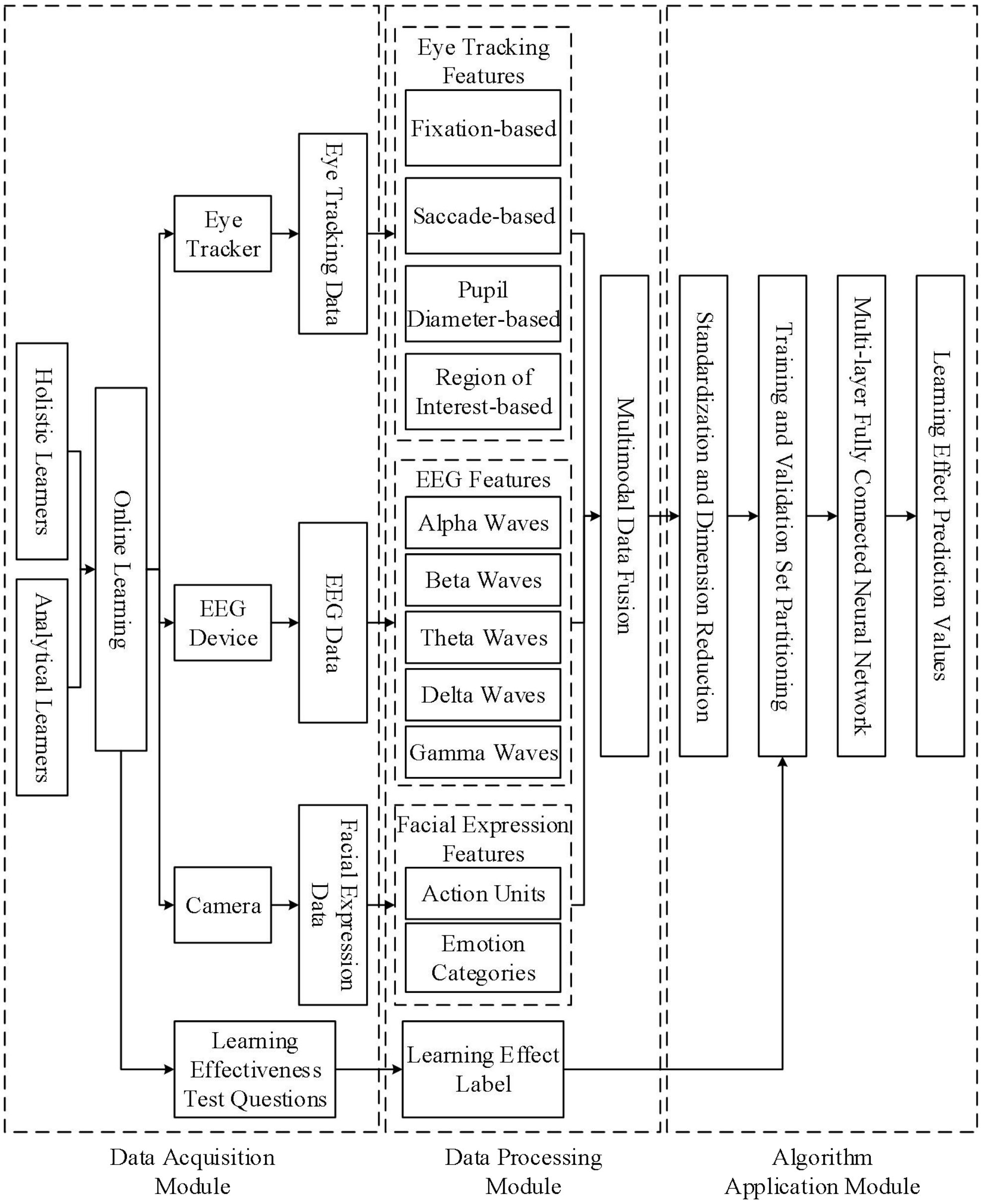


Figure 1. Online learning effect models of different cognitive styles.

2.2. Learner cognitive style identification model

The two recognition algorithms used in this paper are Decision Tree Algorithm and Hidden Markov Algorithm. The reasons are the following 3 points:

(1) Compared with the traditional Bayesian algorithm, the construction process of the decision tree does not rely on any domain knowledge, and the decision process of the decision tree is very intuitive and easy to be understood.

(2) The decision tree algorithm does not consider the sequential learning behavior of the learner, if the learner clicks the three buttons A, B and C sequentially, the decision tree algorithm can only deal with the number of clicks of the buttons and ignores the sequential relationship of the clicks. At this point it is necessary to introduce the HMM algorithm, which is a statistical method that determines the implicit parameters of the process from the observed sequence. So the HMM algorithm can process the sequence of button clicks.

(3) The combination of these two algorithms can be complementary.

2.2.1. Decision tree algorithm and its implementation

(1) C4.5 Decision tree algorithm

Decision tree is a predictive model, it is a tree structure, each node in the tree represents an object, and each forked path represents the value of a possible attribute, and each leaf node corresponds to the value of the object represented by the path from the root node to the leaf node.

The decision tree is split using greedy idea, so the split attribute is selected to find out the attribute that can make all the leaf nodes data the purest, in this paper we use C4.5 decision tree, C4.5 selects the attribute that has the maximum gain rate as the split attribute.

Let D be the division of the training tuple using categories, and the entropy of D is expressed as equation (1):

$$info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i) \tag{1}$$

As shown in Equation (1), where p_i denotes the probability that the i th category occurs in the whole training tuple.

Now suppose the training tuple D is divided by the attribute A, the split information of the division of A to D is Equation (2):

$$split\_info_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \log_2\left(\frac{|D_j|}{|D|}\right) \tag{2}$$

The information gain is equation (3):

$$gain(A) = info(D) - info_A(D) \tag{3}$$

Then the information gain rate is defined as equation (4):

$$gain\_ratio(A) = \frac{gain(A)}{split\_info_A(D)} \tag{4}$$

The higher the information gain rate, the better the split.

(2) Implementation of C4.5 decision tree algorithm

1) Preprocess the behavioral characteristics data of all learners after learning the course to eliminate errors and invalid data.

2) Divide the preprocessed data into two parts, 70% as a training set and 30% as a testing set.

3) Train to construct a decision tree for each cognitive style dimension in each of the four cognitive styles.

2.2.2. Hidden Markov algorithm and its implementation

(1) Hidden Markov model

Hidden Markov model as a statistical analysis model is a type of Markov chain. The difficulty is to determine the hidden parameters of the process from the sequence of observable parameters.

Any Hidden Markov model can be described by the quintuple of equation (5):

$$\lambda = \left(N, M, A, B, \pi\right) \tag{5}$$

N : sequence of observations

M : hidden state

π : initial probability

A : transfer probability

B : emission probability (the probability that a hidden state behaves as a manifest state)

In this study let $\lambda = (\pi, A, B)$ be the parameters of the given HMM, $O = O_1 \cdots O_t$ be the sequence of observed parameters, and the sequence of observed parameters O is the obtained behavioral feature data in this study, since $\lambda = (\pi, A, B)$ parameters are unknown and the sequence of observed parameters is known, the most important step is to train the parameters $\lambda = (\pi, A, B)$ by Baum-Welch algorithm.

(2) Baum-Welch algorithm

As can be seen from the previous section, in the state where only the observation sequence is available, the three most important parameters of the Hidden Markov Model (HMM), $\lambda = (\pi, A, B)$, can only be derived from the training of the Baum-Welch algorithm, which is a special EM algorithm in essence, since there are unobservable data of the hidden variables, i.e., the state sequences, in the HMM, the maximum likelihood estimation cannot be dealt with in the general sense. However, the EM algorithm can be used for maximum likelihood estimation of models containing hidden variables.

(3) HMM algorithm implementation

In this paper, Hidden Markov Models (HMM) are built for the two preferences of each cognitive style, for example, an HMM model is built for the visual preference and the verbal preference when training the visual/verbal dimensions, and the two preferences of the styles are used as the hidden states, and the acquired behavioral data are used as the observation sequences, which are firstly trained using the Baum-Welch algorithm of unsupervised learning to produce $\lambda = (\pi, A, B)$, and then the optimal path is computed by the Viterbi algorithm to derive the probability size of the occurrence of the hidden sequence i.e. the probability size of the occurrence of the two preferences. If the HMM probability of the visual preference is large, the cognitive style is considered to be biased towards the visual type and vice versa.

3. Output-based learning of vocabulary based on learners' cognitive styles

3.1. Mining Analysis of Students' Cognitive Characteristics for Independent Learning

Combined with the analysis above, taking "notification-study course-discussion-quiz-assignment" as the route of learning activities, we propose that the research behaviors in this paper are: (B1) notification, (B2) text material viewing, (B3) video resource viewing, (B4) learning behaviors in exchanging tasks, (B5) communicating behaviors in assigning tasks, (B6) behaviors in testing, and (B7) homework submission and viewing. B6) Behavior in Testing, (B7) Assignment Submission and Viewing.

Behavioral data on the course learning behaviors of the test students were collected and coded according to the behavioral framework of cognitive properties. The coded behavioral data were then subjected to frequency statistics and residual analysis. The behavioral frequency statistics of a randomly selected student are shown in Figure 2, and the results of the behavioral frequency residual analysis are shown in Figure 3.

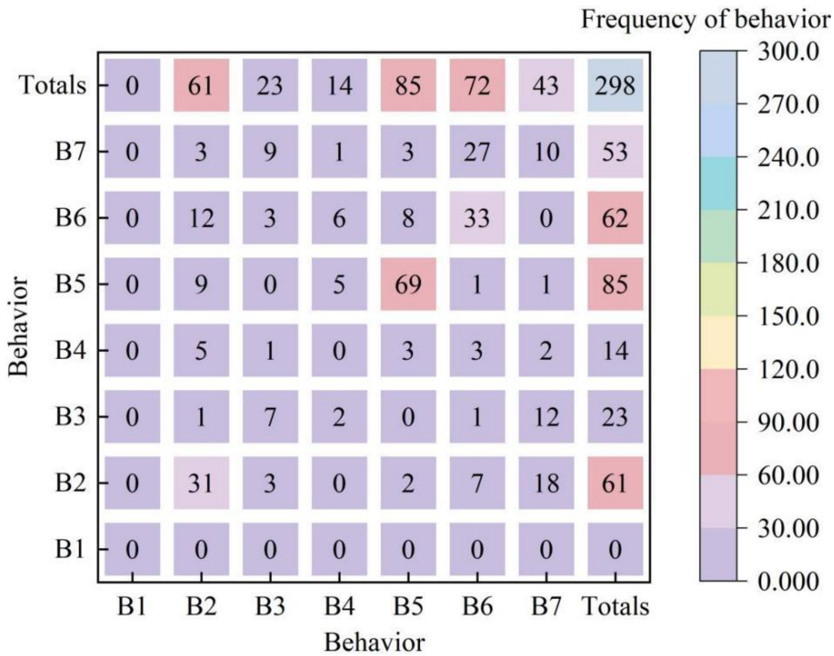


Figure 2. Statistics on behavior frequency.

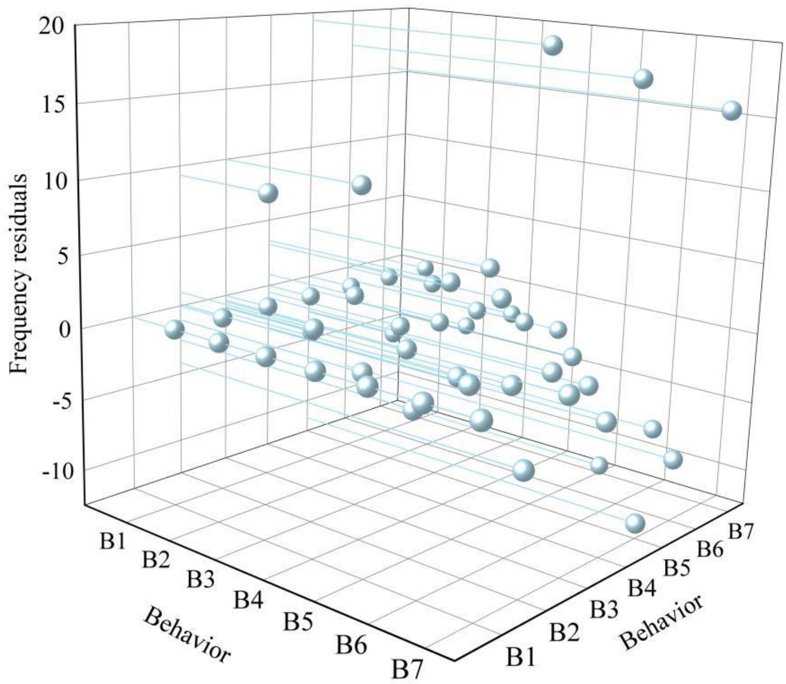


Figure 3. Residual analysis of behavioral frequency.

The adjusted residual values for the overall sequence of student behaviors are shown in Table 1.The paths were statistically significant when the residual values were >1.96, where only the learning behaviors in the (B4) exchange task to the communication behaviors in the (B5) assignment task were statistically significant.

Table 1. The residual value of the overall behavior sequence of the student.

Residual	B1	B2	B3	B4	B5	B6	B7
B1	25.69	18.24	15.81	3.75	-1.71	-1.1	-7.65
B2	5.21	2.75	41.58	3.32	-11.81	-7.06	-36.46
B3	16.36	28.29	29.8	-4.86	-11.95	5.86	-32.15
B4	3.19	5.55	-5.13	35.47	13.35	15.54	-14.22
B5	-1.68	-12.26	-11.76	2.99	14.07	16.49	11.48
B6	-1.39	-6.8	-10.91	-4.4	-11.4	48.4	8.25
B7	-6.59	-35.47	-34.35	-12.06	-35.13	-35.44	4.41

3.2. Performance test of cognitive style recognition model

Based on the experimental dataset, the proposed model is trained for 5000 iterations, and the statistical combing of the network error performance changes during the training process are shown in Fig. 4, where the X-axis is the number of iterations and the Z-axis represents the convergence accuracy (mean square error). After 1000 iterations, the model has begun to converge, and when the number of training times reaches 5000 times its convergence accuracy reaches 0.019874, which is close to the desired goal and the overall accuracy meets the requirements.

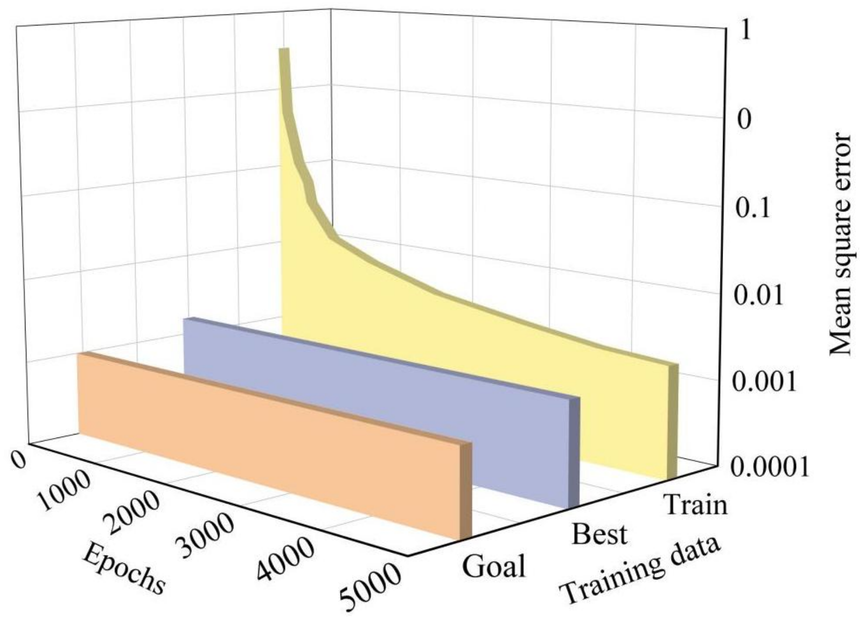
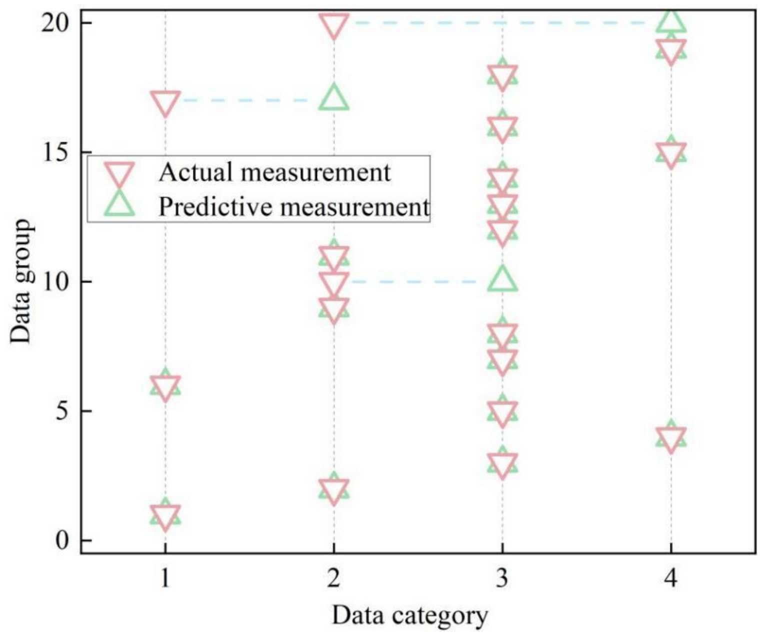
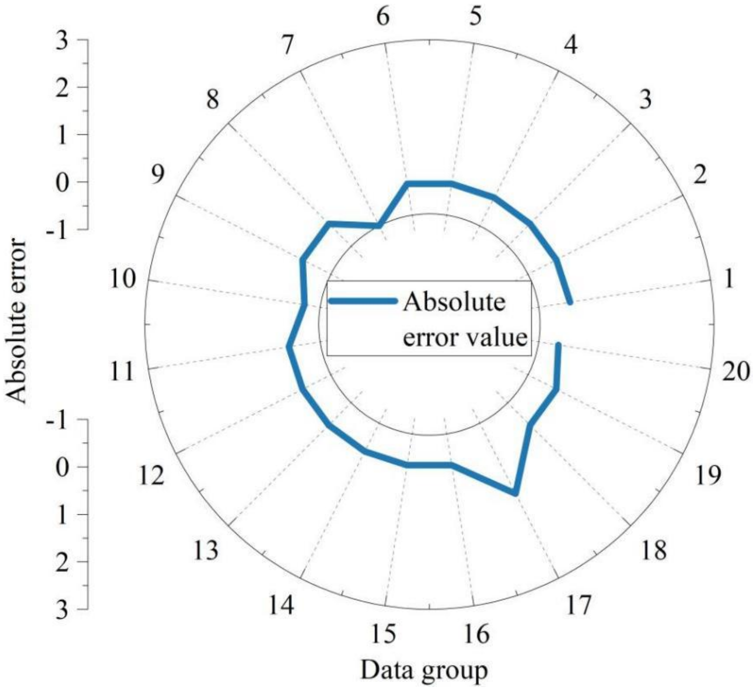


Figure 4. Network error performance during the model training process.

Twenty groups of different actual learning cognitive data in the experimental dataset are selected, and the proposed model is used to classify them based on their attributes, and the comparative validation data are shown in Fig. 5(a), and the absolute errors are shown in Fig. 5(b). The model's classification performance error is centered around 0.00, and the overall control is in the interval of [-1,2], and the classification results can be better close to the learners' actual cognitive style performance.



(a) Data validation



(b) Absolute error

Figure 5. The verification results of the model.

3.3. Application validation of the cognitive style recognition model

3.3.1. Experimental Objects

Two natural classes of first-year college students were used as the subjects of this study, with Class A (45 students) as the experimental class and Class B (45 students) as the control class. Both classes were taught by the same English teacher, and the content and pace of teaching remained basically the same, with vocabulary teaching being the only variable. English vocabulary teaching with the aid of this paper's modeling approach was used in the experimental class A, and conventional teaching was used in the control class B. Through the pre-test, the test results ($P=0.754>0.05$) showed that there was no significant difference in the level of vocabulary productiveness between the two classes; therefore, the students in the two classes were eligible to carry out the teaching experiment.

3.3.2. Experimental tools

In this paper, students' vocabulary outputability level was tested from the perspectives of vocabulary outputability amount and outputability word richness in both pre-test and post-test stages. The vocabulary output level test papers used have a total of five word levels: 2000-word level, 3000-word level, 5000-word level, college word list and 10000-word level. In view of the learning situation of first-year college students and the requirements of the college English curriculum standards, college word level test paper was used in this study.

3.3.3. Analysis of output vocabulary results

Before the beginning of the experiment, the students of the two classes were given a pre-test on vocabulary output measures, the test paper was based on a percentage system, and the descriptive statistics of the test results are shown in Table 2 and the results of the independent samples t-test are shown in Table 3.The mean score of the students in the experimental class in the pre-test test was 55.41, and the mean score of the control class was 56.09, and the difference in the scores of the two classes was only 0.68. Its significance was $0.891>0.05$, $P=0.754>0.05$, and the 95% confidence interval of the difference was [-6.581,4.906] containing 0. That is, the lexical output measures of the experimental class and the control group were almost the same at the pre-test level, and the chi-squaredness of the variance was in accordance with the requirements of the normality test, and there was no significant difference in line with the experimental requirements.

Table 2. Descriptive statistics of the pre-test scores.

Class	Experimental class	Control class
Number of cases	45	45
Average value	55.41	56.09
Standard deviation	13.647	13.025
Average standard error	2.162	2.034

Table 3. Independent sample *t*-test of the pretest results.

Pre-test			Assuming equal variance	Equal variance is not assumed
Levin's equivalence test of variance	F		0.009	
	P		0.891	
t			-0.254	
df			87	84.52
Mean equivalence t-test	Sig. (Two-tailed)		0.754	0.754
	Average value difference		-0.728	-0.728
	Standard error difference		2.189	2.189
	95% confidence interval of difference	Lower limit	-6.581	-6.581
		Upper limit	4.906	4.906

During the semester-long experimental teaching, two post-tests were conducted on the experimental subjects at the mid-term and the end of the semester, and the results of the independent samples t-tests of the results of the two subtests are shown in Table 4. In the mid-term test, the average score of the experimental class students was 68.51, and that of the control class students was 78.49, which showed a certain significant difference (P=0.025<0.05). By the final exam, the mean score of students in the experimental class rose to 78.49, which not only increased by 8.98 points compared with the midterm exam and 23.08 points compared with the preexperimental period, but also separated from the control class students (63.25) by as much as 15.24 points, with a statistically significant level of difference (P=0.000). The Hidden Markov Model-based Learner Cognitive Style Identification method assists in enhancing learners' vocabulary productivity by linking their multimodal behavioral data to predict their learning outcomes, accurately identifying their cognitive styles of learning, and guiding the interaction mode of vocabulary productivity instruction to match learners' cognitive characteristics.

Table 4. Statistics of the results of the two post-tests.

Time	Mid-term exam		Final exam	
Class	Experimental class	Control class	Experimental class	Control class
Sample size	45	45	45	45
Mean	69.51	58.42	78.49	63.25
SD	9.51	10.56	5.03	9.34
t	2.41		0.025*	
p	3.96		0.000***	

Note: * p<0.05,** p<0.01,*** p<0.001.

4. Conclusion

In this paper, by fusing multiple behavioral data features of second language learners in vocabulary productive learning, we realize the accurate prediction of their learning effects. Combining the decision tree algorithm and the Hidden Markov Algorithm, a learner cognitive style recognition model is proposed, which not only has good convergence accuracy (0.019874) and cognitive style categorization error is controlled at [-1,2], but also shows reliable application effects in the vocabulary output teaching interactive mode. In the actual teaching of vocabulary output, the learners under its tutoring not only have a mean score of vocabulary output before and after the experiment of 23.08, but also show a statistically significant difference with the mean score of the learners under the traditional mode of teaching (p = 0.000).

References

1. Bardovi-Harlig, K., & Comajoan-Colomé, L. (2022). The relation of second language acquisition, instructed second language acquisition, and language teaching from the lens of second language tense-aspect. Language Teaching, 55(3), 289-345.
2. Chomsky, N., Gallego, Á. J., & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges. Catalan Journal of Linguistics, 229-261.
3. Dodigovic, M., Jeaco, S., & Wei, R. (2017). Trends in vocabulary research. TESOL International Journal, 12(1), 1-6.
4. Zhao, H., Yau, T. S. H., Li, K., & Wong, N. N. Y. (2018). Polysemy and conceptual metaphors: A cognitive linguistics approach to vocabulary learning. What is applied cognitive linguistics, 257-286.
5. Alahmadi, N. S. (2019). The role of input in second language acquisition: An overview of four theories. Bulletin of Advanced English Studies, 3(2), 70-78.
6. Carroll, S. E. (2017). Exposure and input in bilingual development. Bilingualism: Language and Cognition, 20(1), 3-16.
7. Shahini, A. (2025). Input Hypothesis in Language Learning. In Language Learning Theories: A Student's Guide (pp. 65-80). Cham: Springer Nature Switzerland.
8. Hosseini, S. M. H. (2019). Cognitive socio-political language learning theory, multiple input-output hypothesis and competitive team-based learning. Theory and Practice in Language Studies, 9(4), 411-423.
9. Nelson, K. E. (2021). Some observations from the perspective of the rare event cognitive comparison theory of language acquisition. In Children's language (pp. 289-331). Psychology Press.
10. Dehqan, M., & Mohammadi Amiri, M. (2017). Collaborative output tasks and their effects on learning English comparative adjectives. Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills), 36(1), 1-26.
11. Guo, X. (2022). A Study on Cross-Media Teaching Model for College English Classroom Based on Output-Driven Hypothetical Neural Network. Computational Intelligence and Neuroscience, 2022(1), 5283439.
12. Noroozi, M., & Siyyari, M. (2019). Meaning-focused output and meaning-focused input: The case of passive and active vocabulary learning. Cogent Arts & Humanities, 6(1), 1651443.
13. Allal-Sumoto, T. K., Miyoshi, K., & Mizuhara, H. (2023). The effect of productive vocabulary knowledge on second language comprehension. Frontiers in Psychology, 14, 1049885.
14. Courtney, L., Graham, S., Tonkyn, A., & Marinis, T. (2017). Individual differences in early language learning: A study of English learners of French. Applied linguistics, 38(6), 824-847.
15. Zhang, J., & Zhang, L. J. (2025). Exploring the profiles of foreign language learners' writing self-regulation: Focusing on individual differences. Reading and Writing, 38(6), 1533-1559.
16. Zhuo, H. (2024). The influence of cognitive style differences on foreign language learning strategies. In SHS Web of Conferences (Vol. 200, p. 02027). EDP Sciences.
17. Ujang, S. (2022). Correlational study of vocabulary mastery and cognitive learning style in vocational school. Journal of English Language Teaching Innovations and Materials (JELTIM), 4(1), 102-126.
18. Motallebzadeh, K., & Samadi, F. (2017). The effect of using online collaborative tasks on incidental vocabulary learning of impulsive vs. reflective Iranian EFL learners. International Journal of Applied Linguistics & English Literature, 6(6), 127.
19. Wang, Q. (2019). EFL incidental vocabulary acquisition and retention: Effects of different collaborative output tasks on field independent/dependent learners. Journal of Asia TEFL, 16(4), 1201.
20. Xiaoxiao, L., & Sijia, X. (2020). The Effect of Cognitive Style on Non-English Majors' Incidental Vocabulary Acquisition through Reading. INNOVATION AND MANAGEMENT, 804.
21. Heidari, K. (2022). A Psycholinguistic Look at the Role of Field Dependence/Independence in Receptive/Productive Vocabulary Knowledge: Does it Draw a Line?. Journal of Psycholinguistic Research, 51(6), 1393-1408.
22. Mehdipour-Kolour, D., & Ali, M. B. (2024). Examining the effects of two cognitive styles (field dependence vs. field independence) on learners' mobile-assisted vocabulary acquisition. Journal of China Computer-Assisted Language Learning, 4(2), 225-246.
23. Izmalkova, A., & Blinnikova, I. (2024). Cognitive strategies of second language vocabulary inferencing: An eye tracking study. Education Sciences, 14(5), 490.
24. Lu, W. L., Dostálová, N., Lacko, D., Šašinková, A., & Šašinka, Č. (2023). From reading style to cognitive style and its possible application: an eye-tracking and CFT approach. Asian-Pacific Journal of Second and Foreign Language Education, 8(1), 22.
25. Chuk, T., Chan, A. B., Shimojo, S., & Hsiao, J. H. (2020). Eye movement analysis with switching hidden Markov models. Behavior research methods, 52(3), 1026-1043.
26. Wang, M., Abdelfattah, S., Moustafa, N., & Hu, J. (2018). Deep Gaussian mixture-hidden Markov model for classification of EEG signals. IEEE Transactions on Emerging Topics in Computational Intelligence, 2(4), 278-287.
27. Hanif, M., Sami, F., Hyder, M., & Ch, M. I. (2017). Hidden Markov model for time series prediction. Journal of Asian Scientific Research, 7(5), 196-205.
28. Leopold, H., van der Aa, H., Offenberg, J., & Reijers, H. A. (2019). Using hidden Markov models for the accurate linguistic analysis of process model activity labels. Information Systems, 83, 30-39.

29. Wang, S., Yang, Y., Culpepper, S. A., & Douglas, J. A. (2018). Tracking skill acquisition with cognitive diagnosis models: A higher-order, hidden Markov model with covariates. Journal of Educational and Behavioral Statistics, 43(1), 57-87.
30. Chen, Y. (2022). A hidden Markov optimization model for processing and recognition of English speech feature signals. Journal of Intelligent Systems, 31(1), 716-725.
31. Almutiri, T., & Nadeem, F. (2022). Markov models applications in natural language processing: a survey. Int. J. Inf. Technol. Comput. Sci, 2, 1-16.