

# **Deep Integration of AI and Multimodal Theory in College English Teaching: Design, Practice and Effect Verification**

Haiyan Yin

College of Humanities, Liaoning Communication University, Shenyang, Liaoning, China

## **Abstract**

**In the context of smart education, college English instruction is shifting from knowledge delivery to competency cultivation and humanistic development. Based on multimodal discourse analysis and artificial intelligence (AI) technologies, this paper builds an innovative college English teaching framework supported by intelligent tools, diversified input, interactive output and data-based evaluation. Via theoretical exploration, teaching design, empirical practice and effect analysis, this research explores the implementation path and application value of the integrated model. Results show that the joint application of AI and multimodal theory significantly promotes students' English application skills, autonomous learning capacity, cross-cultural communication awareness and critical thinking competence. This study offers theoretical support and practical references for the deep integration of information technology and college English courses, and facilitates the high-quality advancement of college English education.**

## **Keywords**

**Artificial Intelligence; Multimodal Discourse Analysis; College English Instruction; Smart Teaching; Teaching Reform.**

## **1. Introduction**

Driven by big data, artificial intelligence, virtual reality and natural language processing, global higher education has entered a new phase of digital and intelligent transformation. As a fundamental public course, college English plays a vital role in improving students' comprehensive language ability, shaping cross-cultural understanding and strengthening international communication skills.

In recent years, multimodal discourse analysis has been widely applied in foreign language education[3]. This theory holds that meaning construction is completed through the cooperation of various semiotic resources such as language, pictures, sounds, movements and colors. It provides a new theoretical perspective for optimizing teaching input and designing diversified classroom activities. Meanwhile, AI technologies including intelligent speaking assessment, adaptive learning systems, intelligent Q&A and corpus analysis provide strong technical support for personalized and precise teaching[4].

This research aims to: clarify the theoretical implication and internal connection of AI and multimodal theory in college English teaching; construct a complete and operable intelligent multimodal teaching system; verify the actual effect through teaching experiments; summarize practical problems and put forward targeted improvement strategies.

## **2. Theoretical Foundation and Research Overview**

### **2.1. Multimodal Discourse Analysis**

Multimodal discourse analysis is developed from Halliday's social semiotics. It points out that language is not the only means of meaning-making in modern communication. People often use images, sounds, colors, gestures and other symbolic resources to convey information together. In college English classrooms, teaching discourse includes multiple modes: linguistic mode (vocabulary, sentences, discourse), visual mode (courseware, pictures, videos), auditory mode (listening materials, pronunciation, music), behavioral mode (body language, group interaction) and digital mode (online resources, virtual scenes).

Applied to English teaching, this theory helps break the limitation of the text-only mode, create authentic and rich language situations, stimulate students' multi-sensory participation, enhance memory and deep comprehension, and lay a theoretical foundation for teaching design and activity arrangement[3].

### **2.2. AI-supported English Language Teaching**

Artificial intelligence provides stable technical support for the reform of college English teaching[4]. Natural language processing realizes automatic assessment of listening and speaking, intelligent writing correction and discourse semantic analysis. Big data technology collects and analyzes learning behaviors, tracks learning progress and predicts learning effects, so as to provide personalized resource recommendation. Virtual reality creates immersive cross-cultural communication environments for students to experience real language usage contexts.

Compared with traditional teaching, AI-assisted English instruction has obvious advantages in personalized feedback, large-scale interactive evaluation, adaptive learning paths and low-cost repeated practice. It effectively reduces teachers' repetitive work and improves students' learning initiative and efficiency.

### **2.3. Research Overview of College English Teaching Reform**

In recent years, scholars at home and abroad have carried out extensive research on college English reform. Foreign studies focus on computer-assisted language learning, intelligent tutoring systems and the effect of multimodal input, emphasizing the positive role of technology in language acquisition. Domestic studies mainly involve curriculum structure, teaching model innovation, evaluation system construction and the application of new educational technologies[3].

Nevertheless, most studies focus on a single technology or theory, lacking systematic research on the deep integration of AI and multimodal theory, as well as complete teaching cases and empirical evidence. Therefore, this research takes the integration of the two as a breakthrough point to carry out systematic teaching design and practical research.

## **3. Construction of the Intelligent Multimodal Teaching Model**

### **3.1. Core Design Principles**

#### **3.1.1. Learner-centered Principle**

The model takes students' learning needs, cognitive level and development goals as the starting point. It reduces teachers' long-term lecturing and increases students' independent exploration, cooperative interaction and reflective output.

### **3.1.2. Multimodal Synergy Principle**

Make full use of texts, images, audio, video and virtual scenes to form a complementary and three-dimensional teaching environment, so as to improve the efficiency of meaning construction.

### **3.1.3. Data-driven Intelligence Principle**

Use AI tools to complete data collection, process monitoring, precise evaluation and personalized guidance, making teaching decisions scientific and learning effects continuously optimized[4].

### **3.1.4. Output-oriented Principle**

Take language output as the core goal, design progressive tasks from imitation to creation and from simple expression to complex discussion, so as to cultivate students' comprehensive application ability.

## **3.2. Structural Framework**

The intelligent multimodal teaching model consists of four parts: intelligent pre-class preparation, multimodal in-class instruction, AI-supported after-class expansion and comprehensive evaluation system.

### **3.2.1. Intelligent Pre-class Preparation**

Teachers issue multimodal preview materials such as micro-videos, audio, picture brochures and online articles through the smart teaching platform. The platform collects students' preview data, analyzes key and difficult points and pushes personalized tasks. Students finish vocabulary tests, listening practice and situational perception independently.

### **3.2.2. Multimodal In-class Instruction**

In class, teachers use courseware, videos, situational pictures and real objects to conduct multimodal input[3]. Group discussions, role-plays, speeches and debates are organized to promote active language output. AI tools are used to realize real-time oral evaluation and instant feedback[4].

### **3.2.3. AI-supported After-class Expansion**

After class, the platform pushes personalized expanded resources according to students' classroom performance. Students finish writing with the help of AI correction tools, take part in online cross-cultural communication activities and conduct independent inquiry learning.

### **3.2.4. Comprehensive Evaluation System**

A diversified evaluation system combining process assessment and summative assessment is built. Process assessment includes preview completion, classroom performance, after-class assignments and interactive participation. Summative assessment includes mid-term and final examinations. AI assists in data statistics and effect analysis to realize comprehensive and objective evaluation.

## **4. Teaching Practice and Effect Analysis**

### **4.1. Research Subjects and Experimental Design**

The research subjects are two parallel non-English major classes from a university. The experimental class adopts the intelligent multimodal teaching model, while the control class uses the traditional teaching method. The experiment lasts one semester. Research methods include questionnaire survey, pre-test and post-test comparison, classroom observation and semi-structured interview [1][2].

## **4.2. Data Collection**

### **4.2.1. Test Score Collection**

Pre-test and post-test scores of listening, speaking, reading, writing and comprehensive ability of the two classes are collected[1][2].

### **4.2.2. Questionnaire Survey**

A questionnaire is used to investigate students' learning attitude, autonomous learning awareness, classroom participation and satisfaction with the teaching model[1][2].

### **4.2.3. Classroom Observation**

Students' participation frequency, interaction quality, concentration degree and task completion are recorded in class.

### **4.2.4. Student Interview**

Semi-structured interviews are conducted with some students to understand their real feelings and suggestions.

## **5. Existing Problems and Optimization Strategies**

### **5.1. Existing Problems**

#### **5.1.1. Uneven Digital Literacy of Teachers**

Some teachers are not proficient in using intelligent tools and designing multimodal activities, which affects the overall teaching effect[3].

#### **5.1.2. Insufficient Personalization**

Personalized recommendation only stays at the level of difficulty adjustment, lacking customized design for learning styles and interests.

#### **5.1.3. Resource Redundancy**

Excessive use of videos and pictures leads to information overload and distracts students' attention from key learning points.

#### **5.1.4. Unbalanced Hardware and Network Conditions**

Individual students are affected by poor equipment and network, which reduces their online learning experience[5].

### **5.2. Optimization Strategies**

#### **5.2.1. Strengthen Teacher Training**

Carry out regular training on AI teaching tools and multimodal course design to improve teachers' digital teaching ability.

#### **5.2.2. Deepen Personalized Teaching**

Optimize the platform algorithm to realize personalized learning path planning and resource recommendation based on students' cognitive level and learning preferences.

#### **5.2.3. Standardize Multimodal Resource Application**

Formulate resource usage standards to ensure that multimodal materials serve teaching objectives and avoid blind application.

#### **5.2.4. Improve Infrastructure**

Upgrade campus network and teaching equipment, provide offline resources for students with poor conditions to ensure learning fairness[5].

## 6. Discussion

The integration of artificial intelligence and multimodal theory provides a new path for college English teaching innovation. This model breaks the traditional teacher-centered classroom pattern, creates a rich and immersive language learning environment, and effectively improves students' learning interest, participation and practical language ability. Empirical results confirm the effectiveness and feasibility of the model, and also provide a useful reference for other foreign language courses[1][2].

In practice, technology and theory are only supporting tools. The core of teaching reform lies in people. Teachers' teaching concepts, instructional design ability and students' learning autonomy determine the final effect. Therefore, while promoting the application of technology, we should always adhere to the learner-centered concept, pay attention to the combination of technology and humanism, and avoid over-reliance on technology at the expense of emotional communication and humanistic care.

In addition, the intelligent multimodal teaching model needs continuous iteration and optimization. With the development of large models, metaverse and emotional robots, college English teaching will have more possibilities. In the future, researchers can further explore the application of large model AI in intelligent Q&A, situational dialogue and writing guidance, build metaverse cross-cultural communication spaces, and realize higher-level intelligent and personalized teaching [4].

## 7. Summary

Taking college English teaching innovation as the research goal, this study constructs an intelligent multimodal teaching model integrating artificial intelligence technology and multimodal discourse analysis theory. Through theoretical analysis, framework construction, empirical verification and effect comparison, this paper proves that the model can effectively improve classroom teaching efficiency, and enhance students' comprehensive English application ability, autonomous learning awareness and cross-cultural communication quality. The main conclusions are as follows: first, the integration of AI and multimodal theory conforms to the law of language acquisition and the development trend of education, with solid theoretical rationality; second, the four-in-one teaching framework including pre-class, in-class, after-class and evaluation is complete and operable; third, practical results show that the model has significant effects and is highly recognized by students; fourth, the model still needs to be optimized in teacher training, personalization depth and resource standardization. This study enriches the theoretical research of college English teaching reform, provides practical cases and operation schemes for front-line teachers, and helps promote the deep integration of information technology and foreign language education. In the future, with the continuous development of digital and intelligent education, college English teaching will move towards a more intelligent, personalized, situational and humanized direction, and make greater contributions to cultivating high-quality talents with international competitiveness.

## Acknowledgments

The authors would like to thank all the students and teachers who participated in the teaching practice, as well as the reviewers for their valuable comments and suggestions.

## References

- [1] J. van der Geer, J.A.J. Hanraads, R.A.J. Lupton, The art of writing a scientific article, *J. Sci. Commun.* 163 (2000) 51-59.
- [2] G.R. Mettam, L.B. Adams, How to prepare an electronic version of an article, in: B.S. Jones, R.Z. Smith (Eds.), *Introduction to the Electronic Age*, E-Publishing Inc., New York, 1999, pp. 281-304.
- [3] Li M. *Multimodal Discourse Analysis and English Teaching Innovation*. Foreign Language Education Press, Beijing, 2022.
- [4] Zhang H. Artificial Intelligence-assisted College English Teaching Reform and Practice. *Journal of Foreign Languages*, 45(2023) 78-85.
- [5] Information on: <https://www.moe.gov.cn>.