

# Multimodal Applications in the Teaching of Modern Chinese: An Empirical Study Based on ELAN

Jing Zhang

Lanzhou University of Finance and Economics, Lanzhou 730000, China

**Abstract:** This study is based on multimodal discourse theory and multimodal pedagogy. It selects "Modern Chinese" MOOC courses offered by five universities, establishes a small corpus, and uses ELAN software to annotate and analyze the use of multimodal discourse by teachers in videos and the collaborative relationships between different modalities during teaching segments. The study aims to provide effective application strategies for multimodal teaching methods in offline "Modern Chinese" courses. Through these strategies, issues such as the abstract nature of course content, lack of teacher-student interaction, and low student engagement can be addressed, optimizing teaching effectiveness

**Keywords:** Modern Chinese; multimodal discourse; teaching model; ELAN

## 1. Introduction

With the rapid advancement of technology, human communication methods have fundamentally changed, shifting from single-sensory communication to multimodal interactions involving multiple senses. Multimodal Discourse Analysis (MDA), an emerging interdisciplinary research field, can trace its academic discussions back to the 1970s. French semiotician Roland Barthes systematically analyzed the meaning of images in *Rhetoric of the Image*, laying the theoretical foundation for visual image grammar research. In 1996, Kress and van Leeuwen developed a visual grammar analysis framework based on systemic functional linguistics and social semiotics, formally introducing the concept of "multimodal discourse." [1] Charles Forceville (2009) [2] systematically studied multimodal metaphors, illustrating their applications in advertisements, comics, films, conversations, and music through extensive case studies.

In China, research on multimodal discourse theory began relatively late. Zhanzi Li (2003) [3] introduced the concept of "multimodality" into Chinese academia, providing new perspectives for subsequent studies. Zhuanglin Hu (2007) distinguished between multimodal and multimedia concepts, emphasizing the importance of cultivating multimodal communication competence [4]. Zhang Delu (2009) constructed a theoretical framework for multimodal discourse analysis based on systemic functional linguistics, deeply exploring the interactions and relationships between different modalities [5]. These studies provide essential theoretical support for understanding the complexity and dynamics of multimodal discourse and lay a foundation for further empirical research.

Currently, the application of multimodal theory has expanded into fields such as linguistics, information science, psychology, and the arts, gaining increasing attention in teaching research and practice. Yueguo Gu (2007) [6] was the first to distinguish between multimedia learning and multimodal learning, conducting an in-depth analysis of the two concepts. Delu Zhang (2010) [7] explored the principles of modality selection and application in foreign language classroom teaching design. Ya Li (2018) proposed multimodal improvement strategies from four perspectives: teaching models, teaching methods, teachers, and teaching resources [8]. Shuai Yuan (2020) [9] verified the effectiveness of multimodal teaching models in Chinese reading and writing courses through teaching practice. Namei Li (2022) [10] explored the application of multimodal theory in online teaching, offering targeted suggestions for lesson preparation, pre-study, and interactive methods under the guidance of

multimodal theory. Guoxue Zhang (2023) [11] conducted a comparative analysis of offline and online elementary grammar teaching videos using multimodal discourse analysis and ELAN software. The study found that various discourse modalities collaborate in both settings, though not equally weighted.

Previous research on multimodal teaching design has primarily focused on English language teaching and international Chinese teaching. In terms of methodology, most studies have employed qualitative analysis. However, Modern Chinese, as a core course for Chinese Language and Literature majors, a foundation for linguistic studies, and a practical course for mastering the national language, still faces challenges like abstract content, outdated teaching methods, and low student engagement. There is an urgent need for new teaching models to guide instructional design.

Therefore, this study builds on the research findings of multimodal discourse theory and multimodal pedagogy, selecting Modern Chinese MOOC course resources to establish a small-scale course corpus. using ELAN software to annotate and analyze teachers' use of language, body movements, and multimedia in classroom videos. The study conducts a quantitative analysis of modality types, the collaboration between different modal resources, and their effectiveness, aiming to provide guidance for teachers in effectively using multimodal and multimedia resources and offering references for improving the quality and effectiveness of both online and offline Modern Chinese courses.

## 2. Research Design

### 2.1 Research Subjects and Content

The online video resources for the "Modern Chinese" course are relatively limited. This study selected "Modern Chinese" courses offered by five universities on two MOOC platforms, China University MOOC and Zhihuishu. To facilitate a horizontal comparison of the use of multimodal discourse in different classroom settings, relevant teaching content suitable for the application of multimodal teaching methods was selected from each MOOC. The total duration of the teaching videos included for statistical analysis is approximately 61 minutes. The specific research subjects and content are shown in Table 1.

**Table 1:** Online Teaching Resources for "Modern Chinese".

Number	Offering Institutions	Course Platforms	Course Content	Video Duration/s	Number of Instructors
1	Heilongjiang University	China University MOOC	The physical properties of speech	686	1
2	Shandong University of Aeronautics	Zhihuishu	Consonants	600	1
3	Xi'an University	Zhihuishu	The basic properties of speech	970	1
4	Zhejiang University	China University MOOC	The properties of speech	860	1
5	Zhengzhou Normal University	China University MOOC	The place of articulation of initial consonants	542	1

To facilitate statistical analysis, the following text will use T1-T5 to refer to the Modern Chinese courses from five universities

### 2.2 Research Methods and Tools

This study adopts a mixed-methods approach, combining both quantitative and qualitative research methods. ELAN 6.7 software is used as the analytical tool to segment, annotate, and extract

data from the selected online courses. ELAN (EUDICO Linguistic Annotator) is a multimodal corpus analysis tool developed by the Max Planck Institute for Psycholinguistics in the Netherlands. It allows users to create annotation layers through a customizable coding scheme, enabling synchronized multi-layer annotations of audio and video files. These annotations can be interlinked, facilitating in-depth analysis, and the analysis results support output in various formats.

### 2.3 Design of the Coding Scheme

The use of ELAN software for multimodal discourse analysis requires the coding and annotation of the selected corpus. According to the coding scheme, the language types in the videos are classified and annotated in a hierarchical manner.

Based on Delu Zhang's (2009) [5] multimodal discourse media system, multimodal discourse is divided into two primary categories: linguistic media and non-linguistic media. Linguistic media are further subdivided into pure language and accompanying language. Pure language refers to sound symbols and text symbols, while accompanying language refers to elements such as "volume, pitch, frequency, accent voice, as well as the shape, size, and spatial arrangement of fonts," which serve to assist, complement, and enhance the linguistic media. Non-linguistic media are classified into body media and non-body media. Body media include actions related to the head, arms, legs, trunk, etc., while non-body media include tools and environmental factors. In accordance with Delu Zhang's multimodal classification framework, this study has made adaptive adjustments to suit the online teaching environment, further refining the modal types into nine categories: voice, text, accompanying voice, accompanying text, facial expressions, gaze, gestures, images, and audio-visual media. Due to the limited range of movement during online course recordings, this study excluded the categories of "posture" and "shaking" from the body media in the coding scheme. Additionally, based on cognitive linguistics theory, hand movements were categorized in greater detail. Furthermore, as teachers are unlikely to move frequently during the lecture, the coding scheme does not include environmental factors such as spatial relationships and classroom layout. The specific coding scheme is shown in Table 2.

**Table 2:** Multimodal Coding Table.

Modal Types	Media types	Coding	Specific Meaning
<b>Linguistic</b>	Pure Voice	LV	The teacher's spoken explanation
	Language (L) Text	LT	PPT text, board writing, card text, etc.
	Accompanyin g Language (La)	Accompanying voice LAV	The rise and fall of pitch, stress, pauses, and speech rate
		Accompanying text LAT	Changes in font, color, size, or the addition of other symbols
		Facial expressions BF	Facial expressions such as smiling, frowning, and other changes
<b>Non- linguistic</b>	Body Movements (B)		Looking at the multimedia or the blackboard (whiteboard)
		Gaze BEB	
		BEF	Looking at the students
		GV	Iconic gestures
		GM	Metaphoric gestures
	Gesture	GI	Deictic gestures
		GB	Beat gestures
		TP	Printed images, PPT images, etc.
	Tool (T)	Picture TA	Play audio
		Audio-visual media TV	Play video

### 2.4 Research Objectives

This study primarily aims to address the following questions:

- (1) How is multimodal discourse used by teachers in the Modern Chinese MOOCs?

(2) How do different modalities achieve a synergistic relationship in terms of their primary and secondary roles?

(3) How can multimodal teaching methods be utilized to innovate teaching methods in offline Modern Chinese courses, resolve teaching challenges, and enhance teaching effectiveness?

### 3. Data Analysis

Based on the coding scheme mentioned above, this study annotated five teaching videos and exported the annotation results for saving. The selection of modal types, usage frequency, and the proportion of duration for the five teachers during their lectures are shown in Figure 1:

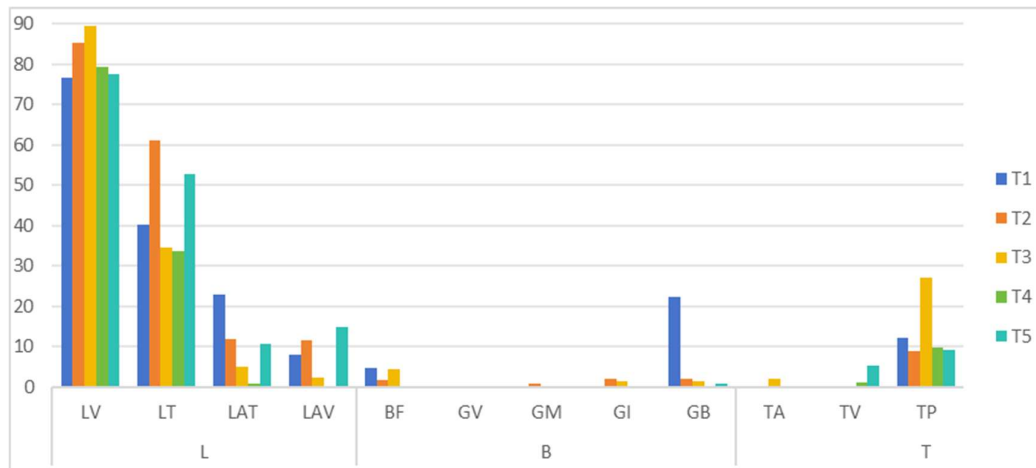


Figure 1: Statistical Chart of Multimodal Usage.

### 3.1 Statistical Analysis of Multimodal Usage

#### 3.1.1 Pure Linguistic modality

The teaching characteristic of MOOCs is that the teacher delivers the course content in front of the camera without interactive segments, and the language used is non-interactive. Moreover, given the high level of difficulty and specialized nature of the content in the Modern Chinese course, the teacher's verbal expression becomes especially crucial. In the selected MOOC videos, the duration of the voice modality for the five instructors accounts for 76%-90% of the total annotated duration, spanning the entire teaching process. Table 3 presents the statistical results for the voice modality:

Table 3: Voice Modality Statistical Table.

Annotation	Introduction (s)	Exemplificatio (s)	Summary (s)	Questioning (s)	Transition (s)	Explanation (s)
T1	/	54.24	51.04	20.52	26.69	373.62
T2	21.33	47.57	6.27	/	24.24	413.98
T3	10.75	90.26	11.79	/	8.98	746.37
T4	2.88	34.59	2.02	/	14.60	629.13
T5	4.82	111.47	27.73	/	1.80	275.35

From Table 3, it can be seen that the teachers' language primarily serves functions such as introduction, explanation, transition, exemplification, questioning, and summarization. The explanation segment is the longest across all five teachers, mainly used to explain and clarify linguistic concepts and knowledge involved in the course. Exemplification is also a crucial part of classroom language, such as demonstrating pronunciation to make abstract concepts more concrete. It is worth noting that Teacher T1 designed a simple questioning segment, whereas the other teachers did not incorporate similar segments in their lessons.

The production standards for MOOCs require that the teacher and the PPT appear simultaneously on the screen, with the teacher typically positioned on one side of the PPT or displayed in a small window, ensuring that the PPT content is not obscured. Therefore, the text modality discussed in this paper mainly refers to the text content within the PPT. This text content runs throughout the entire teaching process, interacting with other teaching modalities, forming reinforcing, overlapping, and other interactive relationships. It effectively showcases the course content and key information, helping to achieve the teaching objectives.

### 3.1.2 Accompanying Linguistic Modality

The accompanying linguistic modality plays a supplementary, complementary, and reinforcing role in conveying the meaning of language. In this paper, paralinguistic modalities are mainly divided into two categories: accompanying voice and accompanying text. Accompanying voice refers to changes in factors such as pitch, intensity, and speed of the voice, while paralinguistic text refers to variations in the font, size, and color of the text in the PPT. Paralinguistic modalities are primarily used to help the teacher attract students' attention, highlight key teaching points, or guide students' thinking during the lesson.

Figure 2,3 presents a comparison of the use of accompanying voice across the five teachers. In the five videos, the primary methods employed by the teachers to alter their speech are increasing volume, enhancing intensity, slowing down speech rate, and using pauses. The comparison reveals that the frequency of accompanying voice usage is closely linked to the individual teaching styles of the teachers. Teachers T1, T2, and T3, who are young female teachers, tend to use changes in their speech to convey teaching information, primarily combining pauses, emphasis, and slowing down the speech rate to highlight key teaching points. Teacher T5, a slightly older female teacher, does not show noticeable changes during the explanation, but when demonstrating phoneme pronunciation, she slows down her speech rate to ensure students can accurately follow along. Teacher T4, a male teacher, has a steady teaching style with no significant variations in speech rate or tone.

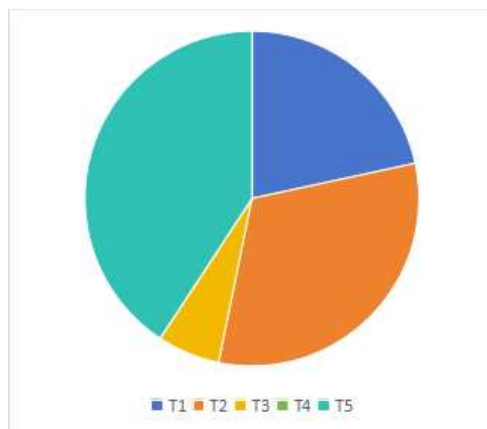


Figure 2: Usage of Accompanying Voice.

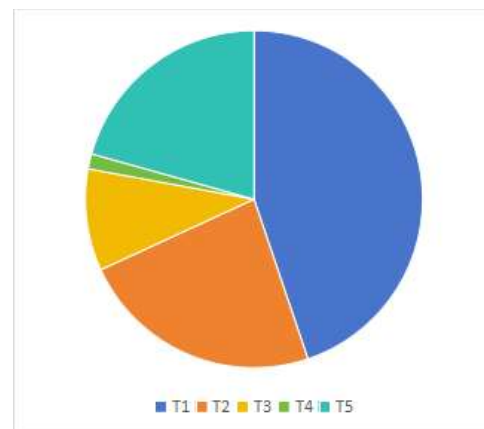


Figure 3: Usage of Accompanying Text.

According to the annotation results, the total duration of pure linguistic text annotations is 1845.944 seconds, while the total duration of paralinguistic text annotations is 343.348 seconds. This indicates that approximately 20% of the time in the multimedia text involves changes in text format. In the five videos, accompanying text is not realized through a single method but is typically achieved through a combination of two or three attributes, such as font, size, and color, to highlight key teaching points. Additionally, Teachers T4 and T5 also used methods such as changing the font background, adding underlines, circles, and dots beneath words to emphasize important content. Among the five teachers, Teacher T1 used paralinguistic text the most frequently, while Teacher T4 used it the least.

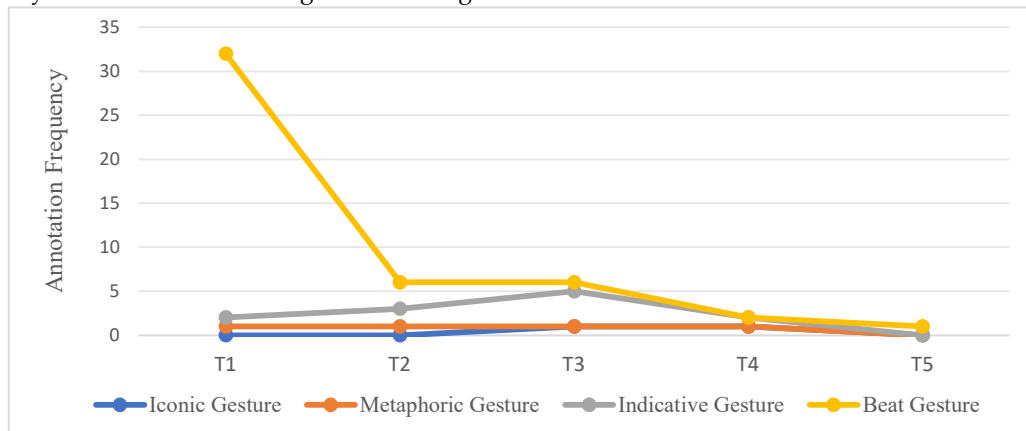
### 3.1.3 Non-linguistic Modality

### (1) Body movements

Classroom teaching relies not only on visual content such as images and text but also on the teacher's gestures, eye contact, facial expressions, and other body language. In the five videos, the teachers maintained eye contact with the camera throughout the entire teaching process, meaning they were looking at the students. There were no instances of the teachers turning their heads to look at the screen or in other directions. This could be due to specific requirements during the MOOC recording, or it could be that a teleprompter was set up in front of the teacher, eliminating the need to look at the screen behind them.

Since MOOCs do not require interaction to provide feedback on students' learning outcomes, the five teachers mostly maintained neutral expressions, with natural facial expressions being dominant. The statistics show that Teachers T1, T2, and T3, who are young female teachers, displayed smiling expressions during the teaching process, but the duration was relatively short, accounting for 4.78%, 1.63%, and 4.51% of the total time, respectively. These smiles primarily occurred at the beginning and end of the course. Teacher T3, while smiling, would also incorporate head-nodding gestures.

The most noticeable variations in body movements are found in hand gestures. This study categorizes hand gestures into four types: iconic gestures, metaphoric gestures, indicative gestures, and beat gestures. Figure 4 presents the statistical data on the frequency of the four types of hand gestures used by the five teachers during their teaching.



**Figure 4:** Gesture Modality Statistics Chart

As shown in Figure 4, beat gestures were annotated the most frequently among all types of gestures, especially by Teacher T1, who had 30 annotations for beat gestures. These gestures are simple, brief, and rhythmic, typically unrelated to the teaching content. Teachers primarily use beat gestures to express personal emotions and energize the classroom atmosphere. The second most frequently annotated gestures are indicative gestures, with Teacher T3 using them most frequently, mainly to indicate the places of articulation when demonstrating the pronunciation of vowels and consonants. In contrast, iconic and metaphoric gestures were used less frequently. This may be because teachers can use PPT to present teaching materials, and the nature of MOOCs somewhat limits the use of gestures.

### (2) Tool

In teaching, tool modalities such as images, audio, and video can provide students with dynamic process demonstrations and create an immersive experience. Among the three types of media tools surveyed in this study, the most frequently used are PPT images, including both dynamic and static images. Several teachers chose to use images to demonstrate places of articulation or speech organs. Audio and video were primarily used to showcase the correct pronunciation of different phonemes. In Teacher T1's course, in addition to the classroom content, an extra video package was provided, demonstrating the vibration of the vocal cords during pronunciation, further enhancing students' visual and auditory experience.

## 3.2 The Collaborative Relationships of Multimodal Discourse in MOOCs



Zhang Delu (2009) [5] elaborated in his study on how different forms within multimodal discourse interact with each other. He pointed out that a typical feature of multimodal discourse is the co-construction of meaning through the interrelation and supplementation of different modalities. The relationships between multimodal discourses are illustrated in Figure 5:

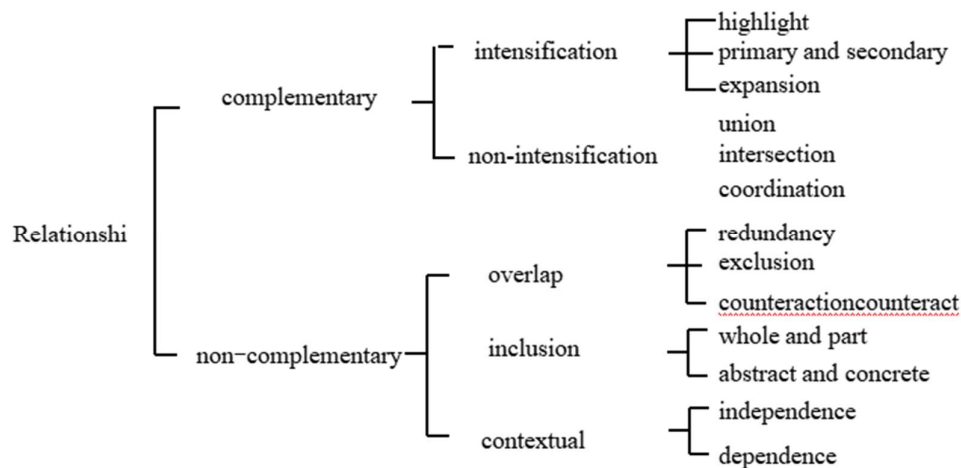


Figure 5: Multimodal Discourse Relations

This study selects course videos that encompass a variety of multimodal combinations. As the auditory modality serves as the primary mode throughout the teaching process, the coordination between modalities is mainly discussed in terms of the relationship between linguistic modality and other modalities. The modal combinations include "linguistic + accompanying linguistic," "voice + text," "voice + gestures," "voice + images," and "images + audio."

### 3.2.1 Linguistic Modality and Accompanying Linguistic Modality

The relationship between sound and non-linguistic speech, as well as between text and accompanying textual elements, mainly reflects a reinforcement relationship. Among these, sound and text modalities are the primary modalities, while non-linguistic speech and accompanying text serve as secondary modalities. Non-linguistic speech and accompanying text strengthen the discourse meaning provided by speech and text through their respective forms of variation. According to the statistics above, in approximately 20% of cases, PPT text highlights content through variations in font, size, color, underlining, and other methods.

### 3.2.2 Voice Modality and Text Modality

The combination of sound modality and text modality is the most common modality pairing in the five MOOC videos. The relationship between them primarily reflects an overlapping relationship, where the content of the PPT text corresponds to the teacher's spoken words, without adding or reducing the meaning provided by the voice. For example, when the teacher demonstrates the pronunciation of phonemes or syllables in Modern Chinese, the PPT behind them displays the content that students need to follow and repeat.

### 3.2.3 Voice Modality and Gesture Modality

In MOOC videos, voice modality and gesture modality frequently appear together. The relationship between them primarily takes the form of either a reinforcement relationship or an overlapping relationship. The relationship between language modality and indicative gestures is a reinforcement relationship. For instance, when explaining the articulation points for consonant sounds, the teacher uses gestures to indicate the position in the mouth where the airflow is obstructed. In this case, voice serves as the primary communicative modality, while the indicative gestures reinforce the meaning conveyed by the sound. The relationship between voice modality and beat gestures, however,

is an overlapping relationship. When explaining a specific concept, the teacher may incorporate hand movements that align with the changes in their language modality. However, beat gestures do not contribute substantively to the construction of discourse meaning; they merely coexist with the language modality as an additional modality, playing a supportive role.

### **3.2.4 Voice Modality and Image Modality**

Images are an indispensable part of Modern Chinese MOOCs, as they help address issues such as the abstract nature of course content and the lack of engagement. In this study, the relationship between images and voice modality can be categorized into reinforcement and conjunction relationships. In the MOOC videos, certain images, such as photographs of performing artist Li Moran and images from live performances of the violin concerto *The Butterfly Lovers*, are secondary elements. These images primarily serve to create context for students and assist in understanding the teacher's verbal explanations, thus forming a reinforcement relationship between sound and images. On the other hand, some images, such as diagrams of the human thoracic cavity, pharynx, and oral cavity, help the teacher to concretize abstract discussions when explaining the physiological attributes of speech sounds. In this case, both sound and images participate in constructing discourse meaning, forming a conjunction relationship.

### **3.2.5 Image Modality and Audio Modality**

In MOOC videos, the relationship between image modality and audio modality is typically complementary. For instance, when an image of a bird is shown, the audio might feature the sound of bird calls, or when an image of a waterfall is displayed, the accompanying audio may present the sound of flowing water. In these cases, the two modalities work together to convey a unified meaning.

## **4. Implications of the Multimodal Teaching Model for Offline Teaching of "Modern Chinese"**

As a novel teaching approach, Massive Open Online Courses (MOOCs) require instructors to deliver content comparable to that of offline courses within a limited time frame (typically no more than 20 minutes), while also fostering a positive classroom atmosphere. Therefore, MOOC instructors must make full use of modern educational technologies and meticulously design multimodal teaching strategies to engage multiple senses of learners, thereby maximizing the transmission and internalization of knowledge. By statistically analyzing the interactive use of various modalities, including language, visual elements, and auditory cues, in MOOC videos, empirical evidence and innovative insights for the reform of the offline teaching model of "Modern Chinese" can be provided, ultimately optimizing teaching effectiveness.

### **4.1 Offline Teaching Should also Embrace a "Multichannel" and "Multimodal" Approach**

In offline teaching, educators can take advantage of technological advancements to incorporate various forms of information—such as text, graphics, sound, video, actions, and gestures—into the learning process. This multimodal teaching approach allows information to be conveyed through multiple sensory channels, activating different areas of the brain, thereby aligning more closely with the natural way humans acquire language, which significantly enhances teaching effectiveness. Moreover, the use of multimodal teaching can address the issue of low engagement in offline classrooms by creating a more enjoyable and dynamic learning environment. This not only helps to motivate students intrinsically but also fosters the development of their creative thinking. In this teaching model, students transition from passive recipients of knowledge to active constructors of knowledge through their engagement and experience.

### **4.2 In Offline Teaching, Instructors can Effectively use Gestures and Eye Contact to Enhance Teaching Effectiveness**

In offline courses, instructors can make full use of modalities that are either unavailable or limited in MOOCs to enhance classroom interactivity.



In the five MOOC videos analyzed in this study, beat gestures were used most frequently, while deictic and metaphoric gestures were used relatively less. In offline classrooms, instructors can consciously design and incorporate pointing and metaphoric gestures. For instance, when explaining the articulation points of consonants, the instructor can use one hand to represent the upper and lower jaws and the other to simulate the tongue, demonstrating the position of the articulatory organs when producing different consonants.

In addition to gestures, eye contact is another modality that is difficult to fully utilize in MOOCs. In MOOC videos, instructors typically face the camera, which limits eye contact with students. In contrast, in offline classrooms, instructors can enhance teaching effectiveness by strategically using eye contact, such as interacting with the PPT content, engaging with students in different parts of the room, or offering affirmations through eye contact when students respond to questions. The teacher's gaze can create a directional focus, conveying meaning and guiding students' attention, helping them stay engaged and follow the flow of the lesson.

#### **4.3 The Multimodal Approach to Course Assessment and Evaluation**

In Chinese language courses at universities, the assessment model is typically unidimensional: "teachers design the questions, students answer them, and scores serve as the sole measure of learning outcomes, failing to reflect students' abilities beyond memorizing fragmented knowledge" [12]. The assessment system for the "Modern Chinese" course could be transformed into a multimodal evaluation system:

(1) Combination of online and offline assessments, where students are required to use online platforms to complete pre-study and in-class tests, while offline assessments include experimental reports and written exams. The grading system should fully consider process-oriented assessments, providing a comprehensive and objective reflection of students' learning outcomes.

(2) Integration of theory and practice in assessments, where the evaluation system should include tasks such as "language fieldwork," "phonetic discrimination," and "language use surveys," ensuring that students can apply theoretical knowledge to solve practical problems.

(3) Diversified assessment content, including submissions of audio, video, and completion of both in-class and extracurricular tasks. This approach can comprehensively and multidimensionally reflect students' learning outcomes and proficiency.

#### **5. Conclusion**

Based on the findings of this study, the linguistic modality occupies a central role in the "Modern Chinese" MOOC, while the paralanguage modality serves an auxiliary and reinforcing function, helping students better understand and grasp abstract linguistic concepts. The inclusion and integration of non-verbal modalities provide insights and ideas for the improvement of offline teaching. Furthermore, multimodal discourse analysis reveals the synergistic relationships between different modalities, such as reinforcement, overlap, and conjunction. Recognizing and applying these relationships is crucial for designing effective multimodal teaching strategies.

The application of multimodal teaching methods holds great promise, and its development potential and practical value will continue to be explored. With the continuous advancement of educational technologies, particularly emerging technologies such as augmented reality (AR) and virtual reality (VR), multimodal teaching is expected to offer more immersive and interactive learning experiences. Future research could explore how to integrate these technologies with multimodal teaching methods to cater to students with diverse learning styles and needs, enabling personalized instruction.

#### **References**

- [1] Kress, G., & Van Leeuwen, T. (1996). *Reading Images: The Grammar of Visual Design*. Abingdon: Psychology Press.
- [2] Forceville, C. J., & Urios-Aparisi, E. (2009). *Multimodal Metaphor*. Berlin and New York: Mouton de Gruyter.
- [3] Zhanzi Li. (2003). Social semiotic analysis of multimodal discourse. *Foreign Language Research*, 5, 1-8.

- [4] Zhuanglin Hu. (2007). Multimodalization in social semiotic studies. *Language Teaching and Research*, 1, 1-10.
- [5] Delu Zhang. (2009). Exploring a comprehensive theoretical framework for multimodal discourse analysis. *China Foreign Language*, 6(1), 24-30.
- [6] Yueguo Gu. (2007). An analysis of multimedia and multimodal learning. *Foreign Language and Digital Teaching*, 2, 3-12.
- [7] Delu Zhang. (2010). Initial exploration of the design and modality invocation of multimodal foreign language teaching. *China Foreign Language*, 5(3), 48-53.
- [8] Ya Li. (2018). The implications of multimodal discourse analysis theory for teaching Chinese as a foreign language. *Ethnic Education Research*, 29(5), 139-144.
- [9] Shuai Yuan. (2019). Teaching design of beginner-level Chinese listening courses based on multimodal teaching models. Shenyang: Shenyang University.
- [10] Namei Li (2022). Research on online Chinese vocabulary teaching based on multimodal discourse analysis theory. *Chinese Character Culture*, 12, 77-79.
- [11] Guoxue Zhang. (2023). A comparative study of multimodal discourse in online and offline grammar classes. *Frontiers in Social Sciences*, 12(6), 2903-2911.
- [12] Fang Wang. (2020). Research on multimodal teaching in modern Chinese courses. *Cultural and Educational Materials*, 28, 165-166.