

ELAN-Based Multimodal Discourse Analysis of Teachers' Talk: A Case Study of EFL Teachers in China

Ansheng Huang, Yue Ma*

School of Foreign Languages, Huizhou University, Huizhou 516007, Guangdong, China

*Corresponding author: Yue Ma, my@hzu.edu.cn

Copyright: © 2024 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This study employs the 13th National Junior High School English Demo Classes as a corpus for multimodal analysis using ELAN, both qualitatively and quantitatively. Its objective is to investigate the structural features of middle school English reading teaching, the utilization of modes, and the realization of the meaning system. The findings are as follows: English reading teaching in Demo Classes comprises eight steps and utilizes multimodality, wherein visual, actional, and auditory modes are extensively employed to execute teaching procedures. The realization of the meaning system relies on these modes themselves, as well as their reinforcement and complementarity.

Keywords: Multimodal discourse analysis; EFL teachers' talk; ELAN

Online publication: March 29, 2024

1. Introduction

The dominance of purely linguistic instruction in traditional classrooms has long fallen short of meeting the needs of students in the new era. In recent years, both domestic and international scholars have increasingly emphasized the importance of multimodal discourse analysis (MDA) in the teaching field. This research aims to apply MDA theories to empirically study teachers' discourse, focusing on English reading teaching in middle schools. The study is based on Visual Grammar by Kress and van Leeuwen^[1], and a Systemic-Functional Synthetic Framework by Delu Zhang^[2,3] and Xueyan Yang^[4], with annotated statistics provided by ELAN.

The paper presents an analytical framework to expand the scope of MDA theory. In terms of practical teaching, by exploring the multimodal distribution features of teachers' discourse in English reading instruction, educators can gain insights into the specific modes used in various teaching stages. This paper offers practical suggestions on employing multimodality for teachers aspiring to participate in Demo Classes competitions.

2. Analytical framework for the MDA

Drawing from systemic functional linguistics by Halliday^[5], the theories of Visual Grammar by Kress and van Leeuwen as summarized by Magret Baker^[6], and the Systemic-Functional Synthetic framework by Zhang^[3], an evaluation model for MDA can be developed.

This study considers three levels: cultural, contextual, and content. The cultural level encompasses ideology and genre; the contextual level involves field, tenor, and mode; and the content level comprises form and discourse systems. The form system includes three modal categories: visual, actional, and auditory, while the discourse system encompasses three meaning categories: representational, interactive, and compositional meanings (**Table 1**).

Table 1. Details of the three levels^[7]

	Content	Instance
Cultural level	Ideology	Relations between participants
	Genre	Discourse mode
Context level	Filed	Teaching content
	Tenor	Teacher and students
	Mode	Teaching methods
		Content
Content level		Visual
	Form	Action
		Auditory
	Discourse	Representational meaning
		Interactive meaning
Compositional meaning		

In MDA, the basic level is instantiated by a higher level. Thus, the cultural level is inter-level realized by the contextual level, and the combination of the cultural and contextual levels aids in the realization of the meaning system at the content level. Moreover, at the content level, the meaning system is realized at the form level.

In this study, the cultural and contextual levels serve as the basis for analyzing the structural features of English reading teaching, while the content level provides possibilities for analyzing the distribution features of modes and meaning constructions.

3. Methodology

3.1. Research subject

In this study, the subject consists of reading teaching videos from the 13th National Junior High School Demo Classes, comprising 32 videos. Four videos were selected to form a corpus based on specific criteria. Firstly, the competition, organized by the School of Foreign Language and Literature at Beijing Normal University, is renowned for its seriousness and authority. The teaching quality in the demonstration classes tends to be higher than that of routine teaching due to the careful selection of both teachers and students. Secondly, the competition features male and female teachers from across the country, minimizing potential gender bias or

regional preferences.

For this study, two male and two female teachers were randomly selected from the contestants to conduct an empirical analysis.

3.2. Research questions

This paper aims to propose suggestions for enhancing the effectiveness and efficiency of classroom teaching by analyzing the application of multimodality in teachers' discourse. The specific research questions are as follows:

- (1) What are the structural features of English reading teaching in middle school?
- (2) Which modes are employed, and what are the distribution features of these modes?
- (3) How do different modes collaborate to realize representational, interactive, and compositional meanings?

3.3. Research tools

To enhance the credibility of this study, the EUDIO Linguistic Annotator (ELAN) will be utilized to annotate the corpus. Developed by the Max Planck Institute of Psycholinguistics, ELAN is specifically designed for Multimodal Discourse Analysis (MDA). It serves as an annotation tool for creating, editing, visualizing, and searching annotations of video and audio data, primarily aiming to provide robust technologies for annotation and to exploit multimedia clips. Since its release in 2001, ELAN has been available in multiple languages and widely utilized in various fields such as discourse analysis, sign language analysis, language archiving, and spoken language corpus construction ^[8].

As a precise quantitative data annotation tool, ELAN facilitates the collection of annotated statistical data for quantitative analysis, thus potentially minimizing the subjectivity of discourse analysis. With powerful annotating and statistical functions, ELAN allows for analysis with a precision of up to 0.01 seconds of video, ensuring the accuracy of the results. Researchers can transcribe, annotate, and modify language and symbols in the video hierarchically to interpret the video from multiple dimensions. ELAN will be employed to annotate teachers' activities, including facial expressions and gestures, as well as elements in multimedia devices such as color, images, and text. Following the annotation process, both qualitative and quantitative analyses will be conducted to clearly illustrate the occurrences and frequency of different modes used by teachers.

3.4. Research procedures

- (1) Establishing the coding scheme: Before annotation, elements in the video corpus are categorized to facilitate annotation and statistical analysis using ELAN. Twelve modal resources of teachers' discourse are annotated through ELAN, and accordingly, twelve codes are designed.
- (2) Defining layers in ELAN: First, create a new media file and save it as an ELAN annotation file (EAF). Next, define the linguistic types. After defining the types, add new tiers, fill in the name of the level in the blank, and select the language type it belongs to.
- (3) Annotation and analysis: Once set up in ELAN, observation and annotation will be conducted simultaneously. The annotated data will then be collected to illustrate in the form of percentages and graphs. The final step is to analyze how these different modes interact and cooperate to construct three meanings from the statistics.

4. Results and Discussion

4.1. The structural features of English reading teaching

As shown in the proposed model, the culture level serves as a fundamental and indispensable lever, greatly aiding in the generation of meaning. According to Zhang Delu, this level is essential for facilitating communication ^[9]. Two aspects are encompassed within this level: ideology and genre. Ideology pertains to people's thoughts, actions, life values, habits, and unwritten rules in their daily lives, while genre is associated with the communicative process or structural potential to realize ideology. Regarding ideology, all students in the videos attend the same school and are of similar age, and the teachers share similar educational backgrounds and are assessed under the same criteria as their competitors. Thus, there are no significant differences in ideology among teachers and students, respectively. Concerning genre, all the researched videos feature the same type of classroom teaching, specifically reading teaching on different themes. Consequently, a general structure of English reading teaching can be summarized.

Regarding the context level, all teaching videos, sourced from the same competition, depict a similar context, namely, a teaching-and-learning context. Thus, the field, tenor, and mode of the corpus can be considered the same, further supporting the summary of a general structure of English reading teaching.

The analysis from the perspective of the culture level and context level confirms the feasibility of summarizing the structural features of English reading teaching. After observing all the videos, the common structure of English reading teaching has been established. All of the English reading teaching sessions in this study consist of eight steps, as outlined in **Table 2**, which align with the genre structure of the teaching and learning process proposed by Zhang Delu ^[2,3]. Among these teaching steps, the Teacher Action (TA) and Student Action (SA) steps account for a larger proportion, indicating a balance between teacher-led and student-centered activities. In the competition setting, the TA step is crucial as it reflects the teachers' ability to control the classroom and deliver effective instruction. During this phase, there is extensive communication between teachers and students through various teaching activities.

Table 2. The common structure of English reading teaching

Number	Teaching step	Detailed description
1	Class begins (CB)	Announcing the beginning of class
2	Introduction topic (IT)	Introducing the teaching topic and triggering students' interest
3	Introduction content (IC)	Introducing the teaching content
4	Teacher's activity (TA)	Teachers conduct certain activities to explicit the teaching content
5	Student's activity (SA)	Students participate in activities by themselves
6	Class summary (CS)	Summarize what they have learned
7	Homework assignment (HA)	Introducing the homework
8	Finish class (FC)	Announcing the ending of class

In English reading teaching, the SA step plays a pivotal role in helping students enhance their reading skills, which is a key teaching objective. During this phase, students engage in reading tasks within a specified time frame, supported by clear instructions and reading tips. This enables them not only to acquire textual knowledge but also to significantly improve their reading proficiency.

Conversely, the Class Begins (CB), Homework Assignment (HA), and Finish Class (FC) steps make up a very small proportion of the teaching process. These steps involve concise language from the teachers, allowing them to efficiently complete the teaching tasks and allocate more time to critical steps.

This well-designed teaching approach not only facilitates students' independent reading and improves their reading skills but also adheres to the principles of effective teaching activity design. Overall, clear time allocation, diverse activities, and focused instruction are the key strengths of the teaching processes in this study, indicating the high reliability of the corpus.

4.2. Annotation occurrence of multimodal modes

The collected statistics reveal that teachers utilize multiple modes to construct and convey meaning. **Table 3** illustrates the abbreviations of the decoding scheme for annotation.

Table 3. The abbreviations of the decoding scheme for annotation

Code	Meaning
[G]	Gesture
[GA]	Gaze
[FES]	Facial expression-simile
[PLT]	Pure language transcription
[P]	Posture
[PI]	Prosody-intonation
[VT]	Visual text
[VI]	Visual image
[TI]	Text and image
[VC]	Color
[VF]	Font
[VS]	Symbol

Among all the modes, [PLT], [GZ], and [FES] are the three most frequently occurring modes, while [VI], [TI], and [VS] are the least frequent. Thus, based on the annotation occurrences, it can be inferred that teachers predominantly employ [PLT], [GZ], and [FES] during the teaching process. This suggests that, in addition to traditional verbal explanations, teachers often make eye contact with students or objects and smile at students.

Furthermore, all annotated modes can be classified into three types: visual modes, auditory modes, and actional modes. Actional modes encompass gesture, gaze, facial expression (smile), pure language transcription, and posture. The categories of pure language transcription and prosody-intonation belong to auditory modes, while visual modes include visual text, visual image, text and image, color, font, and symbol.

A comprehensive analysis reveals that auditory modes constitute the majority (48%), followed by actional modes (39%), and visual modes (13%).

4.3. Distribution of modes in different teaching steps

To summarize the distribution features of modes, this section will further analyze the annotated statistics.

Firstly, the application of visual modes will be examined. From a macro perspective, visual modes are prominent in TA, where teachers utilize various visual texts, images, and resources to emphasize key points and address difficulties. The combination of PowerPoint presentations or writing on the blackboard with verbal explanations is complementary, as neither alone can adequately fulfill teaching tasks. Additionally, SA exhibits fewer modes as teachers aim to create an immersive reading atmosphere, allowing students to

read independently at their own pace. The goal of English reading teaching is to enhance students' reading proficiency. From a micro perspective, among all visual modes, text is the most prevalent, followed by color and font. Teachers segment text into different sections and present them separately in PowerPoint slides, while also writing key information on the board. Furthermore, the use of color helps highlight important points.

Secondly, the application of auditory modes will be analyzed, focusing on pure language transcription and prosody-intonation. It is recognized that the intonation of different sentence types varies. For instance, rising intonation is used for general questions, falling intonation for stating facts, rising-falling intonation for alternative questions, and falling-rising intonation for tag questions. In this study, [PI] marks when the teacher asks a question. From a macro perspective, there is a sharp increase in the frequencies of auditory modes from CB to IT, as teachers use various sentence patterns to engage students and stimulate their interest. From IC to TA, teachers increasingly employ auditory modes to impart knowledge and organize activities, which require more time. Conversely, there is a decreasing trend from SA to FC, possibly due to the shorter duration of these steps. From a micro perspective, the frequency of [PLT] and [PI] is relatively high in TA. The variation in intonation allows teachers to convey different messages to students, enriching communication. Additionally, in SA, teachers may participate in group discussions or assist weaker students, fostering closer relationships with students.

Thirdly, the application of actional modes will be examined. From a macro perspective, TA witnesses the most frequent occurrence of actional modes, similar to visual and auditory modes. From a micro perspective, [G] and [FES] are prominent modes that occur at every step, suggesting that teachers may make eye contact with students to deliver or seek information. Proper use of actional modes helps establish positive relationships between teachers and students.

In conclusion, English reading teaching is multimodal, with teachers employing not only language but also gestures, images, and other modes to construct meaning. Among all steps, TA exhibits the highest frequency of modes, with auditory modes being the most intensive. The distribution of modes is influenced by specific teaching tasks and the duration of teaching steps, a conclusion supported by other researchers in classroom observations^[10-12].

4.4. Qualitative analysis of the meta-functions in English reading teaching

While the quantitative analysis of modes has been conducted separately, English reading teaching cohesively integrates modes. Hence, the subsequent step involves a qualitative analysis of how these modes are integrated to realize the three meanings.

According to Zhang^[4], discourse meaning is instantiated by discourse grammar, which, in turn, is instantiated by discourse entities containing both non-language and language elements. Therefore, understanding how the meaning system is formed entails analyzing how these entities collaborate to generate meaning.

4.4.1 Representation meaning

Representational meaning focuses on the representations of objects and their interactions with others. In other words, it highlights which entity serves as the subject and which one is the object (or target). Representational meaning can be realized through two processes: the narrative process and the conceptual process. The distinction lies in whether vectors are involved in the process. Therefore, vectors are essential indicators of the narrative process.

4.4.1.1. Narrative process

Kress & van Leeuwen^[1] stated that the narrative process expresses changes in actions, events, processes, and instantaneous spatial arrangements, involving the action process, reaction process, and verbal, and mental process. Among these processes, the action process and reaction process can be transitive or non-transitive.

In the action process, actors and goals are significant elements. In the corpus of this study, all videos contain participants and targeted participants, so most processes are transitive. For example, when the teacher guides students to find answers in the text and says, “Let’s look at the PowerPoint,” the teacher is the actor who gives out the vector, and the students are the goals. When asking, “Where can you find those answers?” the teacher is demanding detailed information from the students. The gesture of pointing to the screen forms the vector between the students and the screen, which strengthens the teacher’s discourse.

The action process occurs when the vector is constructed merely by the gaze of one represented participant or more than represented participants. This process can also be transitive or non-transitive. The decisive factor in distinguishing between the two processes is the presence or absence of an actor looking at one or more participants. For instance, the teacher gives out an eye line and students receive his eye line. In this process, the former is the reactor, and the latter is the phenomenon. The teacher is looking at students and guiding them to talk about their hopes for the next reading, during which eye contact can establish a connection between the teacher and students.

In the above two processes, the actional mode strengthens the auditory mode, and both of them construct the representational meaning.

Verbal and mental processes generally occur in the protrusion of the dialogue balloon and thought balloon. As a special vector, it links the thinker or speaker with a balloon. The thought balloon can not only ask students questions but also provoke them to think deeply. It is a classical verbal and mental process, often seen in dialogues or questions in the PowerPoint in the corpus.

In the above process, visual modes complement auditory modes, and both of them construct the representational meaning.

4.4.1.2. Conceptual process

The conceptual process is more stable and general than the narrative process and is characterized by the absence of a vector. This process can reveal certain information, including participants’ statuses, meanings, and structures. There are three categories of this process: classification process, symbolic process, as well as analytical process^[13].

Regarding the classification process, it indicates the relations among participants. One participant is called superordinate while the other participant is called subordinate.

As for the symbolic process, it symbolizes an unusual or abstract thing in visual forms so that many abstract ideas become easier to understand by participants. **Figure 1** is an example. In this case, the teacher reviews the plots of “Two Geniuses” by adopting the story mountain pattern in **Figure 1**. The story mountain, as a tool to present the structure of a story, helps students gain a deeper understanding of the characters, plot, and events in the story. Students can develop a general awareness of the plot and categorize relevant plots from a macro perspective, which helps enhance students’ writing and reading skills.

Review the plots

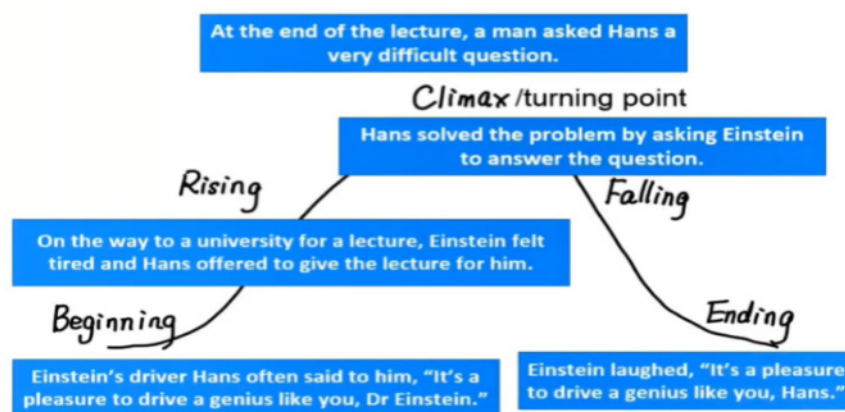


Figure 1. The story mountain pattern of "Two Geniuses" [14]

Regarding the analytic process, it stresses the relation of "part-whole," which means there must be two participants. The "whole" is called the carrier, while the "part" is called the possessive attributes. For instance, the teacher requires students to brainstorm about ways to be a healthy child. In the mind map, the "whole" is to be a healthy child and connects many solutions through lines (equivalent to the "part"). Under this circumstance, there is a 'part-whole' relation. After finishing this mind map, students would have a deeper impression of the solutions for being healthy and learn to construct the mind map.

During the conceptual process, visual modes complement auditory modes to construct the representational meaning.

From the above discussion, the representational meaning in English teaching is mainly constructed by the narrative process, in which the action process and reaction process occur more frequently than verbal and mental processes. Specifically, the action process sees the synergy of the teacher's gesture, posture, and pure language transcript; the reaction process sees the formation of vectors between the teacher and students; the verbal and mental processes are seen in the dialogue balloon and thought balloon.

4.4.2. Interactive Meaning

The interactive meaning reveals the relationship between participants and viewers, which is composed of contact, social distance, perspective, and modality [15]. In this study, the first two elements will be analyzed to understand how modes establish the interactive meaning, as the last two elements are rarely found in the corpus.

Two basic images are contained in the contact system: the demand act and the offer act; they differ in the ways of contacting the viewer. The former refers to the actions of facial expressions, gestures, and gaze, which means the represented participants are demanding from the viewer. While the latter refers to indirect contact, which means the represented participants offer something to the viewer. For instance, when the teacher uses gestures and smiles to encourage the students, she is offering some psychological implications, in which students can be aware of the teacher's encouragement and kindness. From the point of view of the students, she is demanding attention from them to the PowerPoint so that they can act quickly according to her request.

Social distance represents the social relation and intimacy between the represented participants and the viewer. It can be realized through the size of the frame, which means the choice between close-up, medium shot, and long shot. The closer the shot is, the closer their relationship is. Three kinds of social distance are personal distance, social distance, and public distance. In this research, due to the special teaching conditions

where teachers have to maintain a close relationship with students, there is no public distance.

When the teacher stands in front of the students to present the title, we can view it as social distance. While the teacher stands in the middle of the students to participate in their discussion, we can consider it as personal distance. Based on the observation of the corpus, it can be concluded that teachers change their distance for different teaching tasks.

From the above discussion, it can be concluded that the interactive meaning is generated by the reinforcement of actional and auditory modes. In English reading teaching, teachers may as well communicate with students through eye contact because it can encourage students and inspire them to participate in teaching. The appropriate distance between teacher and students is of importance to attract students' attention.

4.4.3. Compositional meaning

According to Kress & van Leeuwen^[1], three elements are contained in compositional meaning: information value, salience, and framing.

Information value is achieved by the placement of specific elements in the image, whose role depends on where it is put, including left and right, top and down. In this research, those specific elements are visual texts and images. Here are two examples.

In the left part of **Figure 2**, the teacher shows the scripts and elicits some questions about the text. For the teacher and students, the left scripts are given information while the right questions can be regarded as new information. In the right one, the teacher shows the theme of a meaningful winter vacation and guides students to brainstorm about it. For them, the center content is the given information while the margin content is the new information. Such placement not only helps students concentrate and grasp important information quickly but also helps explain the topic in detail.

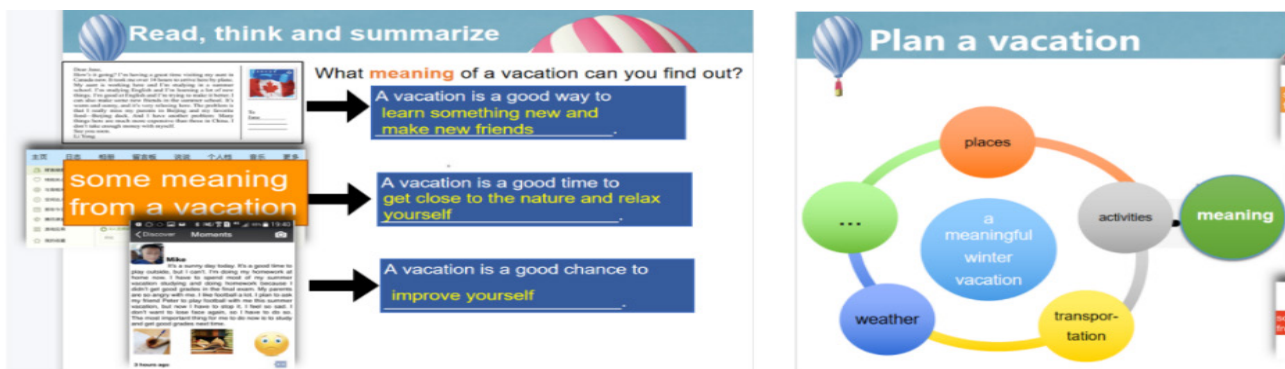


Figure 2. The illustration of the placement of specific elements^[14]

Salience is defined as the degree of attraction to the elements for the viewers. In this study, it is made possible by the placement, sizes, and colors of the participants and the differentiation of sharpness in the PowerPoint. Here are two examples.

To make teaching effective, it is crucial to present information in a clear and eye-catching way. Thus, the salience of information cannot be ignored. The use of colors, slanted fonts, and bold fonts realizes the salience of the keywords. For instance, the student may feel bored and disinclined when he sees black words in large paragraphs. However, words presented with salience to stress the importance of some keywords may not evoke such aversion in students.

Framing refers to the inner connection of images or information in visual design, which is achieved by dividing lines or frames. These lines can indicate whether there exists ownership or not by cutting or connecting

participants.

The device of framing can be a mind map or chart. In this study, framing could be found in the PowerPoint slides and teachers' teaching procedures. The teaching process from CB to FC is a discourse framing, embodying the discourse meaning of a multimodal meaning system.

From the above discussion, it can be concluded that compositional meaning is generated by visual modes, such as visual text and color. In addition, auditory modes usually complement visual modes by using spoken language to explain the visual elements.

5. Findings and implications

The quantitative analysis was based on the multimodal corpus of English reading courses in middle school, revealing the overall annotation occurrences of modes and the application of these modes in different steps.

Firstly, all teachings in the Demo Classes have eight steps, consistent with the genre structure of the teaching and learning process proposed by Zhang in 2021^[3]. Among all the teaching steps, TA and SA steps account for a larger proportion, while CB, HA, and FC steps take a very small proportion. These teaching processes are featured by clear time distribution, abundant activities, and prominent focuses. Additionally, all teaching processes are multimodal, providing students with an authentic environment and a rich source of information for language learning, motivating them to think actively, process various information flexibly, and accurately construct and transmit meaning.

Secondly, teachers mobilize diverse modes in each teaching step to achieve different teaching and communicative objectives. For one thing, the most prominent mode among all modes is the auditory mode, determined by the features of teaching. However, although the auditory mode is the main mode, it cannot fulfill the teaching task on its own due to its drawbacks, such as being temporary and non-visual. Traditionally, the auditory mode is complemented by real objects and gestures. In this corpus, visual and actional modes are adopted to reinforce and complement the auditory mode so that the teaching can be completed. For another, in IT and TA, most modes appear intensively, such as gestures, facial expressions, images, and text. The difference in the distribution features of the different modes may be attributed to the different teaching schedules and teaching tasks of each teaching step. Considering the crucial role of modes in the teaching processes, teachers are expected to identify the specific meaning of each mode and select the most appropriate mode.

Qualitative analysis was conducted to determine how different modes work together to realize representational, interactive, and compositional meanings.

In general, the realization of the meaning system depends on diverse modes themselves and their interaction because different modes have their functions, and different combinations of modes convey different meanings. Representational meaning is primarily constructed by the action process and reaction process in the narrative process, during which actional modes reinforce auditory modes, and auditory modes complement visual modes, respectively. Interactive meaning is mainly realized by the use of social distance and contact. Besides, the flexible matching of auditory and actional modes can effectively convey interpersonal meanings such as equality, kindness, and encouragement, creating a harmonious and pleasant atmosphere. Compositional meaning is mainly realized by the teaching process and visual modes, during which visual modes and auditory modes complement each other. Additionally, the teaching process, as a discourse, also contributes to constructing compositional meaning. In conclusion, different modes cooperate, interact, and complement each other to construct meaning by highlighting or complementing the other.

The implications of this paper are favorable for EFL teachers for teenagers. From a theoretical perspective,

the related theories of multimodal teaching and MDA have been collected, and an analytical framework for MDA has been proposed, enabling them to access it. Modes such as distance, contact, and gesture are fantastic communicative resources used by teachers to construct the meaning system, but these modes are often neglected in traditional teaching. Consequently, the conclusion that teachers are expected to use these resources consciously and efficiently is an important insight into MDA and provides new research directions for future research on classroom discourse. Ideally, from a practical perspective, teachers can effectively learn about these theories and imitate the teaching process to direct their teaching practice precisely. The method and design of multimodal teaching can be applied to daily teaching, thus improving the quality and effectiveness of foreign language teaching. To put it more precisely, teachers are suggested to design courses from a multimodal perspective and take visual mode and actional mode as auxiliaries. Having fully understood the meaning and function of modes, teachers need to make appropriate use of the reinforcement, complement, and synergy across modes to enhance the effectiveness of classroom teaching.

Funding

Philosophy and Social Science Project of Guangdong Province “A Study on the Curriculum-based Ideology Education Competency Model of Chinese EFL Instructors based on Multivariate Analysis” (Project No. GD22WZX01-12)

2022 Higher Education Teaching Quality Project of Guangdong Province / 2021 Huizhou University Teaching Quality Project “Research on English Teacher Education Curriculum”

2023 Higher Education Teaching Quality Project of Guangdong Province “The ‘English+’ Compound Talent Training Program under the Background of New Liberal Arts

2023 Guangdong Province’s New Normal Universities’ Project of Promoting the High-Quality Development of Compulsory Education “Research and Practice on the Reform of Junior Middle School English Comprehensive Practice Activity Course Based on Core Literacy”

Disclosure statement

The authors declare no conflict of interest.

References

- [1] Kress G, van Leeuwen T, 2021, *Reading Images: The Grammar of Visual Design* (3rd ed). Routledge, London.
- [2] Zhang D, 2021, *Multimodality and Multiliteracies in Foreign Language Teaching*. Foreign Language Teaching and Research Press, Beijing.
- [3] Zhang D, 2020, *Research on Systemic Functional Linguistics and Foreign Language Education*. Shanghai Foreign Language Press, Shanghai.
- [4] Yang X, 2023, *Modelling Interactivity of Teacher Discourse in EFL Classrooms: A Systemic-Functional Perspective*. Fudan University Press, Shanghai.
- [5] Halliday MAK, 2000, *An Introduction to Functional Grammar*. Foreign Language Teaching and Research Press, Beijing.
- [6] Baker M, 2021, Book Review: “Reading Images: The Grammar of Visual Design, 3rd ed., by Gunter Kress and Theo van Leeuwen.” *Visual Communication Quarterly*, 28(4): 253–254.
- [7] Kress G, Jewtt C, Ogborn J, 2024, *Multimodal Teaching and Learning: The Rhetorics of the Science Classrooms*.

Foreign Language Teaching and Research Press, Beijing.

- [8] Drude S, TLA, 2012, General Assembly of the Max Planck-Institute for Psycholinguistics, December 18, 2012: The Language Archive (TLA). The Netherlands. <https://doi.org/10.13140/RG.2.2.13330.20167>
- [9] Zhang D, 2018, A Systemic-Functional Synthetic Framework for Multimodal Discourse Analysis. *Modern Foreign Languages*, 41(6): 731–743.
- [10] Chohra H, Hamerlain S, 2018, Classroom Discourse Failure among Algerian Secondary School Learners. *Arab World English Journal*, 9(2): 298–314.
- [11] Feng H, 2021, A Systemic Functional Multimodal Discourse Analysis of English Micro-Lectures. *Journal of Higher Education Research*, 5(2): 293–307.
- [12] Bi M, 2019, Multimodal Discourse Analysis of News Pictures. *Theory and Practice in Language Studies*, 9(8): 1035–1039.
- [13] Cheng X, 2019, *Discourse Analysis of English Teachers*. Shanghai Foreign Language Press, Shanghai.
- [14] Liang Z, 2018, *Selected Cases of Junior High School English Teaching Activity Design*. Beijing University Press, Beijing.
- [15] Morell T, 2018, Multimodal Competence and Effective Interactive Lecturing. *System*, 77: 70–79. <https://doi.org/10.1016/j.system.2017.12.006>

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.