

# A Multimodal Discourse Analysis of College English Classrooms Based on ELAN

## -- Cases Study of Grand Prize of the 13th "SFLEP Cup" College English Comprehensive Course Group

Jing He

School of English, Guilin University of Technology, Guangxi, Guilin, 541006, China

### Abstract

**This paper analyses the classroom discourses of the teacher who won the grand prize in the college English comprehensive course group of the 13th Shanghai Foreign Language Education Press Cup (SFLEP Cup) from the perspective of Systemic Functional Linguistics (SFL), and under the guidance of Zhang Delu's framework for multimodal foreign language classroom teaching through ELAN 5.2 multimodal discourse analysis software, intending to reveal the use of modalities in each stage of the teacher's classroom, the relationship between modalities, and the relationship between teaching objectives, methods, and modal choices. The study shows that the auditory modality of teacher discourse is the primary modality. In contrast, the visual, behavioral, and environmental modalities play a supplementary role. Effective synergy between modalities is crucial to constructing meaning, and inappropriate synergy can have a negative impact. In addition, the choice of teaching methods directly affects the choice of modality, and these two aspects work closely together to achieve the teaching goals. Multimodal teaching has demonstrated significant effectiveness in college English classrooms, and teachers can achieve the teaching goals, enhance the teaching effect, and increase students' participation and interest in learning through the judicious application of various modalities.**

### Keywords

**Multimodal theory, University English classroom, Discourse analysis.**

### 1. Introduction

In the era of the "Internet Plus" and the widespread use of electronic media, monomodal language can no longer keep pace with the rapid flow of information, and discourse analysis is trending towards a shift from monomodal to multimodal. Historically, discourse analysis has primarily focused on a single modality, such as textual analysis of written articles or interpretation of oral speeches. However, with technological advancements and societal progress, multimodal elements like images, sounds, videos, and various interactive media play an increasingly significant role in information transmission. These modalities interweave to construct a more complex and rich information exchange system. Therefore, the inevitable trend towards multimodality in discourse analysis has led to significant changes in the teaching methods of information technology classrooms. Teachers no longer rely solely on language but utilize a variety of non-linguistic symbols, such as facial expressions, gestures, postures, eye contact, and teaching aids, to achieve diverse forms of interaction. Based on multimodal discourse analysis theory and utilizing the ELAN multimodal analysis software, this study selects the teaching video of the special prize-winning entry in the 13th "Foreign Language Teaching Cup" College English Comprehensive Course as case material for quantitative and

qualitative analysis. The aim is to provide a reference for university English teachers to effectively use multimodality in English teaching, thereby enhancing the effectiveness of English language instruction.

## 2. Definition of Key Concepts

### 2.1. Multimodal Discourse Analysis

Discourse takes on various modal forms, with monomodal discourse limited to a single mode, such as auditory modality in listening to news broadcasts or visual mode in reading novels. Multimodal discourse refers to the communication phenomenon that employs multiple senses, including auditory, visual, and tactile, and utilizes various means and semiotic resources such as language, images, sounds, and gestures. It is a type of discourse that integrates multiple semiotic modes for communication (Zhu Yongsheng, 2007), as exemplified by the simultaneous use of visual and auditory modalities in PowerPoint presentations. The multimodal discourse analysis is based on Halliday's Systemic-functional Grammar, with its main theoretical model being the social semiotics theory within System-functional Linguistics. This theory posits that other semiotic systems beyond language are also sources of meaning and that multimodal discourse has the function of expressing meaning. In addition to language, other semiotic systems also possess polyfunctionality, including the ideational metafunction, interpersonal metafunction, and textual metafunction. The framework for multimodal discourse analysis proposed by Martin mainly consists of five aspects: the cultural level, the context level, the meaning level, the form level, and the medium level:

- (1) Cultural level: Ideologies are the main forms of existence of culture and genres, and genre structural potentials are the potential choices for discourse patterns.
- (2) Context level: The context configuration comprises the field of discourse, the tenor of discourse, and the mode of discourse.
- (3) Meaning level: Discourse meaning includes conceptual, interpersonal, and textual meanings.
- (4) Form level: Various form systems realize meaning, and the grammatical relationships of modalities are divided into two categories: complementary (including reinforcement and non-reinforcement) and non-complementary (including inclusion, overlap, expansion, reduction, and situational interaction).
- (5) Medium level: The material manifestations of discourse in the physical world are divided into two categories: language (pure language, paralinguistic) and non-language (physical, such as facial expressions, etc., and non-physical, such as PPT, etc.).

Zhang Delu (2009) proposed an integrated theoretical framework for multimodal discourse analysis based on Halliday's System-functional Linguistics. This framework comprises four levels: cultural context, situational context, content level, and expression level. In college English classroom teaching, based on the similar cultural values and ideological contexts of Chinese students, the teacher's similar classroom discourse situational contexts are an essential means of effectively implementing teaching by developing the modal calls of the content level and expression level for students' language ability. The content level is reflected in both meaning and form, where the meaning level reflects the meaning of discourse: conceptual meaning, interpersonal meaning, and textual meaning. In classroom teaching practices, the conceptual meaning is the development of linguistic potential; the interpersonal meaning reflects the relationship between teachers and students, emotions, and attitudes; the textual meaning reflects the relationships between different semiotic systems in multimodality. The form level is reflected in the interaction of different modalities, such as language, sound, images, and gestures, forming complementary or non-complementary relationships. The expression level mainly refers to the multimodal discourse media system of language (pure language,

paralanguage) and non-language (physical, technological) used to complete classroom teaching activities.

## 2.2. Teacher Talk

In classroom teaching research, teacher talk is often regarded as a “simple code” for providing comprehensible language input to second-language learners, serving as a pedagogical tool for organizing, managing, and guiding second-language classrooms (Hu Xuewen, 2003). From the system-functional linguistics perspective, this view needs to be narrower. Wei Jilan (2011) argues that teacher talk should be a multimodal semiotic system adopted to achieve teaching objectives successfully. In other words, teacher talk refers to all linguistic activities teachers use to teach course materials and organize instruction. Teacher talk invariably involves various teaching stages, such as explaining, questioning, providing feedback, and error correction.

Gu Yueguo (2015) posits that the systematic analysis of discourse that integrates different semiotic resources constitutes multimodal discourse analysis. The teaching process is a quintessential multimodal interactive process. Yang Yi and Chen Changlai (2022) define “multimodal teacher talk analysis” as the process of analyzing how teachers in classroom teaching systems blend different semiotic resources to achieve teaching objectives and the mechanisms of interaction between different categories of semiotic resources.

## 3. Research Design

Due to space limitations, this study selects the classroom video recording of the grand prize winner in the college English comprehensive course group of the 13th SFLEP Cup as the analytical material. The aim is to learn how to effectively integrate and utilize multimodal resources in actual teaching, to explore and reveal the practical application and potential value of multimodal teaching strategies in enhancing the effectiveness of college English teaching, and to investigate the inspirational role of expert teachers’ practices in general college English classroom teaching for other teachers. The total length of the classroom video is 26 minutes, and only the formal class video segment was taken in the study, with a duration of 16.03 minutes, covering four modalities: auditory, visual, behavioral, and environmental. The study explores the following issues: 1) The modalities used by the participating teachers in the classroom and the distribution of modalities; 2) The relationship between different modalities; 3) The relationship between teaching objectives, teaching methods, and the selection of teaching modalities.

In terms of research methodology, this study adopts a quantitative research method and uses ELAN 5.2 video analysis software as a research tool. This software is mainly used for multimodal discourse analysis and can be precise to 0.01 seconds, facilitating segmentation, marking, positioning, and loop playback. According to the time tags, it can achieve synchronized annotation of images and sound, text, and play on demand. The application of the software in teaching video analysis is mainly divided into five steps: First, encode various modalities, as shown in Table 1; Second, define the types of language, which can be defined in the participant or annotator column. This paper chooses to define the modality in the participant column, that is, auditory modality, visual modality, behavioral modality, and environmental modality. Third, the layers, that is, the specific modalities, such as the participants with auditory modality, must be defined. The defined layers are the teacher’s and student’s talk time. Fourth, annotate manually, which requires careful observation and manual annotation; Fifth, statistically analyze the well-annotated video and then refer to the data for discussion and analysis.

**Table 1.** Coding table

Modality	Coding	Modality	Coding
Visual modality	V	Auditory modality	A
Powerpoint	VPPT	Teacher talk	AT
Image	VP	Student talk	AS
Behavioral modality	B	Environmental modality	E
Facial expression	BF	Close distance	EC
Body language	BP	Distance	ED

## 4. Results and Analyses

### 4.1. Utilization of Multimodality in Classroom Teaching and Distribution of Different Modalities

By clicking to view the statistics, the study focuses on analyzing the use and relationships of modalities at each stage, as well as the meanings that the combinations of modalities at each stage attempt to convey. The usage of each modality is shown in Table 2.

**Table 2.** Modality Usage Time Ratios

Layer	Number of Annotations	Min Time	Max Time	Average Time	Median Time	Annotation Time	Percentage of Total Annotation Time
VP	8	0.84	6.83	2.50375	1.01	20.03	1.247
BF	15	0.29	3.282	2.038267	1.938	30.574	1.904
VPPT	52	0.42	9.93	1.683077	0.895	87.52	5.45
BP	31	0.636	24.879	5.468709	4.218	169.53	10.556
AS	33	0.72	27.81	5.280334	3.8	174.251	10.85
EC	8	4.02	64.451	24.213249	15.192	193.796	12.061
AT	83	0.555	34.787	8.098542	6.08	672.179	41.855
ED	-	-	-	-	-	-	-

From Table 2, it can be observed that the dominant modality is the auditory modality of teacher talk (41.86%), followed by close contact (12.06%), student talk (10.85%), body language (10.56%), PPT (5.45%), facial expressions (1.90%), and pictures (1.25%). The dominant modality in English classrooms is the auditory modality, while the visual, behavioral, and environmental modalities reinforce and supplement their discourse.

The use of modalities varies at different stages, and the meanings expressed will also differ: At the beginning part(0.00-0.06), since the earlier part of the content is the teaching segment, the teacher has already greeted the students. At this stage, the teacher smiles and asks the students, "So now, are you ready for taking this adventure?" The primary modality is the teaching discourse, fulfilling the interpersonal meaning. The teacher's facial expression is smiling during the question, and this behavioral modality is a consciously designed modality by the teacher to close the distance with the students and strengthen the teacher-student relationship.

In the warm-up part (0.07-1.05), the teacher's talk under the auditory modality is predominant, followed by student talk, indicating the presence of interactive behavior between teachers and students during this stage. In addition to the auditory modality, the primary modality is visual,

involving PowerPoint (PPT) and images, where the teacher presents pictures and asks students to answer related questions. There is also a notable behavioral modality, albeit brief, that helps to strengthen the teacher-student relationship. When asking, "Imagine you were the man standing on the thin ice, how would you feel?" the teacher acts as if standing on tiptoes. The teacher intentionally designs this behavioral modality, and it is a conscious behavioral modality. The teacher supplements the auditory modality with the visual and behavioral modalities to introduce the topic, allowing students to understand the classroom content and pique their interest. This also reflects both conceptual and interpersonal meanings. The conceptual meaning is manifested in establishing the theme and introducing the term "descriptive language," entering a new field. The interpersonal meaning is reflected in using body language and images to attract students and interactively express the communicative roles of the communicators.

In the review part (1.06-1.50), the primary modality is the auditory, which consists of teacher questioning and explanation, complemented by the visual modality through PowerPoint (PPT). This integrated approach deepens student' comprehension of the material from the previous lesson, effectively bridging the main content of the last class with that of the current one and ensuring the continuity and systematic progression of the teaching content.

In the explanation part (1.51-14.35), as the core part of the course, this stage is how students acquire language knowledge. In this stage, teacher and student talk under the auditory modality is dominant, with PPT as a supplementary tool. The teacher purposefully designs the PPT and pictures to achieve interpersonal functions and enhance students' language abilities. These modalities reinforce and supplement the auditory modality, reflecting the conceptual function. From the annotation results, teacher talk accounts for 41.86%, and student talk accounts for 10.85%, showing a significant difference between the two. However, compared to close contact and body language, etc., student talk ranks third, reflecting that while the teacher delivers the main content, ample opportunities for expression are provided to the students. Frequent interactions between teachers and students in the classroom are conducive to improving students' classroom output expression abilities. In this stage, the content and expression levels of students' language abilities are developed.

In the summary part (14.36-14.50), the teacher skillfully integrates auditory and visual modalities by displaying images related to Chinese proverbs, closely linking visual elements with the teaching theme. The teacher presents the English expressions of these proverbs to the students, thereby deepening their understanding of Chinese culture and its English translations.

In the homework part (14.51-15.57), the synergistic effect of auditory and visual modalities is employed. Regarding conceptual meaning, homework represents a series of actions or steps to be taken. In contrast, from an interpersonal perspective, homework serves not only as guiding advice but also encompasses directive demands on students.

In the closing ceremony (15.58-16.03), auditory and behavioral modalities are used here. The behavioral modality is the teacher bowing to the students while saying 'Thank you!' to express gratitude. This behavioral modality is also intentionally designed by the teacher and achieves conceptual and interpersonal meanings. From the perspective of conceptual meaning, it signifies the end of this classroom process; from the interpersonal meaning standpoint, the teacher's bow conveys gratitude to the students and indicates a change in role relationships, where the teacher and students are no longer in a dominant and submissive relationship.

#### **4.2. Relationships between Different Modalities in Classroom Teaching and Learning**

In multimodal discourse analysis, multiple semiotic systems cooperate to co-construct the meaning of a discourse. How teachers employ multimodality to build discourse meaning in the classroom, the role of different modalities, and which combination of modalities best facilitates

the input being effectively accepted and internalized by students as their knowledge and skills enable creative language output are all critical questions. Zhang Delu (2009) posits that when a single modality is insufficient to express meaning, employing other modalities to strengthen, supplement, adjust, and coordinate can lead to a more complete or complete expression of meaning, allowing the listener to understand the purpose of the discourse. To achieve the best teaching outcomes, teachers actively mobilize various modalities, selectively combine different modalities, and enable multiple modalities to work in concert to maximize meaning construction. He further points out that there are complementary and non-complementary relationships between modalities. Below, based on Zhang Delu's relationships between multimodal discourses, an analysis of the teacher's oral modality and PPT multimodality combinations in the competition video reveals that multiple modalities stimulate students' senses simultaneously, effectively improving communication effects and achieving teaching objectives.

*Example 1:*

*T: Imagine you were the man, standing on the thin ice; how would you feel?*

In Example 1, the teacher presents an image to the students and asks a question. During the questioning process, the teacher performs actions such as tiptoeing and raising their hands above their head. The teacher employs auditory, visual, and behavioral modalities to engage students' multi-sensory participation in learning. This multimodal teaching strategy guides students into a context related to the course theme and clarifies the learning objectives. The visual presentation of the PPT, combined with the teacher's oral explanation and body language, not only reinforces the transmission of information but also arouses students' interest in learning. Moreover, the coordinated use of modalities creates a relaxed and enjoyable learning environment, which helps promote positive teacher-student interaction, reflecting the interpersonal significance of teaching.

*Example 2:*

*T: What we just described is called characterization. Characterization means we use descriptive language to describe emotion, action, and so on to reveal qualities and values. This would be the focus of today's section.*

In Example 2, the teacher uses clear and precise oral discourse modality, supplemented by explicit gestures and PPT displays. The PPT includes well-designed structural diagrams and text highlighted in different colors, using different colors (red, blue, yellow) and arrows to guide students' attention, thereby achieving a synergistic effect of multiple modalities. Coordinating multiple modalities enhances students' ability to capture critical information and promotes their accurate understanding of knowledge points. It also helps students build a systematic knowledge structure, effectively facilitating their cognitive development and improving teaching outcomes.

*Example 3: For now, let's explore the first quality. I want you to underline the descriptions for these questions: Who are the students in paragraphs one to two? You have 30 seconds.*

*Example 4: Again I want you to discuss in group of four and talk about your adventure in daily life, and why do you take adventures?OK, you have one minute.*

In Examples 3 and 4, the teacher organizes student activities, using the auditory modality to require students to discuss in groups, with the PPT displaying problem content and discussion points. In Example 3, the teacher embeds suggestive content in the PPT, which serves as a learning scaffold, helping students complete exercises more effectively and playing an active guiding role. In Example 4, the teacher links the textbook content with students' daily life through images that are close to the students' life and can better exercise their thinking and expressive abilities. In the subsequent group discussions conducted by the students, the points displayed on the PPT provide a framework and direction for discussion, playing a guiding and

supportive role. Students mainly use the auditory modality, supplemented by the visual modality.

In contrast, the teacher mainly uses the environmental modality, walking among the students to guide the discussion during their discussion or problem-solving. In this process, the visual elements of the PPT and the teacher's oral guidance jointly enhance the student's learning experience, increasing their participation and deepening their mastery of the content, thereby improving the overall teaching effect. The teacher successfully inspires students' thinking and promotes comprehensive development by integrating multiple modalities in teaching strategies.

*Example 5:*

*T: Now, I want you all to close your eyes. Come on, close your eyes, and try to feel the jump with me. I reached the soft edge, felt it give beneath my weight, saw the waters, launched myself, soared across, and crashed. How do you feel?*

*S: Panicky.*

*T: Yes! It's so scary and panicky. So, what was the quality needed for this life-or-death jump?*

*S: Courage.*

In Example 5, the teacher guides students to close their eyes and imagine experiencing the action of jumping, mainly relying on the auditory modality. The teacher uses changes in intonation, including rising and falling tones, and infectious language to successfully immerse students in a specific context. While using a falling tone to impart knowledge and control the classroom, an excellent teacher will cleverly insert rising tones to promote interaction, ignite the passion for learning, inspire students' thinking, and thus create a lively classroom atmosphere (Li Wanxuan, 2024). The visual modality is secondary, with the teacher using images and text displayed on the PPT to help students more accurately understand the meaning of the teacher's speech. After the teacher asks a question, students respond based on the teacher's tone and the text and images provided by the PPT. The teacher then gives positive feedback and affirmation. The skillful use of the teacher's infectious words and intonation can significantly enhance the student's immersion, and timely changes in intonation are very effective and necessary for emphasizing the key points of the teaching content.

The relationship between modalities is not static and varies across the different stages of teaching implementation. The role of the same modality can also differ at various points. However, in most cases, the combination of the oral modality and the multimodality of PPT creates a synergistic effect. The PPT provides supplementary textual information for oral discourse, aiding in the content production of the oral modality. At the same time, the oral modality assists students in comprehending the content of the PPT multimodality, especially when explaining more complex or significant concepts. Using multiple modalities helps students learn efficiently and benefits the input and output of knowledge. However, each modality uniquely presents information, requiring students to simultaneously recognize and process these multi-dimensional modalities. If the modalities are well-coordinated, it will facilitate the understanding of meaning; on the contrary, if there is a lack of coherence or the modalities are in conflict, it can have a negative impact, increasing the cognitive load on students or leading to misunderstandings.

### **4.3. The Relationship Between Teaching Objectives, Teaching Methods, and the Selection of Teaching Modalities**

In foreign language classroom teaching, the overall goal of instructional design is to develop students' meaning potential, which means that teachers need to develop students' whole language proficiency through multimodal teaching, including linguistic knowledge, communicative competence, and social interaction skills (Zhang Delu, 2010). Multimodal

teaching refers to teachers using various teaching methods and resources, such as text, images, sound, and movement, during the teaching process to enrich and deepen students' understanding and use of language.

Teaching methods are the main means to achieve teaching objectives in pedagogical practice. The selection of teaching methods will directly affect the choice of teaching modalities, and the selection of teaching modalities serves to realize better teaching objectives (Meng Xiaoping & Zhang Yufeng, 2022). Regardless of the teaching method used, such as Case-based Teaching, Task-based Language Teaching, Production-Oriented Approach, Outcomes-Based Education, etc., the vast majority input information in a multimodal way, and students receive process and output information in a multimodal way to achieve the enhancement of knowledge and skills. Among the various modalities, there are often primary and secondary modalities. For example, the audio-lingual method mainly focuses on the auditory modality, while the reading method mainly focuses on the visual modality. In each specific teaching segment, teachers must construct all the necessary resource symbols for teaching practice and choose the appropriate modality according to the teaching method to fully express meaning and enable students to understand the discourse fully. The teaching methods and modalities of the participating teachers are carefully designed to achieve teaching objectives. For example, the Production-Oriented Approach often introduces courses with audio-visual modalities. In the driving, facilitating, and evaluating teaching stages, teachers use the multidimensionality and dynamics of modalities, using a combination of auditory and PPT-based visual modalities to construct meaning. Various modalities form roles that strengthen and complement each other, promoting the achievement of teaching objectives. In summary, in foreign language classroom teaching, teachers should use multimodal teaching and choose teaching methods and modalities reasonably to enrich and deepen students' understanding and use of language, thereby improving their language proficiency and social interaction skills.

## 5. Conclusion

Using ELAN software and careful observation, a quantitative and qualitative study of the teacher's discourse in the competition found that teachers design different modalities in each stage to achieve different teaching objectives and communicative purposes. Various modalities work in synergy to realize the functions of modality symbols, reflecting the meanings of each symbol and symbol combination. This study provides a detailed analysis of the classroom video of the grand prize winner in the 13th "SFLEP Cup" of the college English comprehensive course group. Multimodal discourse analysis theory explores how teachers use various modalities, such as auditory, visual, behavioral, and environmental, to enhance teaching effect and improve student learning outcomes. The results indicate that teachers skillfully combined multiple modalities in the classroom, conveying knowledge content and successfully establishing interactive relationships with students, increasing their participation and interest in learning. Teacher discourse, as the main auditory modality, dominates the whole teaching process. At the same time, using visual modalities such as PPT and pictures provides visual support for the content and helps students better understand and remember the main points. Furthermore, behavioral modalities such as facial expressions and body gestures were consciously used at specific moments to help reinforce interpersonal meaning and enhance the emotional connection between teachers and students. Environmental modalities, on the other hand, create spatial layouts conducive to learning and communication through close and distant interactions.

Multimodal teaching is a complex and challenging process that requires teachers to have superior instructional design skills and a good site strain capacity. The analyses in this study are expected to bring specific insights and references to university English teachers, helping

them use multimodal resources more effectively in their teaching practice, stimulating students' learning potential, and cultivating linguistic and social-communicative competence.

## References

- [1] Halliday, M.A.K.: *An Introduction to Functional Grammar. Second edition* (Foreign Language Teaching and Research Press, China 2000), p. 3-64.
- [2] Martin, J.R.: *English Text: System and Structure* (John Benjamins Publishing Company, Netherlands 1992), p.493-590
- [3] Xuewen Hu: Features and Functions of Teacher Talk, *Shandong Foreign Language Teaching*, (2003)No.03, p.39-43.
- [4] Yueguo Gu: Research on Multimodal Sensory Systems and Language, *Contemporary Linguistics*, Vol. 17 (2015) No.04, p.448-469.
- [5] Wanxuan Li, Multi-Modal Teaching Analysis Research Based on ELAN: Taking the Video of Winners of the National Vocational College Teaching Skills Competition in 2021 as an Example, *Journal of Shandong Institute of Commerce and Technology*, Vol. 24 (2024) No.04, p.66-70.
- [6] Xiao ping Meng and Yufeng Zhang: Research on College English Classroom Teaching from the Perspective of Multimodal Discourse Analysis: Taking the 2020 FLTRP Star Teacher Contest as an Example, *Journal of Higher Education*, Vol. 08 (2022) No.17, p.92-94+99.
- [7] Jilan Wei: Multi-modal Analysis of Teacher Talk in College English Intensive Reading Class, *Journal of Chengdu Normal University*, Vol. 27 (2011) No.3, p.98-102.
- [8] Yi Yang and Changlai Chen: Multimodal Discourse Analysis: A New Paradigm of Teacher Discourse Research, *Contemporary Foreign Language Studies*, (2022) No.04, p.144-153.
- [9] Delu Zhang: On A Synthetic Theoretical Framework for Multimodal Discourse Analysis, *Foreign Languages in China*, Vol. 6 (2009) No.01, p.24-30.
- [10] Delu Zhang: Preliminary Investigation into the Concept of Design and the Selection of Modalities in Multimodal Foreign Language Teaching, *Foreign Languages in China*, Vol.7 (2010) No.03, p.48-53+75.
- [11] Yongsheng Zhu: Theory and Methodology of Multimodal Discourse Analysis, *Foreign Language Research*, (2007) No.05, p.82-86.