


Open Access Article

 <https://doi.org/10.55463/hkjss.issn.1021-3619.61.38>

Intersemiosis in Classroom Discourse: A Dual Perspective on Systemic Functional Linguistics and Multimodal Interaction Analysis

Cheng Yuan*, Roslina Mamat, Mohd Azidan Abdul Jabar, Ma Xiaoxia

Faculty of Modern Languages and Communication, University Putra Malaysia, Serdang, 43400, Selangor, Malaysia

Received: April 16, 2023 ▪ Reviewed: May 15, 2023

▪ Accepted: June 19, 2023 ▪ Published: July 30, 2023

Abstract:

This study explores how the synergistic relations of multimodal symbolic resources construct classroom discourse. To this end, we combine two perspectives—multimodal interaction analysis and systemic functional linguistics to divide modal integration into two dimensions, meaning and representation, based on which we construct a new intersemiotic analytical framework to explore how different symbolic modes combine to provide meaning in the Chinese as a foreign language classroom. The findings show that teachers in such courses focus more on the visualization process of thinking and that coordinating rich symbolic resources exercises students' Chinese thinking skills and learning initiatives. At the same time, we demonstrate that combining these two research perspectives can strongly contribute to the development of multimodal interaction analysis in teaching and learning, providing practical implications for improving the quality of teaching. Furthermore, the analytical framework proposed in this paper can be applied to more types of teaching practices in the future, offering a valuable reference for multimodal applications of classroom discourse.

Keywords: classroom discourse, modal integration, multimodal interaction analysis, systemic functional linguistics.

课堂话语中的联际现象：系统功能语言学与多模态交互分析的双重视角

摘要：

本研究探讨多模态符号资源的协同关系如何构建课堂话语。为此，我们结合多模态交互分析和系统功能语言学两个视角，将模态整合分为意义和表征两个维度，在此基础上构建新的符号间分析框架，探讨不同符号模式如何结合起来提供意义。汉语作为外语课堂。研究结果表明，此类课程的教师更注重思维的形象化过程，搭配丰富的符号资源锻炼了学生的语文思维能力和学习主动性。同时，我们证明，结合这两种研究

Corresponding Author: Cheng Yuan, Faculty of Modern Languages and Communication, University Putra Malaysia, Serdang, Malaysia; email: GS64116@student.upm.edu.my

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)

视角可以极大地促进教学中多模态交互分析的发展, 为提高教学质量提供实际意义。此外, 本文提出的分析框架可以应用于未来更多类型的教学实践, 为课堂话语的多模态应用提供有价值的参考。

关键词: 课堂话语、模态整合、多模态交互分析、系统功能语言学。

1. Introduction

With the development and use of educational technology resources, classroom teaching gradually shifts from a traditional pedagogical model to a typically multimodal experience with a complex discursive pattern. Although language was previously seen as the primary resource in teaching and learning, from a multimodal perspective, meaning is constructed by combining various symbolic resources (Jewitt et al., 2016, p. 2). After that, it becomes essential to investigate how the modalities interplay. Collaborative relations have been the focus of research in studying interrelations between multimodalities. The reason for this is that multimodal discourse analysis studies are subjective (Zhanzi & Danyun, 2012), and collaborative relations can, to a certain extent, reflect the psychological assumptions of the designer, who, on the one hand, chooses the necessary modalities and, on the other hand, organizes them in an orderly manner to form a modal combination as a whole in order to accomplish the communicative task most effectively (Kress, 2010). Therefore, it is essential to explore the interaction of different symbolic resources in expressing meaning to accurately decode complex multimodal discourse. In past studies of classroom discourse, analysts have focused on the linguistic model of teacher-student interaction (Walsh, 2006). With the intervention of a multimodal perspective, scholars have applied it to teaching and learning. They discovered that language cannot explain the meaning of information made up of various modalities. Language and literacy are partial carriers of meaning (Kress, 2003, p. 35). O'Halloran (2007, p. 79) further noted that the study of linguistic discourse has theoretical constraints and tends to oversimplify and misrepresent the true nature of teaching practice. In recent years, multimodal classroom teaching has received attention from researchers and been applied to discuss the construction of teaching discourse in different lesson types. With the gradual advancement of research, the synergistic relationship between symbols has been discussed by scholars, such as the interplay between gestures and language modalities (Lim, 2019, 2021), classroom power relations embodied under the integration of symbolic resources (Komarawan, 2019), and the construction of the meaning of multiple symbolic resources in children's classroom discourse (Taylor, 2014). Different symbolic resources in the classroom environment carry a rich semantic potential (Halliday & Matthiessen, 1999). Moreover, the teacher's design and integration of these symbolic resources directly influence the impact on teaching effectiveness and objectives. However, there are still some issues with the

existing articles on intersemiotic relations. First, it ignores flexible modality change in different teaching stages and fails to provide targeted guidance for dynamic classroom teaching. Second, the theoretical perspective of modal synergy is single and does not form a systematic theoretical framework, leading to a less in-depth analysis. Finally, the researchers' analysis is only focused on internal teaching activities, ignoring the influence of the social environment on the modal organization. This study adopts a dual perspective based on systemic functional theory (Halliday, 1994) and multimodal interaction theory (Norris, 2004), concentrating on the interactive changes made by teachers in spoken language, image, gestures, and new modalities, systematically analyzing the interaction of symbolic resources in terms of structure and meaning, and building an analytical framework for multimodal interaction in classroom discourse on this basis. Teaching practice is dynamic.

The purpose of this article is to study teachers' use of modalities in different teaching stages through the integration of these two research perspectives, to explore the discursive meaning of modalities in different relationships, and to present the reader with the complete process of constructing modal symbols. Through the analysis and interpretation of teachers' use of symbols, this study can identify problems in teaching and provide valuable references for teachers to coordinate multimodal resources and optimize the efficiency of the teaching classroom.

2. Literature Review

The study of modal collaboration in classroom discourse will be approached from two perspectives: the object of the study of modality and the content of the study of modal collaboration. The former focuses on the categories and characteristics of modality that analysts emphasize in discourse in the classroom. The latter concentrates on the modal synergy research scope and analysis methods.

In terms of objects of study, most previous research on classroom discourse has concentrated on spoken and written language in the classroom (Nystrand, 2006). With the continued development and use of educational technology resources, Unsworth and Cleirigh (2009) contend that "it is now widely accepted that literacy and literacy pedagogy is no longer confined to the domain of language." This understanding has "important epistemological and research methodological implications" (Jewitt, 2008, p. 245). Gradually, multimodal research on teaching in classrooms has emerged in the past few years (O'Halloran, 2007; Greiffenhagen, 2008). According to the multimodal

perspective, language frequently coordinates with other symbolic resources to form meaning and does not always dominate (Kress & Van Leeuwen, 2006). This point makes up for the fact that classroom studies have previously focused on language. Based on the above empirical understanding, scholars have adopted a multimodal viewpoint to create richer symbolic resources and meanings for classroom discourse. Examples include gesture, gaze, and classroom space (Lim, 2019, 2021).

In terms of modal collaboration, multimodal studies of classroom discourse focus on “how different types of symbols are combined into an integrated, multimodal whole” (Jewitt et al., 2016, p. 2). Therefore, exploring the interplay (synergistic relations) of different symbolic resources in meaning is the focus of research on the classroom discourse. For this goal, academics have offered scientific analysis methods (Jewitt, 2013; Taylor, 2014; Takahashi & Yu, 2016). Among these, the most significant analytical approaches focus on the following categories: Royce (1998) proposes a theoretical framework of ‘intersemiotic complementarity,’ emphasizing the integration of meaning between images and language. Lemke (1998) suggests that the superposition of meaning in discourse occurs inter-semiotically. Martinec and Salway (2005) build on Kress and van Leeuwen’s (2006) visual grammar by using logical semantics (Halliday, 1994) to construct image-text relations. O’Halloran’s (2003) theory of symbolic metaphor argues that semantic reconstructions of meaning are produced by inter-symbolic transformations, where the semantics of a functional component is changed when that component is reinterpreted through other symbols. Among the above analytical approaches, scholars have expanded the scope of the synergetic study of modality to involve various aspects such as semantic space, integration, and constituent interface. The theoretical basis of most of these analytical approaches comes from systemic functional linguistics (Halliday, 1994), which extends the linguistic potential to the meaning potential of non-linguistic symbols and focuses on the semantic level to explain the construction of meaning in multimodal discourse. On this basis, analysts apply relevant theories to classroom teaching and learning, with discussions focusing on how relations within symbolic resources are constructed to achieve specific discursive meanings. These meanings include students’ ability to learn (Pinnow, 2011), teachers’ power of discourse (Komarawan, 2019), and reading and writing skills (Karatza, 2020). In summary, the relationship between linguistic and image symbolic resources has been the most discussed topic in previous studies. With the use of multimodal information in classroom teaching, intersemiosis has become the focus and trend of classroom discourse research, providing valuable guidance for teaching practice and widening the field of exploration in discourse construction. However, there are still the following gaps in the relevant investigation: First, there is a lack of exploration of the dynamic

nature of modality, ignoring the flexibility of modality to change through different stages of teaching and learning. To a certain extent, this fails to provide targeted guidance for dynamic and interactive classroom teaching. Second, in analyzing modal integration, most scholars follow the perspective of systemic function theory and focus on the level of semantics. In symbolic systems, discourse meaning promotes the choice of modality and is reflected by modal forms (Zhang & Hu, 2019). This implies that modal forms are the primary way of achieving meaning, and the lack of form-level analysis can lead to an incomplete symbol system. Finally, the reasons for modal integration are not sufficiently explored. The interaction of symbolic systems does not arise spontaneously but is controlled and constrained by certain conditions (Kress, 2010). The lack of exploration by analysts of the needs and motivations of teachers’ choice of modality does not fundamentally offer practical recommendations for teaching practice.

Therefore, this paper will focus on the interactive changes of spoken language, images, actions, and new modalities in classroom discourse and combine the systemic functional linguistics (Halliday, 1994) and multimodal interaction analytical approaches (Norris, 2004) to construct a new analytical framework for multimodal interaction, which is needed not only to reflect the motivational factors of modal integration but also to explore the integration relations at two systemic levels: form and semantics. It is intended to guide teachers using multimodal symbolic resources to design effective classroom instruction.

3. Research Methods

3.1. Data Description

In recent years, Chinese as a Foreign Language (CFL) teachers have been in greater demand worldwide, and there has been a continuous rise in the number of studies on CFL teaching and learning, which has placed higher requirements on the quality and standard of teachers’ teaching (Tsong & Cruickshank, 2011). Unlike other traditional courses, the inflow of international Chinese learners has led to linguistically diverse classrooms, which poses a significant challenge to the single-language teaching model, which requires teachers to change their traditional teaching strategies to achieve cross-language teaching more flexibly and effectively (Yang, 2019). This study provides an example of a comprehensive and excellent model lesson in teaching CFL. In this case, the teacher uses rich pictures, objects, slides, and animations as teaching aids and ultimately achieves the teaching purpose through the cooperative application of symbolic resources, which can serve as a typical model in multimodal classroom discourse.

In teaching CFL, the language points (vocabulary and grammar) are both the focus and the difficulty of classroom teaching. Teachers must ensure that students master the course’s vocabulary and grammar to use

them flexibly in scenario practice. Because of this, teachers need to design many activities to complete the teaching task at this stage, and then they need a wide variety of different modalities to achieve this (Zhanzi & Danyun, 2012). Therefore, this study takes this teaching stage as the object of analysis, focusing on spoken language, images, gestures, additional modalities, and how the modalities interact during each activity step.

3.2. Methodology

3.2.1. Intersemiotic Analytical Framework

Halliday's (1994) systemic functional linguistics mainly initiated massive research on multimodality, in which the context of culture and situational contexts are essential concepts that provide systematic theoretical support for the analytical approach to multimodal discourse. Multimodal interaction analysis (Norris, 2004, 2011) builds on mediated discourse analysis (Scollon, 1998, 2001), which uses the activity as the unit of analysis and provides a scientific approach to deconstructing the complexity of multimodal interactions. These two theories are 'social functional perspectives' (Bateman, 2011) on MDA and have much in common. From the perspective of modal integration, these two theoretical methods demonstrate a significant complementarity and provide a thoroughly analytical approach to explain modal synergies. This view is expressed in two aspects: In terms of synergistic motivation, any communicative activity is facilitated by a specific situational context and cultural context (Zhang & Hu, 2019). Differences in context, communicative purpose, and interest inevitably produce different types and content of activities and, consequently, different discursive meanings, further influencing the modality choice (Bateman et al., 2017). It can be seen that modal collaboration is a complex operational process rather than a simple selection and combination procedure. From the external cultural context to the internal discourse meaning, all produce different levels of constraints on modal integration. SFL goes beyond the internal of the discourse to explore the social environment and symbolic features from an external perspective, which provides a broad analytical view of the dynamics of modal integration. In terms of a synergistic approach, multimodal interaction analysis (Norris, 2004, 2011) focuses on the internal discourse, taking the specific human activity as the unit of analysis and exploring the issue of the structural configuration of multimodality in the activity and the rules for using various modalities (Jewitt et al., 2016), which focuses on the structural features and interactive changes of the symbols. In addition to reflecting the various modalities' formal collocations, there is a need to point out that the meaning created is integrative and synergistic (Matthiessen, 2009, p. 12). At the level of semantics, we can use the logic-semantic relations of SFL, which have the advantage of having the projection system needed to account for projected meanings in

visual symbolic resources, providing analysis of the integration of modality in semantics. Because these two analytical approaches are complementary, it is possible to investigate how modalities interact in terms of formal and semantic symbol systems and modal changes at different stages. This study combines the above two analytical approaches and proposes a new framework for multimodal interaction, as shown in Figure 1. This framework not only constructs the motivation of modal collaboration from an external perspective but also focuses on the internal discourse, dividing modality into two dimensions of meaning and representation and discussing the synergistic relations between semantics and the form of modality at different stages individually. It provides a complete analytical path for how discourse achieves multimodality through modal integration.

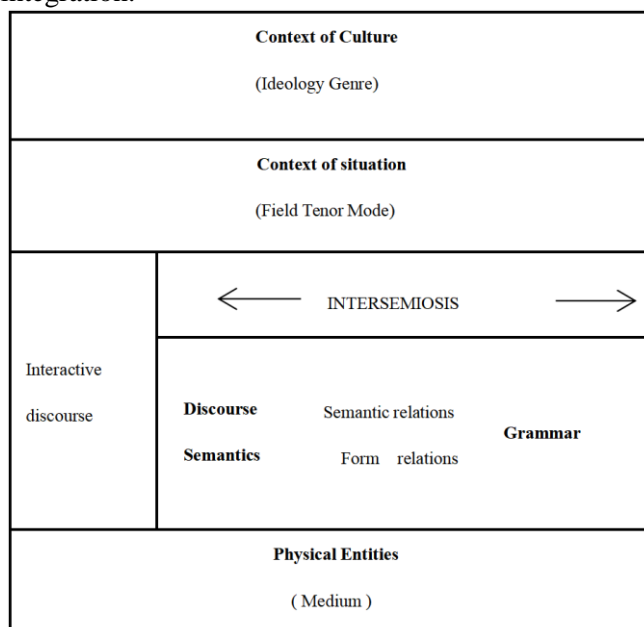


Figure 1. Intersemiotic analytical framework

4. A Multimodal Synergistic Analysis of Classroom Discourse in CFL

4.1. Context of Culture

In multimodal discourse analysis, it is not necessary to describe the complete socio-cultural system of the context but only the discourse, knowledge, and identity associated with the event, as well as the corresponding social space, action system, and multimodal constitutive conditions (Zhang & Hu, 2019). CFL is mainly manifested as classroom-based teaching, in which the teacher stands at the front of the room, and the students sit in their assigned places and complete the teaching content according to the classroom activities organized by the teacher. The overall goal of the course is to improve students' basic Chinese language skills (listening, speaking, writing, and reading) and to acquire some general knowledge of Chinese culture. Genre, as an embodiment of ideology, is a potential action system of behavioral potential (Halliday, 1978) expressed in purposeful, structured activities (Martin,

1992). The teaching case selected for this paper is an integrated elementary Chinese class that aims to develop students' initial listening, speaking, reading, and writing skills to meet the basic communicative needs of daily life, study, and ordinary situations (Jing, 2016). The teaching time of the case was 46 minutes, and the lesson included six genre structure components, each of which had different microlearning objectives, as shown in Table 1. Each of these six genre elements comprised a variety of higher-level actions, and each higher-level action contained a variety of subsidiary activity stages, which together formed the complete classroom teaching model.

Table 1. Genre structure components in CFL

Teaching phase	Micro-learning objectives	Teaching time
Introduction	Guide students into the class	00:09 (9 s)
Review	Consolidate basic knowledge and prepare for learning new ones	00:10—01:49 (1 min 39 s)
Language Point Teaching	Learn vocabulary and grammar and acquire keywords and expressions	01:50—22:50 (21 min)
Text Interpretation	Understand the content of the text and master the usage of vocabulary in the scene	22:51—45:00 (22 min 9 s)
Classroom Summary	Review class content and reinforce the learning points	45:01—45:42 (41 s)
The Assignment of Coursework	Clarify the learning tasks after class	45:43—46:44 (1 min 1 s)

4.2. Context of Situation

Situational context is a specific manifestation of cultural context and includes three variables: field (the social activity involved in the discourse), tenor (the social role relationship between the communicative parties), and mode (the medium and channel of communication) (Halliday, 1978). The teaching case studied in this paper can be described as follows. The teaching topic "Does this red dress look good?," through the discussion of the color of the dress, improves the student's proficiency in the topic of the clothes-buying scenario, especially in listening and speaking skills. The communicators are the teacher and the students. The teacher is the organizer and knowledge provider of this activity; the students have studied Chinese for less than a year and have a low introductory level. The communication method is face-to-face verbal communication, including spoken language, gestures, physical objects, images and animations provided by the PowerPoint equipment, with a combination of audio-visual sensory communication between the participants.

4.3. Interactive Discourse

This study will focus on the phase of language point teaching, which consists of five sub-higher-level actions, under which there are different sub-secondary higher-level actions, and these specific activity steps

often appear to be embedded in each other or a cyclical relation.

In the introduction phase, which aims to guide the students to the start of the new lesson between 01:50 and 02:04, there are two sub-secondary actions: questions and answers, which are embedded in each other and consist of a continuous chain of lower-level activities: spoken language, images, and gestures. The structural configuration is characterized by the dominance of the image and spoken language modalities, which ensure the transmission of information content and the subordination of the gestural modality. In terms of modal density, spoken language is a high-intensity modality in which the teacher immediately asks questions when instructed by the image, and the students respond to the questions. In logical-semantic relations, the image serves as an information clue, elaborating the information in the image in a spoken way, such as "Look, what is in the wardrobe?" or "Whose clothes are these?" This is where the spoken language makes the information more detailed and specific. During this stage, the teacher's spoken language is always the focus of the student's attention, and the discourse moves smoothly from "Mr. Shen's wardrobe" to the content of the new lesson.

The vocabulary teaching phase, which helps students understand the meaning of the words and learn to use them, occurs between 02:05 and 05:47s. This activity consists of four sub-secondary actions: reading aloud, explaining, asking questions, and practicing. These four activities alternate several times and are realized by a chain of lower-level activities: spoken, images, gestures, and physical objects. The structural configuration demonstrates the dominance of spoken and image modalities collaborating for the higher-level action of teaching vocabulary. Gestures and physical objects (textbooks, books) are supplementary to the spoken modality. Regarding modal density, the image modality is the high-intensity modality, and the teacher's teaching activities are sustained as the images emerge, for example, by directing pupils to read aloud based on images, using images to explain vocabulary, and applying images to pose questions. At this stage, the teacher adds a physical modality, mostly along with gestures, to indicate the question's meaning, followed by feedback and correction of the response immediately after the spoken language. At this point, medium-intensity spoken language is subordinate to gestures and physical objects. In terms of logical-semantic relations, the images and spoken language are closely linked during the reading and explanation stages, and the teacher further illustrates the verbal expression based on the colors shown in the images, which have the same generality in content. During the practice stage, the images and objects extend the semantic relations. For example, teachers combine objects and guide them to say "a book" based on their existing knowledge, enabling them to apply their new knowledge flexibly. During questioning, the teacher gestures as a discursive projection of the spoken language, carrying specific

discourse meaning. Specifically, the teacher makes the 'one' gesture with the left hand while holding the book in her right hand to indicate that the student should initiate the full expression 'a book,' followed by the teacher's correction and phonetic demonstration.

In the two phases of the organizing activity, the aim is to have the students to practice conversation based on the vocabulary. It occurs in the periods 05:48-08:23 and 21:39-22:50. There are three sub-secondary actions: assigning tasks, group practice, and performance presentations. In contrast to the first two, the teacher's primary responsibility is to monitor the activity process. The modal structure is configured so that the three modalities, spoken language, pictures, and physical objects, together constitute the completion of a higher-level action, with the spoken modality being of higher intensity and the other modalities having medium intensity. In particular, the teacher's speaking modality acts as a high-intensity modality, focusing the student's attention on the teacher's speaking to make them understand the activity tasks that must be completed. During group practice and presentation, the students' speaking modality changes from weak to strong, and the teacher's focus shifts to the students' practice to examine their mastery of the content. In terms of logical-semantic relations, the teacher's spoken language further illustrates the images when the task is assigned. In practice, students use the information in the pictures and objects to interact with their peers in a question-and-answer process. For example, students commented: "Do you like it?" "I think it looks good," etc. At this point, the spoken language and the objects present additive semantic relations, extending the detailed description of these objects. According to a study of logical semantics, there are some distinctions in the kinds of activities teachers design between these two phases. In the first, the teacher trained through established pictures, and students did not have much opportunity to extend their oral expressions with more focused training on vocabulary. In the second, the teacher allowed students to use objects around them for practice and began to broaden the scope of communication, with students engaging in more flexible and free extension training.

In the grammar explanation phase, teaching enables students to master the structural sentences and express the whole meaning of the discourse. The timing is between 08:24s-22:50s. During this phase, it is composed of three sub-secondary actions: introduction, explanation, and practice. The configuration of the modal structure is shown by the dominance of images and spoken modalities, objects, gestures, actions, and the page layout intricately interacting with each other, presenting a complexity of modalities that enriches the expression of meaning. Specifically, the image modality acts as a high-intensity modality during the introduction stage. In explanation, the image modality remains high, and the students' body posture, eyes, and actions are focused on this. The spoken modality interprets and

asks questions as information from the image appears, the student's focus shifts from the image to the teacher's speech, and they respond accordingly. In addition, the teacher also uses action and object modalities to support the spoken language. In this case, the objects and actions act as high-intensity modalities to construct new content information, with the spoken acting as a secondary position to reinforce and correct it. In practice, the complexity of the modalities is more straightforward, and accordingly, the structural configuration of the modality changes. The students' speaking time increases significantly, while the teacher's speaking intensity is lower, in a lower modal action. The gestural modality is in a secondary position, guiding the students to say the information in conjunction with the images. In addition, this paper notes the layout of the teacher's presentation of images, organizing the different lexical positions according to word categories, combined with gestural pointing, directing the students to read aloud the content, reinforcing the use of vocabulary in grammar in a visual way, and enriching the variety of the training.

Regarding logical-semantic relations, the spoken language further illustrates the image during the introduction and explanation stages with no new discursive information attached. When the physical modality is present, the object is a discursive projection of the spoken, leading to active expression, followed by feedback and correction by the teacher. In addition to the information offered by images, the teacher also uses the action modality. For example, when teaching 'put on' (穿) and 'didn't' (没), 'put on the coat' and 'take off the coat' are illustrated, and students immediately grasp them through visual representation. The teacher then immediately extends this content, using 'didn't' with gestures to lead students to say 'didn't see,' 'didn't write,' and 'didn't think.' "The gesture is a discursive projection of the spoken word. This complete closed-loop training allows for concisely presenting new knowledge but is also integrated with old knowledge and used flexibly. This paper noted the teacher's addition of a page layout at this phrase, further reflecting the semantic relationship between modalities. First, the images presented rich pictures, which were used to project them as discourse, allowing students to articulate what they saw adequately. Second, students can independently evaluate the information in the images, at which point the spoken and the images present an additive semantic relationship that extends the detailed description of the images. Finally, the images are shown more richly, expanding students' communicative options, and oral training changes from single-sentence training to turn-talking, with the spoken word elaborating further on the images. The results are gradual reduction in the students' dependence on knowledge and a progressive approach to training their spoken language and consolidating their discursive thinking.

The analysis of the five teaching phases above

clearly shows how the modalities we are concerned with change continuously, as shown in Figure 2.

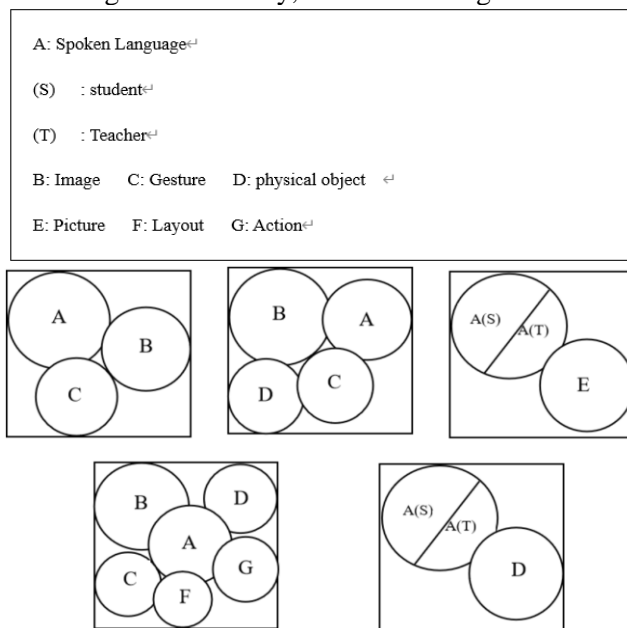


Figure 2. Modal configurations for different teaching phases

5. Discussion and Conclusion

This study found that teaching CFL differs from traditional course teaching; such courses focus more on visualizing students' thinking and exercising their speaking skills and learning initiatives by coordinating rich symbolic resources. In addition, this research provides several significant findings and interpretations through an analysis of language point teaching.

First, in terms of synergistic motivations, the different aims of the teaching phases lead to different configurations when teachers make modal choices. In Figure 2, the modalities are in different hierarchical positions relative to each other and produce changes at each stage due to different target focuses. In general, when explaining new knowledge such as vocabulary and grammar, teachers use objects, actions, pictures, and other modes as primary modalities, which are seen as 'projective sentences' (Halliday, 1994) with the spoken language as the projective object, leading students to take the initiative in oral training. In the organization of the activities and practice phase, students' spoken language, objects, and pictures are used as the primary modality in a mostly semantically extended relationship to achieve a flexible mastery and application of knowledge. In such classes, teachers clearly emphasize the visualizing process of thought, maximizing the meaning potential of each modality and assisting students in understanding the content visually and simply, dramatically developing students' Chinese thinking abilities.

Second, in modal relations, this study discovered that a modality contains the most discursive information when it acts at a high intensity. Other modalities indicate the complement and interpretation of the primary modality when they are in a secondary position. Any modality can be a high-intensity modality (Norris, 2011), and when images or spoken words are

high-intensity modalities, students' attention is more intense and focused on a single modality. When physical objects or actions are high-intensity modalities, modal density occurs mainly in complexity, requiring both or all three to work together to represent meaning entirely. Semantically, the complexity of the modal structure increases, but students' level of comprehension is achieved with simplicity, which is further evidence of the appropriateness designers follow in teaching (Zhang & Hu, 2019) to ensure more effective teaching and learning by selecting the most appropriate modality for achieving communicative purposes (Kress, 2010, p. 27).

Finally, regarding the type of activity, the teacher emphasizes combining teaching and practice with frequent transitions between the two activities. Although this mode of delivery allows for flexibility in the application of knowledge, it demands a high level of concentration from students, who have to complete two to three transitions in a short time and need to keep up with the teacher's pace of the lesson at all times to complete the teaching task. Therefore, teachers need to know the learning level of their students in advance and design activities more suitable for them.

This study explores classroom interaction between teachers and students using two theoretical perspectives: multimodal interaction analysis and systemic functional linguistics. The analysis of examples shows that the two research perspectives can be connected not only in terms of theoretical paradigms but also at the application level of classroom discourse, where there is a great deal of scope for complementarity. Theoretically, combining the two paradigms provides a more complete and systematic analytical approach to explaining the synergistic relationship between modalities. In the application of practical discourse, the complementarity of the two analyzes dynamic classroom discourse more adequately and comprehensively, providing targeted guidance for improving the quality of classroom teaching, which has important practical implications for pedagogy and the professional development of teachers.

6. Limitations and Further Study

Concerning the limitations of this study, the realization of the synergistic meaning of multimodal symbolic resources should be interpreted in a specific context and not generalized. Furthermore, using various visual media to enrich student's learning experiences and develop students' and teachers' multimodal communicative skills could be the focus of further research.

Authors' Contributions

All the authors were involved in the study, working as a research team to conduct the literature review, data collection and analysis of the data, and review of the paper.

References

- [1] BATEMAN, J.A. (2011). The decomposability of semiotic modes. In: O'HALLORAN, L., & SMITH, B. (eds.) *Multimodal Studies: Exploring Issues and Domains*. London: Routledge, pp. 17-38. <https://doi.org/10.4324/9780203828847-10>
- [2] BATEMAN, J., WILDFEUER, J., & HIPPALA, T. (2017). *Multimodality: Foundations, Research and Analysis - A Problem-Oriented Introduction*. Berlin: De Gruyter Mouton. <http://dx.doi.org/10.1515/9783110479898>
- [3] GREIFFENHAGEN, C. (2008). Video Analysis of Mathematical Practice? Different Attempts to "Open Up" Mathematics for Sociological Investigation. *Forum: Qualitative Social Research*, 9(3), 32. <https://doi.org/10.17169/fqs-9.3.1172>
- [4] HALLIDAY, M.A.K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- [5] HALLIDAY, M.A.K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.
- [6] HALLIDAY, M.A.K., & MATTHIESSEN, C.M.I.M. (1999). *Construing Experience through Meaning: A Language-Based Approach to Cognition*. London: Cassell.
- [7] JEWITT, C. (2008). *Technology, Literacy, Learning: A Multimodal Approach*. London: Routledge.
- [8] JEWITT, C. (2013). Multimodality and digital technologies in the classroom. In: SAINT-GEORGES, I., & WEBER, J.J. (eds.) *Multilingualism and Multimodality: The Future of Education Research*. Rotterdam: Sense Publishers, pp. 141-152.
- [9] JEWITT, C., BEZEMER, J., & O'HALLORAN, K.L. (2016). *Introducing Multimodality*. London: Routledge.
- [10] JING, G. (2016). On the construction of a system for teaching spoken Chinese as a foreign language. *Chinese Teaching in the World*, 30(3), 368-378.
- [11] KARATZA, S. (2020). Multimodal literacy and language testing: Visual and intersemiotic literacy indicators of reading comprehension texts. *Journal of Visual Literacy*, 39(3-4), 220-255. <https://doi.org/10.1080/1051144X.2020.1826222>
- [12] KOMARAWAN, Y. (2019). Constructing Interpersonal Meaning in Indonesian Science Classrooms through Language, Space and Gaze. *KEMANUSIAAN: The Asian Journal of Humanities*, 26(1), 1-23. <https://doi.org/10.21315/kajh2019.26.1.1>
- [13] KRESS, G. (2003). *Literacy in the New Media Age*. London & New York: Routledge.
- [14] KRESS, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Routledge.
- [15] KRESS, G., & VAN LEEUWEN, T. (2006). *Reading Images: The Grammar of Visual Design*. 2nd ed. Routledge.
- [16] LEMKE, J.L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In: MARTIN, J.R., & VEE, R. (eds.) *Reading Science: Critical and Functional Perspectives on Discourse of Science*. London/New York: Routledge, pp. 87-133.
- [17] LIM, F.V. (2021). Investigating intersemiosis: A systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse. *Visual Communication*, 20(1), 34-58. <https://doi.org/10.1177/1470357218820695>
- [18] LIM, V.F. (2019). Analysing the teachers' use of gestures in the classroom: A systemic functional multimodal discourse analysis approach. *Social Semiotics*, 29(1), 83-111. <https://doi.org/10.1080/10350330.2017.1412168>
- [19] MARTIN, J.R. (1992). *English Text: System and Structure*. Amsterdam: John Benjamins Publishing Company.
- [20] MARTINEC, R., & SALWAY, A. (2005). A system for image text relations in new (and old) media. *Visual Communication*, 4(3), 337-371. <https://doi.org/10.1177/1470357205055928>
- [21] MATTHIESSEN, C.M.I.M. (2009). Multisemiosis and Context-Based Register Typology: Registerial Variation in the Complementarity of Semiotic Systems. In: VENTOLA, E., & GUIJARRO, A.J.M. (eds.) *The World Told and the World Shown*. London: Palgrave Macmillan, pp. 11-38. https://doi.org/10.1057/9780230245341_2
- [22] NORRIS, S. (2004). *Analyzing Multimodal Interaction: A Methodological Framework*. London: Routledge.
- [23] NORRIS, S. (2011). Three hierarchical positions of deictic gesture in relation to spoken language: A multimodal interaction analysis. *Visual Communication*, 10(2), 129-147. <https://doi.org/10.1177/1470357211398439>
- [24] NYSTRAND, M. (2006). Research on the Role of Classroom Discourse As It Affects Reading Comprehension. *Research in the Teaching of English*, 40(4), 392-412. Retrieved from https://www.troyespier.com/assets/files/bibliographies/teaching/nystrand_classroom_discourse.pdf
- [25] O'HALLORAN, K.L. (2003). Intersemiosis in mathematics and science: Grammatical metaphor and semiotic metaphor. In: SIMON-VANDENBERGEN, A.-M., TAARNIERS, M., & RAVELLI, L. (eds.) *Grammatical Metaphor: Views from Systemic Functional Linguistics*. Amsterdam/Philadelphia, Pennsylvania: John Benjamins, pp. 337-365. <https://doi.org/10.1075/cilt.236.18oha>
- [26] O'HALLORAN, K.L. (2007). Systemic functional multimodal discourse analysis (SF-MDA) approach to mathematics, grammar, and literacy. In: MCCABE, A., O'DONNELL, M., & WHITTAKER R. (eds.) *Advances in Language and Education*.

- London: Continuum, pp. 75–100.
- [27] PINNOW, R.J. (2011). “I’ve Got an Idea”: A Social Semiotic Perspective on Agency in the Second Language Classroom. *Linguistics and Education*, 22(4), 383–392. <https://doi.org/10.1016/j.linged.2011.09.002>
- [28] ROYCE, T. (1998). Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text. *Japan Association Systemic Functional Linguistics Occasional Papers*, 1, 25–49.
- [29] SCOLLON, R. (1998). *Mediated Discourse as Social Interaction*. London: Longman.
- [30] SCOLLON, R. (2001). *Mediated Discourse: The Nexus of Practice*. London: Routledge.
- [31] TAKAHASHI, J., & YU, D. (2016). Multimodality in the classroom: An introduction. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 16(2), i-vi. <https://doi.org/10.7916/salt.v16i2.1247>
- [32] TAYLOR, R. (2014). Meaning between, in, and around words, gestures and postures: Multimodal meaning making in children's classroom communication. *Language and Education*, 28(5), 401–420. <https://doi.org/10.1080/09500782.2014.885038>
- [33] TSUNG, L., & CRUICKSHANK, K. (eds.) (2011). *Teaching and Learning Chinese in Global Contexts: CFL Worldwide*. London: Continuum.
- [34] UNSWORTH, L., & CLEIRIGH, C. (2009). Multimodality and reading: The construction of meaning through image-text interaction. In: JEWITT, C. (ed.) *The Routledge Handbook of Multimodal Analysis*. London & New York: Routledge, pp. 151–163. Retrieved from https://www.researchgate.net/publication/259559003_Multimodality_and_Reading_The_Construction_of_Meaning_Through_Image-Text_Interaction
- [35] WALSH, S. (2006). *Investigating Classroom Discourse*. London: Routledge.
- [36] YANG, J. (2019). Understanding Chinese language teachers' beliefs about themselves and their students in an English context. *System*, 80, 73–82. <https://doi.org/10.1016/j.system.2018.10.014>
- [37] ZHANG, D., & HU, R. (2019). System, Affordances and Choice in Constructing Multimodal Discourse. *Contemporary Rhetoric*, 5, 68–79. <https://doi.org/10.16027/j.cnki.cn31-2043/h.2019.05.007>
- [38] ZHANZI, L., & DANYUN, L. (2012). Multimodal Semiotics: Theoretical Basis, Research Approaches and Development Prospects. *Foreign Language Studies*, 2, 1–8. <https://doi.org/10.4324/9780203828847-10>
- [2] 贝特曼, J., 维尔德费尔, J., & 希帕拉, T. (2017)。多模态: 基础、研究和分析-面向问题的介绍。柏林: 德格鲁特木桐。 <http://dx.doi.org/10.1515/9783110479898>
- [3] 格莱芬哈根, C. (2008)。数学练习视频分析?“开放”数学用于社会学调查的不同尝试。论坛: 定性社会研究, 9(3), 32。 <https://doi.org/10.17169/fqs-9.3.1172>
- [4] 哈利迪, 马克 (1978)。作为社会符号学的语言: 语言和意义的社会解释。伦敦: 爱德华·阿诺德。
- [5] 哈利迪, 马克 (1994)。功能语法简介。伦敦: 爱德华·阿诺德。
- [6] M.A.K. 哈利迪和 C.M.I.M. 马蒂森 (1999)。通过意义解释经验: 基于语言的认知方法。伦敦: 卡塞尔。
- [7] JEWITT, C. (2008)。技术、素养、学习: 多模式方法。伦敦: 劳特利奇。
- [8] JEWITT, C. (2013)。课堂上的多模态和数字技术。载于: SAINT-GEORGES, I. 和 WEBER, J.J. (编) 多语言和多模式: 教育研究的未来。鹿特丹: 感觉出版社, 第 141–152 页。
- [9] JEWITT, C., BEZEMER, J. 和 O'HALLORAN, K.L. (2016)。引入多模态。伦敦: 劳特利奇。
- [10] 景国 (2016)。论对外汉语口语教学体系的构建世界汉语教学, 30(3), 368–378。
- [11] 卡拉扎, S. (2020)。多模态读写能力和语言测试: 阅读理解文本的视觉和符号间读写能力指标。视觉素养杂志, 39 (3–4), 220–255。 <https://doi.org/10.1080/1051144X.2020.1826222>
- [12] 科马拉万 (2019)。通过语言、空间和目光构建印度尼西亚科学课堂的人际意义。克马努西亚安: 亚洲人文杂志, 26 (1), 1–23。 <https://doi.org/10.21315/kajh2019.26.1.1>
- [13] 克雷斯, G. (2003)。新媒体时代的素养。伦敦和纽约: 劳特利奇。
- [14] 克雷斯, G. (2010)。多模态: 当代传播的社会符号学方法。伦敦: 劳特利奇。
- [15] KRESS, G., & VAN LEEUWEN, T. (2006)。阅读图像: 视觉设计的语法。第二版。劳特利奇。
- [16] 莱姆克, J.L. (1998)。多重意义: 科学文本中的视觉和语言符号学。见: MARTIN, J.R. 和 VEE, R. (编辑) 《阅读科学: 科学话语的批判性和功能性观点》。伦敦/纽约: 劳特利奇, 第 87–133 页。
- [17] LIM, F.V. (2021)。研究联间症: 对课堂话语中语言和手势之间关系的系统功能多模态话语分析。视觉传达, 20 (1), 34–58。 <https://doi.org/10.1177/1470357218820695>
- [18] LIM, V.F. (2019)。分析教师在课堂上手势的使用: 一种系统功能多模态话语分析方法。社会符号学, 29(1), 83–111。 <https://doi.org/10.1080/10350330.2017.1412168>

参考文献:

- [1] 贝特曼, J.A. (2011)。符号模式的可分解性。见: O'HALLORAN, L. 和 SMITH, B. (编辑) 多模式研究: 探索问题和领域。伦敦: 劳特利奇, 第 17–38 页。

- [19] 马丁, J.R. (1992)。英文文本: 系统和结构。阿姆斯特丹: 约翰·本杰明出版公司。
- [20] MARTINEC, R. 和 SALWAY, A. (2005)。新(和旧)媒体中图像文本关系的系统。视觉传达, 4(3), 337-371。
<https://doi.org/10.1177/1470357205055928>
- [21] 马蒂森, C.M.I.M. (2009)。多符号学和基于上下文的语域类型学: 符号系统互补性中的语域变化。见: VENTOLA, E., & GUIJARRO, A.J.M. (编) 讲述的世界和展示的世界。伦敦: 帕尔格雷夫·麦克米伦, 第 11-38 页。
https://doi.org/10.1057/9780230245341_2
- [22] 诺里斯, S. (2004)。分析多模式交互: 方法框架。伦敦: 劳特利奇。
- [23] 诺里斯, S. (2011)。指示手势与口语相关的三个层次位置: 多模态交互分析。视觉传达, 10(2), 129-147。
<https://doi.org/10.1177/1470357211398439>
- [24] NYSTRAND, M. (2006)。课堂话语影响阅读理解的作用研究。英语教学研究, 40(4), 392-412。摘自
https://www.troyspier.com/assets/files/bibliographies/teaching/nystrand_classroom_discourse.pdf
- [25] 奥哈洛兰, K.L. (2003)。数学和科学中的符号间关系: 语法隐喻和符号隐喻。见: SIMON-VANDENBERGEN, A.-M., TAARNIERS, M. 和 RAVELLI, L. (编) 语法隐喻: 来自系统功能语言学的观点。阿姆斯特丹/宾夕法尼亚州费城: 约翰·本杰明, 第 337-365 页。
<https://doi.org/10.1075/cilt.236.18oha>
- [26] 奥哈洛兰, K.L. (2007)。针对数学、语法和读写能力的系统功能多模态话语分析(顺丰-丙二醛)方法。见: MCCABE, A., O'DONNELL, M. 和 WHITTAKER R. (编辑) 语言和教育的进展。伦敦: 连续体, 第 75-100 页。
- [27] 平诺, R.J. (2011)。“我有一个想法”: 第二语言课堂中能动性的社会符号学视角。语言学与教育, 22(4), 383-392。
<https://doi.org/10.1016/j.linged.2011.09.002>
- [28] 罗伊斯, T. (1998)。页面上的协同作用: 探索基于页面的多模式文本中的符号间互补性。日本系统功能语言学协会临时论文, 1, 25-49。
- [29] 斯科伦, R. (1998)。作为社会互动的中介话语。伦敦: 朗文。
- [30] 斯科伦, R. (2001)。中介话语: 实践的纽带。伦敦: 劳特利奇。
- [31] 高桥 J. 和 YU D. (2016)。课堂上的多模态: 简介。哥伦比亚大学师范学院对外英语教学和应用语言学工作论文, 16(2), i-vi。
<https://doi.org/10.7916/salt.v16i2.1247>
- [32] 泰勒, R. (2014)。言语、手势和姿势之间、之中和周围的意义: 儿童课堂交流中的多模式意义建构。语言与教育, 28(5), 401-420。
<https://doi.org/10.1080/09500782.2014.885038>
- [33] TSUNG, L., & CRUICKSHANK, K. (编辑) (2011)。全球背景下的汉语教学: 全球节能灯。伦敦: 连续体。
- [34] UNSWORTH, L. 和 CLEIRIGH, C. (2009)。多模态与阅读: 通过图像-文本交互构建意义。见: JEWITT, C. (编) 《劳特利奇多模态分析手册》。伦敦和纽约: 劳特利奇, 第 151-163 页。摘自
https://www.researchgate.net/publication/259559003_Multimodality_and_Reading_The_Construction_of_Meaning_Through_Image-Text_Interaction
- [35] 沃尔什, S. (2006)。调查课堂话语。伦敦: 劳特利奇。
- [36] 杨洁 (2019)。在英语背景下了解中文教师对自己和学生的信念。系统, 80, 73-82。
<https://doi.org/10.1016/j.system.2018.10.014>
- [37] 张丹, 胡瑞 (2019)。构建多模态话语的系统、可供性和选择。当代修辞学, 5, 68-79。
<https://doi.org/10.16027/j.cnki.cn31-2043/h.2019.05.007>
- [38] 展子, L., 丹云, L. (2012)。多模态符号学: 理论基础、研究方法与发展前景。外语研究, 2, 1-8。