

A Review of Empirical Studies on Multimodal Second Language Classroom Teaching in China — Based on the Papers Published in 9 Core Chinese Linguistic Journals

Yanling Pi *

School of Foreign Languages and Literature, Chongqing Normal University, Chongqing 400000, China

* Corresponding Author Email: yanling_pi@163.com

Abstract. This study is based on a statistical analysis of 28 empirical research papers on multimodal second language classroom teaching published in 9 core Chinese linguistic journals from 2009 to 2024. Firstly, this study introduces the development of multimodal language teaching and multimodal discourse analysis, which serve as the theoretical basis for multimodal second language classroom teaching. Secondly, it investigates the research status of the following three main themes: multimodal classroom discourse analysis, learners' multimodal cognitive research, and multiliteracies. Thirdly, it analyzes the major problems in current research from the following three dimensions: research subjects, research instruments, and theoretical basis. Finally, it puts forward some suggestions for the development of empirical research on multimodal second language classroom teaching in China from three aspects: (1) Research on multimodal classroom discourse analysis should diversify research subjects and instruments, shift from static analysis to dynamic analysis, and quantify multimodal interrelationships; (2) Research on learners' multimodal cognition should focus on learners' internal cognitive processes, exploring the intrinsic mechanisms of how learners input and output different modal information; (3) Multimodal second language classroom teaching research needs to be broadened by covering more lesson types and deepened by promoting interdisciplinary methodological innovations and integrating cutting-edge science and technology.

Keywords: multimodal, second language teaching, classroom, empirical studies, review.

1. Introduction

The rapid development of modern information technology has made human communication no longer restricted to language but integrated with multimodal methods of meaning-making, such as images, sounds, colors, and flashes. The use of multimodal means in foreign language teaching has also become a big trend. Multimodal research has blossomed since the late twentieth century, and different scholars have defined multimodality in different ways. Kress & van Leeuwen (1996) point out that multimodality is the simultaneous use of several semiotic modes, as well as the patterns or methods of information exchange between people or between people and machines through various media; Li (2003) states that multimodality refers to composite discourse that includes not only text but also images, charts, and other elements, or any text in which more than one symbol encodes meaning; Gu (2007) defines modality as the way humans interact with the external environment (e.g., people, machines, objects, animals) through their senses (e.g., vision, hearing). Interaction with three or more senses is called multimodality. Thus, modality is not only a social symbol but also a medium of interaction, constructing meaning and embodying cognition. At the same time, classroom-based second language acquisition research has received increasing attention. Many researchers have regarded the classroom as an ideal setting for investigating second language learners and their acquisition process, as classroom second language acquisition research builds a platform that combines theoretical research and teaching practice. (Yuan, 2016).

With the rapid advancement of multimodality research and classroom-based second language acquisition studies, research on multimodal second language classroom teaching has also become a hot topic in recent years, yet relevant reviews remain absent. In order to explore the developing trend

in this field, this paper provides a statistical analysis of the empirical research papers on multimodal second language classroom teaching published in 9 core Chinese linguistic journals from 2009 to 2024.

2. Theoretical Basis of Multimodal Second Language Classroom Teaching Research

The multimodal empirical research on second language classroom teaching discussed in this paper can be broadly categorized into two directions. The first focuses on using multimodality as a teaching method, while the second employs multimodality as a research method. Second language classroom research that uses multimodality as a teaching method is directly related to language teaching. Multimodality, as part of the classroom environment, promotes language acquisition among second language learners. Second language classroom research that employs multimodality as a research method is closely tied to multimodal discourse analysis. Researchers use the theory of multimodal discourse analysis to analyze and interpret classroom discourse. While distinct, these two methods are complementary and ultimately contribute to second language classroom teaching. This paper briefly introduces the development of multimodal language teaching and multimodal discourse analysis, which serve as the theoretical basis for multimodal second language classroom teaching.

The theoretical basis of multimodal language teaching originates from multiliteracies, a concept first proposed by the New London Group in 1994. In 1996, they published *A Pedagogy of Multiliteracies: Designing Social Futures* in Harvard Educational Review, which was a milestone in the study of multiliteracies. In 2000, Kress also pointed out that the centrality of language in the New Media Age was being challenged, and paralinguistic and non-linguistic modalities were gradually gaining significance. The first researchers to study multimodal language teaching in China were Hu and Gu. Hu (2007) distinguished between multimodal semiotics and multimedia semiotics, and introduced computer semiotics with dual characteristics of media and mode. Gu (2007) differentiated multimedia learning from multimodal learning, and constructed a model to analyze the two forms of learning. Zhang proposed a series of frameworks and principles, establishing himself as a leading figure in the field of multimodal language teaching in China. Multimodal language teaching research in China has made many breakthroughs and achievements, and has gradually narrowed the gap with the West.

The surge in multimodal discourse analysis research began in the 1990s. Kress and van Leeuwen, representatives of the social semiotic school, published *Reading Images: The Grammar of Visual Design* in 1996, which is a landmark achievement in multimodal discourse analysis. In this work, they proposed representational, interactional, and compositional meanings in images, based on the three metafunctions of Halliday's systemic functional linguistics. Cognitive linguists (Forceville, 1996; Forceville & Urios Aparisi, 2009) argue that relevance theory is not limited to verbal communication but can also be applied to non-verbal and multimodal discourse analysis. They employ pictorial metaphor to study multimodal discourse. Interactional sociolinguists (Norris, 2004; Norris & Jones, 2005; Scollon, 2001) regard language as a social behavior. They argue that discourse analysis should not only analyze language but also consider the role of interactive processes. In China, Li (2003) systematically investigates the social semiotic analytical framework for multimodal discourse. Later, Hu (2007) and Zhu (2007) began to introduce and review multimodal discourse analysis, investigating its definition and research methods. In recent years, Zhang (2009a, 2018) attempted to build a synthetic theoretical framework for multimodal discourse analysis based on systemic functional linguistics, marking the establishment of a relatively complete theoretical framework for multimodal discourse analysis. Although multimodal discourse analysis started late in China, its theoretical framework has become increasingly independent and comprehensive.

3. Analysis of Empirical Research on Multimodal Second Language Classroom Teaching in China

3.1. Research Methods

This paper adopted the bibliometric method, combining qualitative and quantitative research. This study used the journal navigation feature in CNKI to select 14 linguistic journals indexed in the Peking University Core Journals as the literature sources. Then, "modality" was used as the topic term to search papers in each journal, and papers relevant to multimodal second language classroom teaching were selected from them. Ultimately, 52 valid papers were obtained from 12 core journals, including 24 non-empirical papers and 28 empirical papers. This study selected 28 empirical papers from 9 core Chinese linguistic journals as its research subjects. While the field of multimodal second language classroom teaching is rich in theoretical innovation, there is a scarcity of studies that empirically validate these theories, leaving many theoretical frameworks at the hypothetical stage. Empirical research, characterized by its objectivity and replicability, directly reflects the practical application effect of theories and offers a foundation for their further refinement and optimization. Therefore, conducting a review of empirical studies on multimodal second language classroom teaching is both necessary and valuable.

3.2. Analysis of Research Themes in Empirical Studies on Multimodal Second Language Classroom Teaching in China

When categorizing the themes of the 28 empirical papers, it is observed that some papers covered multiple themes. As a result, the total number of the themes of the papers presented in the table below (34) exceeds the actual number of papers included in the study (28). As shown in the table, research focusing on learners' multimodal cognition constitutes the largest proportion, 11 papers in total, followed by research on multimodal classroom discourse analysis, 9 papers in total. Since most papers under the themes of listening teaching and writing teaching are closely related to learners' multimodal cognition, they are integrated into the analysis of learners' multimodal cognition rather than being analyzed separately.

Table 1. Statistics on Research Themes in Empirical Research on Multimodal Second Language Classroom Teaching in China

Theme	Learners' Multimodal Cognitive Research	Multimodal Classroom Discourse Analysis	Listening Teaching	Multiliteracies	Teaching Tools	Writing Teaching	Others	Total
Number	11	9	5	3	3	2	1	34
Percentage	32.4%	26.5%	14.7%	8.8%	8.8%	5.9%	2.9%	100%

3.2.1. Multimodal Classroom Discourse Analysis

Multimodal discourse refers to the communication that utilizes multiple senses, such as hearing, vision, and touch, and various means and symbolic resources, including language, images, sounds, and actions (Zhang, 2009b). Foreign language classroom teaching has a strong multimodal nature. The study of multimodal discourse in the foreign language classroom can help to analyze how different modes in the classroom work together to construct meanings and achieve communicative objectives. It is of great significance for the effective organization of classroom teaching and the enhancement of students' second language acquisition outcomes. All of the empirical studies on multimodal classroom discourse analysis collected in this study have adopted video recording analysis as the sole research method, and most of them have focused on only one or two teaching recordings. Future research needs to adopt diversified research methods and expand the sample size to enhance the credibility of the findings.

Empirical studies on multimodal classroom discourse analysis focus on modal choice and modal synergy. For example, Zhang and Li (2012) investigated the synergy between the oral mode and the modality set presented by PPT, as well as their realization in lexis and grammar. They found that the same modality plays different roles at different stages of teaching and that the choice of modality depends on the specific goals of each teaching stage. Wu and Zhang (2024) examined and summarized the features of selection and arrangement of modes in live video streaming English classes for children. Some scholars have found that excellent teachers excel at selecting appropriate mode types and synergizing multimodal discourse to enhance their teaching outcomes. This conclusion was drawn from analyzing the multimodal classroom discourse of award-winning teachers in teaching competitions. Three relevant studies are discussed below: Zhu and Chen (2015) studied the multimodal classroom discourse of two award-winning teachers in a teaching competition from the perspective of adaptation theory. Their research revealed that excellent teachers excelled at selecting and designing multimodal discourse to adapt to culture, society, and cognition. Furthermore, these teachers synergized various modal discourses to co-construct meaning and ultimately achieve optimal teaching effects. Ning (2019), based on Zhang's multimodal discourse analysis framework, analyzed the multimodal classroom discourse features of two award-winning teachers in a teaching competition. The study found that both teachers used similar types of modalities and were adept at selecting appropriate paralinguistic and non-linguistic modalities to construct meanings with linguistic modalities. Lei and Zhang (2022), based on multimodal stylistic theory, proposed an analytical approach for the stylistic analysis of multimodal classroom discourse. Through analyzing the foregrounded mode features and their effects in the Grand Prize teaching videos, they found that language served as the primary modality, while audio and image modalities repeated or reinforced linguistic content. Additionally, body language modalities demonstrated high compatibility with other modalities.

Studying the synergy and coordination between modalities is of great significance for foreign language teaching, but the importance of investigating the interaction between different semiotic resources should not be overlooked. Wu (2017), on the basis of visual grammar and systemic-functional image-text relations theory, constructed a multimodal theoretical framework for image-text relations in micro-lessons. She conducted a multimodal discourse analysis of two award-winning micro-lessons, delving into the relationship between images and textual symbols in the visual modality. Guo and Yang (2020), adopting a multimodal social semiotic framework and appraisal theory, analyzed the inters miotic mechanisms between metadiscourse and gestures of an English university teacher. They found that teachers' gestures can embody interactional and interactive meanings of metadiscourse. Moreover, deictic, beating, and metaphorical gestures can combine with teachers' metadiscourse to achieve inters miotic cohesion, constructing inters miotic textual structures.

Some scholars have also investigated scaffolding in the classroom through multimodal analysis. For example, Sun et al. (2021), from the perspective of sociocultural theory, employed multimodal analysis methods to explore teacher scaffolding in Content and Language Integrated English teaching. They investigated the multimodal features of teacher scaffolding from two dimensions: scaffolding types and focus distribution. Ying and Xie (2024), based on the framework of systemic-functional multimodal discourse analysis, explored how teachers' gestures, postures, and facial expressions cooperate with language to generate meaning and serve as scaffolding.

3.2.2. Learners' Multimodal Cognitive Research

Multimodality is essentially a process in which the brain inputs and processes a variety of semiotic resources through various senses. Therefore, multimodal research is inseparable from cognitive research. At present, there are two distinct views on the cognitive impact of multimodal input in the field of second language acquisition. A group of scholars are optimistic about multimodal input, arguing that the integration of multiple semiotic systems can improve concentration, reduce cognitive load, and enhance learning outcomes. However, another group of researchers hold a contrasting perspective, arguing that there may be interference effects among multimodal subsystems, which increase cognitive load and decrease learning outcomes.

Research on learners' cognition in multimodal second language classroom teaching covers a broad scope. Among the papers collected in this research, most empirical research related to listening teaching and writing teaching is also closely tied to cognitive research. Therefore, in this paper, studies focusing on learners' multimodal cognition are categorized into three directions: listening teaching, writing teaching, and other multimodal cognition research in teaching practice.

Listening, an important means of language input, is closely tied to learners' cognitive comprehension and serves as a crucial precondition for communication. The following three research papers explored strategies to enhance listening comprehension in multimodal second language classroom teaching. Long and Zhao (2009) introduced metacognitive strategies into multimodal college English listening teaching. They demonstrated that there is a strong interaction between multimodality and metacognitive strategies, and their combination promotes learners' listening comprehension more effectively than either approach used independently. Hu and Zhang (2013) demonstrated that the rational use of multimodal combinations can significantly improve students' listening comprehension. Wang and Yang (2014) found that a teaching mode emphasizing the multimodal learning of words as chunks could significantly improve listening comprehension and also help overcome vocabulary learning barriers. Other scholars have also explored the synergy of multimodality in listening teaching. For instance, Hui et al. (2013) clarified the interrelationships between auditory and visual modalities at different stages of listening teaching and in different tasks.

Writing is an important way for learners to output information in second language teaching in China. Its prerequisite is the deep processing and integration of input information, which makes it inherently tied to cognitive research. The following two studies explored the impact of multimodal input on second language writing teaching. Qin and Lou (2012) investigated the relationships between input modalities and the development patterns of college English learners' writing system from the perspective of information processing theory. They confirmed that modality-converted input abilities could facilitate both the automatic processing and the restructuring process, thus promoting the development of English learners' target language writing system. Cai (2022) investigated the effects of modality input and task structure on the accuracy of summary writing among second language learners with different language proficiency. The study found that learners achieved higher summary accuracy with visual input. However, the facilitation effect of task structure, as well as the interaction effect between modality input and task structure, were influenced by the learners' second language proficiency.

In multimodal second language classroom teaching, cognitive theory is widely used in listening and writing teaching, but its application in other areas of teaching practice should not be neglected. Many scholars have designed new multimodal teaching models by integrating cognitive theory into multimodal teaching. They implemented these models in practical teaching, confirming their effectiveness in improving learners' second language proficiency. For example, Liu et al. (2013) noticed the influence of multimodal classroom discourse on students' learning enthusiasm, constructed a theoretical framework for using multimodal discourse to stimulate students' motivation, and validated its effectiveness through teaching experiments. Zhang and Yan (2014) analyzed the correlation between multimodal teaching and the law of internalization, designed a multimodal teaching model, and applied it in teaching experiments. They found that the multimodal teaching model could stimulate the learners' motivation, reduce the cognitive load, and promote the internalization of knowledge. Rui and Ji (2017) integrated constructivist learning theory, output-driven hypothesis, and situated cognition theory to design a multimodal teaching model for listening and speaking teaching. They applied this model in college English listening and speaking classes. The study revealed that speaking anxiety could explain classroom silence, and the multimodal listening and speaking teaching design could alleviate students' speaking anxiety and improve classroom silence. Li and Li (2024) constructed a multimodal cognitive contrastive analysis teaching model for English passive sentence acquisition. Through empirical teaching research, they found that the new teaching model could facilitate learners' acquisition of English passive sentences, with more significant effects observed among advanced learners. Other scholars have also examined individual-

level differentiated cognitive factors. For example, Qin and Lou (2018), from the micro-level perspective of individual cognitive differences, explored the impact of different input modalities and multimodal learning abilities on learners' meaning-form acquisition. The study found that multimodal input facilitates learners' acquisition of meaning, but its effectiveness in form acquisition is influenced by the combination of input modalities and modality-converted input abilities. The overall outcomes of meaning-form acquisition depend on the coordination between input modalities and between input modalities and multimodal learning abilities.

3.2.3. Multiliteracies

In China, there is a wide range of papers on multiliteracies in second language classroom teaching, but most of them are non-empirical studies which proposed theoretical frameworks or teaching suggestions. Empirical studies are rare, with the majority focusing on the effectiveness of innovative English teaching models in enhancing students' multiliteracies. Wang (2012) found that a multimodal integrated teaching model, compared with traditional teaching methods, could better improve students' multicultural literacy and boost their learning confidence and interest, ultimately enhancing their comprehensive English level. Song (2013) compared multiliteracies pedagogy with traditional teaching methods in college English teaching, and found that the former could address many problems of the latter in actual teaching and had better teaching outcomes. Lei (2018) constructed an analytical framework based on learning-by-design theories and validated its effectiveness in cultivating students' multiliteracies, particularly critical literacy, media and modality application skills, and media and modality critical abilities.

4. Problems and Prospects in Empirical Research on Multimodal Second Language Classroom Teaching in China

4.1. Problems in Empirical Research on Multimodal Second Language Classroom Teaching in China

4.1.1. Research Subjects

Among the 28 empirical papers collected in this study, 27 focused on university teachers or students. This is because most researchers are university teachers, while primary, secondary, and early childhood teachers are mainly engaged in first-line teaching. Although university foreign language teaching is indeed an important field, limiting the research perspective to universities inevitably overlooks the diversity of language learners and the complexity of cognitive mechanisms within learning groups. At the same time, multimodal foreign language teaching research focusing on college teachers mostly selected only one or two teachers as research subjects, neglecting the impact of individual differences among teachers. Similarly, student samples were mostly limited to the same school and major, resulting in small sample sizes and a lack of cross-dimensional comparisons. The potential effects of differences in institution type, institution level, and academic discipline were overlooked. In summary, the research subjects in empirical studies on multimodal second language classroom teaching are generally limited, raising questions about the generalizability of the findings. If sample selection is more comprehensive, it could significantly enhance the credibility and reliability of the research.

4.1.2. Research Instruments

Table 2. Statistics on the Number and Percentage of Research Instruments Used in the Literature

	One Instrument	Two Instruments	Three Instruments	Four Instruments	Total
Number of Studies	10	7	9	2	28
Percentage	35.7%	25.0%	32.1%	7.1%	100%

Table 2 indicates that among the 28 empirical studies on multimodal second language classroom teaching, 10 studies (35.7%) utilized one research instrument; 7 studies (25.0%) employed two

instruments; 9 studies (32.1%) used three instruments; 2 studies (7.1%) applied four or more instruments. Notably, 9 out of 10 studies using one research instrument focused on multimodal classroom discourse analysis, representing 100% of the studies in this theme. This finding further demonstrates the limitations in analytical dimensions within multimodal classroom discourse analysis. It would be more advisable to apply emerging tools such as eye-tracking and multimodal sensor networks. Additionally, 11 studies utilized SPSS for statistical analysis; 4 studies employed ELAN for multimodal annotation of teaching recordings; 2 studies applied Matlab for modeling.

Table 3. Statistics on the Frequency and Percentage of Research Instruments Used in the Literature

	Teaching Experiments	Tests	Questionnaires	Audio/Video Recordings	Interviews	Classroom Observations	Others	Total
Frequency	14	14	10	10	7	2	4	61
Percentage	23.0%	23.0%	16.4%	16.4%	11.5%	3.3%	6.6%	100%

As shown in Table 3, the frequency of research instrument usage, ranked from highest to lowest, is as follows: teaching experiments (23.0%), tests (23.0%), questionnaires (16.4%), audio and video recordings (16.4%), interviews (11.5%), other methods (6.6%), and classroom observations (3.3%). The high-frequency instruments (experiments, tests, questionnaires, and audio/video recordings) are quantitative instruments, while qualitative instruments (interviews and classroom observations) are utilized less frequently. This indicates an imbalance between qualitative and quantitative instruments. There is a need to incorporate more qualitative research tools such as interviews and think-aloud protocols.

4.1.3. Theoretical Basis

The theoretical basis of empirical research on multimodal second language classroom teaching is still dominated by Halliday's systemic functional linguistics, but many studies have demonstrated a certain degree of theoretical integration and innovation. For instance, research on multimodal classroom discourse analysis has incorporated perspectives from sociocultural theory, adaptation theory, multimodal stylistics, visual grammar, and the theory of image-text relations. Similarly, studies on multimodal listening teaching have integrated the lexical chunk theory, sociocultural theory, cognitive psychology, and metacognitive strategies. Multimodal research is closely connected with cognitive studies of the brain, which enables a deeper understanding of multimodality in second language classroom teaching from a cognitive perspective. Currently, some empirical innovations which integrate cognitive theory and multimodal teaching concepts have been made in the areas of listening and writing teaching. However, in other fields, such as multimodal classroom discourse analysis, there remains considerable potential for further integration and innovation. The theoretical basis of multimodal second language classroom teaching primarily comes from the foreign theory. Although scholars in China have achieved relatively comprehensive reviews and analyses in recent years, researchers should continue striving to localize these theories. This involves integrating these theories with China's educational context and information technology development realities to construct a localized multimodal foreign language teaching theoretical framework. Such a framework can ultimately be applied in practical teaching contexts, contributing to the reform of foreign language education in China. In summary, while the theoretical basis of empirical research on multimodal second language classroom teaching demonstrates a trend toward diversification, there still remains significant potential for further innovation in theoretical integration and localization.

4.2. Prospects in Empirical Research on Multimodal Second Language Classroom Teaching in China

4.2.1. Prospects in Multimodal Classroom Discourse Analysis

Firstly, as mentioned before, research on multimodal classroom discourse analysis commonly faces problems of limited research instruments and constrained sample sizes. Future research should consider combining diversified research tools and expanding sample sizes to conduct

multidimensional analyses of classroom discourse. Secondly, existing empirical studies on multimodal features of classroom discourse primarily focus on quantifying the length and frequency of individual modality usage. However, this approach appears slightly fragmented, as the classroom constitutes an organic whole. Moreover, the majority of those studies remain confined to analyzing the hierarchical relationships among modalities, lacking innovative and comprehensive feature summarization. Since the classroom constitutes a dynamic multimodal discourse, simplistic static statistical approaches prove inadequate for exploring the deeper semantics of multimodality. As a result, subsequent research should progressively transition from static to dynamic analysis to better capture the complex interplay between multimodal elements in classroom discourse. Lastly, modal synergy and modal selection are the focus of multimodal classroom discourse analysis research. However, existing studies predominantly limit their exploration of multimodal interrelationships to modal co-occurrence and descriptive analyses. Furthermore, the interpretation of multimodal interrelationships often relies on subjective assessments, lacking objective data to systematically analyze the processes of modal synergy. The quantification of multimodal interrelationships may become a significant breakthrough in the research on multimodal classroom discourse analysis. Future studies could focus on constructing multimodal corpora and multimodal data models, thus advancing the shift in research paradigms from subjective description to data-driven approaches.

4.2.2. Prospects in Learners' Multimodal Cognitive Research

The input and output of multimodal information in second language classrooms are closely related to brain cognition. However, current empirical research on multimodal second language classroom teaching still lacks depth in cognitive studies. Most studies focus on the effects of external environmental factors, such as teaching models, multimodal input combinations, and metacognitive strategies. However, few studies have explored learners' internal cognitive mechanisms. The key to studying learners' internal cognition lies in investigating how learners coordinate their multimodal abilities and how they identify, decode, internalize, and reconstruct different modal information before producing output. If future studies incorporate advanced neuroscience technologies, such as eye-tracking, electroencephalography (EEG), and physiological sensors, to quantitatively analyze learners' internal cognitive mechanisms, learners' multimodal cognitive research may achieve further development.

4.2.3. Prospects in Research on Multimodal Second Language Classroom Teaching

Multimodal communication exists in various areas of foreign language teaching. However, empirical research on multimodal second language classroom teaching has mainly focused on listening teaching. It is essential to expand research to other teaching areas, such as speaking, multiliteracies, and translation. Moreover, there are still numerous directions worthy of further exploration in multimodal second language classroom teaching research. Research on multimodal listening teaching could incorporate cognitive and metacognitive theories to explore the synergistic mechanisms and interference effects among modalities, while utilizing cognitive neuroscience technologies, such as EEG and fMRI, to investigate the neural foundations of multimodal perceptual transformations. The cultivation of multiliteracies still requires adapting Western theoretical frameworks to the context of Chinese culture. Simultaneously, it must address the dual challenges posed by emerging media forms, such as short videos and metaverse social platforms, as well as the development of multiple intelligence. This involves exploring integrated cultivation mechanisms of literacy skills and digital intelligence competencies. Multimodal research on speaking and interpreting should strengthen its integration with corpus linguistics, utilizing large-scale multimodal data to construct intelligent analysis platforms. To sum up, research on multimodal second language classroom teaching should focus on covering more teaching areas, promoting interdisciplinary methodological innovations, and integrating cutting-edge cognitive science and intelligent technologies into research designs, while strengthening empirical investigations to avoid superficial discussions of teaching models.

5. Conclusion

In recent years, the rapid development of cutting-edge technologies, such as artificial intelligence, virtual reality, and big data, has brought new opportunities and challenges to multimodal second language classroom teaching. Research on multimodal second language classroom teaching needs to utilize cutting-edge science and technology to dynamically analyze multimodal classroom discourse, and quantitatively analyze multimodal interrelationship and learners' internal cognition. Moreover, it is essential to promote the independent innovation of multimodal foreign language teaching theory and practice and take account of China's national conditions. This involves constructing localized theoretical frameworks, teaching models, and evaluation systems for multimodal foreign language teaching. It is hoped that this study will contribute to the further development of multimodal second language classroom teaching research in China.

References

- [1] Cai, C. (2022). The effects of modality input and task structure on the accuracy of summary writing of L2 learners with different language proficiency. *Foreign Languages Research*, (06), 56 - 61+67.
- [2] Forceville, C. (1996). *Pictorial Metaphor in Advertising*. London & New York: Routledge.
- [3] Forceville, C., & Urios-Aparisi, E. (2009). *Multimodal Metaphor*. Berlin: Mouton de Gruyter.
- [4] Gu, Y. (2007). On multimedia learning and multimodal learning. *Technology Enhanced Foreign Language Education*, (02), 3 - 12.
- [5] Guo, H., & Yang, X. (2020). A multimodal study of intersemiosis between teachers' metadiscourse and gestures. *Foreign Language Education*, (04), 52 - 57.
- [6] Hu, Y., & Zhang, D. (2013). An experimental study on the function of multimodality in teaching listening to English majors. *Foreign Language World*, (05), 20 - 25+44.
- [7] Hu, Z. (2007). Multimodalization in social semiotics. *Language Teaching and Linguistic Studies*, (01), 1 - 10.
- [8] Hui, Z., Wang, L., Lin, Y., & Zhao, L. (2013). An empirical study of primary and secondary modality coordination—A case study of English listening teaching. *Shandong Foreign Language Teaching Journal*, (03), 77 - 81.
- [9] Kress, G., & van Leeuwen, T. (1996). *Reading Images: The Grammar of Visual Design*. Routledge.
- [10] Lei, Q. (2018). Developing EFL Majors' multiliteracies through learning by design—A case study of problem-solution argument writing. *Foreign Languages Bimonthly*, (03), 19 - 25+159.
- [11] Lei, Q., & Zhang, C. (2022). On the choice of modes in English classroom teaching: A multimodal pedagogical stylistics perspective. *Foreign Languages and Their Teaching*, (03), 73 - 83+121+148 - 149.
- [12] Li, Y., & Li, H. (2024). An empirical study of English passive sentence teaching based on the cognitive contrastive analysis model in multimodal contexts. *Technology Enhanced Foreign Language Education*, (02), 85 - 91+113.
- [13] Li, Z. (2003). Social semiotic approach to multimodal discourse. *Foreign Languages Research*, (05), 1 - 8+80.
- [14] Liu, X., Zhang, D., & Zhang, Y. (2013). Correlation between the teacher's multimodal discourse and the students' learning enthusiasm. *Technology Enhanced Foreign Language Education*, (03), 3 - 9.
- [15] Long, Y., & Zhao, P. (2009). The interaction study between multi-modality and meta-cognitive strategy in college English listening comprehension teaching. *Technology Enhanced Foreign Language Education*, (04), 58 - 62+74.
- [16] Ning, J. (2019). Features of excellent teachers' multimodal discourse in college English teaching contest: Take the 7th SFLEP National College English Teaching Contest as an example. *Shandong Foreign Language Teaching Journal*, (03), 62 - 77.
- [17] Norris, S. (2004). *Analyzing Multimodal Interaction: A Methodological Framework*. London: Routledge.
- [18] Norris, S., & Jones, R. H. (2005). *Discourse in Action: Introducing Mediated Discourse Analysis*. London: Routledge.

- [19] Qin, Q., & Lou, F. (2012). Study of input modality influences on the development of college English writing system under the multimedia teaching and learning environment. *Foreign Language World*, (06), 58 - 65.
- [20] Qin, Q., & Lou, F. (2018). A study on the mechanism of second language meaning-form acquisition under multimodal input conditions. *Foreign Languages Bimonthly*, (01), 117 - 125.
- [21] Rui, Y., & Ji, H. (2017). The impact of multimodal listening & speaking teaching on English speaking anxiety and classroom reticence. *Technology Enhanced Foreign Language Education*, (06), 50 - 55.
- [22] Scollon, R. (2001). *Mediated Discourse: The Nexus of Practice*. Routledge.
- [23] Song, Q. (2013). An empirical study on multimodality and the cultivation of multiliteracies in college English teaching. *Foreign Languages Research*, (02), 55 - 59.
- [24] Sun, X., Wang, P., & Hong, Y. (2021). A multimodal study of teacher scaffolding in a content and language integrated classroom. *Foreign Languages in China*, (01), 63 - 71.
- [25] Wang, J., & Yang, Y. (2014). The positive impact of multimodal intervention strategies on chunk learning and listening skills. *Foreign Languages in China*, (06), 43 - 52.
- [26] Wang, M. (2012). A multimodality-based study on cultivating students' multi-cultural literacy ability. *Foreign Language Education*, (01), 66 - 69+80.
- [27] Wu, L., & Zhang, D. (2024). A study on the design of modes in live video streaming foreign language classes. *Shandong Foreign Language Teaching Journal*, (03), 60 - 69.
- [28] Wu, T. (2017). A probe into two micro-lecture videos: A systemic-functional approach to intersemiosis analysis in multimodal discourse. *Foreign Language Learning Theory and Practice*, (01), 72 - 77.
- [29] Ying, J., & Xie, C. (2024). A study of meaning coordination between gestures, postures, facial expressions and scaffolding language during English language teaching. *Foreign Languages in China*, (02), 58 - 66.
- [30] Yuan, F. (2016). *Classroom-based Second Language Acquisition Research*. Beijing: The Commercial Press.
- [31] Zhang, D. (2009a). On a synthetic theoretical framework for multimodal discourse analysis. *Foreign Languages in China*, (01), 24 - 30.
- [32] Zhang, D. (2009b). The application of multimodal discourse theory and media technology in foreign language teaching. *Foreign Language Education*, (04), 15 - 20.
- [33] Zhang, D. (2018). A systemic-functional synthetic framework for multimodal discourse analysis. *Modern Foreign Languages*, (06), 731 - 743.
- [34] Zhang, D., & Li, Y. (2012). A study on modal coordination in multimodal classroom discourse. *Foreign Languages and Their Teaching*, (01), 39 - 43.
- [35] Zhang, S., & Yan, C. (2014). An empirical study on the correlation between multimodal teaching and the law of learning internalization. *Foreign Languages in China*, (02), 74 - 79.
- [36] Zhu, J., & Chen, X. (2015). Pragmatic analysis of excellent teacher's multimodal discourse in English classroom. *Shandong Foreign Language Teaching Journal*, (01), 44 - 49+80.
- [37] Zhu, Y. (2007). Theory and methodology of multimodal discourse analysis. *Foreign Language Research*, (05), 82 - 86.