Journal of Contemporary Educational Research, 2026, Volume 10, Issue 10
https://bbwpublisher.com/journal/JCER
Online ISSN: 2208-8474
Print ISSN: 2208-8466

Reforming University Chinese Composition Instruction from a Multimodal Perspective

Jingxuan Yang

School of Liberal Education, Liaoning University of International Business and Economics, Dalian, China

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: University Chinese composition instruction in China has long been beset by a disconnect between classroom writing tasks and real-world communicative purposes, insufficient student motivation, and an overreliance on print-only texts. As digital media have become deeply embedded in everyday communication, the form of writing has expanded beyond single-mode prose to encompass image-text essays, audiovisual scripts, infographics, and other multimodal texts; the rise of generative artificial intelligence has further introduced new variables into the writing process and feedback practices. Grounded in multimodal writing theory and multiliteracies theory, and drawing on process writing pedagogy and genre analysis, this paper proposes a theoretical model for multimodal writing instruction in university Chinese composition courses for the digital age. The paper first reviews the theoretical foundations of multimodal social semiotics, the multiliteracies pedagogy, the cognitive process model of writing, and genre analysis. It then elaborates a four-stage instructional cycle, beginning with the design of authentic multimodal writing tasks, moving on to multimodal input modeling through analysis of exemplary texts, then to iterative composition and revision with AI-assisted feedback, and concluding with multi-dimensional assessment and authentic publication. On this basis, the paper further explores the auxiliary role of generative AI, such as large language models, in writing feedback and revision, arguing that AI should be integrated into the writing assessment system to form a multi-layered feedback chain of "AI preliminary review–teacher-guided discussion–peer feedback–independent revision." The paper argues that university Chinese composition instruction needs to move beyond the frame of single-mode printed texts toward multimodal meaning-making; process guidance deserves more weight than product-oriented grading; and assessment should no longer rest solely on the teacher but should combine human judgment with machine input. These adjustments may offer a theoretically grounded reference for curricular reform.

Keywords: Multimodal writing; University Chinese composition; Pedagogical reform; Multiliteracies; Process writing pedagogy; Generative artificial intelligence

Online publication: November 19, 2026

1. Introduction

Writing competence is a core component of undergraduate education and an essential capacity for

university graduates. Yet the actual state of university Chinese composition instruction in China falls short of expectations. Students often lack genuine motivation to write; many produce essays solely to fulfill assignments and earn grades, rather than to express ideas in authentic communicative contexts. Instructional content also remains disconnected from the media ecology in which contemporary students actually communicate. As smartphones, social media, and digital authoring tools become ubiquitous, students' everyday expression routinely combines text with images, sound, layout, and video—yet university composition courses still revolve almost exclusively around printed prose [1].

The rapid rise of generative artificial intelligence is also profoundly reshaping the writing landscape. Generative AI tools based on large language models can now assist students in brainstorming, language polishing, structural revision, and multiple-round editing. Recent systematic reviews indicate that AI-assisted writing yields significant improvements in textual coherence, lexical richness, and linguistic accuracy, but the risks of over-reliance and weakened higher-order thinking also deserve attention [2]. This technological shift brings both new tools and new challenges: How should writing pedagogy position AI? Should it be prohibited as a threat or integrated as a facilitator? Existing research suggests that a blended model combining AI and teacher feedback outperforms either alone—AI handles surface-level linguistic issues, while teachers remain indispensable for deep logical reasoning, contextual appropriateness, and critical thinking guidance [3].

In the evolution of writing pedagogy itself, scholars have long moved beyond the product approach toward the process approach. Flower and Hayes's cognitive process theory, which characterizes writing as a recursive cycle of planning, translating, and reviewing, transformed writing instruction from a focus on final texts toward attention to composing processes [4]. In recent years, Yan further proposed a "process-oriented writing" pedagogy in the Chinese context, emphasizing that a writing course should cover the entire arc from brainstorming and drafting to revision and publication, with teachers adopting different guiding roles at different stages [5]. Nevertheless, the process approach has encountered practical limitations: multi-drafting can be time-consuming without a clear sense of social purpose, and the quality of peer feedback varies considerably. How to balance process-oriented guidance with the social communicative value of writing remains an open question.

Parallel explorations in digital writing practice are also underway. Rong and Zhang argued that cross-media writing is a new form of expression and communication in the digital age, and that writing instruction needs systematic improvement across instructional awareness, task design, methodological guidance, and assessment standards—rather than simply adding images or video clips to a text draft [6]. This means that university Chinese composition instruction can no longer remain confined to the framework of print-only texts, but must respond to the expressive habits of digital natives.

Against this backdrop, this paper attempts to construct a multimodal writing pedagogy for university Chinese composition, using multimodal writing theory as its core framework, incorporating insights from process writing pedagogy and genre analysis, and positioning generative AI as an auxiliary tool within the feedback loop. The central questions are: How can the theoretical resources of multimodal writing be translated into operational instructional stages? And what auxiliary role should generative AI play within this model?

2. Theoretical foundations

2.1. Multimodal writing and multiliteracies

The theoretical foundation of multimodal writing lies in social semiotics and multimodal discourse analysis. Kress and van Leeuwen argued that contemporary communication is no longer confined to language as a

single semiotic mode; instead, it integrates language, image, sound, action, and space, each with distinct "affordances"—potentials for making specific kinds of meaning [7]. Kress's Literacy in the New Media Age further contends that digital media are fundamentally reshaping the definition of literacy: literacy is no longer merely reading and writing printed words but encompasses interpreting and producing image-based, layout-based, and video-based multimodal texts.

Kress's Multimodality provides a systematic social-semiotic account of this position, emphasizing that the selection and use of semiotic resources is always socially and culturally situated, and that meaning-making is a process of design rather than simple transmission [8]. In 1996, the New London Group published *A Pedagogy of Multiliteracies: Designing Social Futures* in *Harvard Educational Review*, formally introducing the concept of "multiliteracies." Arguing that cultural and linguistic diversity combined with multimedia technology demands a broader conception of literacy, the group proposed a pedagogy comprising four components: situated practice, overt instruction, critical framing, and transformed practice [9].

Chinese scholars have also made their own contributions in this field, providing useful analytical tools for multimodal writing pedagogy. These studies mainly revolve around the theoretical framework and research methods of multimodal discourse analysis, which will be further discussed in the specific sample analysis and instructional design sections later.

2.2. Process writing and genre awareness

If multimodal theory addresses the questions of "what to write" and "what to write with," process writing pedagogy and genre analysis address the question of "how to write."

Flower and Hayes's cognitive process model decomposes writing into three major components: task environment, long-term memory, and the writing process itself—the latter comprising planning, translating, and reviewing as recursive subprocesses. This theory fundamentally transformed the paradigm of writing instruction: writing is no longer a one-time text production but a non-linear, recursive cognitive process. Process writing pedagogy emphasizes pre-writing preparation, drafting, revising, proofreading, and publishing, with the teacher's role shifting from evaluator to guide within the process. Yan Cheng's "process-oriented writing" pedagogy has further developed this idea in the Chinese context, advocating for an instructional sequence covering the entire writing process with differentiated intervention strategies at each stage.

Genre analysis theory complements writing pedagogy from another angle. Swales's concepts of "discourse community" and "genre" reveal that writing is not purely personal expression but purposeful meaning-making governed by genre conventions within specific social contexts [10]. Cultivating genre awareness—understanding how communicative purposes, audiences, and media shape writing choices—aligns closely with multimodal writing's emphasis on selecting appropriate modes for specific audiences and purposes. Whether academic papers, promotional copy, or short-video scripts, each has its inherent genre conventions and reader expectations.

2.3. Generative AI in writing and feedback

The impact of generative AI on writing instruction is a frontier topic in current research. Existing studies provide preliminary empirical grounding for positioning AI within writing pedagogy, and the specific effects and mechanisms will be further discussed in the feedback loop of the instructional model below.

3. A multimodal writing pedagogical model

Drawing on the theoretical resources outlined above, this paper proposes a four-stage multimodal writing instructional model: authentic task design, multimodal input modeling, iterative composition and revision, and multi-dimensional assessment and publication. These four stages are not strictly linear but recursive—problems identified during revision may prompt a return to task adjustment or supplementary input.

3.1. Authentic task design: From "writing for the teacher" to "writing for real audiences"

The starting point of multimodal writing pedagogy is authentic communicative tasks. A core problem with traditional writing instruction is that students lack genuine writing motivation—they write to complete assignments and earn grades, not to express ideas or influence others. Multimodal task design must first address the question of "why write."

Teachers can design multimodal tasks around different course objectives. For campus communication scenarios, students might produce an illustrated promotional post for a student club or event. For public issue expression, students might create a podcast script or short-video script on a social topic. For knowledge dissemination, students might transform an academic article into an infographic. What unites these tasks is that they have real audiences rather than merely the teacher, real distribution channels rather than merely submitting work for a grade, and a range of available modes rather than text alone.

At the task-launch stage, the teacher creates an authentic communicative scenario, explains the social purpose and target audience, and then has students attempt a simplified version of the task. For example, in a "campus event promotional post" task, students first describe an event in roughly 100 words, after which the teacher asks: If this post were to attract classmates' attention on a mobile screen, would plain text suffice? What visual elements would be needed? Students thereby discover gaps in their own competence in layout design, image-text pairing, and mode selection. This "gap discovery" experience is key to motivating learning. It should be emphasized that during task launch and initial drafting, students should work independently without AI assistance; otherwise, the intrinsic motivation generated by discovering one's own gaps is undermined.

Multimodal tasks do not require every assignment to combine video, audio, and animation. Kress emphasizes that different modes have different affordances, and writers should choose the mode combination most appropriate to their expressive needs. Sometimes a well-organized, carefully typeset illustrated article is the right multimodal product; sometimes a short video with visuals and music is more emotionally resonant. Teachers should help students understand the rationale for mode selection rather than equating multimodality with formal complexity.

3.2. Multimodal input modeling: From "reading model essays" to "analyzing multimodal exemplars"

The input stage of multimodal writing instruction should not remain confined to traditional textual model essay analysis, but should introduce multimodal exemplars. Process writing pedagogy emphasizes the importance of pre-writing preparation, but multimodal writing also requires deconstructive analysis of multimodal texts as part of that preparation.

Teachers should first select multimodal exemplars matched to the output task. For a "campus promotional post" task, several well-designed posts with high engagement and effective image-text coordination might be analyzed. For an "infographic" task, outstanding examples of information visualization

serve as models. These exemplars are not for imitation of form but as objects of selective learning: students analyze why a particular mode combination was chosen, how text and image relate to each other, and how layout guides the reader's gaze. Hu Zhuanglin, one of the first scholars to distinguish multimodal semiotics from multimedia semiotics from a social-semiotic perspective, provided basic conceptual tools for this kind of sample analysis [11].

Teacher guidance in the input stage can operate on several levels. At the content level, students analyze how exemplars organize material around a core communicative intention. At the formal level, they attend to how different modes—text, image, color, layout—divide labor and coordinate to make meaning. They also need to develop a feel for genre conventions: promotional posts typically follow a "hook–body–call to action" structure, while infographics tend to organize material as a "central claim–sub-claims–data support" sequence. Bhatia's cognitive structuring model further elucidates the underlying mechanisms of this genre construction, helping students understand the deeper cognitive structure behind different genres [12].

Overt instruction is one of the core components of the New London Group's multiliteracies pedagogy. Zhu systematically reviewed the theoretical foundations and research methods of multimodal discourse analysis, carefully distinguishing concepts such as mode, medium, and modality that are easily confused—these conceptual clarifications form the knowledge base for overt instruction [13]. Teachers should not merely display excellent multimodal works but explicitly teach principles of multimodal design—for example, the representational, interactive, and compositional meanings of visual grammar, and complementary relations between text and image. Such overt instruction helps students move from intuitive aesthetic judgment to operational design knowledge.

Selective learning means students do not need to study input materials comprehensively; instead, they take what they need to accomplish their output goal. In multimodal writing, this selectivity manifests differently across students: those stronger in textual expression may focus on narrative technique, while those more visually oriented may focus on image-layout coordination. Teachers should respect these differences while guiding all students toward basic competence in weaker modes.

3.3. Iterative composition and revision: Multimodal meaning-making and multi-round feedback

The composition stage is the heart of multimodal writing. After clarifying the task and analyzing exemplars, students enter the design and production of their multimodal works. This process is not completed in one pass but unfolds through multiple iterations—precisely the "reviewing" phase of the process writing model.

After completing a first draft, the revision stage can introduce AI-assisted feedback. Students may first use an LLM to evaluate the textual portions of their work, focusing on formalizable dimensions such as linguistic norms, structural logic, and cohesion. Systematic reviews show that AI is fast and efficient in identifying these surface-level issues and can significantly improve students' linguistic accuracy and textual coherence [14]. However, students should not blindly accept all AI suggestions—AI sometimes offers plausible but incorrect revisions, or "corrects" expressions that are actually valid. Teachers should present examples of AI feedback in class, guiding students to critically examine each suggestion.

Following AI feedback, the class moves to teacher-student collaborative evaluation. The teacher selects two or three representative student works before class—not necessarily the best or worst, but those illustrating typical problems or strengths—and leads students in discussion during class: Is the

communicative intention clear? Is the mode combination appropriate? How is the image-text relationship? What could be improved? Students internalize evaluation standards through discussion and then revise their own works accordingly. This process resonates with peer feedback in the process writing tradition but is more structured and criterion-referenced.

Revision of multimodal works is more complex than pure text revision, as it involves design-level judgments about whether mode choices are appropriate, whether modes complement each other, and whether the overall effect is coherent. Teachers should guide students during revision not only to polish language but also to reflect from a holistic design perspective on whether the multimodal meaning-making of the work is effective.

3.4. Multi-dimensional assessment and authentic publication

The assessment stage is what makes the whole model work. Evaluation criteria should not stop at language quality but should expand to several dimensions: communicative effectiveness, i.e., whether the work conveys its core intention and reaches its audience; appropriateness of mode selection, i.e., whether modes are chosen according to their affordances; intermodal coordination, i.e., whether text, image, and sound complement rather than redundantly overlap; and textual quality, i.e., accuracy, fluency, and appropriateness of language. The weighting of these dimensions can vary according to task.

Beyond teacher evaluation and AI feedback, authentic audience feedback should also be incorporated into the assessment system. When students' multimodal works are genuinely published on campus accounts, video platforms, or exhibitions and receive real audience feedback, that "publication" itself serves as the most meaningful form of assessment. The social communicative value of writing is thereby realized. Teachers should create authentic distribution channels so that writing moves from classroom assignment to public expression.

Rong and Zhang similarly emphasized that assessment of cross-media writing needs dedicated criteria, extending beyond textual quality to consider dissemination effects and audience responses across different media. This resonates with the "transformed practice" component of the multiliteracies pedagogy—students transfer what they have learned to new contexts, designing and producing their own multimodal texts.

4. Implementation considerations and reflections

4.1. The changing role of the teacher

In the multimodal writing instructional model, the teacher shifts from knowledge transmitter to "design facilitator" and "process guide." This demands more than traditional writing pedagogy: teachers need a basic understanding of multimodal meaning-making, the ability to select multimodal exemplars, design multimodal tasks, and facilitate multimodal evaluation, as well as the capacity to guide students in critically engaging with AI feedback. For teachers accustomed to conventional composition instruction, this transformation requires sustained professional development support.

4.2. Technology and resource conditions

Multimodal writing instruction requires basic technological infrastructure, including graphic editing, video production, and information visualization tools, as well as access to generative AI platforms for feedback purposes. Contemporary university students generally possess basic digital literacy, and many authoring tools

are free or low-cost. The critical issue is not technology per se but whether teachers guide students to use technology—including AI—in service of meaning-making rather than for its own sake. Institutions should provide necessary software licenses, hardware support, and equitable access to AI tools.

4.3. Norms and risks of AI-assisted writing

When integrating generative AI into writing instruction, clear usage norms must be established and potential risks addressed. The most fundamental question is AI's positioning—it is an "auxiliary feedback tool," not a "ghostwriter." Teachers should guide students to use AI during the self-evaluation and revision phase after completing a first draft, rather than using it to generate full texts. Research indicates that students tend to over-rely on AI, which can weaken independent writing ability and critical thinking. Therefore, teachers should require students to retain multiple draft versions, documenting the revision trajectory from first draft to final version as a way to assess genuine writing competence.

Related to this is the cultivation of students' "AI literacy." Beyond learning how to craft effective prompts to elicit useful feedback, students need the ability to evaluate the quality of AI suggestions. Teachers should present examples of AI feedback in class, guiding students to critically examine each suggestion rather than accepting it uncritically.

Academic ethics and originality boundaries also need clarification. Teachers should specify the scope and limits of AI use in the course syllabus and require students to disclose AI-assisted portions. The ultimate goal of writing instruction is to develop students' own expressive competence, not to train them to direct AI to write on their behalf. During task launch and initial drafting, in particular, students should be prohibited from using AI to complete the output task directly; otherwise, the "gap discovery" that motivates learning is undermined.

4.4. Avoiding formalism

A recurring risk in multimodal writing instruction is reducing multimodality to "adding images" or "adding video"—assuming that inserting pictures into an essay constitutes multimodal writing. Kress emphasizes that multimodality centers on meaningful coordination among modes, not their simple accumulation. Zhang Delu's integrated framework for multimodal discourse analysis similarly notes that modal relations include complementary and non-complementary types; multimodality has genuine pedagogical value only when one mode cannot independently convey the full meaning and requires supplementation [15]. Teachers should therefore anchor all multimodal activities in communicative purpose, avoiding multimodality for its own sake.

4.5. Boundaries of theoretical transfer

The theoretical resources drawn upon in this paper originate partly from social semiotics and foreign language education. Their transfer to university Chinese composition—in which students write in their mother tongue—requires attention to boundaries. The core claims of multimodal social semiotics—the affordances of modes, meaning as design, the collaborative relations among modes in multimodal texts—have cross-linguistic generality, but specific instructional operations, such as the content of overt instruction and the weighting of evaluation criteria, need to be adapted to the Chinese writing context. This paper does not advocate wholesale importation of all details of multimodal theory, but rather takes its core spirit—

that writing is multimodal meaning design, that writing should serve authentic communication, and that the writing process requires multi-round feedback and iteration—and fuses it with process writing pedagogy and genre awareness to construct a model suited to university Chinese composition. The actual effectiveness of this model remains to be tested and refined through ongoing teaching practice.

5. Conclusion

Writing in the digital age is shifting from single-mode printed texts toward multimodal meaning-making, and the rise of generative AI adds a further variable to this transformation. University Chinese composition instruction that remains confined to traditional essays can neither respond to changes in students' everyday expression nor meet the new demands that the AI era places on writing competence. Grounded in multimodal writing theory and multiliteracies, drawing on process writing pedagogy and genre analysis, and positioning generative AI as an auxiliary feedback layer, this paper has proposed a multimodal writing instructional model for university Chinese composition. The model runs on authentic multimodal tasks as the driver, selective analysis of multimodal exemplars as the enabling mechanism, and a multi-layered feedback chain of "AI preliminary review–teacher-student collaborative discussion–authentic publication" as the pedagogical closure. At the end of the day, writing courses still grapple with the same old questions—why write, what to write, and how to evaluate—but in the digital age, the form of "writing" and the methods of "evaluation" have both changed. Whether this model actually works is something that can only be tested through ongoing teaching practice, particularly regarding the effectiveness assessment, ethical norms, and pedagogical boundaries of AI-assisted writing, which warrant further empirical investigation.

Disclosure statement

The author declares no conflict of interest.

References

[1] Kress G, 2003, Literacy in the New Media Age, London: Routledge.

[2] Sanz-Tejeda A, Domínguez-Oller JC, et al., 2026, The Impact of Generative AI on Academic Reading and Writing: A Synthesis of Recent Evidence (2023–2025). Frontiers in Education, 10: 1711718.

[3] Lim JM, Go C, 2025, Integrating ChatGPT into Teacher Feedback: Practical Insights for L2 Writing Instruction. TESL-EJ, 29(3): EJ115.

[4] Flower L, Hayes JR, 1981, A Cognitive Process Theory of Writing. College Composition and Communication, 32(4): 365–387.

[5] Yan C, 2023, On Process-Oriented Writing Pedagogy. Writing, (5): 107–114.

[6] Rong WD, Zhang Y, 2024, Cross-Media Writing: New Forms of Writing in the Digital Age and Instructional Strategies. Secondary School Chinese Language, (34): 10–18.

[7] Kress G, Van Leeuwen T, 2001, Multimodal Discourse: The Modes and Media of Contemporary Communication, London: Arnold.

[8] Kress G, 2010, Multimodality: A Social Semiotic Approach to Contemporary Communication, London: Routledge.

[9] New London Group, 1996, A Pedagogy of Multiliteracies: Designing Social Futures. Harvard Educational Review,

66(1): 60–93.

[10] Swales JM, 1990, Genre Analysis: English in Academic and Research Settings, Cambridge: Cambridge University Press.

[11] Hu ZL, 2007, Multimodalization in Social Semiotics. Language Teaching and Linguistic Studies, (1): 1–10.

[12] Bhatia VK, 1993, Analysing Genre: Language Use in Professional Settings, London: Longman.

[13] Zhu YS, 2007, Theoretical Foundations and Research Methods of Multimodal Discourse Analysis. Foreign Language Research, (5): 82–86.

[14] Jiao Y, Wang Q, 2026, Large Language Models for Formative Feedback in Writing Instruction: A Systematic Review of Classroom Interventions, Feedback Quality, and Student Outcomes. Frontiers in Education, 11: 1834085.

[15] Zhang DL, 2009, Exploring an Integrated Theoretical Framework for Multimodal Discourse Analysis. Foreign Languages in China, 6(1): 24–30.

Publisher's note

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