International Journal of Chinese Language Teaching (2026)
260603, 1-28 https://doi.org/10.46451/ijclt.260603

Article

The Application of ChatGPT in Creating Chinese Cultural Context Scripts: An Activity Theory Perspective

Emma Rao*
Ruohan Zhao
The Hong Kong University of Science and Technology, Hong Kong SAR, China

Received: 1 September 2025/Accepted: 15 March 2026/Published: 6 July 2026

Abstract
From a sociocultural perspective, and based on an activity theoretical framework, this study examines the integration of ChatGPT into a structured, extracurricular program entitled the Lens on China workshop series, investigating how Chinese language learners use this tool to construct cultural context scripts. Accordingly, two research questions are posed: (1) How does learners' perception of ChatGPT's role influence their approach to scriptwriting within a cultural context? (2) What factors influence the creation of cultural scripts with ChatGPT? The results indicate that the perceived role of ChatGPT varies and influences prompting strategies: learners viewing it as an assistant prefer contextual prompts, while those who see it as a tool tend to use procedural prompts. Contextual prompts provide learners with appropriate culturally relevant content, while procedural prompts offer necessary structural elements. A sequential combination of contextual-first and subsequent procedural prompts proved most effective for achieving culturally appropriate and structurally sound scripts. Further analysis through Activity Theory identified several contradiction pairs. The contradiction between rules and community is identified as key to understanding the task and learning outcomes. This tension arose when ChatGPT generated culturally plausible-sounding but unauthentic content, leaving learners unsure of genuine norms. When this contradiction was effectively mediated by the instructors—who articulated explicit cultural and pragmatic rules—it transformed abstract cultural knowledge into performable linguistic expressions, thereby significantly enhancing intercultural communicative competence and speaking skills. Consequently, instructors' role must shift from a resource provider to a mediator of contradictions, harnessing the gap between AI logic and language culture to deepen learners' intercultural competence.

Keywords
ChatGPT, Activity Theory, cultural context script

*Corresponding author. Email: lcemmarao@ust.hk

1 Introduction

In the field of language teaching and learning, several studies on using films as learning resources have proven advantageous for learners in acquiring a target language and its culture. According to multiple experts in the field, English language learners can develop their linguistic competence through films, as they hone their listening, reading, writing, and speaking skills while simultaneously learning English grammar, vocabulary, and pragmatics (Kabooha, 2016; Liu, 2019; Vyushkina, 2016). When learners explore these aspects of the target culture through films, they spot similarities and differences with their own community, which fosters a deeper awareness of their home culture (Xue & Pan, 2012; Yue, 2019; Zoreda, 2005).

The Lens on China workshop series, an extracurricular program, aims to enhance Chinese language learners' speaking skills and intercultural awareness beyond the formal classroom. In selecting films, instructors consider two main criteria: learners' target language proficiency and the content of the films. Targeting learners at an intermediate-high level, the selected films must address cross-cultural issues and be accessible. Two films are selected from the *Lens on China* textbook (Wang & Tian, 2020). One is The *Wedding Banquet*, directed by Ang Lee in the 1990s, and the other is *Beijing Meets Seattle*, released in 2013. Although produced in different eras, both films explore cross-cultural ideologies and controversial themes in contemporary China, thereby stimulating learners' critical thinking and insightful discussion.

During the development of the workshop series, the emergence of generative artificial intelligence (AI) models like ChatGPT has profoundly impacted education and learning experience. For learners, ChatGPT can play various roles in supporting language learning (Huang et al., 2022). Within this context, learners are tasked with using ChatGPT to create short, cultural context scripts based on film roles.

The Performed Culture Approach (namely *PCA*) is employed in this workshop series for its emphasis on constructing a cultural context that enables learners to acquire the target language through target culture immersion. According to Kramsch (1993), language and culture are inseparable; thus, second language learners inevitably become learners of the second culture. As a pedagogical framework, *PCA* seeks to develop both linguistic proficiency and intercultural communicative competence when studying a second language (Yu, 2021). According to Walker (2010), social life is essentially composed of a series of performances, with each specific act containing specific time, place, roles, dialogue, and audience. These fundamental elements constitute the core structure of Performed Culture Approach (*PCA*) pedagogy, which emphasizes that learners must engage in appropriate communication within culturally specific contexts. When learners utilize new technologies such as ChatGPT to create cultural context scripts, they engage in a socio-cultural communicative activity that involves complex interactions among multiple agents (learners, instructors, and ChatGPT) within a defined activity system. From the perspective of Activity Theory, this framework enables us to systematically examine the interactions and emerging contradictions within the script creation activity, particularly how the introduction of ChatGPT as a new mediating artifact reshapes the activity system. Consequently, this paper explores the role of ChatGPT in facilitating learners' construction of typical cultural contexts through the lens of Activity Theory.

2 Literature Review

2.1 Construction of cultural context

Learning and understanding a language within a context have reached a consensus in the academic community, and this is also true in the field of teaching Chinese language as a foreign or second language (Jiang, 2007; Lyu, 2008). Zhao (2008) has highlighted that language learning is a multifaceted process

that involves communication, exchange of ideas and expression of emotions. The ultimate aim of language learning is not simply to grasp grammatical rules, but to enhance communicative competence in a social context. Therefore, Chinese language teaching should emphasize the real-life context and authentic materials. Zhang (2020) points out that communication skills encompass not only linguistic elements but also psychological language and cultural knowledge. Language instructors should not expect that learners can acquire communication skills automatically, and instructors need to strive to incorporate authentic communication and near-authentic communication into the classroom. Through contextualized teaching processes, they intentionally train and develop learners' communication skills so that they can engage naturally and appropriately in real social environments. Furthermore, due to learners' insufficient Chinese proficiency, it is important for them to receive support from their instructors when constructing a cultural context script.

The *PCA* pedagogy is widely used by instructors, as it is more practical and operational to enhance learners' intercultural communicative competence. The *PCA* was developed by Walker and Noda in the 1990s of last century, by a group of Chinese and Japanese educators. It borrows the philosophical concept "zhi xing he yi" (知 行 合 一) from Wang Yangming (1472-1529), a prominent scholar and educator during the Ming Dynasty, meaning "integration of knowledge and practice". The performed-culture pedagogy aims to enable language learners to not only gain advanced linguistic skills, but also develop intercultural communicative competence, successfully navigating in the target culture and effectively communicating with native speakers in culturally appropriate ways (Yu, 2021). Many *PCA* practitioners believe that to know something is to do it. By constructing a typical cultural context, a learner can perform the target language and remember the target culture, which cannot be learned directly from their own culture.

Therefore, setting up a cultural context is crucial for improving learners' communicative competence, such as speaking skills and intercultural competence. Using the fundamental elements of the *PCA*—time, venue, participants, audience, and dialogue (or lines)—learners can construct a cultural context reflective of the target culture. When *PCA* is applied in the Lens on China workshop series, learners are required to create cultural context scripts based on the five fundamental elements mentioned above. They are encouraged to perform different roles using their voices while applying speaking skills like rhythm, logical stress and intonation. The workshop series mainly places a strong emphasis on script reading, where learners utilize speaking techniques to interpret the roles in the script, rather than engaging in physical performance. This represents a focused adaptation of *PCA*, shifting its traditional emphasis from holistic embodiment to the nuanced use of vocal prosody for cultural expression.

2.2 ChatGPT in Chinese language education

The emergence of generative AI has brought a significant impact on language education, especially in Chinese as a second language teaching and learning, presenting both opportunities and challenges for language learners and instructors. On the opportunities side, ChatGPT can help instructors prepare teaching plans, generate text, vocabulary lists, and explanations of words. ChatGPT can be a friendly tool, a peer, even a tutor for learners. On the challenges side, compared with traditional teaching, instructors have to consider the way of their teaching, and their authority may face doubts. Learners' prior learning experience and their creativity of self-regulated learning may face threats as well (Liu et al., 2023).

Previous studies in the field of Chinese language teaching and learning primarily focus on the discussions at the theoretical level or the exploration of the possibilities of ChatGPT in various domains, such as vocabulary learning (Meng, 2024), written corrective feedback (Gao, 2025), and writing for children's reading materials (Song et al., 2024). The recent studies have indicated a deeper examination with more experimental evidence in the application of ChatGPT in Chinese learning environments, such as Chinese sentence grammar teaching (Yu et al., 2025), and Chinese learners' stress appraisal and

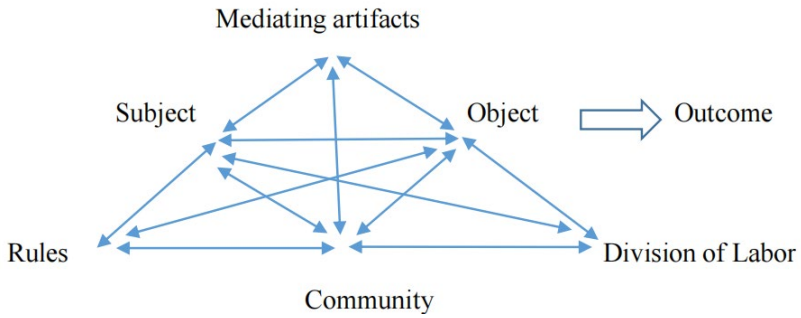
coping strategies (Li et al., 2025). Additionally, several studies involve pedagogies related to teaching and learning, such as task-based learning and teaching approach (Jiang et al., 2023), project-based course design (Liang & Luo, 2024), and differential approaches in teaching and learning (Li et al., 2024). However, research on culturally relevant pedagogy in the use of ChatGPT remains limited. Shi and Lyu (2025) suggest that the few existing empirical studies focus primarily on the overall effectiveness of AI tools, lacking in-depth exploration of specific questions such as which functions are effective, and what conditions they are effective, and for what type of learners. Liu et al. (2023) strongly emphasize that the pedagogical effectiveness of generative AI depends heavily on how users frame inputs to elicit structured and contextually appropriate information. Therefore, for learners, the strategic design of prompts—defined as the textual inputs or instructions used to initiate and direct ChatGPT's responses—is highly crucial. Prompt design serves as a strategic mediation between the learner's intention and the AI's textual output related to culture. This paper seeks to address the identified gaps by refining how ChatGPT is utilized to help learners construct cultural context. Grounded in Activity Theory, this approach recognizes that language use, while situated within specific interactions, is fundamentally shaped by the broader social context.

2.3 Activity Theory as the theoretical framework

Activity theory originated from Vygotsky's (1978) work, which conceptualizes learning as involving a subject, an object, and mediating artefacts. Three main components usually refer to the learner, the task, and a tool such as a computer. The relationship between the subject and object of activity is mediated by a tool. However, the above system cannot deal with the relationship between the learner and the environment in an activity. EngestrÖm (1987) expanded it by including contextual elements of rules, community and division of labor in addition to subject, object and tools. EngestrÖm conceived Activity System as shown in Figure 1.

Figure 1

Activity Theory Framework (EngestrÖm, 1987)

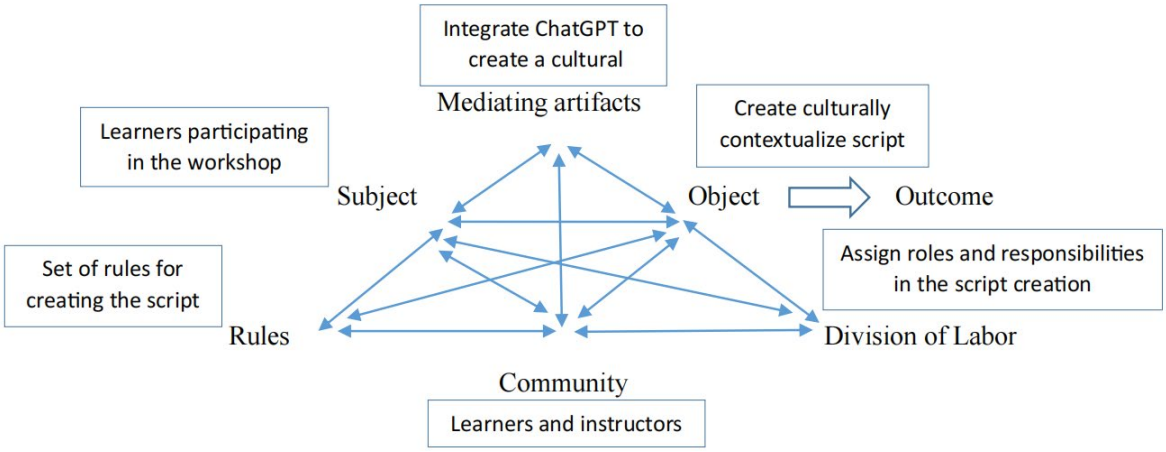


An activity system as above is composed of subject, object, tool (material or psychological), rules, community, division of labor and outcome. The object is the motive and driving force of an activity (ie, outward goal, concrete purpose or objectified motive). The subject is an individual or a group. Rules refer to norms and conventions concerning how certain things are decided within the system and exist to ensure shared responsibility and cooperation between the participants. The division of labor is defined as the organization of processes related to the goal. Community functions as a social group and mediates the interaction between the subject and the rules or the division of labor or the object. Hence, there are always mediating interactions within or between the activity's constituent components. The outcome is the intended (or not) implications of the activity (Barab et al., 2004).

When the Activity Theory framework is applied to the Lens on China workshop series (see Figure 2), the subject refers to the learners who participated in the workshop. The object serves as the driving force that motivated the learners to complete the video recording task by the end of the workshop. Unlike traditional artifacts, ChatGPT, as an AI-driven chatbot, raises the question of whether it functions purely as a tool or adopts more "human" roles, such as assisting learners in writing a script, which adds depth and relevance to this study. The rules refer to the guidelines that learners had to follow when creating the cultural context of their scripts, including the appropriate use of Chinese language and cultural norms. The division of labor involves the interactions and responsibilities of learners within small groups, specifying their tasks and cooperation. The community provides the social environment in which these interactions take place, including those between learners and the instructor when necessary.

Figure 2

AT-based Framework for the Lens on China Workshop Series (Originated from EngestrÖm,1987)



According to previous research, Activity Theory has been widely used as a design framework to specify the computer-supported collaborative learning activities (Collis & Margaryan, 2004; Zurita & Nussbaum, 2007), human-computer interactions (Lazarou, 2011), and mobile learning (Uden, 2007). In the last decade, the application of Activity Theory as a conceptual basis for analyzing educational technology has increased dramatically, including game-based learning (Carvalho et al., 2015) and social media-oriented learning (Rambe, 2012).

Recently, with the trend shifting from computer-assisted language learning to artificial intelligence-assisted language learning, it is worth conducting research in this area. However, there has been little research using Activity Theory as a conceptual framework to further explore the use of generative AI technology in the field of Chinese language teaching and learning. As the role of generative AI differs from that of other tools, Cai (2023) suggests that ChatGPT can serve as both a tool and a member of the community in the reading activity.

Alongside ChatGPT's actual function within the scriptwriting activity, interactions among all components can generate various contradictions. These are not viewed as individual errors or mechanical failures, but as structural tensions accumulated over time that act as the 'motive force' for change and development. As Engeström (2001) further illustrates, these contradictions expand the framework, shifting the analysis of the learning process from a static description to a dynamic one. By adopting Engeström's concept of contradictions, this study will not only identify the components of ChatGPT-mediated learning but also deeply explore how contradictions arise from the tensions, such as between AI-generated outputs (as tools) and authentic cultural norms (as rules or key components) within the framework. The tensions, in turn, can serve as a driving force for learner transformation, pushing learners to move beyond their initial capabilities via expansive learning.

To systematically trace how such contradictions manifest and are resolved, this study integrates Engeström's theoretical framework with the qualitative data analysis approach of Miles, Huberman, and Saldaña (2014). Their structured methods—including iterative coding and pattern recognition—enable detailed tracking of how learners navigate tensions during interactions with ChatGPT.

Consequently, guided by this integrated framework, this study addresses the following questions:

(1) How does learners' perception of ChatGPT's role influence their approach to scriptwriting within a cultural context?

(2) What factors influence the creation of cultural scripts with ChatGPT?

3 Methodology

3.1 Research design

This study adopted a qualitative design with embedded quantitative descriptors within the Lens on China workshop: four sessions over two weeks (six hours total). The pedagogical progression moved from vocal delivery training to instructor-guided small-group film analysis, culminating in the core task where learners used ChatGPT to create, rehearse, and perform original cultural context scripts. Participants formed two small groups, each with a dedicated instructor, concluding with a joint viewing and feedback forum. Activity Theory provided the conceptual framework for analyzing interactions among six activity system components: subject, tool, object, rules, division of labor, and community. This framework guided analysis of how learners' prompting strategies, perceptions of ChatGPT's role, and script cultural appropriateness contributed to the activity system.

3.2 Participants

Ten learners initially participated; seven completed all sessions. All were enrolled in four-year undergraduate programs in business, science, or engineering at a local university. Detailed demographics appear in Table 1. Most completed pre-tertiary education in Hong Kong international schools (one from Guangzhou), attained HSK Level 5, and completed International Baccalaureate Chinese examinations as second-language learners, though several also speak Cantonese. A minority took no formal Chinese examinations. One learner was of Korean origin with international education in Guangzhou prior to higher education.

Table 1

Participants' Demographic Information

Learner ID	Year of study	Major	Educational background prior to higher education
Learner 1	Yr 4	Computer Engineering	International school (Hong Kong)
Learner 2	Yr 4	Computer Engineering	International school (Hong Kong)
Learner 3	Yr 1	Business Management	International school (Guangzhou)
Learner 4	Yr 1	Business Management	International school (Hong Kong)
Learner 5	Yr 2	Business Management	International school (Hong Kong)
Learner 6	Yr 1	Business Management	International school (Hong Kong)
Learner 7	Yr 1	Business Management	International school (Hong Kong)

3.3 Data sources and coding procedures

Data comprised four types. Learners' texts included self-reflections and two semi-structured interviews. Instructors' texts included two self-reflections and two semi-structured interviews as primary evidence. Workshop verbatim transcripts and artifacts—slides with annotations, final scripts, and video performances—provided corroborative evidence.

The coding framework was developed through systematic analysis of learner–ChatGPT interactions, guided by the iterative approach outlined by Miles, Huberman, and Saldaña (2014). The framework was structured around two primary analytical foci: first, characterizing learners' interactions and script outcomes to address RQ1; and second, identifying systemic tensions and their mediation to explain the mechanisms underlying RQ1 outcomes and address RQ2.

The first focus characterized learners' ChatGPT interactions and script outcomes through three dimensions: (1) Prompt Types categorized learner instructions. *PromptProcedural* denotes task-driven prompts specifying structural parameters without context (e.g., "generate a five-minute script"), while *PromptContextual* specifies role identities, relations, settings, and cultural requirements (e.g., "wedding banquet scene; groom with Chinese parents; address as '爸、妈'; respectful, colloquial tone"). These categories analyzed how prompt design correlated with output quality. (2) Role Perception analyzed reflections and interviews for ChatGPT characterizations (e.g., "assistant," "tool"), grouping participants by predominant metaphor: *assistant-perceiving* (n = 3) and *tool-perceiving* (n = 4). A *Tool-Assistant Index (TAI)* quantified emphasis: *TAI = (Role_Assistant − Role_Tool) / (Role_Assistant + Role_Tool)*, with positive values indicating assistant orientation. (3) Script Outcomes evaluated cultural and pragmatic appropriateness. *ScriptFit* labeled segments conforming to norms; *ScriptDeviation* labeled deviations (non-native phrasing, cultural inaccuracy, register misalignment). Coverage—proportion of coded spans weighted by group text length—enabled quality comparison. These dimensions revealed how role perception shaped prompting strategies and script quality.

The second analytical focus employed contradiction and mediation analysis to explain the mechanisms enabling or hindering high-quality, culturally appropriate script outcomes. This approach posits that the differential script quality observed in RQ1 resulted from systemic tensions—structural misalignments within the activity system. These systemic tensions were theorized, through Engeström's (2001) framework, as contradictions—structural points where activity system elements co-occurred but pulled in opposing directions. By identifying these contradictions, analyzing how they were mediated, and examining what outcomes they produced, this focus elucidated the deeper mechanisms influencing cultural script creation.

To operationalize these tensions, this study identified Contradiction Pairs through co-occurrence analysis. A pair was operationalized as two Activity Theory elements co-occurring but pulling in conflicting directions. To quantify the salience of these tensions, a *Strength Index (SI)* was calculated: *SI = co-occurrence frequency / min (code 1 frequency, code 2 frequency)*. Within these identified contradictions, Mediation Modes were coded to analyze resolution strategies: Instructor Mediation (direct intervention and guidance), Peer Collaboration (problem-solving or discussion among learners), and Self-Correction (individual learner attempts to resolve issues). Micro-pathways then traced sequences from tension detection → mediating action → outcome (e.g., *ScriptFit* or *ScriptDeviation*) within a minimal adjacency window (≤5 lines). Key codes within these pathways included *Rule_Interpret* (cultural rule clarification), *C_Mentoring* (instructor cultural guidance), *Obj_Fit* (specification alignment), *ICC_Improve* (intercultural competence enhancement), and *SS_Improve* (speaking skills enhancement). This dual analysis of contradictions and their mediation elucidated the deeper factors influencing cultural script creation.

Coding was conducted in four iterative rounds using MAXQDA with continuous cross-checking, negative-case searches, and comprehensive decision logging. Round 1: open coding with contextual detail preserved. Round 2: consistent naming conventions applied. Round 3: guided by research questions

and Activity Theory's six elements, axial coding (Corbin & Strauss, 2015) linked initial codes into Context–Contradiction–Mediation–Outcome sequences. Concurrently, internalization chains (e.g., *Rule_ConceptBuild* → *Rule_Interpret* → *Rule_Apply*) tracked learners' progression from external mediation to individual mastery (Kaptelinin & Nardi, 2006). Round 4: holistic review merged overlapping labels, unified conventions, and tested code stability through multi-source cross-checks.

4 Results

4.1 Prompting approaches and ChatGPT role perception

This section addresses RQ1 by presenting the analysis of learners' prompting approaches and their perceptions of ChatGPT's role during script creation. The findings illustrate how these two factors are associated with the observed variations in script quality.

This section presents findings related to learners' prompting approaches and ChatGPT role perception. Consistent with the definitions in Section 3.3, prompts were classified as *PromptProcedural* or *PromptContextual*, and learners were grouped into an assistant-perceiving group (n=3) and a tool-perceiving group (n=4) based on their conceptualization of ChatGPT's role, quantified by the *Tool-Assistant Index* (*TAI*). Script outcomes were evaluated for *ScriptFit* or *ScriptDeviation*, as detailed in Section 3.3.

Table 2 summarizes the prompting strategies, ChatGPT role perceptions, and script appropriateness outcomes for both groups.

Table 2
Prompting Strategies and Script Appropriateness by ChatGPT Role Perception

Group (n)	TAI range	Prompt procedural coverage (%)	Prompt contextual coverage (%)	Procedural/ contextual ratio	Script fit coverage (%)	Script deviation coverage (%)	Fit/ deviation ratio
Tool (4)	−1.00 to −1.00	7.89	12.36	0.64	2.95	23.18	0.13
Assistant (3)	+0.29 to +1.00	10.58	22.32	0.47	23.15	15.96	1.45
Δ (Assistant − Tool)	—	+2.69	+9.96	-0.17	+20.20	-7.22	+1.32

Notes. *TAI* = (Role_Assistant−Role_Tool) / (Role_Assistant+Role_Tool). Coverage denotes the proportion of coded spans within a given dimension, weighted by each group's total analyzable text length. Due to separate normalization of Prompt and Script dimensions, row/column values do not necessarily sum to 100%.

From Table 2, the assistant group demonstrated higher contextual prompt usage (22.32% vs. 12.36%) and lower *Procedural*/*Contextual* ratio (0.47 vs. 0.64), indicating a contextual-first approach associated with significantly higher script appropriateness (ScriptFit: 23.15% vs. 2.95%; *Fit*/*Deviation* ratio: 1.45 vs. 0.13). Positive Δ values (+20.20 and +1.32) underscore the assistant group's substantial advantage, suggesting a "contextual prompts → appropriate scripts" pathway.

Qualitative data illustrated these patterns. Assistant-oriented learners followed a "contextual prompt → initial draft → deeper refinement" pathway, treating ChatGPT as a negotiable "assistant" to unlock ideas. One learner noted: "it can be an assistant…at least ChatGPT wrote something, so I didn't need to start from zero" (Learner 5, Interview). This involved iterative refinement to "fit better with the western or Chinese culture," frequently yielding higher ScriptFit where "each line really matches the character."

Tool-oriented learners adopted a "parameter setting → one-shot generation → local patching" pathway. One student reported: "I stated that I wanted six unique characters and the script's length to be five minutes…it took me five attempts before generating a script consisting of six characters" (Learner 4, Self-reflection). Insufficient upfront specification of role background and speech style resulted in outputs that drifted from targets, necessitating extensive revision—consistent with lower *ScriptFit* and higher *ScriptDeviation*.

Qualitative evidence also revealed linguistic and cultural deviations. One learner observed: "some phrases are very awkward and aren't things that Chinese people would use" (Learner 3, Interview). Another noted ChatGPT's weaker Chinese ability: "ChatGPT's Chinese ability was clearly weaker than its English…we generated the script in English first, then translated it into Chinese…resulting in grammatically incorrect sentence" (Learner 2, Self-reflection). Learners also reported stylistic issues: "The first script was unnatural…the wording was often too formal…we needed to make it more down-to-earth" (Learner 7, Self-reflection). These insights support the observed ScriptDeviation patterns and contextualize the subsequent analysis of tensions and revision processes.

4.2 Contradiction pairs and micro-pathway identification

Section 4.1 identified two distinct prompting pathways and their divergent outcomes. To provide a more granular understanding of the mechanisms behind these results, this section (addressing RQ2) analyzes the systemic tensions within the activity system. As theorized in Section 3.3, these tensions manifest as structural contradictions—points where elements co-occur but pull in conflicting directions. By identifying these contradiction pairs and tracing their micro-pathways, this analysis documents the iterative sequences from tension detection → mediating action → outcome (e.g., *ScriptFit* or *ScriptDeviation*) within a minimal adjacency window (≤5 lines) (Engeström, 2001; Miles et al., 2014).

Based on the *Strength Index (SI)* values derived from co-occurrence analysis (detailed in Appendix C), contradiction pairs were selected and ranked. Only pairs with $SI \geq 0.300$ were retained for interpretability.

4.2.1 Dominant contradiction pairs: Results

To locate the primary tensions, we ordered contradiction pairs by *SI*, yielding four salient pairs: Rules–Community (*SI*=0.762), Tools–Object (*SI*=0.360), Community–Division (*SI*=0.352), and Tools–Rules (*SI*=0.313) (see Table 3). These four pairs map onto four distinct activity system tensions: (1) whether cultural rules are commonly understood and enacted within the community; (2) whether ChatGPT outputs meet target script specifications; (3) how task division is coordinated during collaborative work; and (4) how rule implementation is constrained by tool limitations.

Table 3
Dominant Contradiction Pairs in ChatGPT-Mediated Script Creation Activity

Contradiction pair	**Strength index (≥ 0.300)**	**Pair-associated dynamic (non-causal)**
Rules–Community	0.762	Rules—team interaction: consensus-in-practice formation
Tools—Object	0.360	ChatGPT output–target script specification: alignment fit
Community–Division	0.352	Team interaction–task allocation: coordinated workflow
Tools–Rules	0.313	ChatGPT constraints–rule implementation: operational limits

Notes. $N = 992$ coded segments. *SI* = co-occurrence frequency / min (code 1 frequency, code 2 frequency). Pairs are included when $SI \geq 0.300$ and ordered by *SI* (ties broken by co-occurrence frequency, then alphabetically). The third column summarizes observed associations within the workshop data.

4.2.2 Mediation modes in the rules–Community Contradiction Pair

The Rules–Community pair exhibited the highest *Strength Index* (*SI* = 0.762), notably surpassing the other three pairs (Tools–Object: *SI* = 0.360; Community–Division: *SI* = 0.352; Tools–Rules: *SI* = 0.313). This dominance suggests that Rules–Community tensions directly impacted how cultural rules were understood and applied, influencing text appropriateness and quality. To identify how these tensions were addressed, mediation modes were analyzed through three primary strategies: Instructor Mediation, Peer Collaboration, and Self-Correction (Table 4).

Mediation Modes and Resolution Efficacy. Instructor Mediation was predominant (100% frequency, 81.3% resolution rate), substantially exceeding Peer Collaboration (37.5% frequency, 66.7% resolution rate) and Self-Correction (18.75% frequency, 33.3% resolution rate). This hierarchy reflects instructor-led intervention as the primary strategy for addressing Rules–Community tensions.

Outcome Orientation. Instructor-mediated episodes predominantly yielded *ICC_Improve* ($n = 10$ of 13), reflecting cultural rule clarification focus. Peer Collaboration was more frequently associated with *SS_Improve* ($n = 4$ of 4), suggesting peer interactions primarily contributed to fluency rather than cultural appropriateness. Self-Correction yielded limited improvement outcomes ($n = 1$ of 1), indicating constrained efficacy as a standalone mode.

Table 4
Mediation Modes and Resolution Outcomes for Rules–Community Tensions (N = 16 episodes)

Mediation mode	**Episodes (*n*)**	**Episode rate (%)**	**Resolved (*n*)**	**Resolution rate (%)**	**Primary outcome evidence (*n*)**
Instructor Mediation	16	100.0	13	81.3	ICC_Improve ($n = 10$)
Peer Collaboration	6	37.5	4	66.7	SS_Improve ($n = 4$)
Self-Correction	3	18.75	1	33.3	Limited improvement ($n = 1$)

Notes. Episode rate = Episodes (*n*) / 16 total episodes. Resolution rate = Resolved (*n*) / Episodes (*n*) within each mode. Mediation modes can co-occur within a single episode. Outcome counts represent resolved outcomes only and are not mutually exclusive. *ICC* = Intercultural Communicative Competence; *SS* = Speaking Skills. Micro-pathway sequences are detailed in Table 5.

Qualitative evidence further illuminates these patterns. One learner (Learner 3, Self-reflection) noted the limitation of self-correction: "Although we explicitly told GPT this was for a five-minute script, it only generated around two to three minutes." Another learner (Learner 5, Self-reflection) reflected on instructor intervention: "We [then] instructed ChatGPT to extend the script length…[though] we did notice some content repetitions." This sequence illustrates how initial self-directed attempts were insufficient, necessitating instructor clarification of the script length requirement before effective resolution could occur.

As demonstrated in Table 4, peer collaboration and self-correction consistently yielded lower resolution rates and more limited scope in fostering deeper learning outcomes (e.g., *ICC* improvement) compared to instructor-led mediation. Table 5 presents the ordered sequence of micro-pathways within instructor-mediated episodes, tracing tension-detection→mediation→outcome trajectories to provide finer-grained insight into instructor-mediated resolution structures.

Table 5
Ordered Micro-Pathways within Instructor-Mediated Episodes
(Tension Detection → Mediation → Outcome; adjacency window ≤ 5 lines; N = 16 episodes)

Pathway (ordered sequence)	**Episodes (*n*)**	**Share of episodes (%)**
Rule_Interpret → C_Mentoring → ICC_Improve	10	62.5
Rule_Interpret → C_Mentoring → SS_Improve	4	25.0
Rule_Interpret → C_Mentoring → Obj_Fit	2	12.5

Notes. Micro-pathways are traced within an adjacency window of ≤ 5 lines, consistent with the operationalization in Section 3.3. Pathways are assigned by their first realized endpoint outcome; proportions are calculated as Episodes (*n*) / 16. For code definitions, refer to Section 3.3.

As shown in Table 5, the most prevalent micro-pathway was "*Rule_Interpret → C_Mentoring → ICC_Improve*," occurring in 10 episodes (62.5%). This pathway indicates that systematic instructor intervention in cultural rule interpretation predominantly enhanced learners' intercultural communicative competence, substantially exceeding pathways leading to *SS_Improve* (25.0%) or *Obj_Fit* (12.5%). This distribution underscores instructor-led cultural rule mentoring's capacity to advance learners beyond surface-level task compliance toward deeper competence development.

In summary, the analysis of mediation modes and micro-pathways provides empirical evidence of how Rules–Community contradictions were systematically resolved within the workshop activity system. Instructor-led mediation emerged as the dominant and most efficacious strategy, with its characteristic outcome being enhancement of intercultural communicative competence. These findings will be synthesized in relation to the broader research questions in the Discussion section.

5 Discussion

5.1 Interpreting prompting strategies and role effects

Results revealed two distinct prompting orientations. Learners perceiving ChatGPT as an "assistant" predominantly employed contextual prompts, achieving higher *ScriptFit* coverage (+20.20 percentage points) and stronger *Fit/Deviation* ratio (+1.32), while those perceiving it as a "tool" relied on procedural prompts, experiencing lower script appropriateness. This differential performance reflects underlying differences in how learners conceptualize the human-AI interaction and structure their prompting strategies.

The emergence of these two pathways stems from differing conceptualizations of the human-AI interaction itself. When learners frame ChatGPT as a collaborative "assistant," they implicitly adopt a goal of mutual understanding and situated problem-solving; interaction becomes an ongoing negotiation of meaning. This framing naturally prompts learners to externalize context—articulating audience, setting, relationships, tone, and cultural parameters—because the "assistant" is conceptualized as a responsive partner requiring clarification. Conversely, framing ChatGPT as a "tool" invokes a means-end logic: specify input parameters (length, character count, template) and expect conforming output. This instrumental view treats the interaction as a one-directional configuration task rather than collaborative refinement. From an Activity Theory perspective, role perception functions as a shaping force: it channels learners' attention toward different activity elements, leading to divergent prompt design choices and, consequently, divergent production pathways (Engeström, 2015). The difference is not merely stylistic; it reflects fundamentally different assumptions about what the tool is for and how human agency operates within AI-mediated work.

The performance advantage of contextual prompts operates through a metacognitive mechanism: externalization of evaluative standards. When learners specify audience, setting, tone, and cultural background upfront, they are forced to articulate—and thus make explicit—their own tacit criteria for "what counts as appropriate here." This externalization serves a dual function. First, it provides ChatGPT with richer, more constraining input, reducing the likelihood of contextually misaligned outputs. Second, and more importantly for learning, it scaffolds learners' own pragmatic reasoning. By articulating contextual parameters before generation, learners activate sociopragmatic awareness—they bind communicative intent to linguistic form from the outset (Spencer-Oatey, 2008). This pre-generation reasoning heightens attention to pragmatic features during subsequent review and revision cycles, fostering what Ishihara and Cohen (2010) term "metapragmatic awareness": understanding not merely what to say, but why specific forms are appropriate in context. Procedural prompts, by contrast, defer this reasoning to post-generation review. Learners configure parameters reactively—"the script is too short, so extend it"—treating revision as local problem-solving rather than principled pragmatic reasoning. Without upfront externalization of evaluative criteria, post-hoc revisions tend to remain at the surface level (word substitution, structural adjustment) rather than addressing deeper contextual misalignment. This explains why tool-oriented learners, despite multiple revision attempts, continued to encounter "phrases that don't sound native" and required more extensive manual intervention.

These findings suggest a pedagogical implication: role perception and prompt design are teachable dimensions. Rather than treating prompting as a spontaneous skill, instructors can explicitly guide learners toward contextual-first strategies—teaching them to interrogate context (audience, purpose, cultural norms) before issuing generation requests. However, the two strategies need not be antagonistic. A balanced, staged approach sequences them: contextual prompts first establish relationships, setting, and cultural parameters; procedural prompts then fine-tune structural elements (length, pacing, grammar). This sequencing mirrors the learning trajectory from "external scaffolding to internalized application" (Lantolf & Thorne, 2006), progressively transferring regulatory control from the tool back to the learner. Cognitively, this sequence trains the cycle "contextual reasoning → strategy selection → linguistic realization," equipping learners with a metacognitive framework transferable to real intercultural communication scenarios beyond the workshop setting.

5.2 Explaining contradiction pairs and mediation dynamics

Section 5.1 traced how role perception shapes prompting strategies and script quality outcomes. Yet between strategy deployment and final script quality lies a layer of systematic misalignments within the activity system. As introduced in Section 3.3, these misalignments manifest as contradictions among components (tools, rules, community, division of labor, object). Rather than isolated obstacles, they constitute an interconnected resolution sequence: each mediation enables progression toward transforming AI-generated drafts from "structurally complete" to "culturally appropriate and performable." Analysis of contradictions revealed four contradiction pairs in descending order of intensity: Rules–Community ($SI = 0.762$), Tools–Object ($SI = 0.360$), Community–Division ($SI = 0.352$), and Tools–Rules ($SI = 0.313$), each representing a distinct stage of script refinement.

The Rules–Community pair, exhibiting the highest tension, operates through a mechanism of norm specification. Learners encountered outputs containing fabricated or misappropriated cultural references—for instance, "ChatGPT suggested the parents give a 'traditional wooden spoon,' but I'm not sure if such a tradition exists. Instead, instructors suggested adding a scene with ice cream, pregnancy, and cold food taboos" (Learner 2, Self-reflection). The underlying contradiction is that ChatGPT generates culturally plausible-sounding content without encoding genuine cultural norms, leaving learners unable to distinguish authentic from spurious references. Resolution occurs when the community—instructors, peers, native speakers—articulates explicit cultural and pragmatic rules ("what counts as appropriate here and which cultural cues should appear"), transforming tacit knowledge into usable specifications.

This process exemplifies 5.2 mediation mechanism: contradictions surface as rule deviations; community actors clarify rules; learners internalize standards and apply them in revision. Notably, the pathway is not iterative generation, but rather single-round community consultation followed by targeted revision—reflecting that cultural appropriateness cannot be resolved through trial-and-error prompting alone, but requires authoritative cultural-pragmatic input.

The Tools–Object pair reveals a distinct contradiction: misalignment between stated specifications and model outputs. Learners explicitly specified requirements—"five-minute script length"—yet received outputs of only two to three minutes. Attempts to remedy through procedural prompts ("extend the script") triggered repetition rather than coherent elaboration. This contradiction arises from the tool's limited capacity to operationalize temporal or structural constraints consistently. Importantly, this is not a learner-comprehension problem but an architectural limitation of the generative model. Resolution requires translating abstract goals ("five minutes") into quantifiable tool-executable constraints (word count, sentence count) and stating them explicitly and repeatedly in prompts. In Engeström's framework, the tool mediates between learner intent (object) and output (result); when the tool cannot reliably execute the mediation, learners must either decompose goals into finer granularities or accept bounded deviation. Functionally, ChatGPT operates as a "deviation probe": by failing to meet specifications, it exposes under-specified requirements and prompts learners to refine their own planning granularity, progressively narrowing the gap between intention and execution.

The Community–Division pair centers on organizing collaborative labor within the community. Evidence indicates that effective revision occurred when instructors assumed explicit leadership roles—"We went line by line through dialogue; the instructor acted as leader, giving us feedback on whether ChatGPT-generated sentences matched personality and character" (Learner 5, Interview)—and when task division was clarified: "Native speakers guided us; everyone showed up on time" (Learner 3, Interview). The contradiction is not whether collaboration helps, but how to operationalize it: vague aspirations ("let's revise together") fail unless decomposed into concrete role assignments and task sequences. Mediation occurs when instructors and community members establish clear divisions of labor—who provides cultural input, who handles linguistic refinement, who executes final proofreading. This structured collaboration translates abstract collaborative intent into executable processes, reducing coordination overhead and ensuring each community member's expertise is deployed where needed. This mechanism aligns with Engeström's (2001) emphasis on the necessity of structured mediation in resolving activity system contradictions, and reflects principles outlined by Miles, Huberman, and Saldaña (2014) regarding the role of clear role definition in achieving collaborative efficiency.

The Tools–Rules pair addresses a specific pragmatic mismatch: AI-generated Chinese frequently deviates from colloquial and native norms, manifesting as literal translation, overly formal diction, or grammatical infelicities. As one learner noted, "ChatGPT's Chinese ability is significantly weaker than English; we first generated in English, then translated to Chinese, resulting in grammatically incorrect sentences" (Learner 2, Self-reflection). Another observed, "The first script was unnatural; words were too formal; we needed to make it more down-to-earth" (Learner 7, Self-reflection). The contradiction arises because the tool, lacking sufficient corpus or training data for colloquial Chinese pragmatic norms, produces outputs that are structurally sound but pragmatically misaligned. This differs fundamentally from the Rules–Community contradiction: here, the rules are known (colloquiality, grammatical appropriateness), but the tool cannot reliably produce them. Resolution requires human intervention—native speaker review, colloquial editing, pragmatic proofreading—to calibrate outputs against established usage norms. This aligns with Ishihara and Cohen's (2010) observation that pragmatic appropriateness depends critically on contextual specification and authentic model exposure, and with Heift and Schulze's (2007) finding that computer-assisted language learning requires corrective feedback to prevent drift from target norms. In this case, ChatGPT functions not as a solution but as a preliminary generator whose outputs demand human-expert correction—a necessary step in ensuring that final scripts embody authentic pragmatic appropriateness.

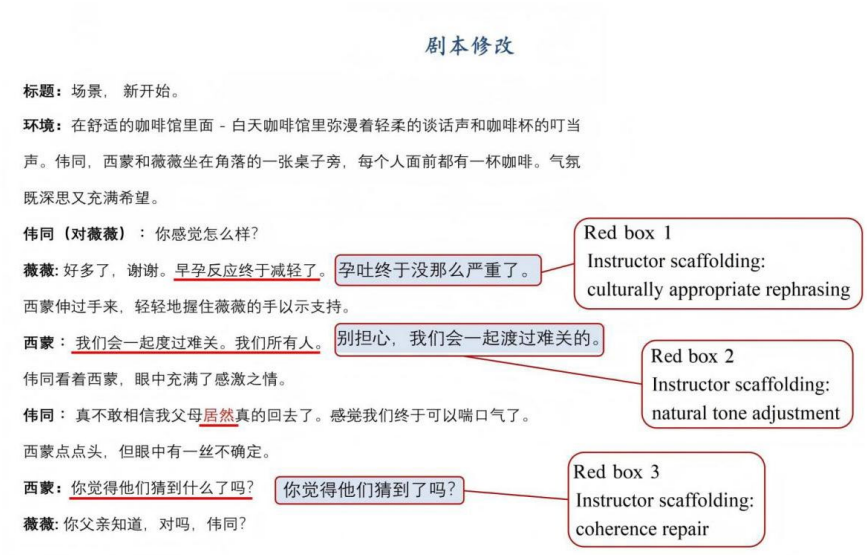
Sequencing the four pairs reveals a coherent resolution trajectory: Rules–Community establishes cultural-pragmatic baseline ("what appropriateness means"); Tools–Object operationalizes this baseline into executable specifications ("how to achieve it with the tool"); Community–Division structures the human effort ("who does what"); and Tools–Rules performs final pragmatic calibration ("ensuring authentic realization"). These are not independent problems but linked stages in a single progressive transformation from draft to performance-ready script. ChatGPT's role shifts across stages: initially, a "draft generator" accelerating idea externalization; subsequently, a "deviation exposer" highlighting areas of cultural, structural, or pragmatic misalignment; finally, a resource requiring human expert correction. This progression reflects Engeström's (2001) model of contradiction resolution through progressive mediation and Vygotsky's (1978) principle of internalization—learners begin with external support (community guidance, tool scaffolding) and progressively internalize standards as they diagnose deviations, make corrections, and refine their understanding of what constitutes appropriate intercultural communication. The learning trajectory is not linear task completion but iterative calibration: each contradiction mediation deepens learners' pragmatic awareness and refines their sense of context-appropriate language use.

5.3 Instructor mediation as a scaffolded performance pathway

Analysis of resolution pathways revealed that instructor-mediated episodes resolved the highest proportion of contradictions (81.3%) and most directly facilitated *ICC* improvement. This differential effectiveness demands explanation beyond resolution rates; it requires understanding the underlying cognitive and sociocultural mechanisms by which instructor mediation transforms abstract cultural norms into performable, contextualized language.

Figure 3

Script Revision with Instructor Annotations



The superiority of instructor mediation rests on a fundamental sociocultural principle: the *Zone of Proximal Development* (*ZPD*) and guided scaffolding. When instructors directly model contextually appropriate expressions and insert them into specific lines aligned with character and setting, they operationalize abstract rules ("what counts as appropriate") into concrete, task-embedded language resources. For example, a learner recalled: "At a dinner table with two arguing people, I said ' 我们快吃饭吧 [Shall we eat now]' to ease the conflict. That phrase was something instructors and native speakers

modeled for us" (Learner 3, Interview). This process exemplifies Vygotsky's (1978) principle that learners appropriate culturally valued forms through guided interaction within the *ZPD*; the instructor does not merely state a rule ("be more appropriate"), but demonstrates a situated use case and embeds the expression in the learner's working text. This externalizes the instructo's tacit pragmatic knowledge into a reusable template, allowing learners to internalize not just the form, but the situational context that justifies its use. Research on explicit corrective feedback and metapragmatic awareness confirms this mechanism: explicit modeling plus contextual embedding produces more stable, transferable pragmatic gains than abstract rule explanation alone (Ishihara & Cohen, 2010; Taguchi, 2011). Figure 3 illustrates this principle through a concrete example of instructor annotations that combine explicit modeling with cultural context.

In contrast, peer collaboration and self-correction operate at different levels of constraint and feedback immediacy. Peer collaboration (66.7% resolution rate) primarily improves delivery naturalness through micro-level, just-in-time correction cycles: peers flag unnatural spots, the group revises line-by-line, and then practices—"they marked where to pause, read faster, and helped teammates present lines naturally" (Learner 7, Self-reflection). This repetition and social rehearsal enhances fluency and coherence, but leaves cultural-pragmatic decisions underexamined; without authoritative cultural or pragmatic input, peers rely on intuition rather than principle. Self-correction (33.3% resolution rate), lacking external reference and authoritative feedback, typically remains at word-level substitution without achieving paragraph-level, contextually coherent rewrites. One learner exemplified this limitation: "Without asking others, I can't tell what's wrong; I would have asked instructors or teammates" (Learner 3, Interview). The critical difference is that instructor mediation provides both clarity of standards (what constitutes appropriate) and immediate embedding (modeling the solution in situ), whereas peer collaboration provides rehearsal but not authoritative calibration, and self-correction provides neither. From a corrective feedback perspective (Lyster & Ranta, 1997; Li, 2010), explicit external feedback with clear targets produces more stable, transferable learning gains than peer-negotiated or self-initiated corrections—particularly in culturally and pragmatically sensitive domains where learners lack confidence or cultural reference points.

It is important to note that learners' linguistic histories, cultural backgrounds, and identity trajectories shape starting points and the trajectory of gains: pathways to improvement in ICC and SS are not equally accessible for all learners, and learners' existing cultural resources and sense of identity condition how rules are understood, adopted, and internalized (Norton, 2000). Instructor mediation scaffolds this process, but cannot wholly override individual variation in prior knowledge, confidence, and cultural positionality.

Instructor mediation operates through a three-stage progression, each addressing a distinct outcome: Stage 1 (Script) ensures structural and plot coherence—instructors flag inconsistencies, operationalize "stage-readiness" into concrete revisions, and align character lines with scene logic. Stage 2 (*ICC*) embeds cultural and pragmatic appropriateness at the sentence and plot level—instructors identify culturally misaligned cues, model rewordings that fit character and context, and transform abstract norms into usable expressions. For instance, after identifying ChatGPT's fabricated "traditional wooden spoon," the instructor guided a rewrite centered on "pregnancy and cold food," anchoring cultural appropriateness in a performable plot element (Learner 2, Self-reflection). Stage 3 (*SS*) refines vocal realization—instructors mark pause points, model intonation patterns, and guide learners to associate linguistic form with paralinguistic features. One learner reported: "I learned to use stress and a slower rhythm to enhance pauses between phrases" (Learner 5, Self-reflection); another noted: "learned where to emphasize certain words in Chinese acting" (Learner 6, Self-reflection).

This three-stage progression illustrates how instructor mediation transforms the role of ChatGPT. Initially, ChatGPT functions as a "cultural-conflict generator"—it accelerates drafting while simultaneously exposing register, cultural, and pragmatic misalignments that human revision cannot

readily overlook. Instructor mediation then systematically repairs these misalignments layer by layer. The cumulative effect is "layered alignment": outputs progress from structurally complete to culturally appropriate to naturally performable. This multi-stage calibration is neither achievable through unrevisited ChatGPT output nor through self-correction alone; it requires instructor expertise to diagnose where deviations occur, model principled solutions, and embed them in learners' working texts. Consequently, instructor-mediated revision pathways are not only high in resolution success and direct *ICC* improvement, but constitute a replicable, systematic workflow—from text coherence to cultural appropriateness to performance fluency—that operationalizes the learning objective of intercultural communication competence.

6 Conclusion

This study examined how Chinese language learners integrate ChatGPT into cultural scriptwriting, analyzed through Activity Theory. Learners' perceptions of ChatGPT's role distinctly shaped their prompting strategies. The most effective pattern combined contextual prompts first—anchoring situation and cultural depth—followed by procedural prompts for form and structure refinement. ChatGPT thus served dual functions: draft generator and deviation exposer, accelerating production while revealing cultural, register, or detail misalignments.

Engeström's contradiction framework proved crucial in explaining learning dynamics. Four contradictions shaped outcomes: Rules–Community, Tools–Object, Community–Division, and Tools–Rules. The Rules–Community contradiction was most decisive. When learners grappled with cultural appropriateness norms, instructor mediation—clarifying rules and providing usable expressions—resolved tensions and catalyzed gains in intercultural communicative competence and speaking skills. This reveals that learning with generative AI is fundamentally social, not individual; contradictions within the activity system drive expansive learning.

Pedagogically, instructors should deliberately design scriptwriting as a community activity mediating the Rules–Community link. Concretely: make role perception explicit, model context-first prompting, convert abstract cultural rules into promptable specifications, and establish frequent feedback touchpoints where misfits are reframed into usable expressions. This approach leverages ChatGPT's speed while anchoring outputs to cultural expectations, producing performable, fluent, culturally appropriate lines.

The small sample limits generalizability, and long-term durability of gains remains uncertain. Future research should examine how learners' diverse language backgrounds shape ChatGPT use and adopt standardized evaluation rubrics for consistent script and video assessment.

Acknowledgements

We would like to express our heartfelt thanks to all the people who have ever helped us with this paper.

Statements on Open Data, Ethics and Conflict of Interest

The participants were protected by hiding their personal information during the research process. They knew that their participation was voluntary and they could withdraw from the study at any time. AI tools were used during the writing process to polish some English expressions. There is no potential conflict of interest in this study.

Appendix A

From learners' perspective

Guidelines for Self-Reflection on Utilizing ChatGPT to Create the Cultural Context Script

1. Overall, are you satisfied with the first script generated by ChatGPT? In which aspects do you think it is effective, and in which aspects do you think it might not be efficient?
2. Did you encounter any difficulties while using ChatGPT to generate your own script? If so, how did you solve those difficulties? If not, what was the reason?
3. What role did ChatGPT play while completing this task? Did it act as a tutor, a tool, or something else?
4. Can you explain how you used ChatGPT to establish the cultural context script? For example, what prompts did you give it, and what interactions did you have with ChatGPT, your peers, or your instructors?
5. Have you applied the speaking skills learned from the first section in the roles you played? Can you give us some concrete examples?
6. After completing this task, do you think you have enhanced your awareness of intercultural communication and can apply it in reality?

From instructors' perspective

Guidelines for Self-Reflection on Utilizing ChatGPT to Create the Cultural Context Script

1. Overall, are you satisfied with the learners' cultural context script generated by ChatGPT? In which aspects do you think ChatGPT was effective, and in which aspects do you think it was inefficient?
2. Did you encounter any difficulties in helping the learners revise their cultural context script generated by ChatGPT? Which main areas did you explore in more depth, and why did you do so?
3. What is your perception of the role of ChatGPT in this learning task? Did it act as a tutor, a tool, or something else?
4. Do you think the learners appropriately applied the speaking skills in the recorded video? Please provide examples.
5. Do you think the learners have a deep understanding of intercultural communication as a result of completing this learning task? Please provide examples.

Appendix B

Based on the AT framework, eight questions are posed as interview guidelines for learners.

Subjects

1. As a Chinese language learner, could you share some of your previous experience of using ChatGPT? When, why and how do you usually use ChatGPT? How about the purpose or motivation for using ChatGPT in this workshop?

Object

2. Are you clear about the goal for conducting this workshop series for the target learners? E.g. the speaking skills you should know and apply them appropriately in your final video recording. The application of ChatGPT in the creation of scripts within the cultural context.

Outcome

3. By the end of the workshop series, do you think learners successfully complete the task and achieve the goal/object for running this workshop? If yes, why; if no, why. Any follow-up steps need to be further explored in the future?

Community

4. Within the teaching and learning community, among all participants and instructors in this workshop series, or even the iLang team, CLE (Center for Language Education), other units/ divisions of HKUST (e.g. CIE, Center for Innovation Education), what do you think their attitudes towards GenAI like ChatGPT? Are these factors influence your thoughts or behaviors in the application of ChatGPT in your future work or professional development (e.g. organizing activities/events, pursuing your academic or employment career, etc.)

Tools (mediating artifacts)

5. For the task in the creation of a script within the cultural context, as a learner, what do you think ChatGPT's real role is? For example, an active assistant, a useful tool? Can they fully understand the characters' roles in the story you created and your utterances in the script for the corresponding roles in the video recording? In your opinion, what do you think about ChatGPT's role in this learning task? Can they give appropriate prompts?

Rules and regulations

6. This is an informal learning event, given in this context, what do you think your performance over a two-week long period? E.g. your attendance and engagement in the classroom, your time management, your groups' cooperation, etc. Are you sure that each participant fully understood the guidance of this task? E.g. The cultural context, their performance in playing different roles.

Division of labor

7. During this workshop series, your main responsibility is to improve your speaking skills and intercultural competence through "performing the roles (refers to voicing the characters through rhythm, stress, pauses, etc.)" in the script your team created with ChatGPT's support.

Others

8. In general, what do you think about the emergence of Gen AI, especially ChatGPT, in assisting learners to facilitate language learning? And any other aspects you think are very important or need to be addressed with more attention by a language learner? (positive or negative, such as plagiarism, ethical issues, etc.)

Appendix C

Ranked Co-occurrence of AT Element Pairs by Strength Index

Rank	Contradiction Pair	Co-occurrence (*n*)	Strength Index (*SI*)	Top contributing code(s)
1	Rules–Community	16	0.762	Rule_Interpret (21); C_Mentoring (67)
2	Tools–Object	18	0.360	Tool_Constraint (50) ; Obj_Deviation (58)
3	Community–Division	24	0.358	C_Mentoring (67); Div_Regulator (69)
4	Tools–Rules	5	0.313	Tool_Constraint (50); Rule_Deviation (16)
5	Subject–Tools	9	0.257	Sub_Lrn_ToolUsage (35); Tool_Constraint (50)
6	Tools–Division	11	0.220	Tool_Constraint (50); Div_Regulator (69)
7	Subject–Community	10	0.217	Sub_Lrn_ConstraintResponse (46); C_Mentoring (67)

Rank	Contradiction Pair	Co-occurrence (*n*)	Strength Index (*SI*)	Top contributing code(s)
8	Subject–Rules	3	0.176	Sub_Lrn_CulturalRule (22); C_Mentoring (67)
9	Subject–Division	6	0.171	Sub_Lrn_ToolUsage (35); Div_Regulator (69)
10	Rules–Division	2	0.125	Rule_Deviation (16); Div_Regulator (69)
11	Division–Object	6	0.104	Div_Regulator (69); Obj_Deviation (58)
12	Subject–Object	2	0.091	Sub_Lrn_CulturalRule (22); Obj_Deviation (58)
13	Tools–Community	3	0.060	Tool_Constraint (50); C_Mentoring (67)
14	Community–Object	1	0.030	C_Mentoring (67); Obj_Fit (33)
15	Rules–Object	0	0.000	Rule_Interpret (21); Obj_Fit (33)

Notes. *SI* (Strength Index) = co-occurrence / min (frequency of code 1, frequency of code 2). Pairs are ranked by *SI*; ties are broken by co-occurrence (*n*), then alphabetically. Values in bold meet the reporting threshold ($SI \geq 0.30$). Top contributing code(s) list, in the same left–right order as the Element Pair, the highest-share code for each element as CodeA (f1); CodeB (f2). $N = 992$.

References

Barab, S. A., Evans, M. A., & Baek, E.-O. (2004). Activity theory as a lens for characterizing the participatory unit. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 199–214). Lawrence Erlbaum Associates.

Cai, W. (2023). ChatGPT 环境下的汉语学习与教学 [Chinese learning and teaching in the ChatGPT environment]. 语言教学与研究 [Language Teaching and Linguistic Studies], 2023(4), 13–23.

Carvalho, M. B., Bellotti, F., Berta, R., De Gloria, A., Sedano, C. I., Hauge, J. B., Hu, J., & Rauterberg, M. (2015). An activity theory-based model for serious games analysis and conceptual design. *Computers & Education*, *87*, 166–181. https://doi.org/10.1016/j.compedu.2015.03.023

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Collis, B., & Margaryan, A. (2004). Applying activity theory to computer-supported collaborative learning and work-based activities in corporate settings. *Educational Technology Research & Development*, *52*(4), 38–52.

Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Sage.

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Ablex.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Orienta-Konsultit.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, *14*(1), 133–156. https://doi.org/10.1080/13639080020028747

Engeström, Y. (2015). *Learning by expanding: An activity-theoretical approach to developmental research* (2nd ed.). Cambridge University Press.

Gao, A. (2025). ChatGPT 辅助汉语二语写作反馈研究 [ChatGPT-assisted written corrective feedback in teaching Chinese as a second language]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 6(2), 30–43.

Heift, T., & Schulze, M. (2007). *Errors and intelligence in computer-assisted language learning: Parsers and pedagogues*. Routledge.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, *38*(1), 237-257.

Ishihara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Pearson.

Jiang, L. (2007). 关于构建“以培养交际能力为目标”的对外汉语教学框架的思考 [Reflections on Constructing a Framework for Teaching Chinese as a Foreign Language Aimed at Cultivating Communicative Competence]. 汉语学习 [Chinese Language Learning], (1), 77-82.

Jiang, X. L., Li, J., & Chen, C. H. (2023). Enhancing critical thinking skills with ChatGPT-powered activities in Chinese language classrooms. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 5(1), 47–73.

Kabooha, R. (2016). Using movies in EFL classrooms: A study conducted at the English Language Institute (ELI) King Abdul-Aziz University. *English Language Teaching*, 9(3), 248–257.

Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design*. MIT Press.

Kramsch, C. (1993). Context and culture in language teaching. London: Oxford University Press.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). Sage.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lazarou, D. (2011). Using cultural -historical activity theory to design and evaluate an educational game in science education. *Journal of Computer Assisted Learning, 27*(5), 424-439.

Li, Q. C., Zhang, J. F., & Cai, W. (2024). Utilizing ChatGPT to implement differentiated instruction. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 5(1), 74–89.

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, *60*(2), 309–365. https://doi.org/10.1111/j.1467-9922.2010.00561.x

Li, Y., Han, J., & Wu, X. (2025). 汉语学习者对 ChatGPT 参与写作反馈的压力评价与应对研究 [Chinese learners’ stress appraisal and coping strategies in response to ChatGPT-involved writing feedback]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 6(3), 67–75.

Liang, X., & Luo, J. (2024). Integration of ChatGPT into project-based learning: A course design framework. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 5(1), 29–46.

Liu, L., Shi, Z., Da, J., Tian, Y., Liang, X., … & Hu, X. (2023). ChatGPT 给国际中文教育带来的机遇与挑战 : 北京语言大学与美国中文教师学会联合论坛专家观点汇辑 [The opportunities and challenges of ChatGPT for international Chinese language education: A view summary of the joint forum of Beijing Language and Culture University and the Chinese Language Teachers Association of America]. 世界汉语教学 [Chinese Teaching in the World], 2023(4), 291–315.

Liu, S. (2019). Using science fiction films to advance critical literacies for EFL students in China. *International Journal of Education & Literacy Studies*, 7(3), 1–9.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*(1), 37–66.

Lyu, Y. (2008). 课堂教学中的教师引入内容分析 [Analysis of Teacher-Introduced Content in Classroom Teaching]. 世界汉语教学 [Chinese Teaching in the World], (1), 108-114.

Meng, K. (2024). ChatGPT 在国际中文词汇教学中的表现与应用策略 [The performance and application strategies of ChatGPT in CSL vocabulary teaching]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 5(1), 3–14.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook (3rd ed.)*. Sage.

Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change*. Longman.

Rambe, P. (2012). Activity theory and technology mediated interaction: Cognitive scaffolding using question-based consultation on Facebook. *Australasian Journal of Educational Technology*, *28*(8), 1333–1361. https://doi.org/10.14742/ajet.775

Saldaña, J. (2021). T*he coding manual for qualitative researchers* (4th ed.). Sage.

Shi, Z., & Lyu, B. (2025). AI 在中文教学中的实证研究：探索与展望 [Empirical research on AI in Chinese language teaching: Exploration and prospects]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 6(3), 1–7.

Song, F., Tan, Y., & Pu, Y. (2024). 生成式人工智能辅助编写国际中文教育少儿读物的应用与实践 —— 以 ChatGPT 与 Midjourney 为例 [Application and practice of generative artificial intelligence for children’s reading materials in international Chinese language education: ChatGPT and Midjourney as examples]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 5(1), 15–28.

Spencer-Oatey, H. (2008). *Culturally speaking: Culture, communication and politeness theory* (2nd ed.). Continuum.

Taguchi, N. (2011). Teaching pragmatics: Trends and issues. *Annual Review of Applied Linguistics*, *31*, 289–310. https://doi.org/10.1017/S0267190511000018

Uden, L. (2007). Activity theory for designing mobile learning. *International Journal of Mobile Learning and Organisation*, *1*(1), 81–102.

Vijaymeena, M. K., & Kavitha, K. (2016). A survey on similarity measures in text mining. *Machine Learning and Applications: An International Journal*, *3*(1), 19–28.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Vyushkina, E. (2016). Legal English through movies: Development of professional communicative competence. *Studies in Logic, Grammar and Rhetoric, 45*(58), 253–263.

Walker, G. (2010). Performed Culture: Learning to participate in another culture. In Walker, G. (Ed.), *The Pedagogy of Performing a Culture,* Columbus, OH: National East Asian Languages Resource Center at Ohio State University.

Wang, J., & Tian, X. (2020). *Lens on China: Intermediate and advanced readings on film for learning Chinese*. Hong Kong University Press.

Xue, J., & Pan, Q. (2012). The effects of film appreciation on improving the students’ intercultural communication competence. *Theory and Practice in Language Studies*, 2(8), 1741–1745.

Yu, L. (2021). The performed culture approach. In Z. Ye (Ed.), *The Palgrave handbook of Chinese language studies* (pp. 843–876). https://doi.org/10.1007/978-981-13-6844-8_7-1

Yu, X., Wang, X., & Qi, F. (2025). 基于 ChatGPT 的国际中文句式语法教学实证探析 [An empirical analysis of sentence grammar teaching based on ChatGPT]. 国际汉语教学学报 [International Journal of Chinese Language Teaching], 6(3), 6–24.

Yue, J. (2019). Use of foreign films in cultivating intercultural communicative competence in ELT—A case study. *Theory and Practice in Language Studies, 9*(2), 198–203.

Zhang, Y. (2020). 体演教学法对语境的认识及在非汉语环境下对语境的设计 [Understanding context in the performed-cultural approach and designing context in a non-Chinese language environment]. 国际汉语教育（中英文）[International Chinese Language Education (Chinese and English)], 2020(2), 60–68.

Zhao, J. (2008). 汉语作为第二语言教学：理念与模式 [Teaching Chinese as a second language: Concepts and models]. 世界汉语教学 [Chinese Teaching in the World], 22(1), 93–107.

Zoreda, M. (2005). Teaching film, culture, and language: An advanced English course in Mexico. I*nterdisciplinary Humanities*, 22(2), 61–68.

Zurita, G., & Nussbaum, M. (2007). A conceptual framework based on activity theory for mobile CSCL. *British Journal of Educational Technology*, *38*(2), 211–235.

Dr. Emma Rao, a lecturer at the Center for Language Education, School of Humanities and Social Science, the Hong Kong University of Science and Technology, has conducted research in the areas of Putonghua learning and cultural identity, international Chinese language curriculum design and the professional development of international Chinese language teachers.

Miss Ruohan Zhao, an MA graduate of the School of Humanities and Social Science at The Hong Kong University of Science and Technology. Her research interests include the use of emerging technologies in teaching Chinese as a second language and the professional development of international Chinese language teachers.

国际汉语教学学报（2025）
260603, 1-28 https://doi.org/10.46451/ijclt.260603

ChatGPT 在构建中文文化语境剧本中的应用——活动理论的视角

饶宇靖
赵若含
香港科技大学，中国

摘要
本研究以社会文化视角为基础，采用活动理论框架，考察了 ChatGPT 在 Lens on China 系列课外工作坊这一结构化项目中的应用，探究中文学习者如何借助该工具构建文化语境剧本。为此，本研究提出两个研究问题：（1）学习者对 ChatGPT 角色的认知如何影响其在文化语境中的剧本创作方式？（2）哪些因素影响了学习者使用 ChatGPT 创作文化剧本的过程？研究结果表明，学习者对 ChatGPT 角色的认知存在差异，并对其提示策略产生影响：将 ChatGPT 视为“助手”的学习者倾向于使用情境性提示，而将其视为“工具”的学习者则倾向于使用程序性提示。情境性提示为学习者提供了文化相关的适切内容，而程序性提示则提供了必要的结构性要素。先使用情境性提示、再结合程序性提示的顺序组合策略，是创作文化内容准确、结构完整的剧本的最有效方式。本研究进一步从活动理论视角识别出文化语境剧本创作中的矛盾对，其中，规则与社区之间的矛盾被认为是理解任务表现与学习成果的核心。当 ChatGPT 生成看似合理却缺乏真实性的文化内容时，这一张力便随之产生，使学习者对真实的文化规范感到困惑。当该矛盾由教师通过明确阐释文化规则与语用规范加以有效调解后，抽象的文化知识得以转化为可表达的语言形式，显著提升了学习者的跨文化交际能力与口语表达能力。因此，教师的角色必须从资源提供者转变为矛盾的调解者，善用人工智能逻辑与语言文化之间的落差，以深化学习者的跨文化能力。

关键词
ChatGPT，活动理论，文化语境剧本

饶宇靖，香港科技大学人文社会科学学院语文教育中心讲师，研究领域为普通话学习与文化认同、国际汉语课程设计、国际汉语教师专业发展等。

赵若含，香港科技大学人文社会科学学院国际中文教育硕士，研究兴趣为新兴科技与国际汉语教学、国际汉语教师专业发展等。