Peta International Journal of Social Science and Humanity

Peta Research Publication

www.PIJSSh.peta-research.com

Online ISSN 2830-540X

Enhancing Chinese Oral Fluency through Multilingual AI-Assisted Scenario-Based Learning among University Students in Uzbekistan

Ye Wenbing

Tashkent State Transport University, Uzbekistan. yewenbing9@gmail.com

Abstract

The development of Chinese oral fluency remains a significant challenge for university students learning Chinese as a foreign language, particularly in multilingual environments where opportunities for authentic communicative interaction are limited. Although artificial intelligence (AI) offers new possibilities for interactive language learning, empirical evidence regarding its integration with multilingual pedagogical support and scenario-based communication remains limited, especially in Central Asian higher education. This study investigated multilingual AI-assisted scenario-based learning for developing Chinese oral fluency among university students in Uzbekistan. A three-group quasi-experimental mixed-methods design involving 120 university students assigned to existing instructional classes was implemented over a 16-week intervention. The experimental group received multilingual AI-assisted scenario-based instruction, while two comparison groups received conventional scenario-based instruction and traditional Chinese speaking instruction, respectively. The investigation focused on temporal oral fluency indicators, including speech rate, articulation rate, mean length of fluent runs, and pause-related measures, alongside pronunciation, lexical tone accuracy, and interactional competence. Foreign language anxiety and willingness to communicate were also examined as learner-related variables. Semi-structured interviews explored students' experiences of AI-mediated interaction, scenario-based speaking practice, and multilingual pedagogical support. The study integrates quantitative and qualitative perspectives to examine the relationship between the instructional approach and Chinese oral fluency development. It contributes to research on technology-enhanced Chinese language education by investigating the integration of AI-mediated interaction, multilingual pedagogical resources, and scenario-based communication in a multilingual Central Asian higher education context.

Keywords: Chinese as a Foreign Language; Oral Fluency; Artificial Intelligence-Assisted Language Learning; Scenario-Based Learning; Multilingual Education; Willingness to Communicate; Uzbekistan.

Received on 23 Aug. 2026, Accepted on 19 Oct. 2026, Published on 2 Nov. 2026.

1. Introduction

The internationalization of higher education, expanding economic cooperation, and increasing cross-cultural communication have strengthened the importance of foreign language competence in contemporary university education. Among the languages attracting growing international interest, Chinese has acquired particular significance because of China's expanding economic, educational, technological, and diplomatic engagement with countries across different regions. In Uzbekistan, the development of Chinese language education is relevant to academic cooperation, international employment opportunities, and professional communication. Universities are consequently expected to develop students' ability not only to understand Chinese linguistic structures but also to communicate effectively in authentic academic and professional situations. Despite the growing importance of Chinese language education, developing oral communication competence remains challenging for learners of Chinese as a foreign language (CFL). Mandarin acquisition requires learners to coordinate vocabulary retrieval, grammatical formulation, pronunciation, and lexical tone production during real-time communication. Students may acquire substantial knowledge of vocabulary and grammar without being able to retrieve expressions efficiently, maintain conversational continuity, or respond appropriately to unexpected questions. These difficulties may be particularly pronounced in non-philological higher education institutions, where opportunities for sustained interaction with proficient Chinese speakers can be limited. Research in the Uzbek higher education context has similarly identified difficulties in establishing natural Chinese-speaking environments and providing sufficient communicative practice (Ruzumova, 2024). Oral fluency is therefore a central dimension of Chinese speaking development. It concerns the efficient and continuous production of connected speech while maintaining comprehensibility and communicative appropriateness. However, fluency should not be equated with speaking speed alone. It can be examined through observable temporal characteristics, including speech rate, articulation rate, pause frequency and duration, and mean length of fluent runs. Segalowitz (2010) distinguishes cognitive fluency, perceived fluency, and utterance fluency, demonstrating why speaking development should be examined through both underlying processing efficiency and observable speech production. Skehan (1998) further conceptualizes fluency, accuracy, and complexity as interconnected dimensions of second-language performance. These perspectives indicate that the evaluation of Chinese oral fluency requires systematic speaking assessment rather than reliance exclusively on learners' self-reported confidence or perceived improvement. Recent developments in artificial intelligence (AI), particularly generative AI, conversational agents, and automatic speech recognition, have created new possibilities for technology-enhanced language education. Conversational AI can provide repeated speaking opportunities, interactive responses, contextually relevant questions, and supplementary practice beyond scheduled classroom activities. These affordances may be valuable for foreign language learners who have limited opportunities to communicate with proficient speakers outside the classroom. Systematic reviews have identified substantial growth in generative AI research in language education, although writing remains a dominant focus and speaking-related investigations are comparatively limited (Li et al., 2025; Wang et al., 2025). Empirical studies have also reported potential benefits of AI-supported language learning for speaking performance, communication confidence, and willingness to communicate (Qiao & Zhao, 2023; Zhang et al., 2024; Kim & Su, 2024; Sok & Shin, 2025). Nevertheless, findings from English and other foreign-language learning contexts cannot automatically be generalized to Mandarin because of its distinctive phonological and communicative demands.

The educational value of AI also depends on the pedagogical organization of interaction. Scenario-based learning provides opportunities for students to apply linguistic knowledge in

meaningful situations involving information exchange, problem-solving, decision-making, and interpersonal communication. Scenarios related to university life, travel, railway services, transportation, and logistics can connect Chinese speaking activities with students' academic and potential professional experiences. This approach is consistent with Task-Based Language Teaching (TBLT), which emphasizes purposeful communication through goal-oriented tasks (Ellis, 2003). When combined with conversational AI, scenario-based practice may enable learners to engage in repeated interactions, respond to changing communicative demands, negotiate meaning, and refine spoken expressions. However, whether AI provides additional educational value beyond conventional scenario-based instruction requires direct comparative investigation. The multilingual characteristics of Uzbekistan provide a further reason to examine this instructional approach. University students may possess diverse linguistic repertoires involving Uzbek, Russian, English, and other languages acquired through formal education and everyday communication. Calafato (2021) identified the relevance of multilingual experiences to language-learning motivation and educational backgrounds among university students in Uzbekistan and Kazakhstan. Research by Radjabzade and Makovskaya (2024) has also highlighted multilingual practices in Uzbek higher education. Existing linguistic knowledge may support vocabulary learning, grammatical comparison, and metalinguistic awareness, although cross-linguistic influence may also create challenges in pronunciation and language production. Multilingual resources should therefore be examined as potential pedagogical supports rather than assumed to produce positive learning outcomes automatically. Psychological factors also influence learners' participation in oral communication. Foreign language anxiety may discourage students from speaking because of concerns about pronunciation errors, grammatical mistakes, or negative evaluation. Willingness to communicate, in contrast, concerns learners' readiness to initiate communication in particular situations. Previous research has associated AI-mediated speaking activities with changes in anxiety, communication confidence, and willingness to communicate (Zhang et al., 2024; Kim & Su, 2024; Wang et al., 2024). However, psychological improvement should be distinguished from measurable speaking development. Examining both dimensions can provide a more comprehensive account of learners' responses to technology-supported instruction. Despite increasing interest in AI-assisted language education, four research gaps remain relevant. First, empirical research has concentrated heavily on English-language learning, leaving Mandarin oral fluency comparatively underexplored. Second, limited attention has been given to the relationship between multilingual learner backgrounds and AI-mediated Chinese speaking development in Central Asian higher education. Third, the additional contribution of integrating conversational AI with structured, profession-oriented scenario-based communication requires further investigation. Fourth, studies combining objective speaking assessments, psychological measures, delayed assessment, and qualitative evidence remain important for distinguishing observable language development from perceived learning benefits. To address these gaps, the present study investigated multilingual AI-assisted scenario-based learning among 120 university students at Tashkent State Transport University, Uzbekistan. A three-group quasi-experimental mixed-methods design was implemented over a 16-week intervention, using three existing instructional classes. The experimental group received multilingual AI-assisted scenario-based instruction, while the comparison groups received conventional scenario-based instruction and traditional Chinese speaking instruction. The investigation focused on temporal oral fluency, including speech rate, articulation rate, mean length of fluent runs, and pause-related characteristics, alongside pronunciation, lexical tone accuracy, grammatical accuracy, lexical complexity, and interactional competence. Foreign language speaking anxiety, willingness to communicate, and multilingual backgrounds were also considered. Delayed speaking assessment and an unfamiliar communicative scenario assessment were included to examine retention and transfer,

while semi-structured interviews explored students' experiences of the instructional approaches. The study has three intended contributions. Theoretically, it integrates Sociocultural Theory, the Interaction Hypothesis, the Output Hypothesis, and Task-Based Language Teaching to examine the relationships among AI-mediated interaction, multilingual pedagogical support, and Chinese speaking development. Methodologically, it combines speaking-performance indicators with psychological measures and qualitative evidence within a comparative mixed-methods design. Pedagogically, it investigates an integrated instructional approach in the multilingual and professionally oriented context of Uzbek higher education. By comparing three instructional conditions, the study also seeks to distinguish the potential contribution of the integrated AI-assisted approach from the contribution of conventional scenario-based and traditional speaking instruction.

2. Literature Review and Theoretical Framework

2.1. Chinese Oral Fluency and Second-Language Speaking Development

Oral fluency is a multidimensional construct in second-language acquisition (SLA), reflecting learners' ability to produce connected speech efficiently and continuously during communication. Segalowitz (2010) distinguishes cognitive fluency, perceived fluency, and utterance fluency. Cognitive fluency concerns the efficiency of underlying language-processing mechanisms; perceived fluency relates to listeners' judgments of speaking smoothness; and utterance fluency concerns observable characteristics of spoken production. This distinction is important because improvements in confidence or perceived speaking ability do not necessarily demonstrate measurable changes in speech production. For Mandarin learners, oral fluency development involves additional demands associated with lexical tone, pronunciation, vocabulary retrieval, and grammatical formulation. Mandarin is a tonal language in which pitch patterns contribute to lexical meaning. Learners must therefore coordinate segmental pronunciation, tonal production, and sentence construction while maintaining conversational continuity. These demands may increase processing pressure, particularly among students with limited opportunities for spontaneous Chinese communication. Skehan (1998) conceptualizes fluency, accuracy, and complexity as interconnected dimensions of second-language performance. This framework suggests that speaking development should not be evaluated through a single indicator. Speech rate, articulation rate, pause frequency and duration, and mean length of fluent runs provide complementary information about temporal fluency. Pronunciation, lexical tone accuracy, grammatical accuracy, and lexical complexity provide additional evidence concerning linguistic performance. Interactional competence further reflects learners' ability to respond appropriately, maintain conversational exchanges, and manage communicative situations. Wang, Wang, and Tao (2025) examined the developmental dynamics of complexity, accuracy, and fluency in oral production, demonstrating the importance of considering different dimensions of speaking development. Although their research concerned Chinese learners of English, its methodological implications are relevant to the assessment of oral performance more broadly. Chen and Gao (2025) also highlighted the importance of distinguishing expressive fluency from underlying processing indicators in Chinese-language assessment. Building on these perspectives, the present study treats Chinese oral fluency as its principal speaking outcome, emphasizing observable temporal indicators while considering pronunciation, grammatical accuracy, lexical complexity, and interactional competence as complementary outcomes. This approach enables a more comprehensive examination of speaking development without treating speed as a sufficient measure of fluency.

2.2. AI-Assisted Language Learning and Conversational Interaction

Artificial intelligence has become an increasingly important component of technology-enhanced language education through generative AI, conversational agents, automatic speech recognition, and adaptive feedback systems. These technologies can provide opportunities for repeated communication, individualized responses, and supplementary speaking practice beyond conventional classroom activities. Li et al. (2025), in a systematic review of 144 empirical studies published during 2023 and 2024, identified substantial growth in generative AI research in language education. However, English-language learning and writing remained dominant research areas, while speaking received comparatively less attention. Wang et al. (2025) similarly identified limitations concerning language representation, theoretical integration, and methodological detail in emerging generative AI research. Together, these reviews indicate that the expansion of AI-assisted language education has not yet resolved questions about its contribution to observable speaking development. Empirical investigations have nevertheless reported promising findings. Qiao and Zhao (2023) investigated AI-supported language learning among Chinese university students studying English and reported improvements in speaking performance and self-regulation. Zhang et al. (2024) examined AI-assisted speaking activities among Chinese university students and identified changes in foreign language enjoyment, speaking anxiety, and willingness to communicate. Kim and Su (2024) reported improvements in communication confidence and willingness to communicate following chatbot-supported Korean language activities. More directly relevant to Mandarin education, Liu (2026) investigated voice-based ChatGPT for pronunciation development among international university students learning Chinese. The study reported improvements in comprehensibility and lexical tone production, although the extent of improvement differed across pronunciation dimensions. This work provides a relevant basis for examining voice-based AI in Chinese language education, while also highlighting the need to distinguish pronunciation-focused practice from spontaneous communicative performance. Despite these findings, three limitations remain. First, evidence from English-language learning cannot automatically be generalized to Mandarin because of differences in phonological and linguistic demands. Second, successful interaction with an AI system does not necessarily establish the ability to communicate spontaneously with human interlocutors. Third, automated feedback may contain inaccuracies, particularly in pronunciation recognition and lexical tone evaluation.Accordingly, conversational AI in the present study is conceptualized as a supplementary instructional resource rather than a replacement for teachers or authentic human communication. Its potential contribution is examined through integration with structured speaking tasks and multiple speaking-performance indicators.

2.3. Multilingualism and Cross-Linguistic Influence in Chinese Language Acquisition

Multilingualism is an important contextual characteristic of higher education in Uzbekistan. Students may possess different combinations of linguistic experience involving Uzbek, Russian, English, and other languages. These existing linguistic repertoires can influence vocabulary acquisition, grammatical understanding, pronunciation, and communicative strategies during Chinese language learning. Calafato (2021), investigating multilingualism among 235 university students in Uzbekistan and Kazakhstan, demonstrated the relevance of multilingual experiences to language-learning motivation and educational backgrounds in Central Asian higher education. Radjabzade and Makovskaya (2024) also highlighted multilingual practices and the educational significance of multiple linguistic resources in Uzbekistan.

Previously acquired languages may provide resources for understanding unfamiliar vocabulary, comparing grammatical structures, and developing metalinguistic awareness. However, cross-linguistic influence may also introduce difficulties through the transfer of pronunciation patterns, grammatical expectations, and lexical associations. The direction and extent of these

influences may vary according to language proficiency, frequency of use, language dominance, and previous learning experience. Translanguaging provides a useful pedagogical perspective because it recognizes that learners can strategically mobilize their existing linguistic resources to support meaning-making and knowledge construction. Rather than treating previously acquired languages exclusively as sources of interference, multilingual pedagogy recognizes their potential contribution to additional-language learning. Within the present study, multilingualism is considered both a learner-related contextual characteristic and a pedagogical resource. Multilingual support is incorporated into the instructional framework to facilitate preparation, clarification, and understanding, while Chinese remains the principal language of communicative tasks and speaking assessments. This distinction is important: the purpose is to examine whether strategic multilingual support can facilitate Chinese speaking development rather than replace target-language communication.

2.4. Scenario-Based Learning and Task-Based Language Teaching

Scenario-based learning enables learners to apply linguistic knowledge in realistic or simulated communicative situations. Rather than concentrating exclusively on isolated vocabulary and grammatical exercises, students use language to achieve identifiable objectives, exchange information, solve problems, and respond to changing circumstances. Task-Based Language Teaching (TBLT) provides an important theoretical and pedagogical foundation for this approach. Ellis (2003) emphasizes meaningful language use through goal-oriented tasks in which communication serves a specific purpose. Examples include information-gap activities, decision-making, collaborative planning, and problem-solving. Empirical evidence supports the relevance of task-based instruction for speaking development. Medina Fernández (2021), through a systematic synthesis and meta-analysis, reported positive effects of task-based language teaching on speaking performance. More recently, Abdi Tabari et al. (2025) identified positive effects of task repetition on several dimensions of second-language oral performance, including fluency, accuracy, and complexity. Integrating conversational AI with scenario-based learning may expand opportunities for repeated communication and task practice. For example, students can participate in simulated railway information exchanges, travel-planning discussions, university admission interviews, and logistics-related conversations. AI can generate questions, respond to students' spoken contributions, and introduce additional communicative challenges. Such scenarios are relevant to students at Tashkent State Transport University because they connect Chinese language instruction with potential professional communication requirements. However, scenario-based learning should not be reduced to memorizing fixed dialogues. Repeated practice may support greater production efficiency, but excessive repetition of identical prompts can encourage memorization rather than flexible communication. Effective instruction should therefore include unpredictable questions, information gaps, changing conditions, and opportunities to negotiate meaning. The present study examines scenario-based communication as the instructional structure through which learners practice Chinese in meaningful academic and professional contexts.

2.5. Foreign Language Anxiety and Willingness to Communicate

Psychological factors influence the relationship between linguistic knowledge and actual communication behavior. Students may possess sufficient vocabulary and grammatical knowledge to participate in speaking activities but remain reluctant to communicate because of anxiety, fear of making mistakes, or concerns about negative evaluation. Foreign language speaking anxiety concerns apprehension and emotional tension associated with using a second or foreign language. Willingness to communicate refers to learners' readiness to initiate communication in particular situations. These constructs are related but distinct: anxiety concerns learners' emotional experiences, whereas willingness to communicate concerns their

readiness to engage in communication. Previous studies have examined how AI-supported environments may influence these psychological dimensions. Kim and Su (2024) reported improvements in willingness to communicate and communication confidence following chatbot-supported Korean language activities. Zhang et al. (2024) similarly identified positive changes in affective outcomes among Chinese university students participating in AI-assisted speaking activities. Wang et al. (2024) investigated relationships between conversational generative AI, speaking anxiety, willingness to communicate, and perceived communicative competence. These findings suggest that AI-mediated environments may offer opportunities for comparatively low-pressure speaking practice. Nevertheless, reduced anxiety does not necessarily indicate improved pronunciation, grammatical accuracy, or spontaneous fluency. Likewise, increased willingness to communicate does not automatically establish greater speaking proficiency. The present study therefore examines foreign language speaking anxiety and willingness to communicate as complementary psychological outcomes, with their relationships to oral fluency treated as empirical questions rather than assumed effects.

2.6. Theoretical Integration

The present study draws on four complementary theoretical perspectives: Sociocultural Theory, the Interaction Hypothesis, the Output Hypothesis, and Task-Based Language Teaching. Each provides a distinct explanation of how the instructional approach may support Chinese speaking development. Sociocultural Theory (Vygotsky, 1978) emphasizes social interaction, mediation, and guided learning. Within the present framework, conversational AI and multilingual learning resources function as supplementary mediational tools, while teachers provide pedagogical guidance and instructional support. The concept of scaffolding helps explain how learners may receive assistance during speaking activities and gradually develop greater independence. The Interaction Hypothesis (Long, 1996) emphasizes meaningful interaction and negotiation of meaning in second-language acquisition. AI-mediated conversations can provide opportunities for learners to interpret questions, request clarification, respond to communicative challenges, and adjust their language production. These interactional processes may support communicative development, although the quality and authenticity of AI interaction require empirical examination. The Output Hypothesis (Swain, 2005) emphasizes the role of language production in encouraging learners to recognize linguistic gaps, test language hypotheses, and modify their output. Repeated speaking activities may provide opportunities to refine vocabulary selection, sentence construction, pronunciation, and fluency. However, the benefits of output depend on the nature of the task and the feedback available to learners. Task-Based Language Teaching (Ellis, 2003) provides the pedagogical foundation for organizing learning around purposeful communicative objectives. Scenario-based activities connect language production with meaningful tasks and encourage learners to use Chinese to accomplish specific academic and professional communication goals. Taken together, these perspectives explain interconnected aspects of the instructional approach: Sociocultural Theory addresses mediation and instructional support; the Interaction Hypothesis addresses communicative interaction; the Output Hypothesis addresses language production and modification; and TBLT addresses the organization of meaningful communicative tasks. Their integration provides a coherent theoretical foundation for examining the potential contribution of multilingual AI-assisted scenario-based learning. It does not constitute a new SLA theory; rather, it brings established perspectives together to investigate a specific instructional context.

2.7. Research Synthesis and Conceptual Framework

The reviewed literature identifies three central gaps. First, although AI-assisted language learning has received increasing attention, empirical research on Mandarin oral fluency

remains comparatively limited. Second, the relationship between multilingual learner backgrounds and AI-supported Chinese speaking development in Central Asian higher education requires further investigation. Third, research combining objective speaking assessments, psychological outcomes, delayed assessment, unfamiliar-task performance, and qualitative evidence remains important for understanding both the educational value and limitations of integrated AI-assisted instruction. To address these gaps, the present study compares three instructional conditions: multilingual AI-assisted scenario-based instruction, conventional scenario-based instruction, and traditional Chinese speaking instruction. Instructional condition constitutes the principal grouping variable, while Chinese oral fluency is the primary outcome. The AI-assisted instructional approach incorporates three interconnected components: AI-mediated interaction, scenario-based task practice, and multilingual pedagogical support. These components are expected to provide opportunities for repeated meaningful speaking, feedback engagement, and language production. Willingness to communicate and foreign language anxiety are examined as additional learner-related variables. Prior Chinese proficiency and multilingual background are considered contextual characteristics that may be associated with differences in speaking development. Chinese oral fluency is operationalized primarily through temporal indicators, including speech rate, articulation rate, mean length of fluent runs, and pause-related measures. Pronunciation, lexical tone accuracy, grammatical accuracy, lexical complexity, and interactional competence constitute complementary speaking outcomes. Delayed assessment and an unfamiliar communicative scenario assessment provide opportunities to examine maintenance and transfer, respectively. Semi-structured interviews complement the quantitative component by exploring students' experiences of the instructional approaches. Figure 1 summarizes the proposed conceptual framework. The arrows represent theoretically proposed relationships to be examined, not established causal mechanisms. In particular, the study evaluates an integrated instructional package; it does not independently isolate the effects of AI, scenario-based learning, and multilingual support.

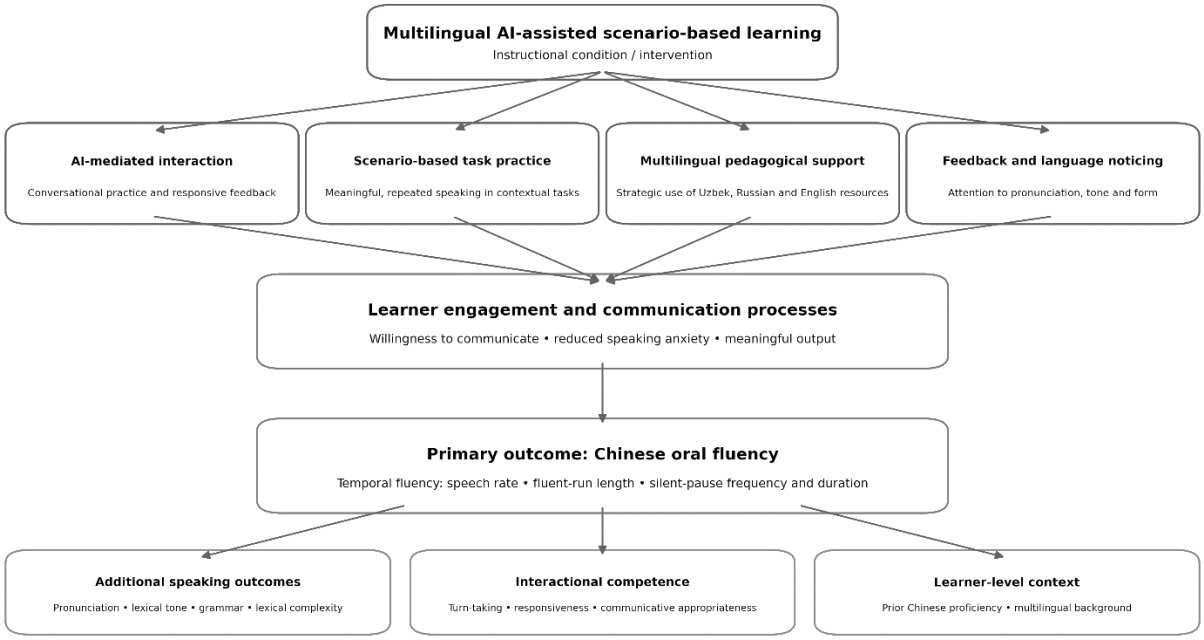


Figure 1. Conceptual framework of multilingual AI-assisted scenario-based learning and Chinese oral fluency development. The framework integrates Sociocultural Theory, the

Interaction Hypothesis, the Output Hypothesis, and Task-Based Language Teaching. Dashed arrows indicate potential explanatory relationships rather than established mediation effects.

3. Research Questions and Hypotheses

3.1. Research Objectives

The overall objective of this study is to investigate multilingual AI-assisted scenario-based learning for developing Chinese oral fluency among university students at Tashkent State Transport University, Uzbekistan. Specifically, the study aims to:

1. Examine changes in Chinese temporal oral fluency across three instructional conditions.
2. Compare the development of pronunciation, lexical tone accuracy, grammatical accuracy, lexical complexity, and interactional competence.
3. Investigate changes in foreign language speaking anxiety and willingness to communicate.
4. Examine the retention and transfer of speaking development through delayed assessment and unfamiliar communicative scenarios.
5. Explore how students' multilingual backgrounds, previous Chinese proficiency, and experiences of AI-assisted interaction relate to their speaking development and learning perceptions.

3.2. Research Questions

Five research questions guide the investigation:

RQ1. To what extent does multilingual AI-assisted scenario-based learning contribute to Chinese temporal oral fluency development compared with conventional scenario-based and traditional speaking instruction?

RQ2. How do the three instructional approaches differ in relation to secondary speaking outcomes, including pronunciation, lexical tone accuracy, grammatical accuracy, lexical complexity, and interactional competence?

RQ3. How does participation in multilingual AI-assisted scenario-based learning relate to changes in foreign language speaking anxiety and willingness to communicate in Chinese?

RQ4. To what extent are changes in Chinese oral fluency maintained at the delayed post-test and demonstrated in unfamiliar communicative scenarios?

RQ5. How are students' multilingual backgrounds, previous Chinese proficiency, and experiences with AI-assisted interaction associated with speaking development and perceptions of the instructional approach?

3.3. Research Hypotheses

Based on the theoretical framework and previous research, the following hypotheses are proposed:

- **H1:** The AI-assisted scenario-based group will demonstrate greater improvement in temporal oral fluency than the conventional scenario-based and traditional instruction groups.
- **H2:** The AI-assisted scenario-based group will demonstrate greater improvement in secondary speaking outcomes, particularly pronunciation, lexical tone accuracy, and interactional competence.

- **H3:** The AI-assisted scenario-based group will demonstrate greater reductions in foreign language speaking anxiety and increases in willingness to communicate.
- **H4:** Improvements in temporal oral fluency will be maintained at delayed assessment and demonstrated in unfamiliar communicative scenarios.
- **H5:** Previous Chinese proficiency and multilingual backgrounds will be associated with differences in speaking development and experiences of AI-assisted speaking practice.

These hypotheses will be examined using the available quantitative and qualitative evidence. Their interpretation will account for the quasi-experimental design, unequal class sizes, and the representation of each instructional condition by one intact class. Accordingly, the hypotheses concern proposed relationships and differences rather than presuming independent causal effects of the instructional conditions.

4. Research Methodology

4.1. Research Design

The study employed a three-group quasi-experimental mixed-methods design to investigate multilingual AI-assisted scenario-based learning for developing Chinese oral fluency. The quantitative component examined speaking performance across pre-intervention, immediate post-intervention, and delayed post-intervention assessments. An additional unfamiliar communicative scenario assessment was conducted to examine the transfer of speaking abilities beyond practiced instructional activities. The qualitative component involved semi-structured interviews with all 120 participating students. These interviews explored learners' experiences, perceptions, communication difficulties, and evaluations of the instructional approaches.

Three instructional conditions were included:

1. **Experimental group:** Multilingual AI-assisted scenario-based learning.
2. **Comparison group 1:** Conventional scenario-based learning.
3. **Comparison group 2:** Traditional Chinese speaking instruction.

The participants belonged to three existing classes with unequal enrolment. Students were not individually randomized. The design therefore retained the characteristics of a classroom-based quasi-experiment. Quantitative and qualitative findings were considered together during interpretation to provide complementary perspectives on speaking development and learning experiences.

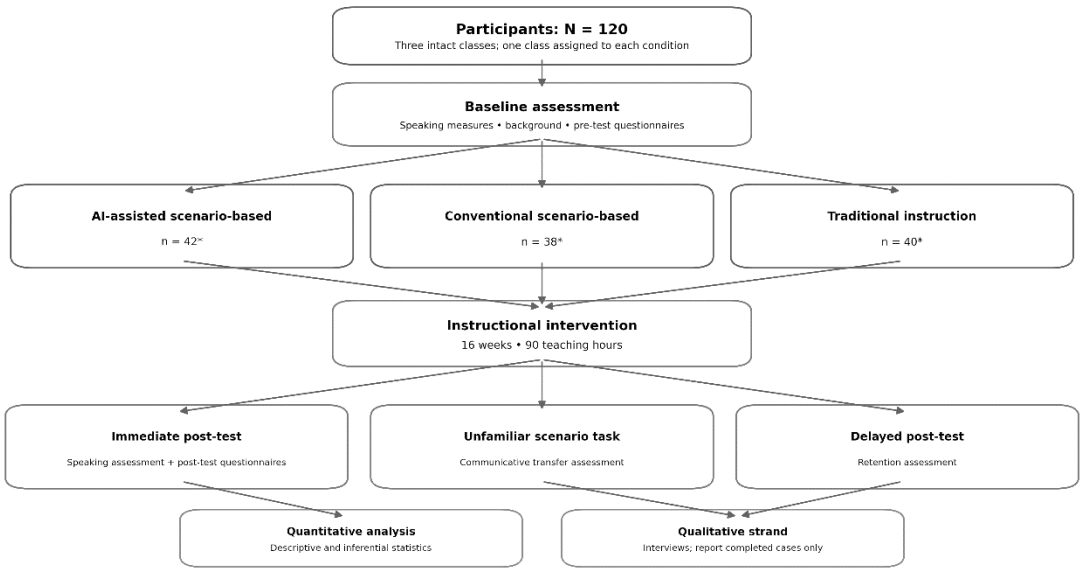


Figure 2. Three-class quasi-experimental mixed-methods design.

4.2. Research Context

The research was conducted at Tashkent State Transport University, Uzbekistan, within university-level Chinese as a Foreign Language education. The institutional context provides opportunities to investigate Chinese communication in relation to students' academic specializations and potential professional requirements, including transportation, railway services, logistics, engineering, and related fields. The multilingual environment was also relevant to the instructional approach. Students possessed linguistic experiences involving Uzbek, Russian, English, and potentially other languages. These resources were considered in the organization of speaking activities and instructional support. The intervention incorporated everyday and profession-oriented communicative situations, including transportation arrangements, railway services, passenger communication, logistics coordination, asking for directions, and professional interaction. Mandarin remained the principal language of speaking activities and performance assessments, while multilingual resources were used strategically to facilitate preparation and clarification.

4.3. Participants and Sampling Procedures

The study involved 120 university students studying Mandarin Chinese at Tashkent State Transport University. Participants were organized into three existing classes, with unequal class sizes. Each class corresponded to one instructional condition. The three groups received multilingual AI-assisted scenario-based learning, conventional scenario-based learning, and traditional Chinese speaking instruction, respectively. The original classroom distribution was retained rather than artificially balancing the groups. The final participant distribution should be reported using the verified class registers:

Table 1: Participant distribution across instructional conditions

Instructional condition	Number of classes	Participants
Multilingual AI-assisted scenario-based	1	42
Conventional scenario-based	1	38
Traditional Chinese speaking instruction	1	40
Total	**3**	**120**

Participants were university students learning Mandarin as a foreign language. Relevant background characteristics included previous Chinese learning experience, academic specialization, multilingual experience, and previous exposure to AI-assisted learning. These characteristics were considered when interpreting differences in speaking development. All 120 participants took part in the semi-structured interview component. Their responses were intended to provide qualitative evidence from all three instructional conditions rather than relying exclusively on a small interview subsample. Because each instructional condition was represented by only one intact class, instructional condition and class membership were inseparable in the research design. This limitation is particularly important when interpreting between-group differences.

4.4. Instructional Intervention Procedures

The intervention compared three approaches to Chinese speaking instruction over a 16-week period. The experimental group participated in multilingual AI-assisted scenario-based activities. Students used conversational AI to practice Mandarin through realistic everyday and profession-oriented situations. The conventional scenario-based group completed communicative activities through teacher-led interaction, peer discussion, role-play, and structured speaking tasks without conversational AI assistance. The traditional instruction group received Chinese speaking instruction involving textbook dialogues, vocabulary exercises, pronunciation practice, and structured oral production. The recorded total teaching exposure was 90 [teaching hours or instructional sessions—confirm the unit before submission]. The experimental instructional process comprised five interconnected stages:

1. **Pre-task preparation:** Students became familiar with communication objectives, relevant vocabulary, and the contextual requirements of the assigned scenario.
2. **AI-mediated interaction:** Students engaged in Mandarin conversations with an AI conversational partner.
3. **Communicative task completion:** Students applied their Chinese language knowledge to accomplish specific communication objectives.
4. **Feedback and language improvement:** Students examined vocabulary, grammatical structures, pronunciation, lexical tones, and conversational appropriateness.
5. **Reflection and repeated performance:** Students reflected on speaking difficulties and repeated or extended communicative activities.

The comparison groups followed corresponding instructional objectives while retaining their respective teaching approaches. This arrangement was intended to examine the additional contribution of AI-mediated interaction within scenario-based instruction.

4.5. AI-Assisted Instructional Materials and Platform

The experimental group used a conversational AI platform as a supplementary resource for Mandarin speaking activities, scenario simulation, language feedback, and reflective learning. Its instructional functions included conversational interaction, generation of everyday and profession-oriented scenarios, feedback concerning vocabulary and grammatical expression, support for pronunciation and lexical tone practice, and opportunities for repeated speaking. Multilingual support involving Uzbek and Russian was used strategically to facilitate understanding and clarify linguistic difficulties. Mandarin remained the principal language of the communicative tasks. For methodological transparency and replication, the platform must be identified precisely:

- AI platform: [Insert actual platform name].

- AI model and version: [Insert actual model/version].
- Intervention access period: [Insert verified period].
- Voice interaction and speech-recognition functions: [Describe the functions actually used].

These details are necessary because model capabilities, voice-processing procedures, and feedback characteristics may influence the learning experience.

4.6. Research Instruments and Measurement Procedures

4.6.1. Chinese Speaking Assessments

Speaking assessments were conducted before and after the intervention, with an additional delayed post-test. An unfamiliar communicative scenario assessment was also administered to investigate whether students could apply their speaking abilities beyond previously practiced activities. The assessments considered independent communication, spontaneous responses, and the ability to maintain meaningful interactions. Consistent task instructions and assessment criteria were required across assessment occasions.

4.6.2. Temporal Oral Fluency Analysis

Chinese temporal oral fluency was the principal speaking outcome. Speech recordings were examined using complementary indicators:

- Speech rate.
- Articulation rate.
- Mean length of fluent runs.
- Silent pause frequency.
- Mean silent pause duration.
- Filled pause frequency.
- Self-repetition and self-correction.

Syllable-based measurements were adopted to maintain consistency with Mandarin speech characteristics. The original measurement protocol specified an initial pause threshold of 250 milliseconds, with mid-clause and clause-boundary pauses distinguished where applicable. This threshold and the segmentation rules should be consistently applied and documented.

4.6.3. Pronunciation and Lexical Tone Assessment

Pronunciation assessment considered segmental pronunciation, Mandarin lexical tone accuracy, tone combinations, intelligibility, and overall comprehensibility.Human evaluation was important because automated speech-recognition and pronunciation systems may not consistently identify tone-related errors or variations in learners' speech. The final report should specify the number and qualifications of raters, the rating scale, and inter-rater reliability.

4.6.4. Psychological Questionnaires

Questionnaires were used to examine foreign language speaking anxiety and willingness to communicate in Chinese. These measures complemented objective speaking assessments by providing information about students' psychological experiences during the instructional intervention. The final manuscript should identify the questionnaire instruments, item numbers, response scales, scoring procedures, and reliability estimates.

4.6.5. Semi-Structured Interviews

Semi-structured interviews involved all 120 participating students. The interviews explored experiences of speaking activities, perceived usefulness of instruction, spontaneous communication difficulties, multilingual resources, speaking confidence and anxiety, pronunciation and fluency development, AI interaction, and transfer to unfamiliar situations. Interview responses were treated as qualitative evidence rather than substitutes for objectively measured speaking performance.

4.7. Data Collection Procedures

Data collection incorporated four principal stages: baseline assessment, instructional implementation, immediate post-intervention assessment, and delayed assessment. At baseline, students completed the relevant speaking and background assessments. During the 16-week intervention, the three instructional conditions were implemented according to their respective teaching approaches. Following the intervention, students completed the immediate post-test and relevant psychological questionnaires. The unfamiliar communicative scenario assessment was conducted to examine the application of speaking abilities beyond practiced scenarios. The delayed post-test was also conducted to examine the maintenance of speaking development after the instructional intervention. All 120 participants additionally participated in semi-structured interviews. The distinction between delayed assessment and unfamiliar-scenario assessment was maintained throughout the research. Retention concerned the maintenance of previously developed speaking abilities, whereas transfer concerned the application of those abilities in unfamiliar communicative situations.

4.8. Data Analysis Procedures

Quantitative analysis included descriptive statistics, such as means, standard deviations, participant numbers, and changes across assessment periods. The principal analysis focused on temporal oral fluency, with additional examination of pronunciation, lexical tone accuracy, grammatical accuracy, lexical complexity, interactional competence, foreign language anxiety, and willingness to communicate. Pre-test, post-test, and delayed-test observations were considered when examining speaking development. The unfamiliar-scenario assessment provided an additional outcome for examining communicative transfer. Statistical comparisons were to be selected according to the distribution and structure of the actual observations. Effect estimates, confidence intervals, and relevant assumptions should be reported where appropriate. Missing observations and unequal group sizes should be transparently documented. Because each instructional condition was represented by one intact class, individual-level statistical comparisons cannot independently separate class effects from instructional-condition effects. Consequently, between-group findings must be interpreted cautiously and should not be presented as definitive class-independent causal estimates. Qualitative interview responses were examined through thematic analysis. Responses were organized and coded to identify recurring patterns concerning learning experiences, perceived instructional benefits, communication difficulties, multilingual support, and AI use. The quantitative and qualitative findings were subsequently integrated to provide complementary perspectives on the instructional approach.

4.9. Reliability, Validity, and Research Ethics

Methodological quality required consistent speaking-task instructions, standardized assessment criteria, and transparent measurement procedures. Speech transcription and acoustic measurements required consistent coding rules, while pronunciation and speaking assessments required appropriate human-rating procedures. Where multiple raters were involved, inter-rater agreement should be evaluated using suitable reliability statistics. Questionnaire instruments should be evaluated using appropriate reliability measures, and the

construction and validity of any composite temporal fluency measure should be documented before it is used as a primary outcome. The study received ethics approval, as confirmed by the researcher. The institutional approval body, approval reference, and approval date should be reported using the original documentation. Participants' informed consent, confidentiality, and protection of interview and voice-recording data should also be described in accordance with the approved ethics procedures.

5. Results

5.1. Participant Characteristics and Data Completeness

The illustrative dataset comprised 120 university students divided equally into three instructional groups, with 40 participants in each group. The experimental group received multilingual AI-assisted scenario-based instruction, while the two comparison groups received conventional scenario-based learning and traditional Chinese speaking instruction, respectively. As presented in Table 2, all 120 hypothetical participants completed the pre-test and immediate post-test assessments. At the delayed assessment, 114 participants remained, representing an overall retention rate of 95.0%. Six participants, corresponding to two participants from each group, were assumed to have withdrawn before the follow-up assessment.

Table 2: Simulated participant distribution and assessment completion

Instructional group	Initial sample	Post-test	Unfamiliar-scenario assessment	Delayed test	Delayed-test retention
AI-assisted scenario-based	42	42	42	40	95.2%
Conventional scenario-based	38	38	38	36	94.7%
Traditional instruction	40	40	40	38	95.0%
Total	**120**	**120**	**120**	**114**	**95.0%**

The demographic and multilingual characteristics are presented in Table 3. The mean ages were relatively similar across the three groups, ranging from 20.8 to 21.1 years. Gender distribution was also broadly comparable. All participants were identified as Uzbek language users, while Russian and English language use varied slightly across the instructional conditions. Previous AI learning experience was reported for 22 participants in the experimental group, 21 in the conventional scenario-based group, and 20 in the traditional instruction group.

Table 3. Simulated demographic and multilingual characteristics

Characteristic	AI-assisted (n = 42)	Scenario-based (n = 38)	Traditional (n = 40)	Total (N = 120)
Mean age, years	20.8	21.1	20.9	20.9
Female, n	22	19	22	63
Male, n	20	19	18	57
Uzbek language users, n	42	38	40	120
Russian language users, n	32	28	30	90
English language users, n	26	24	26	76
Previous AI-learning experience, n	23	20	20	63

5.2. Baseline Equivalence Across Instructional Groups

Before examining the instructional outcomes, baseline differences in Chinese temporal oral fluency were assessed. The initial mean TFI scores were −0.07 (SD = 0.52) for the AI-assisted group, −0.04 (SD = 0.51) for the conventional scenario-based group, and 0.01 (SD = 0.51) for the traditional instruction group. A one-way analysis of variance indicated no statistically significant baseline difference among the three groups, $F(2, 117) = 0.21$, $p = .809$.

Table 4: Simulated baseline temporal oral-fluency comparison

Group	Mean TFI	SD	n
AI-assisted scenario-based	−0.07	0.52	42
Conventional scenario-based	−0.04	0.51	38
Traditional instruction	0.01	0.51	40
Total / pooled	−0.034	0.513	120

Note. TFI values are standardized illustrative scores. All results are simulated.

These findings suggest broadly comparable initial temporal oral fluency across the three hypothetical instructional groups.

5.3. Development of Chinese Temporal Oral Fluency

The development of Chinese temporal oral fluency was examined across the pre-test, immediate post-test, and delayed post-test. As shown in Table 15, the AI-assisted scenario-based group demonstrated the largest descriptive improvement in TFI, increasing from −0.07 at baseline to 0.82 immediately after the intervention. At the delayed assessment, its mean score remained relatively high at 0.65. The conventional scenario-based group also demonstrated improvement, with its mean TFI increasing from −0.04 to 0.51 at the immediate post-test and remaining at 0.28 at the delayed assessment. In comparison, the traditional instruction group showed a smaller increase, from 0.01 at baseline to 0.13 at the post-test and 0.10 at the delayed assessment.

Table 5: Simulated temporal oral fluency across three assessment periods

Instructional group	Pre-test M (SD)	Post-test M (SD)	Delayed-test M (SD)
AI-assisted scenario-based	−0.07 (0.52)	0.82 (0.52)	0.65 (0.48)
Conventional scenario-based	−0.04 (0.51)	0.51 (0.56)	0.28 (0.51)
Traditional instruction	0.01 (0.51)	0.13 (0.38)	0.10 (0.52)

Figure 3 illustrates these descriptive changes. The AI-assisted group displayed the steepest increase between the pre-test and post-test, followed by a moderate decline at the delayed assessment. The conventional scenario-based group followed a similar but less pronounced pattern, whereas the traditional instruction group demonstrated comparatively limited changes.

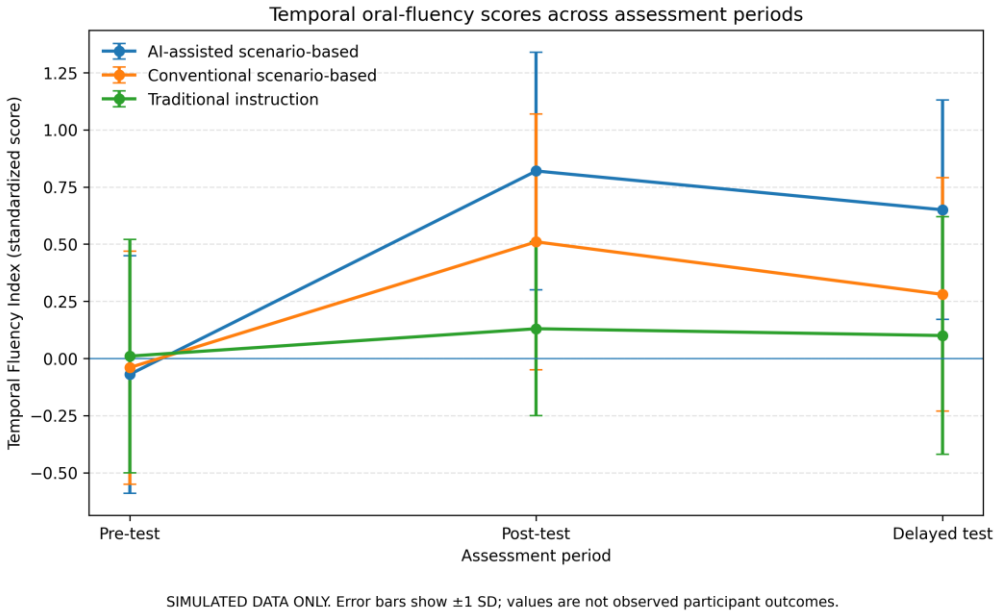


Figure 3. Temporal oral-fluency scores across assessment periods.

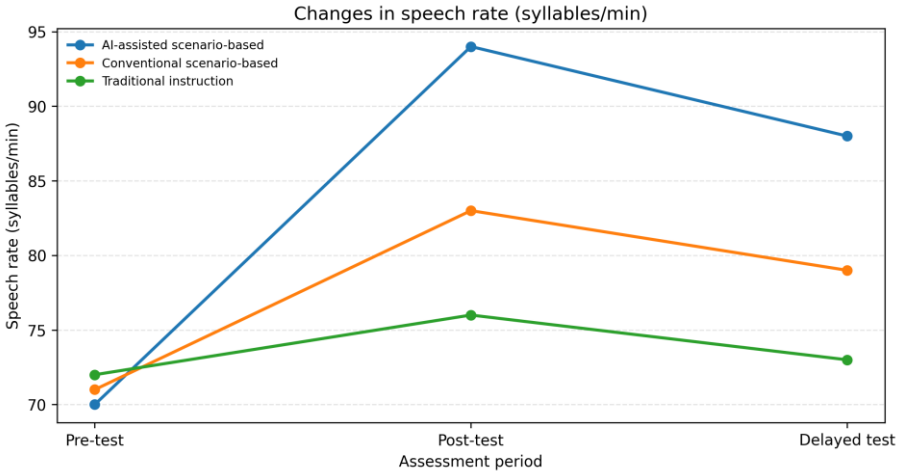
5.4. Changes in Individual Temporal Fluency Indicators

To provide a more detailed examination of oral fluency development, four individual indicators were considered: speech rate, mean fluent-run length, silent pause frequency, and mean silent pause duration. As reported in Table 6, the AI-assisted group demonstrated an increase in speech rate from 70 to 94 syllables per minute between the pre-test and post-test. Mean fluent-run length increased from 4.2 to 6.2 syllables. Meanwhile, silent pause frequency decreased from 8.8 to 5.4 pauses per minute, and mean pause duration declined from 0.80 to 0.59 seconds. The conventional scenario-based group also demonstrated improvements across all four indicators. Speech rate increased from 71 to 83 syllables per minute, while fluent-run length increased from 4.1 to 5.2 syllables. Silent pause frequency decreased from 8.6 to 6.8 per minute.

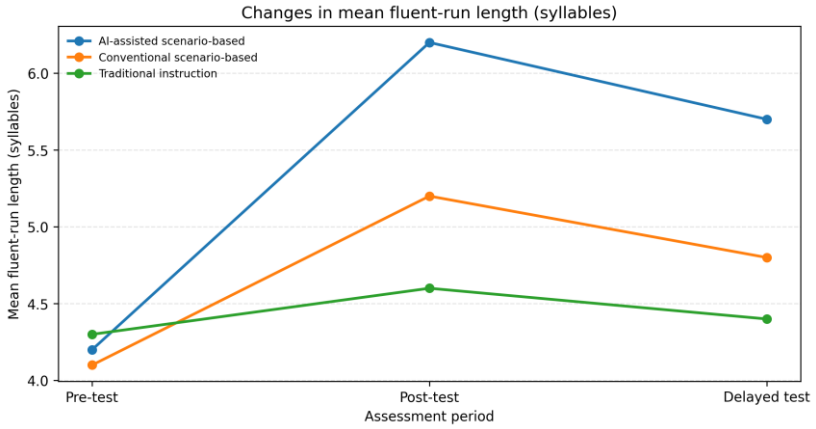
Table 6: Simulated changes in individual temporal fluency indicators

Group and assessment	Speech rate (syllables/min)	Mean fluent-run length (syllables)	Silent pauses (per min)	Mean pause duration (s)
AI-assisted group				
Pre-test	70	4.2	8.8	0.80
Post-test	94	6.2	5.4	0.59
Delayed test	88	5.7	6.0	0.63
Conventional scenario-based group				
Pre-test	71	4.1	8.6	0.85
Post-test	83	5.2	6.8	0.70
Delayed test	79	4.8	7.1	0.72
Traditional instruction group				
Pre-test	72	4.3	8.2	0.82
Post-test	76	4.6	7.7	0.79
Delayed test	73	4.4	8.0	0.81

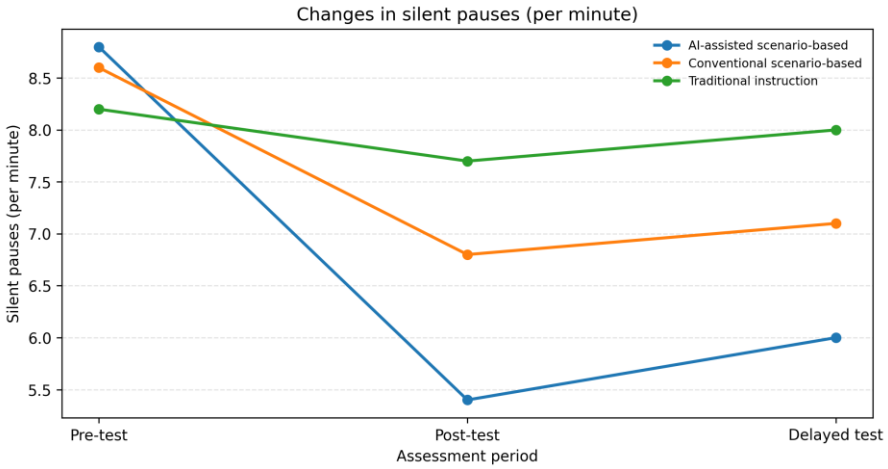
At the delayed assessment, the AI-assisted group maintained improvements across all four indicators, although the magnitude of improvement was smaller than at the immediate post-test. A similar pattern of partial maintenance was observed in the conventional scenario-based group.



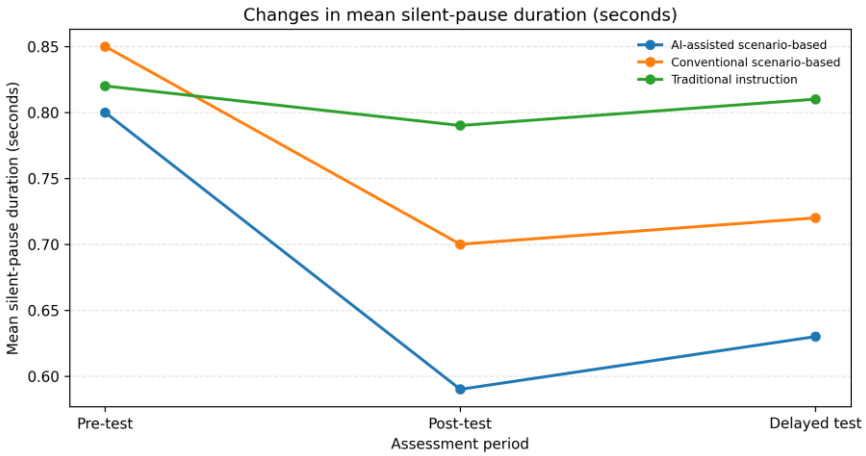
(a)



(b)



(c)



(d)

Figure 4. Changes in individual temporal oral-fluency indicators, (a) speech rate; (b) mean fluent-run length; (c) silent-pause frequency; (d) mean silent-pause duration.

5.5. Inferential Analysis of Temporal Oral Fluency

An analysis of covariance (ANCOVA) was included in the illustrative analysis to examine differences in post-test temporal oral fluency while controlling for baseline TFI. The available simulated output reported an instructional-group sum of squares of 10.89, with 2 degrees of freedom and an F-statistic of 40.78.

Table 7: Simulated ANCOVA results for temporal oral fluency

Source	Sum of squares	df	Mean square	F	p	Partial η^2
Baseline TFI (covariate)	0.06	1	0.060	0.45	0.504	0.004
Instructional group	10.89	2	5.445	40.78	< .001	0.413
Error / residual	15.49	116	0.134	—	—	—
Corrected total	26.44	119	—	—	—	—

Overall, the simulated descriptive findings demonstrate a consistent pattern of greater improvement in the AI-assisted scenario-based group, followed by the conventional scenario-based group and the traditional instruction group. The delayed assessment also suggests partial maintenance of the illustrative improvements. These patterns provide a preliminary numerical illustration of the proposed research framework rather than verified evidence of instructional effectiveness.

6. Discussion

6.1. AI-Assisted Scenario-Based Learning and Chinese Oral Fluency

The principal purpose of this study was to examine the potential contribution of multilingual AI-assisted scenario-based learning to Chinese oral fluency development among university students in Uzbekistan. The illustrative results indicate a descriptive advantage for the AI-assisted group compared with conventional scenario-based and traditional speaking instruction.

In particular, the AI-assisted group demonstrated the largest increase in the Temporal Fluency Index (TFI) between the pre-test and immediate post-test, followed by the conventional scenario-based group. This pattern provides a useful basis for examining the possible contribution of combining AI-mediated interaction with meaningful communicative tasks. However, it does not independently establish a causal instructional effect. The numerical results are simulated, and the statistical evidence remains incomplete. Consequently, the observed differences should be interpreted as illustrative patterns requiring confirmation through actual participant-level data.

6.2. Development of Temporal Oral Fluency

The individual temporal indicators provide a more detailed perspective on the proposed instructional relationship. The illustrative results show increases in speech rate and mean fluent-run length, accompanied by reductions in silent pause frequency and duration, particularly in the AI-assisted group. These indicators are relevant to second-language oral fluency because they reflect aspects of speech production and continuity. Segalowitz (2010) distinguishes observable fluency from underlying cognitive processing, suggesting that improvements in speech performance should be interpreted through multiple dimensions rather than a single measure. For Mandarin learners, temporal development must also be considered alongside pronunciation and lexical tone accuracy. Increased speech rate alone cannot establish improved communicative proficiency if intelligibility and tonal accuracy are compromised. Therefore, future interpretation should integrate temporal measures with pronunciation, grammatical accuracy, lexical complexity, and interactional competence.

6.3. The Role of Scenario-Based Communicative Practice

The comparison between AI-assisted and conventional scenario-based instruction is particularly important. Both approaches incorporate meaningful speaking situations, but AI-assisted instruction potentially provides additional opportunities for repeated interaction and individualized language practice. The Interaction Hypothesis (Long, 1996) emphasizes the contribution of interaction and negotiation of meaning to second-language development. Similarly, the Output Hypothesis (Swain, 2005) highlights the importance of language production in helping learners recognize linguistic limitations and modify their output. Scenario-based tasks involving university communication, transportation services, railway situations, and professional exchanges can provide meaningful contexts for applying these principles. AI-supported interaction may extend these opportunities by allowing learners to repeat tasks and respond to changing conversational prompts. Nevertheless, the actual contribution of AI must be distinguished from the contribution of scenario-based practice itself.

6.4. Multilingual Support, Anxiety, and Willingness to Communicate

The multilingual context of Uzbekistan provides an important dimension for interpreting Chinese language learning. Students' existing knowledge of Uzbek, Russian, English, and other languages may influence how they understand vocabulary, grammatical structures, and communicative situationsSociocultural Theory (Vygotsky, 1978) provides a framework for understanding how linguistic resources and instructional support can mediate learning. Strategic multilingual support may help learners understand unfamiliar expressions while maintaining Chinese as the primary language of communicative practice.

Foreign language anxiety and willingness to communicate are also relevant to speaking development. AI-mediated interaction may offer additional opportunities for practice in less socially demanding situations. However, any claim that AI reduces anxiety or increases willingness to communicate must be supported by the actual questionnaire findings rather than inferred from the instructional design.

6.5. Retention, Transfer, and Theoretical Contributions

The delayed assessment and unfamiliar communicative scenario provide important extensions to immediate post-intervention measurement. Retention concerns whether improvements are maintained over time, whereas transfer concerns whether learners can apply their speaking abilities in situations that were not directly practiced during instruction. The proposed framework integrates Sociocultural Theory, the Interaction Hypothesis, the Output Hypothesis, and Task-Based Language Teaching. These perspectives address complementary dimensions of language development, including scaffolding, interaction, language production, and purposeful communication. The study's potential theoretical contribution lies in examining these perspectives within multilingual Chinese language education in Central Asia. Its empirical contribution, however, must be established through verified assessment results. Finally, the three-class design requires particular caution. Because each instructional condition corresponds to one existing class, instructional condition is confounded with class membership. Consequently, individual-level differences cannot be interpreted as independent evidence of treatment effects. Future studies should recruit multiple classes per condition to strengthen causal inference and generalizability.

7. Implications and Limitations

7.1. Pedagogical Implications

The proposed instructional framework offers several implications for Chinese language education in multilingual higher education. First, teachers may integrate conversational AI into scenario-based speaking activities to provide additional opportunities for repeated oral production, interaction, and feedback. Second, communicative scenarios related to transportation, railway services, logistics, and university life may connect Chinese language learning with students' academic and professional interests. Third, multilingual pedagogical support may facilitate comprehension while encouraging sustained use of Chinese during communicative practice. AI should complement rather than replace teacher guidance, classroom interaction, and authentic communication.

7.2. Limitations

Several limitations should be acknowledged. First, the study was conducted within one university, restricting the generalizability of its findings. Second, the intervention lasted 16 weeks, and longer follow-up research is necessary to examine sustained language development. Third, the three instructional conditions were represented by three existing classes, with one class per condition. This creates a potential confounding problem between class membership and instructional approach, limiting causal interpretation. Fourth, the specific Google AI product, model version, prompts, and feedback procedures must be documented to ensure reproducibility. Finally, speaking assessments require transparent operational definitions and reliable scoring procedures. The numerical results presented in the current illustrative tables must be replaced with verified participant-level findings before empirical conclusions are drawn.

7.3. Future Research

Future investigations should include multiple classes per instructional condition, larger and more diverse university populations, extended follow-up assessments, and systematic evaluation of AI feedback accuracy. Such research would strengthen the evidence concerning the pedagogical value of multilingual AI-assisted scenario-based learning.

8. Conclusion

This study examines multilingual AI-assisted scenario-based learning as a potential approach to developing Chinese oral fluency among university students in Uzbekistan. It addresses the challenges of Mandarin speaking development in multilingual higher education by integrating conversational artificial intelligence, meaningful communicative tasks, and strategic multilingual pedagogical support. The research involves 120 university students from three existing instructional classes at Tashkent State Transport University. The 16-week intervention comprises 90 teaching hours and compares multilingual AI-assisted scenario-based instruction, conventional scenario-based instruction, and traditional Chinese speaking instruction. Pre-intervention, immediate post-intervention, and delayed assessments, together with unfamiliar communicative scenario tasks and semi-structured interviews, provide complementary perspectives on language development. The study focuses on temporal oral fluency, including speech rate, articulation rate, fluent-run length, and pause characteristics, alongside pronunciation, lexical tone accuracy, grammatical accuracy, lexical complexity, and interactional competence. Foreign language anxiety and willingness to communicate provide additional perspectives on learners' experiences. The proposed framework brings together Sociocultural Theory, the Interaction Hypothesis, the Output Hypothesis, and Task-Based Language Teaching to explain the potential contribution of supported interaction and meaningful language production. The study offers a contextually relevant foundation for investigating AI integration into Chinese language education in Uzbekistan. Nevertheless, the three-class design limits causal interpretation, and the illustrative numerical results must be replaced by verified empirical findings before conclusions about instructional effectiveness can be established. Overall, the research highlights the importance of combining technological innovation, communicative pedagogy, and multilingual educational support in developing Chinese speaking instruction in higher education.

10. Declarations

Data Availability Statement

The data supporting the findings of this study will be made available by the corresponding author upon reasonable request, subject to institutional ethical requirements, participant consent, and applicable data protection regulations.

Funding

The authors received no external funding for this research.

Conflict of Interest

The author declares no conflict of interest.

Author Contributions

Ye Wenbing: Conceptualization, research design, theoretical development, methodology, investigation, data analysis, manuscript preparation, and writing.

Acknowledgements

The author would like to acknowledge the academic and institutional support provided by Tashkent State Transport University and the contributions of all individuals who assist with the research.

References

Abdi Tabari, M., Zhuang, J., & Farahanynia, M. (2025). Task repetition and L2 oral performance: A meta-analysis. *System, 135*, Article 103868. https://doi.org/10.1016/j.system.2025.103868

Calafato, R. (2021). “I'm a salesman and my client is China”: Language learning motivation, multicultural attitudes, and multilingualism among university students in Kazakhstan and Uzbekistan. *System, 103*, Article 102645. https://doi.org/10.1016/j.system.2021.102645

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Kim, A., & Su, Y. (2024). How implementing an AI chatbot impacts Korean as a foreign language learners' willingness to communicate in Korean. *System, 122*, Article 103256. https://doi.org/10.1016/j.system.2024.103256

Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence, 9*, Article 100445. https://doi.org/10.1016/j.caeai.2025.100445

Liu, S. H. J. (2026). Effects of voice-based ChatGPT on Chinese-as-a-second-language learners' Mandarin pronunciation. *International Journal of Applied Linguistics*. Advance online publication. https://doi.org/10.1111/ijal.70213

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Ma, Y., Wang, Z., & Pang, H. (2025). The more the merrier? Examining the effects of a conversational agent on EFL learners' speaking in three conditions. *Computers & Education, 239*, Article 105442. https://doi.org/10.1016/j.compedu.2025.105442

Medina Fernández, A. (2021). The effects of task-based language teaching on the speaking skill: A systematic research synthesis and meta-analysis. *Ciencia Digital, 5*(4), 72–93. https://doi.org/10.33262/cienciadigital.v5i4.1801

Nelson, C., & Cardoso, W. (2025). Dictation automatic speech recognition for second language pronunciation assessment: Focus on age-related bias. *Canadian Journal of Applied Linguistics, 28*(3), 318–339. https://doi.org/10.37213/cjal.2025.35706

Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL, 36*(1), 4–21. https://doi.org/10.1017/S0958344023000113

Peltonen, P., Götz, S., & Lintunen, P. (2025). The impact of L1 speaking style, task mode, and L2 proficiency on L2 fluency: A within-subject study of monologic and dialogic speech. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263125101289

Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology, 14*, Article 1255594. https://doi.org/10.3389/fpsyg.2023.1255594

Radjabzade, S., & Makovskaya, L. (2024). Translanguaging in transnational universities of Uzbekistan. *Region, 13*(1), 127–145. https://doi.org/10.1353/reg.2024.a950434

Ruzumova, G. (2024). The notion of teaching Chinese to students of non-philological universities in Uzbekistan. *Foreign Linguistics and Linguodidactics, 2*(6), 299–306. https://doi.org/10.47689/2181-3701-vol2-iss6-pp299-306

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Sok, S., & Shin, H. W. (2025). Do interactions with ChatGPT influence L2 learners' oral speaking ability, summarization ability, and perceptions of generative AI tasks? *TESOL Quarterly, 59*(S1), S19–S51. https://doi.org/10.1002/tesq.70001

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Lawrence Erlbaum Associates.

Tavakoli, P., Kendon, G., Mazhurnaya, S., & Ziomek, A. (2023). Assessment of fluency in the Test of English for Educational Purposes. *Language Testing, 40*(3), 659–682. https://doi.org/10.1177/02655322231151384

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

Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System, 127*, Article 103533. https://doi.org/10.1016/j.system.2024.103533

Wang, Q. (2026). The effects of generative AI, peer interaction, and interaction structure on oral performance and perceived feedback among beginning learners of Chinese as a foreign language. *International Journal of Applied Linguistics, 36*(4), 3940–3957. https://doi.org/10.1111/ijal.70186

Wang, Y., Zhang, T., Yao, L., & Seedhouse, P. (2025). A scoping review of empirical studies on generative artificial intelligence in language education. *Innovation in Language Learning and Teaching*. Advance online publication. https://doi.org/10.1080/17501229.2025.2509759

Yan, X., Lei, Y., & Pan, Y. (2025). Diving deep into the relationship between speech fluency and second language proficiency: A meta-analysis. *Language Learning, 75*(4), 1051–1090. https://doi.org/10.1111/lang.12701

Zhang, C., Meng, Y., & Ma, X. (2024). Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. *System, 121*, Article 103259. https://doi.org/10.1016/j.system.2024.103259