Koh, E., et al. (Eds.) (2026). Proceedings of the 2nd International Conference on Learning Evidence and Analytics. Asia-Pacific Society for Computers in Education.

Instructional Interventions for Improving CFL Speaking:

A Systematic Review Revealing the Need for a Learning Analytics Perspective

Na LI[a,b*] & Masanori YAMADA[c]
[a] *Center for Language Education, Ritsumeikan Asia Pacific University, Japan*
[b] *Graduate School of Human-Environment Studies, Kyushu University, Japan*
[c]*Data-Driven Innovation Initiative, Kyushu University, Japan*
*nanali@apu.ac.jp

Abstract : This review synthesizes 21 empirical studies on instructional interventions for improving Chinese as a Foreign Language (CFL) speaking proficiency identified through ERIC. Interventions include task-based curricula, communicative training programs, and ICT-mediated approaches. While most studies report significant gains, speaking improvement is primarily measured through aggregated pre–post scores and rubric-based ratings. Process-level and trace-based analyses remain limited. The findings suggest that, despite demonstrated effectiveness, CFL speaking research would benefit from greater integration of process-oriented analytic approaches, particularly those aligned with a Learning Analytics perspective

Keywords: Chinese speaking, Chinese as a Foreign Language (CFL), instructional intervention, ICT in language learning, learning analytics

1. Introduction

Developing speaking proficiency remains one of the most challenging aspects of Chinese as a Foreign Language (CFL) learning. As a tone language, Mandarin employs lexical tone as a phonemic contrast, such that tonal inaccuracies may directly alter meaning (Duanmu, 2007). Moreover, Chinese is characterized as an analytic language in which grammatical relationships rely primarily on word order and contextual inference rather than inflectional morphology (Li & Thompson, 1981). It is also described as topic-prominent, where discourse organization and information structure play a central role in sentence construction (Li & Thompson, 1976). These structural properties require learners to coordinate tonal accuracy, lexical retrieval, syntactic sequencing, fluency, and discourse organization simultaneously during interaction, rendering CFL speaking both pedagogically demanding and methodologically complex to assess.

Research in second language acquisition further highlights these challenges. Tone perception and production have consistently been identified as major difficulties for learners from non-tonal language backgrounds. Studies show that learners' tonal perception prior to training is characterized by non-categorical distinctions and high rates of inter-tone confusion (Wang et al., 1999). Oral fluency and accuracy are sensitive to task conditions and the availability of planning time, suggesting that L2 performance reflects dynamic allocation of attentional resources rather than uniform improvement across dimensions (Yuan & Ellis, 2003). Together, these findings suggest that CFL speaking development may involve fluctuating and context-sensitive patterns of change.

To address the challenges inherent in CFL speaking development, instructional interventions—particularly task-based approaches—have received increasing empirical attention as mechanisms for enhancing oral proficiency. At the same time, existing reviews suggest that CFL research has predominantly emphasized instructional effectiveness and

measurable performance gains, typically operationalized through standardized test scores or rubric-based ratings (Gong et al., 2018). In contrast, systematic process-oriented investigations that trace how speaking competence develops over time appear comparatively limited.

Meanwhile, the growing integration of digital technologies into language learning environments has created new opportunities for capturing fine-grained behavioral data. CALL systems and online platforms create digital learning environments in which learner interactions can be recorded and analyzed. (Golonka et al., 2014). Within educational technology, according to the 1st International Conference on Learning Analytics and Knowledge, Learning Analytics (LA) is defined as the measurement, collection, analysis, and reporting of data about learners and their contexts to understand and optimize learning. By emphasizing temporal modeling and behavioral pattern detection (Gašević et al., 2015), LA enables learning to be conceptualized as a process unfolding over time rather than as isolated performance outcomes.

Against this backdrop, the present review moves beyond identifying effective instructional interventions in CFL speaking. Instead, it examines how speaking improvement has been defined, operationalized, and supported by evidence in intervention research. Given the increasing use of digital tools in CFL classrooms, there is a need to synthesize existing studies and clarify the analytic approaches underlying claims of improvement.

Accordingly, this review addresses the following research questions.

RQ1. What types of instructional interventions have been implemented to improve CFL speaking proficiency?

RQ2. How is speaking improvement operationalized and measured in these intervention studies?

RQ3. What forms of learning evidence are used to support claims of speaking development?

RQ4. What methodological gaps emerge in current ICT-supported CFL speaking research, and how might a Learning Analytics perspective address them?

2. Methods

2.1 Literature Search and Selection Process

Various approaches have been adopted in previous review studies to identify and select relevant literature. In the present review, the ERIC database was used as the primary source for locating peer-reviewed research on CFL speaking.

The initial literature search was conducted using the keywords “Chinese as a foreign language” and “speaking.” This search yielded 23 articles. A preliminary examination of these studies revealed that many addressed speaking-related constructs—such as strategy use, assessment frameworks, or linguistic features—but only a small number explicitly implemented instructional interventions aimed at improving speaking proficiency.

In order to broaden the scope while maintaining relevance to speaking development, a second search was conducted using the keywords “Chinese” and “speaking skill.” This search produced 46 additional records. The titles and abstracts of these studies were examined to determine their relevance to instructional efforts targeting speaking improvement.

Across both searches, duplicate records were removed. The remaining studies were then reviewed in greater detail. Articles were retained for inclusion if they met the following criteria:

1. The primary focus of the study was the improvement or development of CFL speaking proficiency.
2. An instructional intervention (e.g., curriculum implementation, training program, or technology-mediated activity) was described.
3. Empirical data were collected to evaluate speaking performance.
4. The study was published in a peer-reviewed academic venue.

Studies that focused solely on lexical processing, character recognition, assessment validation, strategy description, or cross-sectional proficiency comparisons—without implementing an instructional intervention—were excluded from the final sample.

After applying these criteria, 21 empirical studies were retained for analysis. These studies form the evidence base for the present review.

2.2 Analytic Framework

To enable systematic comparison across studies with diverse instructional designs and methodological approaches, the selected articles were examined along four analytic dimensions.

(1) instructional orientation,
(2) conceptualization of speaking proficiency,
(3) type of learning evidence, and
(4) analytic focus.

This analytic framework enabled comparison across intervention studies with different instructional designs and technological implementations, while foregrounding how speaking improvement was operationalized and evidenced. In particular, the framework allowed for examination of whether studies incorporated process-oriented and trace-based analytic approaches consistent with a Learning Analytics perspective.

3. Results

To provide a structured overview of the literature, the 21 selected studies were organized according to the primary type of evidence used to substantiate speaking development. Four evidence categories emerged as shown in Table 1. It should be noted that these categories are not mutually exclusive. Some studies incorporated multiple forms of evidence (e.g., both aggregated scores and feature-level analyses) and therefore appear across categories.

Table 1. *Summary of Evidence Categories in CFL Speaking Intervention Research*

Evidence Category	Focus of Study	Type of Evidence Used	Analytic Granularity
Category A: Outcome-Oriented Instructional Interventions	Instructional improvement of speaking	Pre–post test scores; rubric-based ratings; statistical comparisons	Aggregated performance measures
Category B: Feature-Level / Process-Oriented Studies	Linguistic or behavioral features in speech	Transcript-based analysis; feature counts; metacognitive reflection; disfluency analysis	Fine-grained but small-scale manual analysis
Category C: Strategy and Learner-Perception Studies	Speaking strategies and learner beliefs	Validated questionnaires; self-report surveys	Self-reported perceptual data
Category D: Assessment and Program-Level Studies	Speaking assessment frameworks; proficiency comparison	Standardized test scores; construct review; program comparison	Institutional-level performance indicators

Category A: Outcome-Oriented Instructional Interventions.

Studies implemented instructional interventions explicitly aimed at improving CFL speaking and evaluated effectiveness primarily through pre–post comparisons or aggregated speaking scores (Tan & Tan, 2010; Dai et al., 2024; Hill & Tschudi, 2011; Hsiao et al., 2016; Kelly & Klein, 2016; Wu et al., 2024; Phupunna, 2025). These interventions included task-based curricula, communicative training programs, blended learning models, diagnostic CALL systems, podcast-based learning, and audio blog-supported practice. Speaking improvement was typically operationalized as higher posttest performance, increased rubric scores, or statistically significant gains.

Category B: Feature-Level and Process-Oriented Studies.

A smaller group of studies examined specific aspects of speech production rather than relying solely on overall scores. These investigations analyzed disfluency patterns, self-repair

behaviors, discourse feature uptake, or metacognitive reflection during speaking tasks (Tan & Tan, 2010; Hill & Tschudi, 2011; Tsai & Chu, 2017; Liao, 2023). Although these studies provided more granular insights into speaking behavior, analyses were generally based on manual transcription and feature counting.

Category C: Strategy and Learner-Perception Studies.

Studies in this category focused on learners' reported speaking strategies, attitudes, confidence, and self-efficacy (Sun et al., 2016; Sae-thung & Boonsuk, 2022). These studies highlighted important cognitive and affective dimensions of speaking development but relied predominantly on self-report instruments, limiting their capacity to support process-oriented analysis.

Category D: Assessment and Program-Level Studies.

Another group examined speaking assessment frameworks or compared proficiency outcomes across instructional programs without targeting intervention-based improvement (Xu et al., 2015; Li, 2023). These studies contributed to understanding how speaking proficiency is measured at institutional levels but did not investigate developmental processes within specific instructional designs.

4. Discussion

This review addressed four research questions concerning instructional interventions for improving CFL speaking proficiency and the forms of evidence used to support claims of improvement. While the majority of reviewed studies were not exclusively ICT-based, the methodological patterns observed in technology-mediated contexts illuminate broader analytic gaps within CFL speaking research. The findings provide several insights.

RQ1: Types of Instructional Interventions

With regard to RQ1, the reviewed literature demonstrates that CFL speaking improvement has been pursued through multiple instructional orientations. These include task-based and curriculum-oriented interventions (Hill & Tschudi, 2011; Dai et al., 2024; Wu et al., 2024), communication-focused training programs (Phupunna, 2025), and ICT-mediated approaches such as diagnostic systems, podcasts, and audioblogs (Tan & Tan, 2010; Hsiao et al., 2016; Kelly & Klein, 2016). Although diverse in pedagogical design, these interventions share a common goal of enhancing communicative speaking performance through structured practice and guided feedback. The studies indicate that speaking improvement is an established, though relatively concentrated, area within CFL research.

RQ2: Conceptualization of Speaking Improvement

In response to RQ2, speaking improvement is most frequently conceptualized as an increase in overall communicative competence, typically operationalized through composite speaking scores or rubric-based performance ratings (Tan & Tan, 2010; Hsiao et al., 2016; Kelly & Klein, 2016; Wuttphan & Klinhom, 2022; Wu et al., 2024; Dai et al., 2024; Phupunna, 2025). In some cases, improvement is further specified in terms of discourse feature uptake or metacognitive engagement (Tan & Tan, 2010; Hill & Tschudi, 2011), or examined through disfluency and repair behavior (Tsai & Chu, 2017; Liao, 2023). Nevertheless, the dominant tendency is to treat speaking development as a measurable performance outcome rather than as a temporally unfolding process.

RQ3: Forms of Learning Evidence

Addressing RQ3, the primary forms of evidence across intervention studies consist of pre–post test comparisons and aggregated rating scales. Statistical analyses are commonly employed to establish significant gains following instructional implementation (Tan & Tan, 2010; Hsiao et al., 2016; Kelly & Klein, 2016; Wu et al., 2024; Phupunna, 2025). Feature-level analyses and transcript-based investigations appear less frequently and are typically limited

to small samples (Tan & Tan, 2010; Hill & Tschudi, 2011; Tsai & Chu, 2017 ; Liao, 2023). Strategy-based and assessment-oriented studies further contribute to the conceptual landscape of speaking proficiency (Xu et al., 2015; Sun et al., 2016; Sae-thung & Boonsuk, 2022; Li, 2023), yet they operate largely independently from intervention-based developmental evaluation.

RQ4: Methodological Gaps in ICT-Supported CFL Speaking Research

In relation to RQ4, the review reveals a notable imbalance between instructional innovation and analytic depth. Although ICT-supported interventions frequently report positive speaking outcomes (Tan & Tan, 2010; Hsiao et al., 2016; Kelly & Klein, 2016), the data generated within these digital environments are generally summarized into overall performance scores. Relatively few studies systematically examine how speaking behaviors evolve across instructional cycles or how feedback and task repetition influence developmental trajectories. Consequently, while the effectiveness of interventions is often demonstrated, the underlying processes of speaking growth remain less explicitly modeled.

Taken together, the findings indicate that CFL speaking research has established a foundation of intervention-based effectiveness, yet opportunities remain for more process-sensitive approaches to understanding development.

5. Conclusion

This review examined 21 empirical studies addressing instructional interventions for improving Chinese as a Foreign Language (CFL) speaking proficiency. The findings indicate that a range of pedagogical approaches—task-based curricula, communicative training programs, and ICT-mediated interventions—have been implemented with measurable success. Across these studies, statistically significant gains in speaking performance were frequently reported, demonstrating that structured instructional design can effectively support oral development.

At the same time, the review reveals a consistent pattern in how improvement is conceptualized and evidenced. Speaking development is most commonly represented through aggregated pre–post score comparisons and rubric-based performance ratings. Although a small number of studies incorporated feature-level or transcript-based analyses, such approaches remain limited in scale and are rarely integrated into longitudinal models of developmental change. As a result, the literature provides strong evidence of instructional effectiveness but comparatively limited insight into the mechanisms through which speaking proficiency evolves over time.

Given the increasing integration of digital tools in CFL speaking instruction, these findings highlight an opportunity for closer alignment between instructional innovation and analytic methodology. ICT-supported environments have the potential to generate rich interactional and behavioral data that could be leveraged to examine speaking development as a dynamic, process-oriented phenomenon. A Learning Analytics perspective—emphasizing trace-based modeling, temporal analysis, and feedback-sensitive design—may therefore offer a productive direction for future research. At the same time, this perspective should be more explicitly connected with broader discussions in second language acquisition, so that learner development can be examined not only as an outcome, but also as a dynamic and mediated process.

This review also has several limitations. First, the search was limited to studies indexed in the ERIC database. Second, the search strategy relied on a relatively limited set of keywords, which may have narrowed the scope of the studies retrieved. Future review studies should address these limitations by expanding database coverage and refining search strategies. Such efforts would help build a more comprehensive understanding of CFL speaking research and further clarify how Learning Analytics can contribute to the study of speaking development.

Acknowledgement

This research is supported by Okawa foundation and Smart Learning Design Research Unit, Kyushu University.

References

Dai, Z., Thipatdee, G., & Metjiranont, M. (2024). The development of a task-based Chinese speaking instructional model for CFL learners in Thailand. *Journal of Education and Learning, 13*(1), 161-171.

Duanmu, S. (2007). *The phonology of standard Chinese (2nd ed.).* Oxford University Press.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70-105.

Gong, Y., Lyu, B., & Gao, X. (2018). Research on teaching Chinese as a second or foreign language in and outside mainland China: A bibliometric analysis. *Asia-Pacific Education Researcher, 27*(4), 277–289.

Hill, Y. Z., & Tschudi, S. L. (2011). Exploring task-based curriculum development in a blended-learning conversational Chinese program. *International Journal of Virtual and Personal Learning Environments, 2*(1), 19-36.

Hsiao, H.-S., Chang, C.-S., Lin, C.-Y., Chen, B., Wu, C.-H., & Lin, C.-Y. (2016). The development and evaluation of listening and speaking diagnosis and remedial learning system. *British Journal of Educational Technology, 47*(2), 372-389.

Kelly, W. Q., & Klein, J. D. (2016). The effect of type of podcasts and learning styles on language proficiency and confidence. *Journal of Educational Technology Systems, 44*(4), 421-429.

Li, A. W. (2023). Assessing the speaking proficiency of L2 Chinese learners: Review of the Hanyu Shuiping Kouyu Kaoshi. *Language Testing, 40*(4), 1007-1021.

Li, C. N., & Thompson, S. A. (1976). Subject and topic: *A new typology of language. In C. N. Li (Ed.)*. Academic Press.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar.* University of California Press.

Liao, J. (2023). Disfluency and self-repair in presentational and interpersonal speech modalities. *Foreign Language Annals, 56*(2), 401-427.

Phupunna, S. (2025). Chinese communication training course for upgrading economy and tourism towards improving quality of life in secondary urban society, Thailand. *International Journal of Education & Literacy Studies, 13*(2), 485-491.

Sae-thung, J., & Boonsuk, Y. (2022). Chinese speaking strategies as a foreign language: Success stories from Thai higher education. *Anatolian Journal of Education, 7*(2), 157-172.

Sun, P. P., Zhang, L. J., & Gray, S. M. (2016). Development and validation of the speaking strategy inventory for learners of Chinese (SSILC) as a second/foreign language. *The Asia-Pacific Education Researcher, 25*(4), 593-604.

Tan, Y. H., & Tan, S.-C. (2010). A metacognitive approach to enhancing Chinese language speaking skills using audioblogs. *Australasian Journal of Educational Technology, 26*(7), 1075-1089.

Tsai, P., & Chu, W. (2017). The use of discourse markers among Mandarin Chinese teachers, and Chinese as a second language and Chinese as a foreign language learners. *Applied Linguistics, 38*(5), 638-665.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106*(6), 3649-3658.

Wuttphan, N., & Klinhom, N. (2022). The role of project-based learning in promoting preservice teachers' communicative capabilities. *Journal of Education and Learning, 12*(1), 52-60.

Wu, T., Jermtaisong, R., & Lhongsap, P. (2024). The training curriculum development based on the language teaching for communication approach to develop basic Chinese speaking for graduate diploma program in teaching profession students. *The Turkish Online Journal of Educational Technology, 23*(4), 222-229.

Xu, X., Padilla, A. M., & Silva, D. M. (2015). Learner performance in Mandarin immersion and high school world language programs: A comparison. Foreign Language Annals, 48(1), 26-38.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency and accuracy in L2 oral production. *Applied Linguistics, 24*(1), 1-27.