

Developing Strategic Competence through Multimodal Strategy-Focused Instruction in EFL Speaking Classrooms

Feng Hongtu^{1*}, Nur Salwa Binti Abd. Wahid², Gao Ying³

¹Language Academy, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

²Language Academy, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

³School of Chemistry and Chemical Engineering, Yulin University, Yulin, China

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.100800523>

Received: 24 August 2026; Accepted: 29 August 2026; Published: 10 September 2026

ABSTRACT

Strategic competence enables EFL learners to sustain spoken communication when linguistic resources are insufficient, yet it remains unclear whether multimodal strategy-focused instruction offers benefits beyond conventional explicit communication strategy instruction. This study examined the effects of multimodal strategy-focused instruction on communication strategy use among Chinese first-year English majors and explored learners' perceptions of the instructional approach. A sequential mixed-methods design was employed. The final sample comprised 30 students: the experimental group (EG; $n = 16$) and the control group (CG; $n = 14$). Over eight weeks, both groups received explicit instruction in the same eight communication strategies; the EG received systematic multimodal support through visual, auditory, linguistic, and gestural resources, whereas the CG received conventional strategy instruction. Communication strategy use was assessed through pre- and post-test speaking tasks and analysed using descriptive statistics, paired-samples t -tests, independent-samples t -tests, and Cohen's d . Semi-structured interviews with eight EG participants were analysed thematically. Significant increases in fillers and hesitation devices and self-initiated repair were found in both groups. At post-test, the EG showed significantly higher use of circumlocution, meaning negotiation strategies, approximation, and non-linguistic means, while no significant between-group differences were found for fillers and hesitation devices, self-initiated repair, restructuring, or comprehension checks. Interview findings indicated that learners generally perceived multimodal strategy-focused instruction as engaging, cognitively supportive, and conducive to contextualised communication, while also identifying linguistic, affective, and practice-related constraints. Overall, the findings suggest that explicit strategy instruction can support strategic development under both instructional conditions, while systematic multimodal integration may provide additional benefits for strategies involving flexible meaning construction, interaction, and embodied communication.

Keywords: strategic competence, communication strategies, multimodal strategy-focused instruction, EFL speaking, English majors

INTRODUCTION

With the increasing demand for English communication skills in globalised educational and professional

contexts, developing learners' speaking ability has become an important objective of English as a Foreign Language (EFL) education. However, despite years of formal English learning, many university students in China still encounter difficulties in engaging in effective spoken interaction, particularly when they need to express ideas spontaneously, respond to interlocutors, or manage unexpected communication problems (Wang & Soo, 2024). These challenges suggest that successful oral communication involves more than the mastery of grammatical knowledge and vocabulary; learners also need the ability to manage communication effectively in authentic interactional contexts. Although EFL instruction has increasingly emphasised communicative development, many teaching contexts continue to place considerable attention on linguistic accuracy and examination performance, with comparatively less focus on learners' strategic abilities for managing spoken communication.

Among the components of communicative competence, strategic competence plays a particularly important role in helping learners sustain spoken interaction when linguistic resources are insufficient. Canale and Swain (1980) conceptualised strategic competence as the ability to employ verbal and non-verbal strategies to compensate for communication difficulties, while Celce-Murcia et al. (1995) further emphasised the role of communication strategies in maintaining interaction and achieving communicative goals. In EFL speaking contexts, learners may encounter difficulties when they cannot immediately access appropriate vocabulary, formulate intended meanings, or respond effectively under real-time processing demands. In such situations, strategies such as circumlocution, approximation, meaning negotiation strategies, and non-linguistic means can help learners compensate for linguistic limitations and sustain interaction (Dörnyei & Scott, 1997). Despite the recognised importance of these strategies, explicit communication strategy instruction has not always been systematically incorporated into EFL speaking classrooms, indicating a need for instructional approaches that provide learners with greater opportunities to develop and practise strategic competence (Nakatani, 2010).

In recent years, multimodal approaches have received increasing attention in language education for their potential to support interaction and meaning-making through the integration of multiple semiotic resources, including linguistic, visual, auditory, and gestural modes (Kress, 2010). In EFL speaking classrooms, the combination of these resources may provide learners with richer communicative support and more varied opportunities to participate in spoken interaction. Recent research has also indicated that multimodal pedagogies can contribute to English language development across different skills and support learners' communicative abilities in higher education contexts (Rahmanu & Molnár, 2024). These affordances suggest that multimodal approaches may be particularly relevant to speaking instruction, where learners need to draw on multiple resources to construct meaning and sustain interaction.

Although communication strategy research has provided substantial insights into the types and patterns of strategies used by EFL learners, classroom-based studies have also shown that explicit strategy instruction can improve learners' strategy use and communicative performance (Caprario, 2025; Milliner & Dimoski, 2024; Nakatani, 2010; Peker & Erdemir, 2021). However, comparatively less is known about instructional approaches that integrate communication strategy development with multimodal resources. Recent review evidence on multimodal English language learning in higher education has primarily synthesised outcomes related to language-skill development and communicative ability (Rahmanu & Molnár, 2024), leaving the specific relationship between multimodal instruction and communication strategy development less clearly established. Within the literature reviewed for the present study, empirical research examining multimodal strategy-focused instruction in Chinese university EFL speaking classrooms remains limited.

To address these gaps, the present study investigates the effects of multimodal strategy-focused instruction on the development of strategic competence among Chinese first-year English majors in EFL speaking classrooms. Specifically, the study examines learners' use of eight communication strategies derived from Celce-Murcia et

al. (1995): fillers and hesitation devices, self-initiated repair, circumlocution, approximation, non-linguistic means, meaning negotiation strategies, comprehension checks, and restructuring. Using a sequential mixed-methods design, the study combines quantitative evidence of changes in learners' communication strategy use with qualitative insights into their perceptions of the instructional approach. In doing so, the study seeks to examine whether multimodal strategy-focused instruction can support learners' strategic competence development and how learners perceive its role in their spoken communication.

Accordingly, the study addressed the following research questions:

RQ1. What effects does multimodal strategy-focused instruction have on Chinese EFL learners' use of communication strategies?

RQ2. How do learners perceive multimodal strategy-focused instruction and its role in developing their strategic competence during English speaking activities?

LITERATURE REVIEW

Strategic Competence and Communication Strategies

Strategic competence has long been recognised as an important component of communicative competence in second language learning. Canale and Swain (1980) conceptualised strategic competence as learners' ability to employ verbal and non-verbal strategies to compensate for communication difficulties and enhance communicative effectiveness. Building on this view, Celce-Murcia et al. (1995) further emphasised the role of strategic resources in maintaining interaction, negotiating meaning, and overcoming limitations in learners' linguistic repertoires. From this perspective, strategic competence enables learners to draw on available resources to sustain interaction when their linguistic knowledge is insufficient for expressing intended meanings.

Communication strategies can be viewed as observable manifestations of strategic competence during spoken interaction. Celce-Murcia et al. (1995) proposed a range of strategic competence components, including achievement or compensatory strategies, stalling or time-gaining strategies, self-monitoring strategies, and interactional strategies. Drawing on this framework, the present study focuses on eight communication strategies: circumlocution, approximation, non-linguistic means, restructuring, fillers and hesitation devices, self-initiated repair, meaning negotiation strategies, and comprehension checks. These strategies represent different ways in which learners may compensate for linguistic limitations, gain planning time, repair problematic utterances, negotiate meaning, check understanding, and maintain interaction when communication difficulties arise.

Although communication strategies may emerge naturally during second language interaction, learners may not always use them effectively or systematically without instructional support. Strategy-focused instruction can therefore play an important role in raising learners' awareness of available communicative resources and providing opportunities to practise their use in spoken interaction. Nakatani (2010) showed that strategy-oriented instruction could facilitate learners' use of oral communication strategies. More recent classroom-based research has similarly demonstrated the pedagogical value of explicit communication strategy instruction. Peker and Erdemir (2021), for example, found that compensation strategy instruction helped EFL learners develop communicative competence in speaking activities, while Milliner and Dimoski (2024) reported that communication strategy training improved task accuracy and communicative efficiency in information-gap speaking tasks. More recently, Caprario (2025) provided further evidence that explicit communication strategy instruction can facilitate learners' strategic and communicative development. Taken together, these findings support the pedagogical value of explicit strategy instruction, awareness raising, and strategy-based practice.

Existing intervention studies, however, have generally approached communication strategies primarily from the perspective of strategy training rather than explicitly examining the instructional process through a multimodal framework. Some interventions have incorporated resources such as video modelling and interactive speaking tasks (Milliner & Dimoski, 2024), demonstrating that strategy instruction may already involve resources beyond verbal explanation alone. Nevertheless, how linguistic, visual, auditory, and embodied resources can be systematically integrated with communication strategy instruction remains less clearly understood. This issue provides an important point of connection between research on strategic competence and the growing body of work on multimodal language pedagogy.

Multimodal Approaches in EFL Speaking Instruction

Multimodal approaches have received increasing attention in language education because meaning-making is not limited to linguistic forms but involves the coordinated use of multiple semiotic resources. From a social semiotic perspective, multimodality concerns the ways in which meaning is constructed through different modes, including language, images, sound, gesture, and other visual or embodied resources (Kress, 2010). This perspective is particularly relevant to EFL speaking instruction because spoken interaction involves not only verbal language but also contextual and non-verbal resources that may assist learners in expressing and interpreting meaning.

Different modes may provide complementary support during spoken interaction. Visual materials such as images and videos can provide contextual information and prompts for meaning construction, while auditory resources can offer models of spoken language. Gestures, facial expressions, and other embodied resources may also contribute to communication when linguistic expression alone is insufficient (Wang et al., 2023). Rather than functioning as isolated teaching aids, these resources can be combined within speaking activities to broaden the means through which learners construct, convey, and interpret meaning. This view is consistent with recent evidence showing the increasing use of digital and non-verbal modes, including gesture, facial expression, gaze, and other multimodal resources, in higher education English learning (Rahmanu & Molnár, 2024).

Such multimodal affordances may be particularly relevant to strategic competence. Several of the communication strategies examined in the present study involve resources and interactional processes that extend beyond the production of linguistic forms alone. Non-linguistic means directly involve embodied resources, while meaning negotiation strategies and comprehension checks require learners to respond dynamically to interlocutors and contextual cues. From this perspective, multimodal speaking environments may provide learners with additional opportunities to draw on available communicative resources when difficulties arise. However, this theoretical connection does not in itself demonstrate that multimodal instruction leads to greater communication strategy use; the relationship requires empirical investigation.

Recent research has provided growing evidence for the pedagogical value of multimodal approaches in English language education (Lee & Azman, 2022; Rahmanu & Molnár, 2024). Rahmanu and Molnár's (2024) systematic review of 34 studies in higher education found that multimodal pedagogies had been applied across vocabulary, reading, writing, and speaking development and were associated with learners' communicative ability. More specifically, Chen (2025) reported improvements in verbal fluency and visual-resource use following a multimodal skill-building intervention in English oral presentation instruction. However, the comparative benefits of multimodal input do not appear to be uniform. Pu and Chang (2025), for example, found that both bimodal and multimodal input supported EFL speaking development, while no significant between-group differences were observed in overall speaking performance or learning attitudes. Taken together, these studies suggest that multimodal pedagogy can support different dimensions of language and communicative development, although its advantages may vary across instructional conditions and learning outcomes.

Nevertheless, existing evidence has largely addressed language skills and broader communicative outcomes rather than learners' use of communication strategies as a specific manifestation of strategic competence.

Research Gap

Taken together, previous research provides evidence for both the teachability of communication strategies and the pedagogical potential of multimodal language instruction. Communication strategy studies have demonstrated benefits of explicit instruction and strategy-based practice (Milliner & Dimoski, 2024; Nakatani, 2010; Peker & Erdemir, 2021), while multimodal research has shown that integrating multiple semiotic resources can support language learning and communicative development (Chen, 2025; Rahmanu & Molnár, 2024). However, these two lines of research have largely developed separately, and the specific relationship between multimodal instruction and communication strategy development remains less clearly established.

Furthermore, within the literature reviewed for the present study, relatively limited empirical attention has been given to multimodal strategy-focused instruction in Chinese university EFL speaking classrooms. In particular, further classroom-based evidence is needed to examine whether the systematic integration of multimodal resources and communication strategy instruction can support learners' strategic competence, as reflected in their observable communication strategy use. Accordingly, the present study investigates multimodal strategy-focused instruction among Chinese first-year English majors and combines quantitative evidence of changes in communication strategy use with qualitative evidence of learners' perceptions of the instructional approach.

METHODOLOGY

Research Design

This study employed a sequential mixed-methods design to investigate the development of strategic competence through multimodal strategy-focused instruction. The quantitative phase examined changes in learners' communication strategy use, treated as an observable indicator of strategic competence, while the subsequent qualitative phase explored learners' perceptions of the instructional approach.

The quantitative component adopted a quasi-experimental pre-test–post-test control-group design. Participants came from one intact class of 40 first-year English majors that had already been divided into two teaching groups of 20 under the university's small-group spoken English arrangement. The two intact groups were subsequently randomly assigned at the group level to the experimental group (EG) and control group (CG).

Both groups were taught by the same instructor, received the same amount of instructional time, and focused on the same target communication strategy each week. The EG received multimodal strategy-focused instruction, whereas the CG received conventional strategy instruction. Pre- and post-test speaking assessments were administered before and after the eight-week intervention. Following the quantitative phase, eight EG participants took part in semi-structured interviews.

Participants

The study was conducted at a university in northwestern China. Initially, 40 first-year English majors participated, with 20 students in each teaching group.

Participant attrition occurred during the study. In the EG, one student did not complete the pre-test because of illness, while three students who completed the pre-test subsequently transferred to other majors and did not complete the post-test. In the CG, six students completed the pre-test but transferred to other majors before the

post-test. Students without complete pre- and post-test data were excluded from the final analysis.

The final analytic sample therefore consisted of 30 participants, including 16 in the EG and 14 in the CG. All included participants completed the eight-week intervention and both the pre- and post-test assessments. They were native speakers of Chinese, aged 18–19 years, and had received several years of formal English instruction before entering university.

Instructional Intervention

The intervention lasted eight weeks and was embedded in the participants' regular spoken English course. Each group attended one weekly class consisting of two 50-minute sessions. The first session followed the university's prescribed spoken English syllabus and was identical for both groups. The second session focused on explicit communication strategy instruction.

Both groups were explicitly taught the same eight strategies, adapted from Celce-Murcia et al. (1995), in the following sequence: fillers and hesitation devices, self-initiated repair, circumlocution, approximation, non-linguistic means, meaning negotiation strategies, comprehension checks, and restructuring. One strategy served as the principal focus each week, although previously introduced strategies could recur in later activities.

The main difference concerned instructional delivery. In the EG, visual, auditory, linguistic, and gestural resources were systematically incorporated into strategy explanation, modelling, guided practice, communicative tasks, and reflection. Materials included video clips, images, audio, visual prompts, gestures, role plays, and interactive speaking activities. In the CG, the same strategies were taught explicitly through conventional strategy instruction without systematic multimodal integration.

Several procedures were used to support intervention fidelity across the eight-week instructional period. The same course instructor taught both groups, and detailed lesson plans and weekly activity guidelines were prepared in advance. Each week followed a predefined instructional sequence specifying the instructional objectives, target communication strategy, classroom activities, and time allocation. Implementation was monitored through regular researcher–instructor discussions and the instructor's weekly reflective records, which were used to document lesson delivery and any instructional adjustments. These procedures were intended to maintain consistency of implementation while preserving the planned distinction between the two instructional conditions.

Table 1. Overview of the Instructional Conditions

Component	EG	CG
Duration	8 weeks	8 weeks
Weekly class	2 × 50 min	2 × 50 min
First 50 min	Regular syllabus instruction	Regular syllabus instruction
Second 50 min	Multimodal strategy-focused instruction	Conventional strategy instruction
Target strategy	Same weekly strategy	Same weekly strategy
Strategy instruction	Explicit	Explicit
Instructor	Same	Same
Multimodal integration	Systematic	Not systematic

Instruments

Two instruments were used for data collection: a speaking-based strategic competence assessment for the quantitative component and semi-structured interviews for the qualitative component.

Speaking-Based Strategic Competence Assessment

The assessment was designed to elicit learners' observable use of the eight target communication strategies. It consisted of a concept-description task and IELTS-based speaking tasks.

The concept-description task was adapted from Mei and Nathalang (2010). Thirty-two concrete and 32 abstract concepts were selected from their task materials and organised into parallel pre- and post-test sets. At each assessment stage, participants randomly selected one concrete and one abstract concept and described them to the researcher for approximately two to three minutes. The task was particularly suited to eliciting strategies such as circumlocution, approximation, and non-linguistic means.

The second component was modelled on IELTS Speaking Parts 2 and 3. Parallel pre- and post-test versions were used to reduce practice effects. In Part 2, participants received one minute of preparation time and then spoke for one to two minutes. Part 3 involved a three- to four-minute follow-up discussion with the researcher. These tasks provided opportunities for extended production and interaction and were suited to eliciting strategies such as self-initiated repair, restructuring, meaning negotiation strategies, and comprehension checks.

The same researcher administered both the pre- and post-test assessments to the two groups following identical procedures. All performances were video-recorded and transcribed verbatim. Fillers, pauses, repairs, and communicatively meaningful non-verbal behaviours were retained or annotated for subsequent coding.

Semi-Structured Interviews

Following the intervention, eight EG participants were selected through stratified purposive sampling to capture variation in classroom participation, engagement, and strategic performance. The purpose was not statistical representativeness but to obtain a range of learner perspectives on the multimodal instructional experience. The interviews explored learners' perceptions of multimodal strategy-focused instruction, their experiences of learning and applying communication strategies, and perceived benefits and challenges.

Interviews were conducted in Chinese, lasted approximately 30–60 minutes, and were audio-recorded and transcribed verbatim. Quotations reported in English were translated using a meaning-oriented approach, with minor linguistic adjustments for clarity without altering participants' intended meanings. Participants were also invited to clarify their responses where necessary, and member checking was used to enhance credibility.

Prior to the main study, the assessment procedures were piloted over a two-week period with 10 first-year English majors from a university with a comparable educational context. These students were not included in the main study. Participants completed both the concept-description and IELTS-based speaking tasks, and the resulting recordings were transcribed and preliminarily coded to examine whether the tasks elicited the intended range of communication strategies and whether the coding procedures were feasible and sufficiently clear. In addition, the instructional design and sample materials were independently reviewed by two university academics with expertise in applied linguistics and English language teaching. Feedback from the pilot and expert review informed minor refinements to task instructions, instructional materials, and coding definitions before the main study.

Data Analysis

For the quantitative analysis, all speaking performances were transcribed verbatim and relevant non-verbal behaviours were annotated. Communication strategy use was coded using an a priori framework adapted from Celce-Murcia et al. (1995), comprising circumlocution, approximation, non-linguistic means, restructuring, fillers and hesitation devices, self-initiated repair, meaning negotiation strategies, and comprehension checks. Each identifiable instance was counted as one occurrence, and the frequency of each strategy was calculated for each participant at pre- and post-test.

Approximately 20% of the transcripts were randomly selected and independently coded by a second trained coder. Inter-rater agreement, assessed using Cohen's kappa, was $\kappa = .87$. Disagreements were resolved through discussion before the dataset was finalised.

Means and standard deviations were calculated to summarise communication strategy use. Baseline comparability between the final EG and CG samples was examined using Welch's *t*-tests to accommodate the unequal group sizes, with Hedges' *g* reported as the corresponding effect-size estimate. Paired-samples *t*-tests examined pre-test–post-test changes within each group, while independent-samples *t*-tests compared the groups at post-test. Cohen's *d* was calculated to estimate the magnitude of the observed differences. Analyses were conducted in SPSS Version 27 with the significance level set at $p < .05$.

Interview data were analysed thematically following Braun and Clarke (2006). After repeated familiarisation with the transcripts, initial codes were generated inductively and subsequently compared, refined, and organised into broader themes and subthemes. Themes were iteratively reviewed for coherence and distinctiveness. ATLAS.ti Version 8.4.4 was used to support data organisation and coding.

Ethical Considerations

Ethical approval was obtained from the relevant institutional authorities before data collection. Participants received information about the study and provided written informed consent, including consent for audio- and video-recording. Participation was voluntary, and students could withdraw without penalty or consequences for their academic standing.

Identifying information was removed from the dataset, participant codes were used during analysis and reporting, and all recordings were securely stored for research purposes only.

RESULTS

Changes in Communication Strategy Use

Pre-test comparisons indicated no statistically significant between-group differences across any of the eight target communication strategies (all $p > .05$), with generally small Hedges' *g* values, suggesting that the final EG and CG samples were broadly comparable in strategic competence at baseline. Table 2 presents the descriptive statistics for the eight communication strategies used by the two groups at the pre- and post-test stages.

Table 2. Descriptive Statistics for Communication Strategy Use in the EG and CG

Communication Strategy	EG Pre-test <i>M (SD)</i>	EG Post-test <i>M (SD)</i>	CG Pre-test <i>M (SD)</i>	CG Post-test <i>M (SD)</i>
Fillers and hesitation devices	63.00 (33.56)	80.00 (37.31)	64.93 (33.79)	79.57 (37.40)
Self-initiated repair	6.25 (5.11)	7.81 (6.26)	7.00 (4.98)	8.29 (5.90)
Circumlocution	3.12 (1.78)	6.88 (3.20)	2.93 (1.54)	2.79 (1.63)
Meaning negotiation strategies	2.31 (3.07)	4.31 (4.30)	1.93 (3.08)	1.71 (1.90)
Restructuring	0.19 (0.40)	0.12 (0.34)	0.07 (0.27)	0.00 (0.00)
Comprehension checks	0.00 (0.00)	0.19 (0.40)	0.07 (0.27)	0.00 (0.00)
Approximation	1.19 (1.56)	2.94 (3.55)	1.07 (1.38)	0.71 (0.99)
Non-linguistic means	0.19 (0.40)	1.00 (0.52)	0.21 (0.43)	0.00 (0.00)

As shown in Table 2, the two groups displayed different patterns of change across the eight communication strategies. In the EG, increases were observed in six strategies: fillers and hesitation devices, self-initiated repair, circumlocution, meaning negotiation strategies, approximation, and non-linguistic means. Circumlocution increased from 3.12 to 6.88, meaning negotiation strategies from 2.31 to 4.31, approximation from 1.19 to 2.94, and non-linguistic means from 0.19 to 1.00. By contrast, restructuring decreased slightly from 0.19 to 0.12, while comprehension checks remained infrequent despite increasing from 0.00 to 0.19.

The CG showed a different pattern. Fillers and hesitation devices increased from 64.93 to 79.57, and self-initiated repair increased from 7.00 to 8.29. However, the remaining strategies showed little change or decreased over the intervention period. Circumlocution decreased slightly from 2.93 to 2.79, meaning negotiation strategies from 1.93 to 1.71, approximation from 1.07 to 0.71, and non-linguistic means from 0.21 to 0.00. Restructuring and comprehension checks were also rarely observed at either assessment stage.

Descriptively, therefore, increases in fillers and hesitation devices and self-initiated repair were evident in both groups, whereas the more distinct EG–CG differences were observed in circumlocution, meaning negotiation strategies, approximation, and non-linguistic means. These patterns were examined further through inferential statistical analyses.

Inferential Analysis of Communication Strategy Use

To determine whether the descriptive changes in communication strategy use were statistically significant, paired-samples *t*-tests were first conducted to examine pre-test–post-test changes within the EG and CG. Independent-samples *t*-tests were then used to compare the two groups at the post-test stage.

Within-Group Comparisons

Table 3 presents the results of the paired-samples t -tests for the EG and CG.

Table 3. Within-Group Changes in Communication Strategy Use

Communication Strategy	EG t	p	Cohen's d	CG t	p	Cohen's d
Fillers and hesitation devices	6.29	< .001	1.57	5.87	< .001	1.57
Self-initiated repair	4.16	.001	1.04	4.22	.001	1.13
Circumlocution	9.30	< .001	2.33	-1.47	.165	-0.39
Meaning negotiation strategies	5.16	< .001	1.29	-0.51	.620	-0.14
Restructuring	-1.00	.333	-0.25	-1.00	.336	-0.27
Comprehension checks	1.86	.083	0.47	-1.00	.336	-0.27
Approximation	3.47	.003	0.87	-2.69	.019	-0.72
Non-linguistic means	8.06	< .001	2.02	-1.88	.082	-0.50

In the EG, significant increases were observed in six of the eight communication strategies: fillers and hesitation devices ($t = 6.29, p < .001, d = 1.57$), self-initiated repair ($t = 4.16, p = .001, d = 1.04$), circumlocution ($t = 9.30, p < .001, d = 2.33$), meaning negotiation strategies ($t = 5.16, p < .001, d = 1.29$), approximation ($t = 3.47, p = .003, d = 0.87$), and non-linguistic means ($t = 8.06, p < .001, d = 2.02$). Restructuring showed no significant change ($p = .333$), and the increase in comprehension checks did not reach statistical significance ($p = .083$).

The CG showed a more limited pattern of change. Significant increases were found in fillers and hesitation devices ($t = 5.87, p < .001, d = 1.57$) and self-initiated repair ($t = 4.22, p = .001, d = 1.13$). Circumlocution, meaning negotiation strategies, restructuring, comprehension checks, and non-linguistic means showed no significant pre-test–post-test changes. Approximation, however, decreased significantly ($t = -2.69, p = .019, d = -0.72$).

These results indicate that fillers and hesitation devices and self-initiated repair developed under both instructional conditions, whereas significant gains in circumlocution, meaning negotiation strategies, approximation, and non-linguistic means were observed only in the EG.

Between-Group Comparison at Post-test

To further examine differences between the two instructional conditions, independent-samples t -tests were conducted using the post-test scores. The results are presented in Table 4.

Table 4. Post-test Comparison Between the EG and CG

Communication Strategy	EG M	CG M	t	p	Cohen's d
Fillers and hesitation devices	80.00	79.57	0.03	.975	0.01
Self-initiated repair	7.81	8.29	-0.21	.833	-0.08

Circumlocution	6.88	2.79	4.49	< .001	1.61
Meaning negotiation strategies	4.31	1.71	2.19	.040	0.78
Restructuring	0.12	0.00	1.46	.164	0.52
Comprehension checks	0.19	0.00	1.86	.083	0.66
Approximation	2.94	0.71	2.40	.028	0.85
Non-linguistic means	1.00	0.00	7.75	< .001	2.74

The post-test comparisons revealed significant between-group differences in four communication strategies. The EG showed significantly higher use of circumlocution ($t = 4.49, p < .001, d = 1.61$), meaning negotiation strategies ($t = 2.19, p = .040, d = 0.78$), approximation ($t = 2.40, p = .028, d = 0.85$), and non-linguistic means ($t = 7.75, p < .001, d = 2.74$). The largest between-group differences were observed for non-linguistic means and circumlocution.

No significant post-test differences were found for fillers and hesitation devices, self-initiated repair, restructuring, or comprehension checks. In particular, the almost identical post-test means for fillers and hesitation devices (EG: $M = 80.00$; CG: $M = 79.57$) and the similar means for self-initiated repair (EG: $M = 7.81$; CG: $M = 8.29$) indicate that improvements in these two strategies were not specific to the multimodal instructional condition.

Taken together, the inferential results show that multimodal strategy-focused instruction was associated with greater development than conventional strategy instruction in circumlocution, meaning negotiation strategies, approximation, and non-linguistic means. However, this advantage was not observed consistently across all eight target strategies.

Learners' Perceptions of Multimodal Strategy-Focused Instruction

Semi-structured interviews with eight participants from the EG provided further insight into how learners experienced multimodal strategy-focused instruction. The thematic analysis generated six interrelated themes: engagement and participation, perceived development of strategic competence, cognitive support and learning efficiency, authenticity and contextualised learning, challenges and constraints, and suggestions for improvement. Overall, participants evaluated the instructional approach positively, although they also identified several factors that constrained its effectiveness. The major themes are summarised in Table 5.

Table 5. Themes Emerging from the Semi-Structured Interviews

Theme	Main Perceptions
Engagement and participation	Greater interest, anticipation, and classroom interaction
Perceived development of strategic competence	Greater willingness to speak and more flexible ways of expressing meaning
Cognitive support and learning efficiency	Improved attention, comprehension, memorability, and idea generation

Authenticity and contextualised learning	More realistic, contextualised, and learner-centred speaking experiences
Challenges and constraints	Linguistic limitations, task-related difficulties, limited transfer, and persistent anxiety or low confidence
Suggestions for improvement	More practice time, clearer scaffolding and modelling, and richer multimodal resources

A prominent perception concerned greater engagement and participation. Participants described videos, audio materials, role plays, games, and interactive activities as making speaking lessons more lively and encouraging greater involvement. For example, P4 commented that the approach could make “the class more lively and the atmosphere more active.” Learners also reported greater anticipation of multimodal activities and perceived the classroom as more interactive than lessons relying predominantly on verbal explanation. These responses suggest that multimodal strategy-focused instruction created conditions in which learners felt more willing to participate in speaking activities and experiment with different communicative resources.

Participants also perceived changes in their strategic expression. They reported becoming more aware that communication did not need to stop when precise vocabulary was unavailable and described greater willingness to use alternative ways of conveying meaning. As P5 explained, “When you come across a word you can’t express, being able to describe it in other ways is useful.” Learners also referred to using fillers to gain planning time and maintain speech. These accounts correspond with learners’ broader perception that multimodal activities provided additional resources for managing communication difficulties rather than requiring exclusive reliance on accurate linguistic forms.

A related theme concerned cognitive support and contextualised learning. Learners reported that visual, auditory, and contextual information helped them maintain attention, understand speaking content, remember classroom material, and generate ideas more efficiently. The interview analysis also showed that learners associated multimodal instruction with more vivid and situationally meaningful communication and perceived some tasks as closer to real-life speaking than conventional textbook routines. Thus, multimodal resources were valued not simply because they made lessons more interesting, but because participants perceived them as helping them organise meaning and participate in communication within recognisable contexts.

At the same time, learners’ perceptions were not uniformly positive. Participants continued to report difficulties arising from limited linguistic resources, task design, insufficient opportunities for out-of-class practice, and persistent anxiety or low confidence. For example, P7 commented, “Sometimes I really don’t know how to express it in English.” This illustrates that multimodal support did not eliminate the linguistic challenges associated with spontaneous EFL speaking. Learners also suggested that some activities would benefit from more practice time, clearer teacher modelling and scaffolding, and richer or more relevant multimodal resources.

Taken together, the interview findings indicate that learners generally perceived multimodal strategy-focused instruction as supporting their engagement, strategic expression, cognitive processing, and experience of contextualised communication. However, these perceived benefits were moderated by learners’ linguistic resources, confidence, task design, and opportunities for continued practice. The qualitative findings therefore provide a more differentiated account of the instructional experience than a uniformly positive evaluation of multimodal instruction.

DISCUSSION

The present study examined whether multimodal strategy-focused instruction was associated with greater development in communication strategy use than conventional strategy instruction and explored learners' perceptions of the multimodal instructional experience. The findings reveal a differentiated pattern rather than a uniform advantage for multimodal instruction. Both groups demonstrated significant gains in fillers and hesitation devices and self-initiated repair, whereas the EG showed additional advantages in circumlocation, meaning negotiation strategies, approximation, and non-linguistic means. Restructuring and comprehension checks, by contrast, remained infrequent and showed no significant between-group differences. These findings suggest that the effects of multimodal strategy-focused instruction were dependent, at least in part, on the type of strategy being developed.

Differential Effects of Multimodal Strategy-Focused Instruction

One important finding is that fillers and hesitation devices and self-initiated repair developed significantly under both instructional conditions. Because both the EG and CG received explicit instruction in the same target communication strategies, these shared gains cannot be attributed specifically to multimodal resources. Rather, they suggest that explicit strategy instruction, awareness raising, and repeated speaking practice may themselves support the development of strategies that learners can readily apply during oral production. This interpretation is consistent with previous research showing that communication strategies can be developed through explicit instruction and strategy-oriented practice (Caprario, 2025; Milliner & Dimoski, 2024; Nakatani, 2010; Peker & Erdemir, 2021).

This shared development is important for interpreting the contribution of the multimodal condition. The present study did not compare strategy instruction with no strategy instruction; it compared multimodal strategy-focused instruction with conventional strategy instruction. The absence of significant post-test differences in fillers and hesitation devices and self-initiated repair therefore indicates that systematic multimodal integration did not provide an additional advantage for every strategy. Instead, its added value appeared to be concentrated in circumlocation, meaning negotiation strategies, approximation, and non-linguistic means, for which the EG demonstrated significantly greater post-test use.

The pattern may be partly explained by the relationship between these strategies and the multimodal resources used during instruction. Circumlocation and approximation require learners to find alternative routes to meaning when precise lexical items are unavailable. Visual prompts, images, and contextual information can make relevant semantic features more salient and provide learners with material from which to construct descriptions or substitute expressions. In this sense, multimodal support may have broadened the resources available for meaning construction rather than requiring learners to depend exclusively on immediate lexical retrieval. This interpretation is compatible with Kress's (2010) social semiotic perspective, which views meaning-making as involving the coordination of multiple semiotic resources rather than language alone.

The strong development of non-linguistic means is also readily interpretable in relation to the instructional design. Gesture and other embodied resources are integral to this strategy, and the EG was explicitly exposed to gestural modelling and activities in which linguistic and non-linguistic resources could be used together. The instructional mode was therefore closely aligned with the communicative behaviour being targeted. Recent multimodal research similarly emphasises that gesture, visual information, and other non-verbal modes can contribute meaningfully to language learning and communicative performance (Rahmanu & Molnár, 2024; Wang et al., 2023). The present findings extend this argument by suggesting that multimodal support may be especially

relevant when the target communication strategy itself draws directly on non-verbal or semiotic resources.

Meaning negotiation strategies also showed a significant advantage for the EG. Unlike purely individual compensatory strategies, negotiation is inherently interactional because learners must recognise communication difficulty and respond to an interlocutor. The multimodal condition included role plays, collaborative activities, and other interactive tasks that created opportunities for learners to clarify, respond, and adjust meaning during communication. Such activities may have made negotiation strategies functionally necessary rather than leaving them as strategies that learners merely understood conceptually. This interpretation accords with research emphasising the importance of interaction and strategy-oriented communicative practice in communication strategy development (Milliner & Dimoski, 2024; Nakatani, 2010).

However, restructuring and comprehension checks did not show comparable development. Comprehension checks were particularly infrequent. In the CG, minimal use was observed at pre-test ($M = 0.07$, $SD = 0.27$), whereas no occurrences were recorded at post-test ($M = 0.00$, $SD = 0.00$), resulting in zero post-test variance. This pattern may partly reflect the interaction-dependent nature of comprehension checks, which are typically elicited when speakers perceive a need to verify interlocutor understanding; such opportunities may not arise consistently across structured speaking tasks. Their low frequency suggests that awareness of a strategy does not necessarily translate into spontaneous use during real-time interaction. Both require learners to monitor ongoing communication and decide rapidly whether an utterance should be reformulated or whether understanding should be explicitly checked. An eight-week intervention may have provided initial awareness and practice without being sufficient for these strategies to become readily available during spontaneous speech. This interpretation should nevertheless remain cautious, because the present study did not directly measure the cognitive complexity or proceduralisation of individual strategies. The low frequency of comprehension checks should therefore not be interpreted as evidence that learners were unable to use the strategy.

Taken together, the quantitative findings suggest that the value of multimodal strategy-focused instruction was strategy-sensitive rather than uniform. Explicit strategy instruction appeared beneficial under both conditions, while systematic multimodal integration was associated with additional development in strategies closely connected to flexible meaning construction, interaction, and embodied communication. This differentiated pattern is broadly consistent with recent evidence indicating that multimodal input does not necessarily yield superior outcomes across instructional conditions (Pu & Chang, 2025). This distinction provides a more precise interpretation than treating multimodal instruction as generally superior across all dimensions of strategic competence.

Learners' Perceptions and Possible Explanatory Mechanisms

The interview findings help explain how learners experienced the instructional conditions associated with these quantitative patterns. Three interconnected features were particularly evident in participants' accounts: greater engagement and participation, cognitive and strategic support, and more contextualised opportunities for communication. At the same time, learners also identified linguistic, affective, and task-related constraints, indicating that multimodal resources did not automatically eliminate the difficulties involved in spontaneous EFL speaking.

First, learners commonly perceived the multimodal activities as engaging and interactive. Videos, audio materials, role plays, games, and collaborative tasks were described as helping sustain attention and encouraging participation. Engagement is relevant to strategic competence because communication strategies become meaningful only when learners enter interaction, encounter communicative problems, and attempt to resolve them. A classroom environment that encourages more active participation may therefore increase opportunities

for learners to experiment with strategic behaviour. This does not mean that engagement itself caused the observed gains, but it provides one plausible condition under which strategy use could be practised more frequently and purposefully. This interpretation is consistent with multimodal oral communication research highlighting the role of systematically scaffolded multimodal resources in supporting learner engagement and oral performance (Lee & Azman, 2022).

Second, participants emphasised the cognitive support offered by visual, auditory, and contextual information. They reported that multimodal resources helped them understand tasks, generate ideas, organise expression, and recognise alternative ways of conveying meaning. This perceived support is particularly relevant to circumlocution and approximation, because both depend on learners' ability to move beyond an unavailable lexical item and reconstruct meaning through other linguistic resources. Learners' accounts of describing unknown words in alternative ways are consistent with the quantitative increase in circumlocution, although the qualitative data should be treated as explanatory perceptions rather than independent evidence that a particular multimodal resource directly produced a statistical effect.

Third, learners valued activities that they perceived as contextualised and closer to meaningful communication. Scenario-based and interactive tasks allowed them to respond to others and construct messages within recognisable communicative situations rather than simply reproduce predetermined language. Such conditions are particularly relevant to meaning negotiation and non-linguistic means because both depend heavily on interactional context. Multimodal instruction may therefore have supported these strategies not merely by adding media to the classroom, but by creating communicative situations in which multiple resources could be coordinated to achieve meaning.

This interpretation is also compatible with broader multimodal research showing that multimodal pedagogies can support communicative development by expanding the resources learners use to understand and express meaning (Chen, 2025; Rahmanu & Molnár, 2024). The present findings suggest, however, that the usefulness of those resources may depend on how closely they are aligned with the communicative function of the strategy being taught. Images may support descriptive meaning construction, gesture can directly model embodied communication, and interactive tasks may create a need for negotiation. The study cannot isolate the independent contribution of each individual mode, but the overall pattern indicates the potential importance of purposeful strategy–mode alignment.

Importantly, the interview data also qualify an overly positive interpretation of multimodal instruction. Some learners continued to report limited linguistic resources, low confidence, difficulty transferring classroom learning beyond the course, and insufficient practice time. These experiences help explain why strategy awareness may not have developed into consistent spontaneous use for all learners or all strategies. Participants' requests for more practice, clearer modelling, and stronger scaffolding further suggest that multimodal resources function most effectively when accompanied by sustained instructional support.

The mixed-methods findings therefore point to an interaction among participatory, cognitive, and contextual affordances. Multimodal activities may encourage learners to participate, provide additional resources for constructing meaning, and create contexts in which communication strategies become relevant. However, these affordances remain dependent on learner resources, task design, instructional scaffolding, and sufficient opportunities for repeated use. This helps explain why the quantitative effects were substantial for some strategies but limited for others.

Pedagogical Implications

Several pedagogical implications follow from these findings. First, communication strategy instruction itself

should be considered an important component of EFL speaking pedagogy. The development of fillers and hesitation devices and self-initiated repair in both groups suggests that learners can benefit from explicit attention to the functions of communication strategies and repeated opportunities to use them during speaking practice. Multimodal instruction should therefore complement rather than replace explicit strategy teaching.

Second, multimodal resources should be selected according to the communicative function of the target strategy. The findings do not support adding images, videos, gestures, or digital materials simply to make lessons more varied. Instead, teachers should consider how a particular mode can facilitate a particular strategic behaviour. Visual prompts may scaffold circumlocution or approximation, gesture modelling may support non-linguistic means, and interactive information-gap or role-play tasks may create genuine opportunities for meaning negotiation. Such purposeful strategy–mode alignment is likely to be more pedagogically meaningful than multimodal variety for its own sake.

Third, strategies that remain infrequent after short-term instruction may require more sustained scaffolding. Restructuring and comprehension checks, for example, may benefit from explicit modelling, guided rehearsal, feedback, and repeated recycling across different communicative contexts before learners are expected to use them spontaneously. Participants' own calls for additional practice and clearer teacher support reinforce this point.

More broadly, strategic competence may be most effectively developed through a progressive instructional sequence in which strategies are introduced explicitly, modelled through appropriate linguistic and multimodal resources, practised in scaffolded tasks, and subsequently recycled in increasingly spontaneous interaction. Such an approach treats communication strategies not as isolated techniques to be taught once, but as a communicative repertoire that develops through repeated and contextually meaningful use.

CONCLUSION

The present study examined the development of strategic competence among Chinese first-year English majors by comparing multimodal strategy-focused instruction with conventional strategy instruction. Using a sequential mixed-methods design, the study combined pre-test–post-test evidence of learners' communication strategy use with semi-structured interviews to examine both instructional outcomes and learners' perceptions of the multimodal approach.

The findings revealed a differentiated pattern of strategy development. Fillers and hesitation devices and self-initiated repair increased significantly in both groups, suggesting that explicit communication strategy instruction and repeated speaking practice may support the development of these strategies under both instructional conditions. In contrast, the EG demonstrated significantly higher post-test use of circumlocution, meaning negotiation strategies, approximation, and non-linguistic means. These results indicate that systematic multimodal integration may provide additional support for strategies closely associated with flexible meaning construction, interaction, and embodied communication. However, no significant between-group differences were found for fillers and hesitation devices, self-initiated repair, restructuring, or comprehension checks, indicating that the benefits of multimodal strategy-focused instruction were not uniform across all eight strategies.

The qualitative findings further showed that learners generally perceived multimodal strategy-focused instruction as engaging, cognitively supportive, and conducive to contextualised communication. Participants reported that multimodal resources helped them maintain attention, generate ideas, explore alternative ways of expressing meaning, and participate more actively in speaking activities. At the same time, they continued to experience difficulties related to limited linguistic resources, confidence, transfer beyond the classroom, and

insufficient opportunities for sustained practice. These perceptions suggest that multimodal resources can support strategy learning, but their effectiveness depends on appropriate task design, scaffolding, and repeated opportunities for use.

Several limitations should be considered. The final sample was relatively small and drawn from a single university context, which limits the generalisability of the findings. Participant attrition also resulted in unequal final group sizes. Although pre-test comparisons indicated no statistically significant differences between the final EG and CG samples across the eight target communication strategies, the possibility of attrition-related bias cannot be completely excluded. The small group sizes may also have limited statistical power to detect small or moderate effects. Accordingly, non-significant findings, particularly for infrequently observed strategies such as restructuring and comprehension checks, should not be interpreted as definitive evidence of no instructional effect. The eight-week intervention may also have been insufficient for some strategies to become readily available during spontaneous communication. In addition, strategic competence was operationalised primarily through the frequency of communication strategy use in the speaking assessments. These frequency measures indicate the extent to which learners deployed particular strategies but do not directly establish whether individual strategy uses were appropriate, effective, or successful in resolving communication difficulties.

Although the stratified purposive sampling procedure was designed to capture varied learner experiences, interviews with eight EG participants may not represent the full range of perspectives within the experimental group. In addition, because no CG interviews were conducted, the qualitative findings cannot support direct comparisons of learner perceptions across the two instructional conditions.

Future research could therefore examine multimodal strategy-focused instruction with larger and more diverse samples, longer intervention periods, and a wider range of communicative tasks. Combining frequency-based measures with discourse-oriented or interactional analyses of strategy effectiveness and communicative success would provide a more comprehensive understanding of strategic competence. Further research could also investigate the contribution of specific multimodal resources to particular strategies rather than treating multimodal instruction as a single instructional package.

Overall, the study suggests that explicit communication strategy instruction can support strategic development under both instructional conditions, while systematic multimodal integration may provide additional benefits for particular forms of strategic behaviour. Rather than viewing multimodal instruction as uniformly superior, the findings highlight the importance of aligning instructional resources with the communicative functions of specific strategies. This strategy-sensitive perspective contributes to a more nuanced understanding of how strategic competence can be developed in EFL speaking classrooms.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the participation and cooperation of the students involved in this study. Their engagement throughout the research process was essential to the successful completion of the study.

Conflict Of Interest

The authors declare no conflict of interest.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to applicable ethical and confidentiality considerations.

REFERENCES

1. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
2. Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/I.1.1>
3. Caprario, M. (2025). ELF communication strategies: A pedagogical intervention pilot study. *TESOL Quarterly*, 59(1), 359–377. <https://doi.org/10.1002/tesq.3309>
4. Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6(2), 5–35. <https://doi.org/10.5070/L462005216>
5. Chen, W.-C. (2025). Verbal-visual skill-building and perceptual changes in English presentation. *English for Specific Purposes*, 77, 71–85. <https://doi.org/10.1016/j.esp.2024.10.002>
6. Dörnyei, Z., & Scott, M. L. (1997). Communication strategies in a second language: Definitions and taxonomies. *Language Learning*, 47(1), 173–210. <https://doi.org/10.1111/0023-8333.51997005>
7. Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
8. Lee, S. S., & Azman, H. (2022). Introducing a responsive multimodal oral presentation pedagogy: Integrating TED videos with Web 2.0, collaborative learning and teacher feedback. *RELC Journal*, 53(3), 703–711. <https://doi.org/10.1177/0033688220945426>
9. Mei, A., & Nathalang, S. S. (2010). Use of communication strategies by Chinese EFL learners. *Chinese Journal of Applied Linguistics*, 33(3), 110–125.
10. Milliner, B., & Dimoski, B. (2024). The effects of communication strategy training on speaking task performance. *RELC Journal*, 55(2), 344–363. <https://doi.org/10.1177/00336882221085781>
11. Nakatani, Y. (2010). Identifying strategies that facilitate EFL learners' oral communication: A classroom study using multiple data collection procedures. *The Modern Language Journal*, 94(1), 116–136. <https://doi.org/10.1111/j.1540-4781.2009.00987.x>
12. Peker, B. G., & Erdemir, N. (2021). Does compensation strategy instruction work? An action research study. *ELT Journal*, 75(3), 351–361. <https://doi.org/10.1093/elt/ccaa073>
13. Pu, P., & Chang, D. Y.-S. (2025). Effects of different input modes on blended EFL speaking instruction: A quasi-experimental study. *Computer Assisted Language Learning*, 38(5–6), 1223–1248. <https://doi.org/10.1080/09588221.2023.2273853>
14. Rahmanu, I. W. E. D., & Molnár, G. (2024). Multimodal immersion in English language learning in higher education: A systematic review. *Heliyon*, 10(19), e38357. <https://doi.org/10.1016/j.heliyon.2024.e38357>
15. Wang, J., Gao, Y., & Cui, Y. (2023). Classroom gesture instruction on second language learners' academic presentations: Evidence from Chinese intermediate English learners. *Journal of English for Academic Purposes*, 66, 101304. <https://doi.org/10.1016/j.jeap.2023.101304>
16. Wang, Y., & Soo, R. S. (2024). Identifying Chinese college EFL students' needs for an oral communication instruction module. *Asian Journal of University Education*, 20(3), 598–611. <https://doi.org/10.24191/ajue.v20i3.27860>