Teaching Speaking in Kazakhstani EFL Classrooms: Negotiating Teacher Beliefs and Assessment Constraints

Askat Tleuov
KIMEP University, Kazakhstan

Abstract

Despite communicative goals in English language instruction, speaking skills remain underemphasized in Kazakhstani secondary schools where assessment policies prioritize grammar and reading proficiency. This study examined how four experienced EFL teachers conceptualized and implemented speaking instruction within an exam-oriented educational context, focusing on their instructional approaches, responses to contextual constraints, and reconciliation of beliefs with practice. Employing a multiple-case qualitative design, data were collected through (POIs), classroom observations, and stimulated-recall interviews. Thematic analysis revealed variations in instructional approaches, with teachers navigating tensions between beliefs about oral proficiency and institutional pressures. Some participants adopted direct, structured methodologies emphasizing controlled output, while others incorporated communicative, learner-centered activities. Contextual factors, including curriculum mandates, high-stakes testing, and student proficiency levels, influenced instructional choices, though teachers demonstrated agency through localized adaptations. The study highlights the affective dimensions of speaking instruction, particularly strategies for mitigating learner anxiety and balancing error correction with motivation. Findings point to the need for assessment reforms that integrate speaking components and targeted professional development to equip teachers with effective oral skills pedagogy. This research examined the interaction between teacher cognition and contextual factors, contributing to the understanding of teacher intervention within assessment-driven educational systems in EFL contexts.

Keywords: assessment washback, contextual constraints, EFL teaching, speaking instruction, teacher agency, teacher cognition

English language education has become increasingly important in Kazakhstan's drive toward international integration and economic development. As the country positions itself in the global arena, proficiency in English – particularly speaking skills – has emerged as a determinant of academic success and professional advancement. Within this evolving landscape, the implementation of English Medium Instruction (EMI) across 44 universities exemplifies a substantial institutional push toward English proficiency, catalyzing innovations in teaching methodologies and student engagement strategies (Gaipov et al., 2024).

The current status of English language teaching in Kazakhstan reflects an intricate interplay between aspirational policies and practical constraints. While English holds a dominant position within the country's language education policy, alongside Kazakh and Russian, its implementation faces substantial challenges, resulting in generally low proficiency levels among students (Khassanov et al., 2024). Teachers navigate multiple barriers ranging from limited language proficiency to psychological impediments, often exacerbated by teacher-student dynamics and evaluative learning environments (Yessenbekova, 2024). These challenges are noticeable in speaking instruction, where the disconnect between policy aspirations and classroom realities becomes evident.

At the secondary level, the foreign-language strand of the State Compulsory Education Standard (Order #348, 2022) lists speaking outcomes but allocates summative marks almost exclusively to reading comprehension and grammatical accuracy; L2 speaking remains a formative-only target (Ministry of Education of the Republic of Kazakhstan, 2022). The Unified National Test (UNT) – Kazakhstan's principal high-stakes school-exit and university-entrance exam – contains 120 multiple-choice items worth up to 140 points, assessing history, mathematical literacy, reading literacy, and two profile subjects; no oral component is included. The optional English profile block likewise measures lexico-grammatical recognition through multiple-choice tasks (Ministry of Education of the Republic of Kazakhstan, 2017). Because school rankings hinge on these reading literacy and subject-content scores, teachers prioritize test-relevant sub-skills. School-level English-only classroom directives as reported by teachers (Ismail et al., 2018) further curtail strategic L1 scaffolding. Collectively, these conditions shape the instructional space for L2 oral skills teaching and learning (Yessenbekova, 2024).

Speaking competence is an important element in today's globalized context, extending its significance beyond classroom settings. Research indicates that effective oral communication can facilitate professional engagement across cultural boundaries, particularly in specialized English language contexts where communicative ability combines with cultural awareness and pragmatic knowledge (Saptiany & Prabowo, 2024). Some educational approaches incorporating technology and task-oriented methodologies show potential for enhancing speaking instruction, though evidence suggests their application varies considerably across different educational environments (Raj & Baisel, 2024).

A close examination of current research reveals theoretical and practical gaps in the understanding of speaking instruction within Kazakhstan's unique educational context. While

studies have extensively documented EMI adoption at higher education levels (Gaipov et al., 2024), speaking instruction at the secondary school level remains underexplored, creating a blind spot in the literature on how younger learners develop speaking proficiency. This chasm is noteworthy, as school years represent a critical period for language skill development. Furthermore, existing research has failed to adequately examine the complex relationship between assessment-driven educational systems and speaking instruction. Though studies acknowledge the impact of standardized testing on teaching practices (Begimbetova et al., 2023), they have not systematically investigated how teachers navigate these challenges while attempting to develop students' oral proficiency. This limitation suggests that effective pedagogical strategies within assessment-driven contexts remain insufficiently examined in the current literature.

Present-day research approaches have primarily focused on documenting surface-level challenges without examining in depth the underlying mechanisms through which contextual realities affect teaching practices. While studies highlight various difficulties faced by teachers (Tajik et al., 2023), they have not adequately explored how teachers' beliefs and practices evolve in response to these issues. This lack of theoretical engagement with the relationship between teacher cognition and institutional constraints poses challenges for developing interventions that support teachers in balancing communicative goals with evaluation-driven requirements.

To address these gaps, this study examined how four experienced EFL teachers in Kazakhstani secondary schools conceptualized and implemented speaking instruction within an assessment-driven educational context. Through a multiple-case qualitative design, the research investigated how teachers navigated contextual constraints while attempting to develop students' oral proficiency.

The study addressed three primary research questions:

1. How do Kazakhstani EFL teachers conceptualize and operationalize the teaching of speaking skills in their classrooms?
2. In what ways do contextual factors shape or constrain teachers' speaking instruction?
3. How do teachers reconcile their stated beliefs with their classroom practices when faced with contradictory pressures?

The findings reveal variations in instructional approaches, with teachers adopting diverse strategies to balance communicative goals with institutional requirements. Some participants favored direct, structured methodologies emphasizing controlled output, while others incorporated more communicative, learner-centered activities. Contextual factors, including curriculum mandates, student proficiency levels, and washback from standardized testing, influenced these instructional choices, often leading to compromises between fluency and accuracy.

Literature Review

Speaking Instruction in EFL Contexts

Communicative Language Teaching (CLT) emphasizes practical communication skills and positions speaking proficiency as central to language competence, with research demonstrating its effectiveness in enhancing students' oral performance through interactive activities that simulate authentic communication (Wei et al., 2024). While studies consistently report increased fluency and confidence through CLT implementation, its success varies based on teacher preparedness and resource availability (Salam & Luksfinanto, 2024). In Kazakhstan's language education context, traditional teaching methods often persist alongside CLT principles. In practice, these "traditional" lessons still mirror Soviet-era grammar-translation and audiolingual routines in which teachers front-load rules, drill memorized dialogues, and assign decontextualized gap-fills, with minimal spontaneous learner talk (Ismail et al., 2018; Tuspekova et al., 2020). Related approaches such as Content and Language Integrated Learning reflect a hybrid stance that combines local and global educational epistemologies (Bedeker et al., 2024). This adaptation highlights the need to understand how communicative approaches function within assessment-driven educational systems. Most existing studies adopt a purely CLT lens. Few engage sociocultural theory, which can better examine teacher agency under systemic constraints – a perspective the present study adopted.

Task-based language teaching (TBLT), particularly when combined with communicative approaches, effectively facilitates meaningful interaction through activities like role-plays and discussions, improving both interaction patterns and pronunciation fluency (Herrera et al., 2024). However, Bedeker et al. (2024) report only marginal gains from similar tasks in Kazakhstani EMI contexts, suggesting that institutional assessment culture—rather than task type *per se*—moderates instructional efficacy. Digital resources enhance vocabulary acquisition while maintaining student engagement (Alharbi et al., 2024), while storytelling strategies provide contextually rich opportunities for communication practice (Nair & Yunus, 2021). These findings indicate that effective speaking instruction requires combining interactive elements with context-rich learning experiences—especially valuable in EFL settings where authentic language use opportunities may be limited.

Teacher Cognition and Practice

Teacher cognition—comprising beliefs, knowledge, and decision-making processes—fundamentally shapes EFL speaking instruction approaches (Borg, 2015). Tensions frequently emerge between stated beliefs and actual teaching practices due to contextual challenges and traditional methods, markedly evident in corrective feedback approaches where implementation varies between novice and experienced teachers (Borg, 2017). These disparities highlight how contextual factors shape teaching practices and underscore the need for reflective professional development.

Although the field of teacher cognition research is steadily expanding, its application to speaking instruction remains comparatively limited. While studies examining teacher beliefs and practices regarding reading, grammar, and vocabulary instruction are abundant in the literature (Li, 2020), comparatively little attention has been given to how teachers conceptualize and implement speaking instruction, particularly in EFL contexts. This lack of attention is especially notable given the centrality of oral proficiency in communicative language teaching approaches and the unique challenges speaking instruction presents in assessment-driven educational environments.

Challenges in Speaking Instruction

In Kazakhstani EFL classrooms, speaking instruction encounters barriers across curricular and institutional dimensions. Rigid curriculum requirements restrict the implementation of interactive activities that enhance speaking skills (Cao et al., 2024; Ismail et al., 2018; Tuspekova et al., 2020; Yessenbekova, 2024). While Tuspekova et al. (2020) relied on a self-report questionnaire distributed to 42 teachers, Ismail et al. (2018) triangulated survey data with eight classroom observations, revealing a systematic gap between declared priorities and actual practice. This disparity underlines the need for triangulated designs such as the present study. Inadequate professional development opportunities further limit teachers' capacity to adopt interactive approaches (Xuan Mai et al., 2024). Sociocultural factors, particularly deep-rooted norms around authority and limited learner autonomy, affect student participation and willingness to engage in speaking activities. As in other high-context cultures, students may defer to teacher authority and avoid initiative in speech. They may view oral risk-taking as inappropriate or anxiety-inducing (Noor, 2024). This dynamic may be compounded by speaking anxiety, which may manifest as fear of public error, low confidence – psycholinguistic barriers shown to negatively affect performance and oral engagement (Hussein et al., 2019). These factors may contribute to conditions in which students are reluctant to act autonomously or take ownership of oral language practice in class.

Assessment systems influence teachers' ability to prioritize speaking skills in Kazakhstan's high-stakes testing environment. The emphasis on standardized tests like the UNT has narrowed the curriculum, with teachers prioritizing test-related content over oral proficiency development (Fleming & Shinjee, 2024). This focus on quantifiable outcomes marginalizes speaking skills, which are challenging to measure through standardized formats (Borger, 2019). The pressure to meet external targets results in teaching approaches emphasizing rote learning over communicative practices (Bhattacharya, 2022), primarily affecting rural areas with limited educational resources (Iddings et al., 2021).

The concept of assessment washback – the influence of testing on teaching and learning – is well-established in applied linguistics (Cheng, 2005). High-stakes, grammar-focused exams often generate negative washback, shifting instructional attention away from L2 productive skills like speaking and toward testable sub-skills such as grammar and reading comprehension (Dong et al., 2021; Kılıçkaya, 2016). In Kazakhstan, where L2 oral communication is absent from major national exams, the potential for negative washback remains high. However, little

research has examined how EFL teachers locally interpret or resist this influence when attempting to develop students' L2 oral proficiency.

When clear, measurable goals for oral communication are absent, lesson planning often defaults to a narrow and repetitive set of drill-type tasks, producing monotonous activity cycles that do little to extend students' communication range (Pan, 2024). Moreover, students with limited productive language tend to disengage, particularly toward the end of the school day, which reduces the effectiveness of speaking activities (Husnaini et al., 2024). Large class sizes complicate individual attention and effective pronunciation correction (Sharma, 2024), while time constraints pose challenges in finding suitable tasks for diverse student backgrounds (Albino, 2017). In Kazakhstan, teacher-centered norms in public state schools rooted in post-Soviet educational traditions limit students' opportunities for spontaneous oral production in EFL classrooms (Tuspekova et al., 2020). Likewise, hierarchical teacher-student relationships and a lack of emotional support further inhibit learner autonomy and speaking confidence (Ismail et al., 2018). These culturally embedded issues may be mediated through increased use of student-centered speaking tasks, greater teacher attention to emotional support during oral activities, and gradual incorporation of classroom routines that promote learner autonomy.

Methods

Study Design

This inquiry adopted a qualitative collective (multiple-case) instrumental explanatory case-study design with embedded units of analysis (Stake, 2013; Yin, 2018). Each of the four in-service EFL teachers working in a Kazakhstani state secondary school constituted a bounded case, selected to address the study's research questions concerning teacher cognition, pedagogical practice, and contextual constraints in the teaching of speaking. Within each case, four embedded units were analyzed: a) teachers' espoused conceptions of teaching speaking; b) enacted classroom practice; c) contextual constraints the teacher perceived or negotiated; and d) the strategies used to reconcile belief-practice tensions. Examining these sub-units separately enabled within-case explanation building. A subsequent cross-case synthesis then identified convergent and divergent explanatory patterns. A collective design was selected because the phenomenon of interest – how speaking instruction is negotiated under contextual pressures – can be illuminated more convincingly by contrasting several theoretically replicating cases than by a single exemplar. This facilitated an "analytic generalization" rather than "statistical inference" (Yin, 2018, pp. 40-42).

The study was situated within a social-constructivist interpretive framework (Creswell & Poth, 2018). This perspective assumes that understandings of classroom practice are co-constructed through interaction. Accordingly, semi-structured POIs, naturalistic lesson observations, and stimulated-recall dialogues were combined to elicit participants' multiple realities. An interpretive orientation aligned with the research questions, which probed *how* teachers made meaning of speaking pedagogy, *how* contextual factors influenced that meaning, and *how*

teachers reconciled belief-practice relationships. These issues necessitated subjective, context-dependent inquiry rather than objective measurement.

Participants and Context

This study involved four in-service EFL teachers—Peter, David, Adam, and Mary—recruited from a state comprehensive secondary school in Almaty, Kazakhstan. Participants were selected through purposive sampling (Creswell & Poth, 2018) based on three criteria: a) a bachelor's degree at minimum in English language teaching (ELT) or equivalent; b) at least one year of continuos EFL teaching experience in state schools; c) engagement in EFL-related professional development. These criteria were selected to ensure that participants possessed sufficient formal training, contextual familiarity, and professional enagement to provide informed and reflective insights into the conceptualization and enactment of speaking instruction in state school settings. All four met these criteria. Three were male and one female; each held a BA in ELT from a Kazakhstani university and had a minimum of three years of total EFL teaching experience in state schools.

Participant recruitment involved approaching six secondary schools in Almaty. In four cases, access was denied by principals citing practical concerns such as scheduling, renovations, or curriculum changes. A fifth principal allowed recruitment emails to be sent by the secretary to EFL teachers, but no teachers volunteered. Access was ultimately secured at a sixth school, led by a principal with a PhD in Education and research experience. She introduced the study to the English department and facilitated an in-person presentation. Seven of eight EFL teachers attended; five initially volunteered. After reviewing the study details and consent forms, one withdrew due to discomfort with audio-recorded lessons, leaving four participants teaching across grades 7 to 10.

Although drawn from a single institutional context, the four teachers represented a range of perspectives and pedagogical approaches to speaking instruction. This internal diversity enabled meaningful comparisons within a shared policy environment. Consistent with a bounded case-study design (Yin, 2018), the single-site focus allowed in-depth exploration of how EFL teachers interpret and respond to contextual factors.

Data Collection

Pre-Observation Interviews

The semi-structured interview protocol was developed by the researcher based on established teacher cognition literature (Borg, 2015) and qualitative interviewing guidelines (Adams, 2015). The guide was then piloted with two master's students in foreign language education at the researcher's university who were concurrently teaching English in private language schools and had prior experience in state schools. Piloting led to several refinements. For example, a double-barreled item on fluency versus accuracy was split into two discrete questions, the term

washback was replaced with a clearer phrase – exam influence, and potentially leading questions about the role of L1 were reformulated.

Following these adjustments, each participant completed two POIs. The first elicited professional background, training history, and personal language-learning experience; the second probed beliefs about teaching speaking, perceived contextual constraints, and learner needs. Interviews were conducted in English, although teachers occasionally used Kazakh or Russian terms for added clarity.

Classroom Observations

Following the interviews, ten observations of each teacher's EFL classes (CO1-CO4) were conducted, for a total of forty observed lessons. Each class lasted 45 minutes. All classes were audio-recorded, and field notes were gathered as well. Observations allowed for a real-time examination of:

- The selection and structure of speaking tasks such as scripted versus open-ended.
- Patterns of interaction such as pair work, small groups, or whole-class discussions.
- Teacher interventions such as error correction strategies, use of L1, and student engagement.

Observation records were expanded soon afterward to capture contextual details. This included adding descriptive notes about classroom atmosphere, non-verbal cues, seating arrangements and any spontaneous teacher or student behaviors that were not captured in the audio-recordings. Although the presence of a researcher can influence natural behaviors (Ciesielska et al., 2018), the study's multi-phase design – interviews both before and after the lessons – aimed to mitigate potential observer effects by allowing teachers to discuss how they felt about being observed and whether they had adapted their usual style.

Stimulated-Recall Interviews

Five stimulated-recall interviews (SRI) were conducted with each teacher – one after every second observed lesson – to elicit the cognitive rationales underlying instructional decision-points related to speaking instruction (Gass & Mackey, 2016). Within 48 hours of each lesson pair, three or four audio extracts of 30–90 seconds each were selected. The excerpts captured a decision made by the teacher, such as initiating or terminating a speaking task, choosing to correct or ignore a spoken error, switching between L1 and English, or reallocating learners between pair, group, and whole-class interaction. These moments, identified from field-note annotations of critical incidents, were replayed to teachers on a laptop. Playback was paused after each clip and teachers were invited to reflect on decisions regarding the teaching of speaking, contextualize those choices, and evaluate how these aligned or conflicted with their personal pedagogical beliefs (Tondeur et al., 2017). This technique yielded a rich understanding of how teachers rationalize spontaneous or planned modifications to their

speaking activities and how they perceive the influence of contextual factors. Twenty SRI transcripts were generated (11 hours of talk) which were coded for thematic analysis.

Data Analysis

Data collection lasted for four months. All data – POIs, field notes from classroom observations, and transcripts of SRIs – were analyzed using thematic analysis (Clarke & Braun, 2016). The process began with repeated readings of transcripts and observation records to gain familiarity with each teacher's narrative and classroom events. During initial coding, passages were labeled according to salient categories, such as direct vs. indirect instruction, speaking anxiety, contextual constraints, and error correction stance. These codes were refined through focused coding to reveal overarching themes that corresponded to patterns observed across participants. Triangulation (Renz et al., 2018) was achieved by comparing teacher statements, classroom behaviors, and their subsequent reflections in SRIs, capturing both alignment and dissonance between stated beliefs and observed practices.

Ethical Considerations, Reflexivity, and Trustworthiness

At the time of data collection (January–April 2023) the researcher's university had not yet established a formal institutional review board (IRB). Consequently, ethical oversight followed the British Educational Research Association (BERA) Ethical Guidelines for Educational Research (2018). Participation in the study was voluntary. Written gatekeeper permission was obtained from the participating school's principal, and all teacher-participants provided written informed consent. The participants were assigned pseudonyms (Peter, David, Adam, and Mary) to protect their identities. Classroom observations minimized disruption to teaching activities, and student privacy was maintained by avoiding collection of identifying information.

A reflexive journal was maintained by the researcher throughout the study (Meyer & Willis, 2019), with ongoing notes on personal assumptions about Kazakhstani EFL contexts and reflecting on how these might shape interpretations.. Member-checking strengthened both ethical practice and trustworthiness, as teachers reviewed interview summaries to verify accurate representation of their views. To ensure dependability, all transcripts were coded manually using a stable coding frame developed through iterative readings, with ongoing memo-writing to document analytic decisions. The coding framework was applied systematically across all data types and revisited as themes were refined. Data triangulation through multiple collection phases – POIs, classroom observations, and SRIs – ensured a robust perspective on teacher cognition and practical constraints, aligning with recognized standards for trustworthiness in qualitative inquiry (Creswell & Poth, 2018). All research data were stored securely and will be destroyed within three years of study completion.

Findings and Discussion

Teachers' classroom practices around speaking instruction reveal how pedagogical choices are shaped by a continuous negotiation between individual beliefs, institutional pressures, and student needs. The analysis situates classroom observations within relevant theoretical frameworks and demonstrates how pedagogical practices reflect and respond to broader disciplinary perspectives.

The study used a multi-phase approach. Initial (POIs) established teachers' backgrounds and beliefs, followed by ten classroom observations per teacher (labeled CO1-CO10). Stimulated-recall interviews (SRIs) after every two observations captured teachers' rationales for instructional decisions and their navigation of contextual constraints.

The following subsections integrate data from all three sources (POIs, SRIs, and classroom observations) to demonstrate how teacher cognition, institutional factors, and student characteristics shape speaking instruction in Kazakhstani secondary schools.

Approaches to Speaking Instruction

In the observed classrooms, the teachers balanced teacher-led techniques with more open-ended, communicative tasks to differing degrees. All four participants—Peter, David, Adam, and Mary—recognized the importance of oral proficiency, yet they adopted varying instructional approaches informed by their beliefs, experiences, and teaching contexts. Despite working within comparable teaching conditions, the four teachers employed distinct methods for developing oral proficiency, shaped by their positions along the continuum between direct (controlled) and indirect (transfer) approaches.

Direct and Indirect Methodologies

The findings reveal that David consistently favored a direct, teacher-centered approach, whereas Mary combined explicit instruction with more communicative tasks. Peter and Adam leaned more heavily on indirect, learner-centered approaches, though they occasionally turned to controlled exercises for less proficient learners.

David believed that mastery of grammar and vocabulary must precede genuine fluency. As he explained, "All the activities that we do, such as dialogues, pair work, memorization and extra reading, are designed to prepare students for that aim [speaking freely later]" (SRI6). In practice, David's lessons often included grammar drills and text recitations, reflecting a skill-getting perspective (Cook, 2016). During multiple observations, half of his class time was spent on these controlled activities (CO3, CO4). He recalled learning in a similar way: "An English teacher would always ask us to memorize texts … It was good for learning vocabulary" (POI2). This personal history influenced his view that well-structured tasks lead to "the rest," namely freer speaking.

Mary employed direct instruction before moving quickly to communicative activities, blending structure with fluency practice. "Grammar knowledge is internalized faster because it was practiced through speaking activities immediately after the grammar work … the instruction should be there," she noted (SRI9). In lessons, she introduced a specific grammatical point and then transitioned to small-group discussions, reflecting Hughes and Reed's (2016) recommendation of pairing explicit teaching with functional practice. Although she retained a teacher-led segment to ensure accuracy, Mary's approach confirms the possibility of merging direct and indirect methods within a single lesson cycle.

In contrast to David's consistent structure, Peter favored freer speech once students demonstrated basic competence. During observations of his more advanced classes, he devoted extended segments to open-ended group projects, reserving memorized dialogues for only the weakest groups (CO6, CO7). He explained that learners at higher levels "already have enough words and grammar to speak freely, so I don't push them to follow a script" (SRI2). This flexibility aligned with a scaffolding process in which early controlled stages give way to more spontaneous interaction (Azir, 2019).

Adam identified with an indirect, communicative method. He believed "it's all about getting them to speak no matter what" (SRI4), frequently allowing students to explore topics in small groups or whole-class discussions with minimal teacher intervention. Although Adam acknowledged moments of unstructured "only practicing," his emphasis on continuous oral exchange supported the argument that authentic conversation fosters more naturalistic language acquisition (Burns, 2019). In his classes, pair or group talk accounted for the majority of the speaking time (CO5, CO8), indicating that work on accuracy, while present, was less extensive than in David's lessons.

To sum up, these participants demonstrate how classroom discourse can move along a direct–indirect continuum depending on teacher beliefs, students' language readiness, and time constraints. This diversity matches theories that argue for dynamic, context-tailored approaches rather than strict adherence to a single method (Willis, 2015).

Context-Sensitive Instructional Choices

Although teaching took place in the same state secondary school, each participant molded speaking instruction to address learner needs, institutional expectations, and their own personal teaching philosophies (Horwitz, 2020). Identical or similar external contexts did not produce uniform practices, demonstrating how teachers' perceived environments shape choices differently.

David's classes featured controlled tasks and memorization—"the sooner they master all the grammar content the sooner they begin to do the rest" (POI2)—whereas Adam swapped textbook themes for more locally relevant topics. "If I see the text topic is about 'British on Holiday,' I replace it with local examples," Adam explained, so students could connect personal experiences to the conversation (SRI5). Observational data show these more

spontaneous discussions often lasted longer than planned, an outcome Adam viewed positively (CO5). Their contrasting methods highlight how individual perceptions of students' proficiency and engagement underpin context-sensitive practice (Shin et al., 2020).

Peter also demonstrated sensitivity to mixed-ability groups. He noted that "some students already have enough words and grammar … so I don't push them to follow a script" (SRI2), preferring localized scaffolding for struggling learners and freer dialogues for more confident ones. In one lesson, he allowed advanced students to conduct peer-led Q&A sessions, while novices relied on written prompts and word banks (CO7). This dual approach maximized classroom participation and seemed to reduce student anxiety.

Similarly, Mary recognized that a blanket English-only guideline might heighten stress for beginners, particularly in extended classes. At more advanced levels, however, she encouraged exclusive L2 use to stretch students' linguistic capacities (CO6). Her subtle calibration reinforced the notion that teachers respond to local conditions—like fatigue, scheduling, and class size—by adjusting the relative amounts of control and freedom in speaking tasks (Villada et al., 2018).

Taken together, these teachers' evolving, context-sensitive strategies revealed how on-the-ground realities mediate methodological choices. That four educators in the same school could employ such distinct approaches testifies to the complexity of speaking instruction and reveals the role of teacher cognition in shaping classroom discourse (Borg, 2015). By adopting or adapting direct techniques, indirect tasks, or a combination, participants forged practices suited to their particular learner groups, thus demonstrating a pragmatic synthesis of pedagogical ideals and real-world constraints.

Speaking Tasks and Interactional Dynamics

The observed speaking tasks varied across classrooms, informed by teachers' instructional priorities and mediated through different interactional modes. These tasks reflected an interplay between teachers' stated beliefs, classroom practices, and anticipated student learning outcomes. Rather than offering a mere inventory of activities, the analysis considers how these practices correspond with established perspectives on oral language pedagogy (Burns, 2019; Richards, 2017).

Task Types and Pedagogical Implementation

All four participants employed a range of speaking tasks, from scripted memorization and monologic presentations to more open-ended or localized activities. David's instruction centered on controlled, memorized tasks. He reasoned that "when they memorize short texts – like 'Boat Race' or 'Sydney Olympics' – they build the habit of speaking with fewer grammar mistakes" (SRI6). Observational data (CO4, CO5) confirmed multiple recitation activities, supporting a skill-getting perspective (Newton & Nation, 2020). However, students appeared

to spend limited time producing their own utterances, suggesting a potential trade-off between grammatical accuracy and communicative fluency (Vercellotti, 2017).

Adam, by contrast, placed learners in contextualized, open-ended tasks. Instead of following the textbook theme "British on Holiday," he encouraged discussions of local travel experiences (SRI5), and in one observed session (CO7) students spontaneously extended the task, indicating strong engagement. While this approach resonates with research that emphasizes authenticity and personal relevance (Richards, 2017), some students appeared hesitant when they lacked sufficient vocabulary (CO7). These moments highlighted the tension between rich, communicative opportunities and the need for targeted language support (Pakula, 2019).

Mary and Peter adopted hybrid methods that connect structured and open-ended elements. Mary incorporated time-limited yet student-selected presentations, explaining, "They can choose any topic they want to present, as long as they speak in English for at least three minutes" (POI2). Observations (CO2) showed students tackling diverse subjects, and several reported feeling more motivated because they could speak about personal interests (Mary's informal follow-up notes). Peter took a further step by embedding interactive Q&A into presentations – "After each presentation, you need to ask at least one question" (SRI4) – thereby shifting "talk as performance" (Richards, 2017) toward more dialogic, meaning-focused interactions. While observational data (CO3, CO6) confirmed higher levels of student engagement, it also revealed variability in the depth of peer questioning, suggesting that teacher guidance on formulating questions could further enhance the task's learning potential (Newton & Nation, 2020).

These findings illustrate how teachers balance accuracy, fluency, and complexity through different task configurations. Importantly, all participants prioritized speaking but differed in how they scaffolded or localized activities based on their perceptions of student ability and institutional pressures.

Interaction Choices: Patterns, Beliefs, and Rationale

Just as the choice of task reflected differing beliefs about speaking development, so did teachers' decisions regarding pair work, small groups, and whole-class discussions. David primarily utilized pair work: "Pair work is easier to manage and helps me check each pair's progress quickly" (SRI4). Observations (CO1, CO5) revealed frequent dyadic drills aimed at reinforcing textbook material. While these activities allowed David to monitor accuracy more directly, some quieter students contributed minimal utterances, raising questions about turn-taking equity (Greer & Potter, 2015).

Peter's group-based approach, conversely, fostered peer scaffolding: "When they are in groups, even shy students have to participate" (SRI3). Classroom data (CO4, CO7) indicated that small-group problem-solving tasks – such as designing class trips – encouraged broad interaction and negotiation of meaning (Zhu & Carless, 2018). This more sociocultural perspective (Sun & Zhang, 2021) suited Peter's goal of maximizing student talk, though it

sometimes required extra guidance to ensure all voices were heard. Meanwhile, Adam often opted for whole-class discussions, believing "sometimes, we all just talk together about a single topic. It feels more natural if everyone's involved at once" (SRI6). Observation (CO6) confirmed high levels of spontaneous participation but also revealed that stronger students occasionally dominated, aligning with the ongoing tension between authenticity and balanced opportunity (Burns, 2019).

Mary flexibly combined these modes by arranging desks "in four corners" (SRI7), rotating students between pairs, small groups, and quick plenary sessions (CO2, CO3). She aimed to capitalize on each interaction type's affordances – dyadic for immediate feedback, group work for collaboration, whole-class for community building. This structured variety suggests a nuanced understanding of how shifting interaction patterns can reinforce communicative, cognitive, and affective outcomes (Namaziandost et al., 2020).

Overall, the varied interactional patterns aligned with each teacher's beliefs about how best to foster speaking development in a Kazakhstani secondary school context. Pair work and memorization supported tight control over linguistic forms, while group tasks and open discussions promoted a higher incidence of spontaneous output. These interactional choices, particularly those employed by Peter and Mary, resonate with calls in the literature for student-centered approaches to mitigate the constraints of teacher-dominated classroom norms (Ismail et al., 2018; Tuspekova et al., 2020). Their use of group-based and rotating formats provided more opportunities for learners to initiate speech, thereby fostering conditions more conducive to learner autonomy.

One finding stands in contrast to concerns raised in earlier research that Kazakhstani students, shaped by post-Soviet educational traditions and high-context cultural norms, tend to defer authority and avoid risk-taking (Hussein et al., 2019; Noor, 2024). Although those dynamics were observable in David's classes, teachers like Peter and Adam appeared to challenge these tendencies by creating interactional routines that required initiative, negotiation of meaning, and peer-led participation. A notable and thus somewhat unexpected observation was that some shy or previously disengaged students appeared to participate more actively when provided with structured yet flexible speaking roles in group settings. In this way, the study both supports and complicates prior research. While hierarchical constraints persist, certain instructional formats can soften their effects. This may suggest that learner autonomy is not fixed by culture but can be incrementally cultivated through pedagogical design. The differentiated interaction patterns, while informed by local contextual factors and individual teacher cognition, seemingly mirror broader pedagogical tensions identified in the literature. Namely, the balance between authenticity and control, fluency, and accuracy, and student comfort and linguistic stretch (Hughes & Reed, 2016).

Affective Dimensions in Speaking Instruction

Emotional security, teacher–student rapport, and learner confidence appeared to play an influential role in shaping the teaching of oral skills across the observed classrooms. In

alignment with humanistic perspectives (Byram et al., 2023) and sociocultural frameworks (Sun & Zhang, 2021), the data illustrate how teachers constructed supportive environments and navigated L1 use alongside error correction.

Fostering Safe and Motivating Environments

An overriding concern among the four teachers was learner well-being, especially in contexts where speaking English can provoke anxiety or self-consciousness. Peter's account vividly reveals how a student's negative experience shaped his avoidance of immediate error interruption: "A student once told me he hated English because the teacher always corrected him right away. That was a turning point for me" (POI2). Observations (CO2, CO4) corroborated his minimal on-the-spot correction strategy, signaling a consistent alignment between belief and practice – an example of methodological triangulation (interviews plus classroom evidence). This pattern echoes concerns reported in Kazakhstani secondary schools, where immediate public correction has been associated with heightened learner anxiety (Ismail et al., 2018) and reluctance to speak (Yessenbekova, 2024).

From a theoretical standpoint, Peter's emphasis on respecting learners' emotional comfort resonates with Byram et al.'s (2023) humanistic approach, wherein psychological safety is paramount to encouraging risk-taking. Moreover, one could interpret his practice as an application of Krashen's Affective Filter Hypothesis (Moriña, 2022), as he actively minimizes stressors to promote comprehensible output. However, some off-record comments by advanced-level students (field notes) indicated a desire for more immediate correction to refine their spoken accuracy, hinting at a subtle tension between emotional reassurance and targeted feedback (Moriña, 2022).

Mary similarly structured her lessons to build confidence incrementally. She reported letting "them practice in smaller groups first" (SRI6) before transitioning to higher-stakes speaking tasks. Observational data (CO3, CO6) showed that after these low-pressure warm-ups, more reticent learners tended to participate willingly in open-class speech. This scaffolding approach—reminiscent of Vygotsky's Zone of Proximal Development (Xi & Lantolf, 2021)—aims to strike a balance between challenge and emotional safety. One group of students spontaneously reflected, "We feel braver when we've tested our ideas in smaller groups" (field notes), suggesting a positive affective impact. Yet Mary also admitted occasionally running short on time for the whole-class stage, illustrating how institutional pacing requirements might constrain fully realized supportive measures. Her staged approach thus illustrates how confidence building can partially offset culturally rooted hesitation without eliminating time-pressure constraints identified in prior local studies (Tuspekova et al., 2020).

David approached affective concerns by adjusting his expectations according to the lesson's timing, acknowledging that "if they've done five or six lessons already that day, I see no problem in letting them use L1 to clarify a point" (SRI5). In CO5, this pragmatism led to more relaxed speaking activities, albeit with limited grammar correction. Some students reported feeling relieved that "he doesn't push grammar too hard at the end of the day," though one

particularly ambitious learner quietly mentioned wanting more rigorous error feedback "to improve faster" (field notes). This discrepancy reveals how an affectively sensitive approach may at times conflict with certain students' desire for higher challenge and more robust feedback, thus underscoring the importance of teacher awareness of diverse learning preferences.

Adam consistently highlighted risk-taking, claiming that "if they have fear, they'll never speak up" (SRI1). Observations (CO2, CO7) bore out his immediate praise of any attempt at spoken output, reinforcing Burns' (2019) argument that positive reinforcement boosts learners' willingness to communicate. Nonetheless, a few students appeared to plateau, repeating the same basic sentence structures without further linguistic development (CO7). An unexpected observation was that the absence of form-focused feedback occasionally coincided with a plateau in syntactic complexity which seems to diverge from studies that equate increased willingness to communicate with proportional linguistic growth (Liu & Jin, 2024). This outcome points to a trade-off: wholeheartedly encouraging risk-taking can promote spoken fluency but may insufficiently scaffold accuracy or complexity if more systematic feedback is lacking (Vercellotti, 2017).

Cases illustrate that affect-oriented strategies can lower the perceived affective filter yet benefit from balance with calibrated feedback to sustain linguistic development. This is an issue of relevance in contexts where foreign language speaking accuracy is seldom formally tested in state schools. Data indicate that while psychological comfort typically expands learners' willingness to speak, it may also leave gaps in targeted intervention for those seeking rapid improvement or more advanced error correction. This tension sets the stage for a deeper look at how teachers manage two elements—L1 usage and error treatment—that carry strong affective implications. Classroom-level evidence from the current study refines the earlier, anxiety-focused work by Hussein et al. (2019) and suggests a practical pathway for professional development programs that aim to integrate affective scaffolding with targeted oral feedback.

Balancing L1 Use and Error Correction for Learner Confidence

A further dimension of affective pedagogy lies in how teachers navigate students' mother tongue and corrections of spoken errors—both of which can profoundly shape learners' confidence (Bremner, 2021). Mary, for instance, articulated a differentiated L1 strategy: "Beginners need the option to express difficult ideas in L1 sometimes. But for advanced students, I say 'only English'" (SRI4). CO1 and CO2 illustrated that she indeed permitted limited L1 for novice learners, enabling them to contribute more fully to discussions rather than remaining silent. This pragmatic translanguaging (Shin et al., 2020) helped build a sense of inclusion, though she reported feeling pressure from the administration, who expected a largely English-only environment. Balancing these sometimes-conflicting directives highlights the "so what?" of her approach: despite potential institutional constraints, she prioritizes emotional security, especially for those at lower proficiency levels.

David's stance toward L1 similarly reflected a belief that "maintaining positive attitudes" in fatigued students overrode strict language separation (SRI5). However, in CO3 and CO7, once students clarified a confusing term in L1, he nudged them back into English quickly. This gentle pivot, corroborated by direct observation, demonstrates how his approach tactically uses L1 as a scaffold without derailing class-wide English use. Some advanced learners expressed a preference for a stricter no-L1 approach, claiming that it forced them to "think harder in English" (field notes). Such varied preferences confirm the complexity of teacher decisions around language policy, underscoring the influence of student identity, motivation, and skill level on an optimal mix of L1 and L2.

Peter adopted a delayed-correction model that intersected with his minimal on-the-spot strategy: "I prefer to take notes during speaking and correct them afterward, so they don't lose confidence" (POI2). Observations (CO4) showed him quietly jotting down repeated errors, which he would address in a short whole-class session. This method resonates with immediate vs. postponed feedback debates in the literature (Zhu & Carless, 2018). Most students apparently appreciated his approach; one noted feeling "safer to talk" (field notes) because the teacher wasn't "jumping in" at every small slip. Yet a couple of more advanced students found the general, post-task feedback less relevant to their individual issues. This discrepancy points to a tension between ensuring a low-stress environment for the majority and providing higher-level learners with the precision they desire.

Adam, by contrast, occasionally corrected "major mistakes" mid-speech but tried "to be friendly about it" (SRI3). Observational data (CO5) showed him interrupting only when confusion impeded comprehension, consistent with a partial focus-on-form approach (Newton & Nation, 2020). While this seemed to keep discourse relatively fluid, at least one shy student appeared startled during the first correction mid-sentence, suggesting that even gentle interruptions can momentarily raise anxiety for learners unaccustomed to such direct intervention (field notes). Adam's method highlights the perennial balancing act: protect learners' confidence but uphold sufficient correctness to maintain comprehensibility and progress.

Overall, these findings point to the affective dimension of L1 usage and error feedback. Teachers' choices reflect both personal beliefs in fostering a "safe place" (Kisfalvi & Oliver, 2015) and pragmatic adaptations to class conditions and external pressures. Data reveal areas of convergence and dissonance. Even teachers who espouse a tolerant L1 stance sometimes revert to stricter English-only norms under specific circumstances, while those who claim minimal correction occasionally find themselves offering in-the-moment feedback. This variation suggests that error treatment and language policy may serve both linguistic objectives and emotional or motivational considerations (Bremner, 2021).

In examining how the teachers create supportive environments and balance L1 use with error correction, a clearer picture emerges of the complex interplay between teacher cognition, institutional forces, and learner affect. The synergy—or tension—among these factors can

shape learners' oral participation, as supportive approaches appear to ease anxiety yet can also leave advanced students wanting more pointed, individualized correction.

These findings speak directly to the need for professional development sessions that equip teachers to gauge and differentiate their affective strategies according to learner profiles, and to do so under sometimes contradictory institutional demands. Such discussions also set the stage for exploring how teachers reconcile their stated beliefs with classroom realities over time – an issue important to language teacher cognition research (Borg, 2015) and one that interlocks with broader contextual variables, such as high-stakes testing or curriculum mandates, explored in subsequent sections.

Contextual Constraints and Opportunities for Speaking Instruction

Although all four teachers expressed an intention to promote oral skills, the scope of their speaking activities appeared contingent on external assessment structures. This type of pattern is consistent with curriculum-narrowing effects reported at the national level for the UNT (Fleming & Shinjee, 2024).

Exam-Driven Pressures and Teacher Responses

One recurrent theme was the absence of a speaking component in major examinations, which affected instructional priorities. Adam conceded that "we can't realistically grade everyone's speaking individually. There's no oral test in the end-of-term exams or the Unified National Test, so we focus on reading, grammar, and vocabulary" (SRI4). During lessons observed for CO3 and CO4, Adam consistently opted for exam-aligned exercises rather than extended oral tasks, a pattern reflecting the classic washback effect whereby high-stakes tests determine classroom emphases (Dong et al., 2021). This is in line with Borger's (2019) conclusion that skills lacking a quantifiable metric are routinely sidelined in test-oriented settings.

David, likewise, acknowledged parental expectations for superior grammar scores: "Most parents want good grammar scores for the exam; they're less worried about how fluently their kids speak" (POI2). In CO5, he concentrated the entire lesson on grammar drills rather than more interactive speaking tasks – an alignment that indicates how broader societal and familial pressures may reinforce a limited test-driven agenda (Tondeur et al., 2017). A handful of students confided informally that they enjoyed speaking games but recognized grammar mastery as "the real ticket to passing exams," illustrating how learners, too, internalize test-centric norms.

Mary, on the other hand, appeared to express frustration with this assessment framework: "The system doesn't test speaking, but the administration expects us to show good test results. I wish they'd include an oral part" (SRI6). Observations from CO3 confirmed her attempts to integrate short presentations – activities that received no formal scoring but nonetheless aligned with her belief in communicative practice. Such actions mirror Spratt's (2017) argument that a test's

scope or omission of key skills can curb teacher autonomy and hamper skill development, even when teachers value oral competence.

Peter's approach, however, showcased a degree of individual agency, as he introduced unassessed dialogue sessions to "bridge the gap" (SRI3) left by official written tests. For instance, in CO2 students shared spoken summaries after reading a text. Although these tasks did not count toward final marks, several students indicated they found the exercise "useful for practicing pronunciation" (field notes). Peter's initiative underscores the potential for teacher-level innovation despite assessment-induced constraints – a nuance that complicates the notion of washback as purely deterministic. It was notable that students themselves judged these ungraded dialogues as beneficial, suggesting that teacher-initiated work-arounds can create perceived value even when institutional rewards are absent.

Institutional Policies and Professional Development Gaps

Teachers also reported limited institutional backing for speaking-focused strategies and few opportunities for specialized professional development. Mary explained that "the only official policy I keep hearing is 'No L1 use.' There's nothing specific about how to teach speaking or how to improve it" (SRI5). Observations during CO6 reflected this top-down English-only mandate in classroom signage, yet no official guidelines on structuring oral practice were evident. The result, aligning with Ismail et al. (2018), is a policy void that leaves teachers navigating speaking pedagogy largely on their own.

David reinforced this view, remarking that "we have short workshops on grammar testing or standardized assessments, but no real training on how to run speaking lessons" (SRI3). This institutional emphasis on tested components, documented in both interviews and his lesson plans (CO2, CO4), may explain why David defaulted to exam-oriented drills at the expense of extended oral tasks. Even though he had expressed an interest in more interactive techniques, the lack of targeted PD perpetuated a grammar-first paradigm (Johnson & Golombek, 2020). Adam highlighted another structural hurdle: "I would attend more PD courses, but we're juggling so many lessons. There's no extra time to learn new methods for speaking" (SRI4). The class schedule logs (CO1–CO5) indicated that he taught a heavy load without sufficient breaks, limiting his ability to experiment with or reflect on speaking-based innovations. Student feedback further suggested that Adam's enthusiastic attempts at open discussions often lacked methodical guidance, reflecting a need for training that would blend communicative ideals with systematic scaffolding.

By contrast, Peter's evolution in error-correction techniques after attending an external Teaching Knowledge Test (TKT) course abroad (SRI7) demonstrated the transformative potential of specialized development. CO6 showed a more nuanced approach to oral feedback compared to CO2, signaling that teacher cognition can shift meaningfully with the right support (Li, 2020). However, this shift arose from self-motivated professional learning rather than any institutional initiative, underscoring a systemic shortfall in fostering speaking skills at school level.

These patterns indicate that high-stakes assessment policies, combined with a grammar-centric PD culture, seemingly continue to constrain oral skills instruction. Nonetheless, instances of teacher agency such as Mary's short, ungraded presentations and Peter's informal dialogue sessions demonstrate that localized adaptations can partially mitigate washback. These initiatives illustrate Golombek and Johnson's (2021) argument that teacher cognition can develop when educators are willing to negotiate or counteract external pressures.

Recommendations

One of the recurrent patterns in the data is the imbalance between oral-communication goals and grammar-centered instruction. Stimulated-recall interviews indicated that teachers recognized how the absence of an oral component in the end-of-term examinations steers instruction toward reading, grammar, and vocabulary; they also suggested that adding a speaking section could provide a more accurate measure of students' progress in spoken English. Classroom observations supported this pattern: exam-aligned drills frequently replaced extended interaction, illustrating a local washback mechanism in which assessment criteria channel instructional time toward the skills that receive marks. In light of this evidence, a feasible school-level response could be to introduce a modest speaking component into the summative English examination and to monitor its curricular effects. An initial allocation of approximately 10% of the total examination score to a short, rubric-based oral feature can be piloted in one grade before wider adoption. Systematic monitoring of washback can then proceed through scheduled lesson-plan reviews, targeted classroom observations, and teacher reflections collected at the start, midpoint, and end of the academic year. These data have the potential to show whether the altered assessment encourages instructors to increase the frequency and variety of interactional tasks and whether students demonstrate gains in spontaneous speech during regular lessons. Evidence from a recent intervention in Chinese primary schools suggests that the introduction of an oral English test can produce measurable changes in both teaching practices and learner engagement when accompanied by appropriate support structures (Liu & Chen, 2022).

Future research would benefit from broadening the evidential base. A multi-site longitudinal design that follows teachers through curriculum or assessment reform could capture changes in cognition and practice that a single-school study cannot (Wu, 2023). Parallel data from students, especially on anxiety and perceived utility of oral work, would help refine understanding of the affective calculus that shapes class participation (Al-Khotaba et al., 2020). Future studies could also correlate specific patterns of instructional diversification with pre- and post-measures of students' spoken accuracy, fluency, and complexity to determine pedagogical impact.

Conclusion

This study examined how Kazakhstani EFL teachers navigate speaking instruction within an institutional landscape that prioritizes grammar and reading over oral proficiency. The findings reveal distinct patterns related to each research question.

Regarding how teachers conceptualize and operationalize speaking instruction, the four participants—Peter, David, Adam, and Mary—demonstrated a spectrum of approaches ranging from controlled, accuracy-focused activities to more communicative, learner-centered tasks. David consistently favored structured exercises and memorization to build foundational skills, while Adam embraced spontaneous discussions prioritizing fluency over form. Peter and Mary adopted hybrid approaches, balancing explicit instruction with opportunities for authentic communication. These variations reflect different conceptualizations of what constitutes effective speaking instruction, with some viewing accuracy as prerequisite to fluency and others seeing natural communication as the primary learning mechanism.

Concerning contextual factors, high-stakes assessments emerged as an important constraint, with the absence of speaking components in national examinations such as the unified National Test – both a school-leaving and a university entrance exam – limiting time allocated to oral skills. Institutional policies emphasizing grammar and reading proficiency appeared to create pressure to focus on testable content, while class sizes, student proficiency levels, and available resources further influenced instructional choices. Yet within these restrictions, teachers demonstrated agency by creating informal speaking opportunities, localizing content, and strategically incorporating L1 to build student confidence.

In reconciling beliefs with practice, teachers showed varying degrees of alignment. While all four acknowledged the importance of speaking instruction, their classroom enactments revealed tensions between stated beliefs and institutional realities. Adam maintained the strongest consistency between his communicative philosophy and teaching approach, whereas David more readily adapted his practices to meet institutional demands despite expressing interest in more interactive methods. These patterns reflect how contextual factors mediate the extent to which teachers can translate espoused beliefs into classroom reality (Li, 2020), highlighting the complex interplay between teacher cognition, institutional pressures, and pedagogical decision-making.

These findings illustrate how teachers navigate pedagogical decisions within institutional limitations (Borg, 2015; Cook, 2016). This observation has potential implications for teacher education, assessment design, and language policy development in Kazakhstan.

Implications

This study has implications for curriculum reform, teacher education, and language policy in Kazakhstan. An important consideration is the absence of oral assessment in high-stakes examinations, which reinforces grammar and reading over speaking skills. Policymakers might

consider including speaking components in assessment frameworks to encourage oral skills instruction, as adjustments in testing policies could potentially influence pedagogical practices in exam-oriented contexts (Dong et al., 2021). While this case study is grounded in Kazakhstan, similar tensions between teacher cognition and high-stakes assessment have been documented in other systems, notably in New Zealand (East, 2015) and the United States (Fives & Buehl, 2016). This indicates that the insights offered here are transferable to exam-oriented contexts worldwide.

Building on this broader relevance, professional development requires attention as teachers often lack structured support for speaking instruction. Resource-rich training programs should provide systematic strategies for integrating oral skills within curricular constraints, focusing on:

- Scaffolding extended discourse for sustained conversations
- Strategic L1 use as a facilitative tool in early-stage oral production (Macaro et al., 2020)
- Affect-sensitive methodologies to mitigate learner anxiety (Tajik et al., 2023)

Teacher agency plays a notable role in sustaining oral proficiency instruction despite curricular limitations, yet not all teacher-initiated adaptations are structurally sustainable. This underscores the need for institutionalized mechanisms—such as speaking-focused teaching communities and collaboration models – to ensure adaptations become systemic rather than remaining isolated initiatives.

Limitations

Although the four teachers varied in their instructional styles and backgrounds, they all worked in the same secondary school, potentially limiting the transferability of results to other Kazakhstani contexts or beyond. Additionally, the researcher's presence in the classroom may have shaped how teachers conducted lessons (Tarusha & Bushi, 2024), although repeated observations and SRIs sought to reduce these reactivity effects by building rapport and discussing any unusual classroom behaviors. Finally, the ten lessons observed per teacher, while offering valuable depth, still represent only a fraction of their broader teaching routines. Despite these issues, the multiple-case qualitative design, supported by triangulated data and reflexive safeguards, yielded a nuanced view of how EFL teachers navigate diverse pressures to foster oral proficiency within an exam-oriented secondary school environment in Kazakhstan.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

In the development and review of the manuscript, I utilized Grammarly as the sole AI-assisted editing tool. No other AI tools were used at any stage of the research or manuscript preparation process. Grammarly was used for proofreading purposes to improve the quality of my writing by checking for grammatical and punctuation errors, identifying redundancies and repetitive phrasing, improving the coherence and flow of text, and enhancing clarity of expression. This usage was primarily concentrated in the abstract, introduction, and literature review sections of my manuscript.

I hereby confirm that the use of AI was strictly limited to language improvement and proofreading functions. AI technology was not used for generating research ideas or questions, collecting or analyzing research data, interpreting findings, drawing conclusions, creating original content or arguments, or analyzing and synthesizing literature. All intellectual contributions, including the conception of the research, methodology design, data collection, analysis, interpretation of results, and discussion of implications, represent my original work as the author.

References

Adams, W. C. (2015). Conducting semi-structured interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of practical program evaluation* (4th ed., pp. 492–505). Jossey-Bass. https://doi.org/10.1002/9781119171386.ch19

Aisyah, M., Setiawan, S., & Mustofa, A. (2023). EFL teachers' perception self-directed professional development toward oral language maintenance. *Loquen: English Studies Journal*, *16*(1), 39–54. https://doi.org/10.32678/loquen.v16i1.8402

Albino, G. (2017). Improving speaking fluency in a task-based language teaching approach: The case of EFL learners at Puniv-Cazenga. *SAGE Open*, *7*(2). https://doi.org/10.1177/2158244017691077

Alharbi, B. S., Alharbi, M. A., & Alharbi, A. S. (2024). Unveiling the pedagogical potential: An in-depth analysis of EFL instructors' perspectives on YouTube for speaking skill enhancement. *Forum for Linguistic Studies*, *6*(5), 698–711. https://doi.org/10.30564/fls.v6i5.7119

Al-Khotaba, A. H., Alkhataba, E. H. A., Abdul-Hamid, S., & Ibrahim, B. (2020). Foreign language speaking anxiety: A psycholinguistic barrier affecting speaking achievement of Saudi EFL learners. *Arab World English Journal, 11*(4), 338–353. https://doi.org/10.2139/ssrn.3512637

Andoh, R. P. K., Mensah, D. Y., & Owusu, E. A. (2022). Trainers' pedagogical competencies and trainees' assimilation of training content. *Journal of Workplace Learning*, *34*(2), 133–149. https://doi.org/10.1108/JWL-02-2021-0024

Azir, I. D. A. (2019). Applying peer scaffolding to enhance the EFL vocational students' speaking skills. Ethical Lingua: Journal of Language Teaching and Literature, 6(2), 149–157. https://doi.org/10.30605/25409190.v6.149-157

Bedeker, M., Ospanbek, A., Simons, M., Yessenbekova, A., & Zhalgaspayev, M. (2024). 'I can easily switch to the Kazakh language, also to the Russian language': reimagining Kazakhstani CLIL implementation as a third space. *Language, Culture and Curriculum*, *37*(2), 121–138. https://doi.org/10.1080/07908318.2023.2245832

Begimbetova, G., Kassymova, G., & Abduldayev, Y. (2023). Criteria-based assessment model in the education system of Kazakhstan. *Iasaýı Ýnıversıtetiniń Habarshysy, 127*(1), 276–287. https://doi.org/10.47526/2023-1/2664-0686.23

Bhattacharya, U. (2022). "I am a parrot": Literacy ideologies and rote learning. *Journal of Literacy Research, 54*(2), 113–136. https://doi.org/10.1177/1086296X221098065

Borg, S. (2015). *Teacher cognition and language education*. Bloomsbury Publishing.

Borg, S. (2017). Teachers' beliefs and classroom practices. In P. Garrett & J. M. Cots (Eds.), *The Routledge handbook of language awareness* (pp. 75–91). Routledge. https://doi.org/10.4324/9781315676494-5

Borger, L. (2019). Evaluating a high-stakes EFL speaking test: Teachers' practices and views. *Moderna Sprak, 113*(1), 25–57. https://doi.org/10.58221/mosp.v113i1.7621

Bremner, N. (2021). The multiple meanings of "student-centred" or "learner-centred" education, and the case for a more flexible approach to defining it. *Comparative Education, 57*(2), 159–186. https://doi.org/10.1080/03050068.2020.1805863

British Educational Research Association [BERA]. (2018). *Ethical guidelines for educational research* (4th ed.).

Burns, A. (2019). Concepts for teaching speaking in the English language classroom. *LEARN Journal: Language Education and Acquisition Research Network, 12*(1), 1–11.

Byram, M., Porto, M., & Yulita, L. (2023). Beyond teaching languages for communication – Humanistic perspectives and practices. *Languages*, *8*(3), 166. https://doi.org/10.3390/languages8030166

Cao, Y., Jeyaraj, J. J., & Razali, A. B. (2024). Challenges in promoting learner autonomy in blended learning: Perspectives from English as a foreign language teachers in China. *International Journal of English Language Education, 12*(2), 122–142. https://doi.org/10.5296/ijele.v12i2.22272

Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.

Ciesielska, M., Boström, K. W., & Öhlander, M. (2018). Observation methods. In M. Ciesielska & D. Jemielniak (Eds.), *Qualitative methodologies in organization studies* (pp. 33–52). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-65442-3_2

Clarke, V., & Braun, V. (2016). Thematic analysis. *The Journal of Positive Psychology*, *12*(3), 297–298. https://doi.org/10.1080/17439760.2016.1262613

Cook, V. (2016). *Second language learning and language teaching* (5th ed.). Routledge. https://doi.org/10.4324/9781315883113

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications Inc.

Dong, M., Fan, J., & Xu, J. (2021). Differential washback effects of a high-stakes test on students' English learning process: Evidence from a large-scale stratified survey in China. *Asia Pacific Journal of Education, 43*(1), 252–269. https://doi.org/10.1080/02188791.2021.1918057

East, M. (2015). Coming to terms with innovative high-stakes assessment practice: Teachers' viewpoints on assessment reform. *Language Testing, 32*(1), 101–120. https://doi.org/10.1177/0265532214544393

Fives, H., & Buehl, M. M. (2016). Teachers' beliefs, in the context of policy reform. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 114–121. https://doi.org/10.1177/2372732215623554

Fleming, K., & Shinjee, B. (2024). English high-stakes testing and constructing the "international" in Kazakhstan and Mongolia. *Applied Linguistics Review*, *15*(4), 1253–1275. https://doi.org/10.1515/applirev-2022-0067

Gaipov, D., Kozhakhmet, S., Kassymova, G., & Tulepova, S. (2024). English medium instruction in the context of internationalisation of universities in Kazakhstan: A systematic review of literature. *Bulletin Series of Pedagogical Sciences*, *82*(2), 256–274. https://doi.org/10.51889/2959-5762.2024.82.2.015

Gass, S. M., & Mackey, A. (2016). *Stimulated recall methodology in applied linguistics and L2 research*. Routledge. https://doi.org/10.4324/9781315813349

Golombek, P. R., & Johnson, K. E. (2021). Recurrent restorying through language teacher narrative inquiry. *System*, *102*, 102601. https://doi.org/10.1016/j.system.2021.102601

Greer, T., & Potter, H. (2015). Turn-taking practices in multi-party EFL oral proficiency tests. *Journal of Applied Linguistics and Professional Practice, 5*(3), 297–320. https://doi.org/10.1558/japl.v5i3.297

Herrera, D. J. N., Quinto, R. P., Castro, M. D. C. R., Peñafiel, C. G. A., & Moposita, V. M. G. (2024). Enhancing speaking proficiency: Innovative pedagogical approaches for third-year Bachillerato students at Bicentenario 'D7' High School. *Ciencia Latina Revista Científica Multidisciplinar, 8*(4), 10315–10337. https://doi.org/10.37811/cl_rcm.v8i4.12184

Horwitz, E. K. (2020). *Becoming a language teacher: A practical guide to second language learning and teaching* (2nd ed.). Castledown Publishers. https://doi.org/10.29140/9780648184416

Hughes, R., & Reed, B. S. (2016). *Teaching and researching speaking*. Routledge. https://doi.org/10.4324/9781315692395

Husnaini, H., Syam, A. T., Niken, N., & Soares, A. (2024). Teachers' strategies in teaching speaking at high schools: Obstacles and challenges. *Journal of Foreign Language and Educational Research, 4*(2), 243–262. https://doi.org/10.53696/27753719.42165

Hussein, H. A. A.-K., Ahmed Alkhataba, E. H., Abdul-Hamid, S., & Bashir, I. (2019). Foreign language speaking anxiety: A psycholinguistic barrier affecting speaking achievement of Saudi EFL learners. *Arab World English Journal, 10*(4), 313–329. https://doi.org/10.24093/awej/vol10no4.23

Iddings, A. C. D. S., Combs, M. C., & Moll, L. (2021). In the arid zone: Drying out educational resources for English language learners through policy and practice. In H. R. Milner & K. Lomotey (Eds.), *Handbook of urban education* (2nd ed., pp. 452–463). Routledge. https://doi.org/10.4324/9780429331435-33

Ismail, K., Tuspekova, A., & Mustaffa, R. (2018). English oral communication in public secondary schools in Kazakhstan: Understanding its practice and challenges. *3L: Language, Linguistics, Literature, 24*(2), 118–131. https://doi.org/10.17576/3L-2018-2402-09

Johnson, K. E., & Golombek, P. R. (2020). Informing and transforming language teacher education pedagogy. *Language Teaching Research, 24*(1), 116–127. https://doi.org/10.1177/1362168818777539

Khassanov, G. K., Utegenova, K. T., Darisheva, T. M., Bozbayeva, A. G., Akhmetova, G., Sarkulova, D. S., & Utegenova, A. (2024). Implementing multilingual education in Kazakhstan: Students' perceptions and attitudes towards the status of English. *Theory and Practice in Language Studies*, *14*(8), 2316–2325. https://doi.org/10.17507/tpls.1408.04

Kılıçkaya, F. (2016). Washback effects of a high-stakes exam on lower secondary school English teachers' practices in the classroom. *Lublin Studies in Modern Languages and Literature, 40*(1), 116–134. https://doi.org/10.17951/lsmll.2016.40.1.116

Kisfalvi, V., & Oliver, D. (2015). Creating and maintaining a safe space in experiential learning. *Journal of Management Education, 39*(6), 713–740. https://doi.org/10.1177/1052562915574724

Lestari, M., & Margana, M. (2024). Communicative language teaching (CLT) implementation in *Kurikulum Merdeka*: A lesson from English teachers' voices. *Journal of Languages and Language Teaching, 12*(4), 1657–1672. https://doi.org/10.33394/jollt.v12i4.11266

Li, L. (2020). Teacher cognition about teaching and learning. In *Language teacher cognition* (pp. 87–134). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-51134-8_4

Liu, Y., & Chen, J. (2022, June). A study of the washback effect of oral English sampling test on primary school English teaching. In *2022 8th International Conference on Humanities and Social Science Research (ICHSSR 2022)* (pp. 1477–1480). Atlantis Press. https://doi.org/10.2991/assehr.k.220504.267

Liu, H., & Jin, J. (2024). Research trends in communicative language teaching: A bibliometric analysis from 2014 to 2023. *English Language Teaching and Linguistics Studies, 6*(5), Article 256. https://doi.org/10.22158/eltls.v6n5p256

Macaro, E., Tian, L., & Chu, L. (2020). First and second language use in English medium instruction contexts. *Language Teaching Research*, *24*(3), 382–402. https://doi.org/10.1177/1362168818783231

Meyer, K., & Willis, R. (2019). Looking back to move forward: The value of reflexive journaling for novice researchers. *Journal of Gerontological Social Work*, *62*(5), 578–585. https://doi.org/10.1080/01634372.2018.1559906

Meyers, R. (2023). English additional language student teachers' development of oral strategic competence and confidence during lectures. *Journal for Language Teaching, 57*(2), Article 5755. https://doi.org/10.56285/jltvol57iss2a5755

Ministerstvo obrazovaniya i nauki Respubliki Kazakhstan. (2017, 2 maya). *Prikaz Ministra obrazovaniya i nauki Respubliki Kazakhstan № 204 «Ob utverzhdenii Pravil provedeniya Edinogo natsionalnogo testirovaniya i vydachi sertifikata»* [Order of the Minister of Education and Science of the Republic of Kazakhstan No. 204 "On approval of the Rules for conducting the Unified National Testing and issuing the certificate"]. Ādilet.

Ministerstvo prosveshcheniya Respubliki Kazakhstan. (2022, 3 avgusta). *Prikaz Ministra prosveshcheniya Respubliki Kazakhstan № 348 «Ob utverzhdenii gosudarstvennykh obyazatelnykh standartov obrazovaniya vsekh urovney obrazovaniya»* [Order of the Minister of Education of the Republic of Kazakhstan No. 348 "On approval of the State compulsory educational standards for all levels of education"]. Ādilet.

Moriña, A. (2022). Faculty members who engage in inclusive pedagogy: Methodological and affective strategies for teaching. *Teaching in Higher Education, 27*(3), 371–386. https://doi.org/10.1080/13562517.2020.1724938

Nair, V., & Yunus, M. M. (2021). A systematic review of digital storytelling in improving speaking skills. *Sustainability, 13*(17), Article 9829. https://doi.org/10.3390/su13179829

Namaziandost, E., Homayouni, M., & Rahmani, P. (2020). The impact of cooperative learning approach on the development of EFL learners' speaking fluency. *Cogent Arts & Humanities, 7*(1), 1780811. https://doi.org/10.1080/23311983.2020.1780811

Nather, E. H. (2021). An action research study of promoting teachers' confidence and competence in teaching English meaningfully. *Journal of Education and Training, 8*(1), 11-37. https://doi.org/10.5296/jet.v8i1.18336

Newton, J. M., & Nation, I. S. (2020). *Teaching ESL/EFL listening and speaking*. Routledge. https://doi.org/10.4324/9780429203114

Nisha, P. R. (2024). Communicative language teaching (CLT) in improving speaking skills of tertiary level EFL students of Bangladesh. *FOSTER: Journal of English Language Teaching*, *5*(2), 120–130. https://doi.org/10.24256/foster-jelt.v5i2.174

Noor, F. (2024). Cultural variations affecting English language instruction for EFL learners and teachers. In *Research trends in language, literature & linguistics* (pp. 38–59). Iterative International Publisher. https://doi.org/10.58532/v3bflt6p1ch6

Pakula, H.-M. (2019). Teaching speaking. *Apples - Journal of Applied Language Studies*, *13*(1), 95–111. https://doi.org/10.17011/apples/urn.201903011691

Pan, J. (2024). Inquiry into teaching listening and speaking in high school English classes based on an activity perspective. *Academic Journal of Management and Social Sciences, 8*(2), 114–117. https://doi.org/10.54097/0wcm9z22

Panggabean, C. I. T., & Triassanti, R. (2020). The implementation of metacognitive strategy training to enhance EFL students' oral presentation skill. *English Education: Journal of English Teaching and Research, 5*(1), 32–40. https://doi.org/10.29407/jetar.v5i1.14324

Raj JR, K. A., & Baisel, A. (2024). Mobile learning and verbal proficiency: A cross-sectional study investigating English speaking skills enhancement. *Journal of Language Teaching and Research*, *15*(5), 1409–1418. https://doi.org/10.17507/jltr.1505.03

Renz, S. M., Carrington, J. M., & Badger, T. A. (2018). Two strategies for qualitative content analysis: An intramethod approach to triangulation. *Qualitative Health Research, 28*(5), 824–831. https://doi.org/10.1177/1049732317753586

Richards, J. C. (2017). Teaching English through English: Proficiency, pedagogy and performance. *RELC Journal, 48*(1), 7–30. https://doi.org/10.1177/0033688217690059

Rozimela, Y., Fatimah, S., Adnan, A., & Tresnadewi, S. (2024). EFL teachers' perceived productive skills for effective teacher professional development program. *Studies in English Language and Education*, *11*(3), 1493–1514. https://doi.org/10.24815/siele.v11i3.37277

Salam, M. Y., & Luksfinanto, Y. (2024). A comprehensive review of communicative language teaching (CLT) in modern classrooms. *Lingeduca: Journal of Language and Education Studies, 3*(1), 58–70. https://doi.org/10.70177/lingeduca.v3i1.1338s

Saptiany, S. G., & Prabowo, B. A. (2024). Speaking proficiency among English specific purpose students: A literature review on assessment and pedagogical approaches. *LITERACY: International Scientific Journals of Social, Education, Humanities, 3*(1), 36–48. https://doi.org/10.56910/literacy.v3i1.1392

Sharma, R. L. (2024). Exploring the landscape of challenges and opportunities in teaching speaking skills. *International Journal of Advanced Multidisciplinary Research and Studies*, *4*(3), 74–78. https://doi.org/10.62225/2583049x.2024.4.3.2745

Shin, J. Y., Dixon, L. Q., & Choi, Y. (2020). An updated review on use of L1 in foreign language classrooms. *Journal of Multilingual and Multicultural Development, 41*(5), 406–419. https://doi.org/10.1080/01434632.2019.1684928

Spratt, M. (2017). CLIL implementation in foreign language contexts: Exploring challenges and perspectives. *Research Papers in Language Teaching and Learning, 8*(1), 44–61.

Stake, R. E. (2013). *Multiple case study analysis*. Guilford press.

Sun, Q., & Zhang, L. J. (2021). A sociocultural perspective on English-as-a-foreign-language (EFL) teachers' cognitions about form-focused instruction. *Frontiers in Psychology*, *12*, 593172. https://doi.org/10.3389/fpsyg.2021.593172

Tajik, M. A., Namyssova, G., Shamatov, D., Manan, S. A., Zhunussova, G., & Antwi, S. K. (2023). Navigating the potentials and barriers to EMI in the post-Soviet region: Insights from Kazakhstani university students and instructors. *International Journal of Multilingualism*. https://doi.org/10.1080/14790718.2023.2265428

Tarusha, F., & Bushi, J. (2024). The role of classroom observation and its impact on improving teachers' teaching practices. *European Journal of Theoretical and Applied Sciences, 2*(2), 718–723. https://doi.org/10.59324/ejtas.2024.2(2).63

Tondeur, J., Van Braak, J., Ertmer, P. A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Educational Technology Research and Development, 65*, 555–575. https://doi.org/10.1007/s11423-016-9481-2

Tran, N. H. N., Amado, C. A. da E. F., & Santos, S. P. dos. (2023). Challenges and success factors of transnational higher education: a systematic review. *Studies in Higher Education*, *48*(1), 113–136. https://doi.org/10.1080/03075079.2022.2121813

Tuspekova, A., Mustaffa, R., & Ismail, K. (2020). Understanding English speaking practice in public schools in Kazakhstan: A case study in Almaty. *3L: The Southeast Asian Journal of English Language Studies, 26*(1), 171–185. https://doi.org/10.17576/3L-2020-2601-12

Vercellotti, M. L. (2017). The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied Linguistics, 38*(1), 90–111. https://doi.org/10.1093/applin/amv002

Villada, C., Hidalgo, V., Almela, M., & Salvador, A. (2018). Assessing performance on an evaluated speaking task: The role of self-efficacy, anxiety, and cardiac autonomic reactivity. *Journal of Psychophysiology*, *32*(2), 64–74. https://doi.org/10.1027/0269-8803/a000185

Wei, L., Wan Mustapha, W. Z., & Binti Awang, S. (2024). An experimental study of communicative language teaching (CLT) in university comprehensive English class. *International Journal of Linguistics, 16*(3), 83-97. https://doi.org/10.5296/ijl.v16i3.21844

Willis, D. (2015). Conversational English: Teaching spontaneity. In R. Hughes (Ed.), *Issues in teaching, learning and testing speaking in a second language* (pp. 3–18). Springer. https://doi.org/10.1007/978-3-642-38339-7_1

Wu, X. (2023). A longitudinal study of EFL teacher agency and sustainable identity development: A positioning theory perspective. *Sustainability, 15*(1), 48. https://doi.org/10.3390/su15010048

Xi, J., & Lantolf, J. P. (2021). Scaffolding and the zone of proximal development: A problematic relationship. *Journal for the Theory of Social Behaviour, 51*(1), 25–48. https://doi.org/10.1111/jtsb.12260

Xuan Mai, L., Ngoc, L. K., & Thao, L. T. (2024). Factors hindering student participation in English-speaking classes: Student and lecturer perceptions. *SAGE Open*, *14*(3). https://doi.org/10.1177/21582440241266297

Yessenbekova, K. (2024). Understanding the dynamics of English-speaking challenges for students in Kazakhstan. *Asian Education and Development Studies*, *13*(4), 373–385. https://doi.org/10.1108/AEDS-04-2024-0083

Yin, R. K. (2018). *Case study research and applications* (6th ed.). Sage Publications.

Zhu, Q., & Carless, D. (2018). Dialogue within peer feedback processes: Clarification and negotiation of meaning. *Higher Education Research & Development, 37*(4), 883–897. https://doi.org/10.1080/07294360.2018.1446417

Corresponding author: Askat Tleuov
Email: A.Tleuov@kimep.kz