



Language Choice in Pre-Speaking Discussions and EFL Learners' Oral Performance, Willingness to Communicate, and Self-Confidence

Masoume Ahmadi^{1*}

Shahab Moradi²

Abstract

This explanatory sequential mixed-methods study examined how the language used in pre-speaking discussions (PSDs) relates to Iranian EFL learners' oral language performance (OLP), willingness to communicate (WTC), and linguistic self-confidence (LSC). Thirty-six male B2-level learners in a university speaking workshop took part. We implemented three PSD language modes across five intact discussion groups: L1-only (Turkish $n = 7$, Kurdish $n = 6$, Arabic $n = 7$), L2-only (English, $n = 9$), and mixed L1–L2 ($n = 7$). Over 10 sessions, each group completed a 10-minute PSD in its assigned mode before a 20-minute English-speaking task. We measured WTC, LSC, and OLP before and after the intervention. One-way ANCOVAs on post-test scores, controlling for pre-test scores, showed significant between-group effects for WTC, $F(4, 30) = 4.56$, $p = .005$, partial $\eta^2 = .38$, and LSC, $F(4, 30) = 6.21$, $p = .001$, partial $\eta^2 = .45$, but not for OLP, $F(4, 30) = 1.84$, $p = .148$. Exploratory within-group comparisons showed significant LSC gains in all four groups whose PSDs involved the first language ($g_{av} = 0.34$ – 0.55) and a significant WTC gain in the mixed-language group ($g_{av} = 0.34$); the English-only group changed significantly on neither affective outcome, and no significant OLP change was detected in any group. Interviews with 25 participants indicated that L1-involving PSDs were experienced as reducing anxiety, supporting brainstorming, and increasing readiness to speak. Because the groups were intact and participants were not randomly assigned, the findings are interpreted as associational rather than causal.

Keywords: Linguistic self-confidence (LSC), Oral language performance (OLP), Pre-speaking discussion (PSD), Willingness to communicate (WTC)

* Review History:

Received: 24/07/2026

Revised: 26/08/2026

Accepted: 09/09/2026

1. Assistant Professor, Department of English Language Teaching, Farhangian University, P.O. Box 14665-889, Tehran, Iran. (Corresponding Author) masoumeahmadi@cfu.ac.ir
2. PhD Candidate and Visiting Lecturer, Department of English Language and Literature, Arak University, Arak, Iran. mshahabmoradi@gmail.com

How to cite this article:

Ahmadi, M. & Moradi, Sh. (2026). Language Choice in Pre-Speaking Discussions and EFL Learners' Oral Performance, Willingness to Communicate, and Self-Confidence. *Teaching English as a Second Language Quarterly*, 45(4), 109-131. <https://doi.org/10.22099/tesl.2026.56641.3522>



A central objective of second language learning is to achieve proficiency in oral communication, enabling learners to share information, beliefs, and ideas effectively and confidently. Achieving this objective involves developing oral proficiency, a critical yet challenging endeavor because it relies on an interplay of linguistic and psychological factors (Derwing & Munro, 2015; Richards, 2008) and encompasses complex processes such as conceptualization, formulation, and articulation (Levelt, 1993). Consequently, many language learners hesitate to participate in oral communication activities, inhibited by a fear of making errors, often compounded by limited fluency, anxiety, and low self-confidence (Clément et al., 2003; Hanifa, 2018; Horwitz, 2010). Recent studies have emphasized the need for socio-affective support in oral language instruction to reduce anxiety and boost learners' engagement and confidence (Ebadi & Azizimajd, 2026). Addressing these challenges requires effective pedagogical strategies to build learners' confidence and willingness to engage in oral communication.

Pre-speaking discussions (PSDs) are a valuable pedagogical strategy that positively prepares learners for oral tasks by giving them the chance to practice language use and build confidence (Ellis, 2009). Within task-based language teaching (TBLT), such preparation plays a particularly vital role, where pre-task activities support performance and alleviate communication anxiety (Skehan, 1998). Recent research supports using guided pre-task planning to improve speaking performance in EFL learners (Cheng & Sun, 2026; Khezrlou, 2023). Recent process-oriented studies indicate that the language used during planning shapes preparation quality, with L1 facilitating idea generation and L2 supporting rehearsal for delivery (Aubrey & Philpott, 2025). These findings align with the potential of mixed-language PSDs to balance conceptual exploration and target-language practice.

The role that learners' first language (L1) plays in EFL contexts has been debated for decades. Proponents of exclusive L2 use argue it maximizes exposure and interaction opportunities (Turnbull & Dailey-O'Cain, 2009), while others advocate strategic L1 use for scaffolding and affective support (Hall & Cook, 2012). In task-based and task-supported contexts, exclusive L2 policies have been linked to increased oral production and fluency (Thompson & Harrison, 2014). In contrast, judicious L1 use can aid pre-task planning and idea organization without reducing target-language output (Izquierdo et al., 2016). This study focuses on one specific dimension of this debate: the language used during PSDs.

Beyond language choice, a learner's participation in oral communication is also vitally determined by their willingness to communicate (WTC). This factor is shaped by a mix of immediate situational factors and more stable factors such as the classroom environment, anxiety and self-confidence (Li et al., 2025; MacIntyre et al., 1998; Peng & Woodrow, 2010). LSC significantly affects learners' readiness to engage in communicative tasks because it directly influences their perception of their ability to succeed in such tasks (Clément et al., 2003). Sadoughi and Hejazi (2024) show that strengthening LSC through task-based planning and confidence-building activities significantly enhances learners' WTC in EFL classrooms.

Thus, pedagogical interventions targeting emotional and cognitive readiness are vital for supporting active oral communication.

While the advantages of pre-task planning are well-documented within the context of L2 writing (e.g., Neumann & McDonough, 2015), particularly the use of L1 to generate ideas and reduce cognitive load (Adams et al., 2015; DiCamilla & Antón, 2012; Wang & Wen, 2002)—its application to oral communication remains comparatively less researched. Specifically, no prior study has simultaneously compared L1-only, L2-only, and mixed-language PSDs in multilingual Iranian EFL classrooms using a mixed-methods design. This combination of context, design, and focus on WTC, LSC, and OLP represents a unique contribution to both SLA research and pedagogical policy debates on English-only versus translanguaging practices. To address this gap, the current study examined the effects of PSD language choice on Iranian EFL learners' OLP, WTC, and LSC. Accordingly, the present study set out to answer the following research questions:

- (1) To what extent is the language mode used during PSDs — L1-only (implemented in Turkish, Kurdish, or Arabic), L2-only (English), or mixed-language (L1 and L2) — associated with differences in learners' OLP, WTC, and LSC, and do outcomes vary across the three L1 implementations?
- (2) How do learners perceive and experience the use of different PSD language modes in relation to their OLP, WTC, and LSC?
- (3) How do the qualitative findings complement and contextualize the quantitative patterns observed across the different PSD language modes?

Literature Review

Pre-speaking discussions (PSDs) can be conceptualized as a form of pre-task planning in which learners are given time to generate ideas, organize content, and rehearse relevant language before completing an oral task. Such preparation may reduce the processing demands of real-time speech production by allowing learners to allocate more attentional resources to aspects of performance such as fluency, accuracy, and complexity (Ellis, 2009; Skehan, 1998). Theoretically, PSDs align with Vygotsky's (1978) sociocultural theory; they act as a scaffold where peer interaction aids the internalization of thought processes, thereby supporting oral performance.

Empirical research supports these benefits. For instance, Gao and Min (2021) reported that L1 and L2 prewriting discussions differed in the cognitive and affective support they afforded. Research on related task-based interventions, such as task repetition, has also shown benefits for aspects of L2 oral production (Fukuta, 2016). This general trend suggests that various pre-task planning forms can positively affect fluency, accuracy, and complexity in oral performance, as confirmed in recent large-scale reviews of the literature (Johnson & Abdi Tabari, 2022). However, these studies treat PSDs as a general pre-task strategy and do not address how the language of the PSD itself might shape these outcomes.

First Language (L1) and Second Language (L2) in Language Learning

The role of learners' first language, the second language, or a combination of both in classrooms has long been debated. For decades, traditional methods like the Direct Method promoted reliance on L2 to maximize exposure to the target language and facilitate acquisition (Krashen, 1982). Yet, more recent approaches, including translanguaging and code-switching, have recognized the cognitive and affective benefits of L1 as a tool for learning L2 (Cook, 2001). An increasing number of studies support the use of L1 in the classroom. Shin et al. (2020) concluded that intentionally incorporating L1 in teaching can maximize L2 learning and should be integrated as a curricular resource. García Mayo and Hidalgo (2017) found that L1 served a facilitative role among young EFL learners, particularly for vocabulary acquisition, although its use varied across different instructional contexts. According to advocates of inclusive L1 use, learners' L1 serves as a scaffolding tool, helping them process complex concepts, brainstorm and organize ideas, and overcome anxiety during challenging tasks (de la Fuente & Goldenberg, 2022; Hall & Cook, 2012).

Despite strong arguments for L1 use, many EFL contexts have adopted an English-only policy that mandates L2 as the exclusive medium of instruction. This perspective was promoted by Krashen (1982), arguing for the importance of maximum exposure to L2. He argued that lessons should be conducted primarily in L2 because of the hypothesized strong relationship between comprehensible L2 input and language proficiency. Following Krashen's influential work, additional empirical studies have explored the benefits of L2 immersion for communicative competence. For example, studies have reported better learning outcomes when vocabulary is taught exclusively in L2 in Spanish EFL classrooms (Shabaka-Fernández, 2021). Debreli (2016) investigated teacher perceptions under strict English-only policies, highlighting that exclusive L2 use enhanced students' communicative abilities but required careful balancing to avoid instructional challenges. Institutionally mandated L2 use has also been associated with greater engagement and more authentic language use (Ford, 2009; Turnbull, 2018).

Empirical research has also pointed to the positive contribution of L1 use in pre-task activities in the EFL context. Aubrey et al. (2020) demonstrated that pre-task planning in L1 generated more comprehensive ideas among Japanese EFL learners. By comparison, subsequent OLP in L2 showed limited benefits, suggesting the prominent role of L1 planning in conceptual organization. Similarly, Galdeano Chasco and García Mayo (2024) indicated that L1 planning encouraged greater attention to L2 forms and increased engagement in language-related episodes compared to L2 planning. Consistently, Ellis (2021) also found that pre-task planning, especially when collaborative and L1-based, could improve learners' fluency and writing complexity.

Given the respective strengths and limitations of L1 and L2-only approaches, some researchers advocate a mixed approach that combines the advantages of both languages. The logic behind this dual approach is that it can reduce cognitive load and provide a low-pressure context for practice. For example, Scott and de la Fuente (2008) showed that brainstorming in

L1 during the initial stages of pre-task planning could help learners conceptualize ideas more effectively, while transitioning to L2 allowed for practice in the target language. Research into such mixed strategies has shown promising results; a study by Jouibar et al. (2022) found that controlled learner code-switching did not significantly affect speaking accuracy or fluency but was associated with significantly higher speaking complexity scores. Recent empirical research also shows that incorporating translanguaging into collaborative pre-task planning may lead to more complex and efficient speaking performance in EFL learners (Luo & Sun, 2025). Complementary evidence from Li and Qu (2024) shows that allowing translanguaging in collaborative talk enhances the depth and exploratory quality of interaction while it maintains L2 output levels, and supports the pedagogical viability of mixed-language PSDs. Leeming et al. (2022) also found that specific planning activities, such as flexible language choice, mediate how much planning benefits transfer to task performance. They emphasized the importance of examining planning processes alongside outcomes. By leveraging the cognitive and affective benefits of L1 while maintaining a focus on L2 immersion, this mixed approach may offer a practical solution to the ongoing debate.

Willingness to Communicate (WTC) and Linguistic Self-Confidence (LSC)

Willingness to communicate (WTC) has developed as a fundamental construct in SLA research, particularly in understanding what drives learners to initiate and sustain oral interaction. In L2 learning, success largely depends on a learner's WTC, defined as a learner's readiness to initiate communication when an opportunity arises. According to the foundational model from MacIntyre et al. (1998), some psychological elements, including LSC, anxiety, and perceived competence, can affect WTC. Other researchers, such as Dörnyei (2005), found that learners with higher WTC are more likely to take risks during communication and, subsequently, practice more frequently; as a result, they reach greater overall proficiency. This makes WTC a critical component for language educators, particularly in oral communication tasks.

LSC, an important antecedent of WTC, is defined as learners' confidence in their ability to use the target language effectively (Clément, 1980; see also MacIntyre et al., 1998). While WTC focuses on learners' readiness to initiate communication, LSC underpins this willingness by shaping their understanding of communicative competence and comfort in using the target language. Previous research reveals the strong correlation between higher levels of LSC and reduced anxiety and enhanced participation in communicative activities (Clément et al., 2003; Pyun et al., 2014). The central role of anxiety is reinforced by recent findings from the Iranian EFL context, which show that L2 anxiety significantly mediates the relationship between a learner's motivation and their ultimate WTC (Sadoughi & Hejazi, 2024). As Liu and Jackson (2008) explain, pre-task activities that alleviate anxiety and give learners a sense of control can significantly enhance LSC and WTC.

Recent research in the Iranian EFL context further highlights the importance of structured

interaction and instructional support in promoting learners' WTC. In the Iranian context, structured peer interaction (Zohrabi & Jafari, 2020) and teacher-mediated strategies in online classes (Zohrabi & Bimesl, 2022) have both been linked to increased WTC. Together, these findings suggest that interaction organization and the support provided before or during communication can meaningfully shape learners' readiness to participate in L2 communication.

Despite abundant research on the cognitive and affective benefits of pre-task planning, its specific impact on affective variables like WTC during oral communication has been explored less systematically. This gap is significant because MacIntyre et al. (1998) report that self-confidence and reduced anxiety are likely to enhance WTC. A key tension exists in the literature. From one perspective, Liu and Jackson (2008) reveal that using L1 in PSDs might provide a psychological safety net that reduces anxiety and enhances WTC, especially in learners who struggle with L2 fluency. Another perspective argues for L2 immersion, highlighting that continuous L2 exposure builds confidence in the target language, which is essential for increasing WTC over time (Clément et al., 2003).

Bringing these perspectives together, Vygotsky's sociocultural theory (1978) positions PSDs as scaffolding that externalizes thought and reduces anxiety; Clément's (1980) LSC model highlights how such scaffolding strengthens competence perceptions; and MacIntyre et al.'s (1998) WTC model shows that these perceptions influence communicative readiness. Within this integrated framework, language choice in PSDs becomes a crucial variable: L1 may lower anxiety, L2 may maximize practice, and mixed-language may combine both. The present study is therefore grounded in an integrated sociocultural-affective framework that links scaffolding, confidence, and WTC.

Although the literature offers a solid theoretical framework for the use of PSDs, empirical studies investigating the direct effects of different PSD language modalities on multilingual learners' affective states, such as WTC and LSC, remain limited. This study aimed to address this specific gap. In the following section, details of the methodology employed to do so are discussed.

Method

Participants

This study involved 36 male EFL learners enrolled in a speaking workshop at a teacher education university in Iran. They were first-year TEFL majors, aged 20–23 years ($M = 21.2$). Their English proficiency was established as B2 (upper-intermediate) using the Speaking module of the Oxford Test of English (OTE; Oxford University Press, 2019) prior to the intervention. Persian was the common language of all participants, while their first languages were Turkish ($n = 14$, 38.9%), Kurdish ($n = 10$, 27.8%), or Arabic ($n = 12$, 33.3%). Table 1 summarizes participant characteristics.

Table 1
Participant Characteristics

Characteristic	Category	<i>n</i>	%
Gender	Male	36	100.0
Age	20–23 years	36	100.0
Academic status	First-year undergraduate TEFL students	36	100.0
English proficiency	B2 (upper-intermediate)	36	100.0
Common language	Persian	36	100.0
First language	Turkish	14	38.9
	Kurdish	10	27.8
	Arabic	12	33.3
PSD language mode	L1-only	20	55.6
	English-only (L2-only)	9	25.0
	Mixed L1–L2	7	19.4

Note. First-language counts refer to the full sample. The English-only and mixed-language groups were linguistically mixed.

The instructional factor in this study was PSD language mode, conceptualized at three levels: L1-only, L2-only (English), and mixed L1–L2. Because the L1-only mode can only be enacted within linguistically homogeneous discussion groups, this level was necessarily implemented through three intact groups defined by the participants' shared first language: Turkish (TG, $n = 7$), Kurdish (KG, $n = 6$), and Arabic (AG, $n = 7$). The L2-only mode was implemented in a single English-only group (EG, $n = 9$) and the mixed mode in a single bilingual group (MG, $n = 7$), giving five intact groups and a total of 36 participants. Within each group, participants worked in subgroups of three to four members during the intervention.

The three L1 groups are best understood as three implementations of a single language mode rather than as three theoretically distinct treatments; no hypothesis was advanced concerning differential effects of Turkish, Kurdish, or Arabic per se. They were nevertheless retained as separate units in the statistical model for two reasons. First, the three groups were not interchangeable at baseline: pre-test LSC ranged from 52.6 (TG) to 60.0 (AG) and pre-test OLP from 119.8 (KG) to 126.6 (TG), so treating them as exchangeable replicates could not simply be assumed and was better examined empirically. Second, retaining them allows a descriptive comparison of whether the L1-only mode operated comparably across the three languages, as reported below.

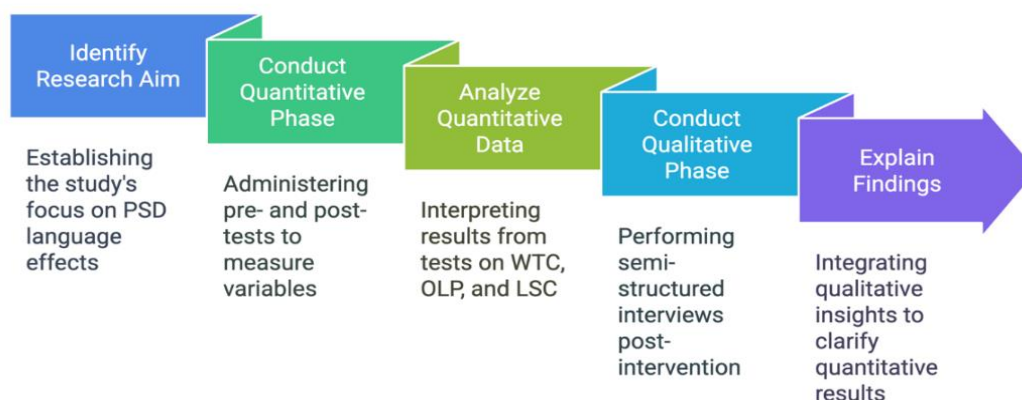
Because data collection was complete and the sample was fixed by workshop enrollment, a sensitivity analysis was more informative than a post hoc power calculation. With $N = 36$ in five intact groups and one covariate, the omnibus ANCOVA had approximately 80% power to detect $f \approx 0.60$ at $\alpha = .05$ and $f \approx 0.66$ at the Bonferroni-adjusted $\alpha = .017$. The design was therefore sensitive only to large between-group effects; non-significant results, particularly for OLP, are inconclusive rather than evidence of no effect, and comparisons among the three L1 groups are correspondingly underpowered and exploratory.

Research design

Following Creswell and Plano Clark (2017), we used an explanatory sequential mixed-methods design (Figure 1). In the quantitative phase, participants completed pre- and post-test measures of OLP, WTC, and LSC. In the subsequent qualitative phase, a randomly selected subsample took part in semi-structured interviews intended to explain and contextualize the quantitative patterns.

Figure 1

Research flowchart for the explanatory sequential mixed methods study design procedure



Instruments

Oxford Test of English (Oxford University Press, 2019): To assess participants' English proficiency, the Speaking module of the Oxford Test of English (OTE) (Oxford University Press, 2019) was administered. The Speaking module comprises four parts: an interview involving questions on everyday topics, two voicemail tasks, a short talk on a given issue or scenario, and follow-up questions related to the short talk. Participants' speaking proficiency was evaluated in terms of pronunciation, fluency, grammar, and vocabulary, and their performance was used to establish their proficiency level according to the Common European Framework of Reference for Languages (CEFR). The same module was re-administered after the intervention, and its score served as the measure of oral language performance (OLP)

WTC Questionnaire: To assess WTC, the study used a 10-item WTC questionnaire developed by Peng and Woodrow (2010) for classroom use. The questionnaire used a 5-point Likert scale, with 1 representing almost never willing and 5 almost always willing; total scores ranged from 10 to 50, and higher scores reflected greater WTC. Iranian EFL studies previously confirmed the reliability of this tool (e.g., Zarrinabadi & Tanbakooei, 2016), and this study reaffirmed it ($\alpha = .84$).

LSC Scale: The LSC was measured by the 16-item scale developed by Clément (1980) in four dimensions: comfort in speaking ($n = 4$), perceived communicative ability ($n = 4$), anxiety levels ($n = 4$), and self-perceived linguistic competence ($n = 4$). Responses were rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree); total scores ranged from 16 to 80, with higher values reflecting greater LSC.

Semi-Structured Interviews: Semi-structured interviews were conducted after the intervention to provide qualitative explanations for the quantitative findings and explore participants' perceptions of how PSD language modes influenced their WTC and LSC. The interview protocol comprised seven open-ended questions addressing participants' feelings about speaking English in class, the perceived role of PSD language in reducing speaking-related anxiety, and the advantages and disadvantages of using L1, L2, or both during PSD. We used follow-up questions to elicit further clarification when needed. Interviews were conducted in Persian, audio-recorded with participants' consent, transcribed verbatim, and thematically analyzed. The researchers randomly selected 25 participants (five from each of the five groups) from the quantitative sample for the qualitative phase. The interviews aimed to provide deeper insight into participants' experiences and help explain the quantitative findings concerning OLP, WTC, and LSC.

Procedure

Before the study began, the researchers explained the study's objectives to all participants, guaranteed that their data would remain confidential, and obtained their written consent to participate. The initial quantitative data collection consisted of a pre-test phase in which participants completed the WTC and LSC scales in printed form, and the OTE determined their level of English-speaking proficiency. In the second session, participants were assigned to one of five discussion groups: the three L1-only groups (TG, KG, AG) were formed based on a shared first language, whereas the English-only (EG) and mixed-language (MG) groups were linguistically mixed.

The syllabus of this general speaking course emphasized communicative English skills, and all classroom communication occurred predominantly in English (L2). The second researcher was the workshop instructor. She focused primarily on task-based language teaching, presenting real-life speaking tasks in three phases: pre-speaking, while-speaking, and post-speaking. The intervention content was based on the first five units of *Q: Skills for Success: Listening and Speaking 4*, using themes such as personal experiences and global issues to encourage active discussion (Freire & Jones, 2015). The workshop spanned 10 sessions, with four hours of instruction per week. Each session began with a 10-minute PSD in the group's assigned language, where each group was divided into smaller subgroups of 3–4 members, and rotating peer observers monitored language use. A 20-minute group discussion task immediately followed the PSD in English, in which all groups applied ideas from the preparatory phase in activities such as role-plays, presentations, and debates. After the 10-session intervention, participants completed the WTC and LSC scales again as post-tests to assess changes in WTC and LSC, and they retook the speaking module of the OTE.

Subsequently, the qualitative phase involved semi-structured interviews with 25 participants, with five participants selected from each group. Each interview lasted 10–20 minutes (358 minutes in total) and was conducted in Persian to facilitate participants' expression

of their experiences. The interviews were transcribed verbatim, and the second researcher translated the transcripts into English; the first researcher then checked the translations against the original Persian transcripts to ensure accuracy and consistency.

Data analysis

Quantitative data were analyzed using SPSS version 20.0. We calculated descriptive statistics (means and standard deviations) for OLP, WTC, and LSC at pre-test and post-test across the five groups. Paired-samples t-tests examined within-group changes, and one-way ANCOVAs compared post-test scores across groups (TG, KG, AG, EG, MG), with the corresponding pre-test score entered as a covariate. Tukey-adjusted post-hoc comparisons followed significant ANCOVA effects.

Because the L1-only mode was implemented through three intact groups, we examined the comparability of those three implementations descriptively by comparing their post-test scores and within-group change estimates. Given six to seven participants per group, this comparison is exploratory and is not treated as a formal test of differential L1 effects.

We conducted separate univariate ANCOVAs, rather than a single MANCOVA, for OLP, WTC, and LSC. This choice reflects the small and unequal group sizes (as few as six participants per group), which would have limited the power and stability of a multivariate model given the additional covariance-matrix assumptions MANCOVA requires (Stevens, 2009; Tabachnick & Fidell, 2019), as well as the research questions' focus on each outcome individually. To guard against inflated Type I error from conducting three separate ANCOVAs, a Bonferroni-adjusted significance threshold ($\alpha = .05/3 \approx .017$) was applied when interpreting the results. We checked normality (Shapiro–Wilk), homogeneity of variance (Levene's test), and homogeneity of regression slopes; results are reported below. Effect sizes are reported as partial eta squared (η^2p) for the ANCOVAs and, for the paired comparisons, as d_{av} (mean change over the average of the pre- and post-test standard deviations), its Hedges-corrected counterpart g_{av} , and d_z (mean change over the standard deviation of the difference scores; Lakens, 2013). Because every cell had $n < 20$, g_{av} is the preferred estimate.

Qualitative interview data were analyzed thematically using a hybrid deductive–inductive approach, with initial codes drawn from the research questions and literature and additional codes emerging from the data. Two researchers coded the dataset independently, resolved discrepancies through discussion, and maintained an audit trail of coding decisions to support credibility and trustworthiness. The resulting themes were used to interpret and contextualize the quantitative findings.

Results

The results of the quantitative and qualitative phases are presented according to the three research questions.

RQ 1: The effects of PSD language mode on OLP, WTC, and LSC

Differences across groups

We used descriptive and inferential statistics to evaluate differences associated with PSD language mode. Table 2 presents the unadjusted descriptive statistics (M and SD) for pre- and post-test scores across the five groups. Shapiro–Wilk tests did not indicate significant departures from normality for the residuals of any dependent variable ($p > .05$), and Levene's test indicated no violation of homogeneity of variance for WTC ($p = .23$), LSC ($p = .11$), or OLP ($p = .17$). The covariate $\times$ group interaction was non-significant for all three outcomes ($p > .05$), supporting the assumption of homogeneous regression slopes. Given the small cell sizes ($n = 6$ – 9), these assumption tests had very limited statistical power, and non-significant assumption checks should not be read as positive evidence that the assumptions hold.

Table 2

Means and Standard Deviations for Pre- and Post-Test Measures across Groups

Group	n	Pre WTC M (SD)	Post WTC M (SD)	Pre LSC M (SD)	Post LSC M (SD)	Pre OLP M (SD)	Post OLP M (SD)
TG	7	39.9 (13.7)	43.0 (13.7)	52.6 (15.0)	58.4 (15.3)	126.6 (8.9)	129.1 (9.2)
KG	6	41.5 (10.3)	41.5 (10.4)	58.2 (13.8)	66.5 (11.7)	119.8 (6.7)	122.5 (8.3)
AG	7	41.4 (8.5)	41.6 (8.8)	60.0 (11.9)	67.1 (11.1)	121.7 (6.4)	123.3 (7.5)
EG	9	36.6 (7.7)	37.0 (7.7)	49.0 (11.3)	51.4 (11.6)	126.1 (8.6)	126.9 (8.6)
MG	7	42.3 (8.7)	45.7 (8.7)	60.3 (11.8)	66.4 (13.4)	121.4 (9.2)	123.4 (9.1)

Note. TG = Turkish L1; KG = Kurdish L1; AG = Arabic L1; EG = English-only; MG = mixed-language. WTC = willingness to communicate (possible range 10–50); LSC = linguistic self-confidence (possible range 16–80); OLP = Oxford Test of English Speaking module score (reported range 0–140; CEFR B2 band = 111–140). Values are unadjusted raw scores reported to one decimal place.

We conducted a one-way ANCOVA for each dependent variable, with the corresponding pre-test score entered as a covariate (Table 3). Because three ANCOVAs were conducted on outcomes measured in the same participants, a Bonferroni-adjusted threshold of $\alpha = .05/3 \approx .017$ was applied to the three omnibus tests. Pairwise comparisons within a significant model were Tukey-adjusted and additionally evaluated against the same .017 threshold; this double adjustment is deliberately conservative. Because participants were allocated to conditions by linguistic requirement rather than at random, and because each condition was implemented in a single intact group, we interpret all between-group results as associational and not causal.

For WTC, the ANCOVA yielded a significant main effect of group, $F(4, 30) = 4.56$, $p = .005$, partial $\eta^2 = .38$, which remained significant under the adjusted threshold. Tukey-adjusted pairwise comparisons showed significantly higher post-test WTC for MG than for EG ($p = .008$). Because EG had the lowest pre-test WTC of any group (Table 2), baseline adjustment would raise EG's estimated marginal mean relative to its raw mean; the adjusted difference is therefore expected to be smaller than the raw 8.7-point gap. This single pairwise result should not be read as evidence that the mixed mode outperformed the L1-only mode.

The ANCOVA also showed a significant main effect of group on LSC, $F(4, 30) = 6.21$, $p =$

.001, partial $\eta^2 = .45$. As shown in Table 2, unadjusted post-test LSC was higher in all four groups whose PSDs involved the first language (TG, KG, AG, MG; $M = 58.4$ – 67.1) than in the English-only group ($M = 51.4$), and the four L1-involving groups differed little from one another. Because pre-test LSC also differed appreciably across groups, this pattern is described descriptively and is not tested with specific pairwise hypotheses. The data support a distinction between L1-involving modes and the exclusively L2 mode, not between any individual L1 and the mixed mode.

In contrast, no significant main effect of group was observed for post-test OLP, $F(4, 30) = 1.84$, $p = .148$, partial $\eta^2 = .20$.

Table 3

Analysis of Covariance (ANCOVA) for Post-Test Scores Controlling for Pre-Test Baseline

Dependent Variable	df	F	p	Partial η^2
WTC	4, 30	4.56	.005	.38
LSC	4, 30	6.21	.001	.45
OLP	4, 30	1.84	.148	.20

Note. Pre-test scores served as the covariate in each model. Bonferroni-adjusted significance threshold: $\alpha = .017$.

For the second component of RQ1, the three implementations of the L1-only mode produced broadly comparable results. Turkish, Kurdish, and Arabic groups showed similar unadjusted post-test scores on all three outcomes (LSC: 58.4, 66.5, 67.1; WTC: 43.0, 41.5, 41.6; OLP: 129.1, 122.5, 123.3) and within-group LSC gains of comparable, small-to-medium magnitude ($g_{av} = 0.34, 0.55, 0.54$). These three groups therefore behaved as comparable implementations of a single L1-only language mode rather than as distinct treatments. However, each L1 was implemented in a single intact group of six to seven learners, so any comparison among them is severely underpowered and fully confounded with group composition and peer dynamics. The similarity observed is an absence of evidence for differential L1 effects, not evidence of equivalence; no equivalence test was possible at these sample sizes.

Differences within groups

Paired-samples t -tests evaluated within-group changes from pre- to post-test (Table 4). Fifteen within-group comparisons were possible (five groups $\times$ three outcomes) and are reported without alpha correction; they are therefore treated as exploratory and secondary to the between-group models, and individual p -values close to .05 should not be read as confirmatory.

For WTC, MG showed a significant increase, $M_{diff} = 3.42$, 95% CI [0.40, 6.44], $t(6) = 2.77$, $p = .032$, $g_{av} = 0.34$. TG showed a positive change that did not reach significance, $M_{diff} = 3.14$, 95% CI [−0.31, 6.59], $t(6) = 2.23$, $p = .067$, $g_{av} = 0.20$. WTC was essentially unchanged in KG ($M_{diff} = 0.0$) and changed only marginally in AG ($M_{diff} = 0.2$) and EG ($M_{diff} = 0.4$); none of these approached significance.

For OLP, scores increased by between 0.8 and 2.7 points in all five groups, and no within-group change reached significance ($p > .05$).

For LSC, significant increases were observed in all four groups whose PSDs involved the L1: TG ($p = .024$, $g_{av} = 0.34$), KG ($p = .022$, $g_{av} = 0.55$), AG ($p = .018$, $g_{av} = 0.54$), and MG ($p = .039$, $g_{av} = 0.43$). The English-only group showed no significant change.

Table 4

Within-Group Pre- to Post-Test Changes (Significant and Near-Significant Comparisons)

Group	Outcome	<i>n</i>	Pre <i>M</i> (<i>SD</i>)	Post <i>M</i> (<i>SD</i>)	<i>Mdiff</i> [95% CI]	<i>t(df)</i>	<i>p</i>	<i>d_{av}</i>	<i>g_{av}</i>	<i>d_z</i>
TG	WTC	7	39.86 (13.72)	43.00 (13.74)	3.14 [−0.31, 6.59]	2.23 (6)	.067	0.23	0.20	0.84
MG	WTC	7	42.29 (8.68)	45.71 (8.72)	3.42 [0.40, 6.44]	2.77 (6)	.032	0.39	0.34	1.05
TG	LSC	7	52.57 (15.03)	58.43 (15.33)	5.86 [1.08, 10.64]	3.00 (6)	.024	0.39	0.34	1.13
KG	LSC	6	58.17 (13.83)	66.50 (11.70)	8.33 [1.82, 14.84]	3.29 (5)	.022	0.65	0.55	1.34
AG	LSC	7	60.00 (11.94)	67.14 (11.11)	7.14 [1.76, 12.52]	3.25 (6)	.018	0.62	0.54	1.23
MG	LSC	7	60.29 (11.78)	66.43 (13.35)	6.14 [0.43, 11.85]	2.63 (6)	.039	0.49	0.43	0.99

Note. d_{av} = mean change divided by the average of the pre- and post-test SDs; $g_{av} = d_{av}$ with Hedges's small-sample correction, $J = 1 - 3/(4df - 1)$; d_z = mean change divided by the SD of the difference scores (Lakens, 2013). Because $n < 20$ in every cell, d_{av} is upwardly biased and g_{av} should be preferred; d_z is reported for transparency but is not comparable to between-subjects benchmarks. The remaining nine within-group comparisons (KG, AG, and EG for WTC; EG for LSC; all five groups for OLP) were non-significant ($p > .05$).

In summary, PSD modes involving the learners' first language—whether L1-only or mixed L1–L2—were associated with significant within-group gains in linguistic self-confidence of small-to-medium magnitude, and the mixed-language mode was additionally associated with a significant within-group gain in willingness to communicate. The English-only mode showed no significant within-group change on either affective outcome. Because the within-group tests are uncorrected and the groups were not randomly formed, this pattern is best read as a consistent descriptive trend across four independent groups rather than as a set of confirmatory results. Oral language performance increased by 0.8 to 2.7 points in every group, with no comparison reaching significance. Two features of the OLP measure limit what can be concluded: participants were selected into the study based on B2 performance on the same instrument, producing substantial restriction of range and a likely regression-to-the-mean component, and post-test means of 122.5–129.1 sit close to the scale ceiling of 140. Given also the sensitivity of the design (detectable $f \approx 0.66$ at $\alpha = .017$ with 80% power; $f \approx 0.60$ at $\alpha = .05$) relative to the observed effect (partial $\eta^2 = .20$, $f \approx 0.50$), the OLP result is treated as inconclusive rather than as evidence of no effect.

RQ2: Learners' perceptions of different PSD language modes

The qualitative findings from the semi-structured interviews provided a deeper understanding of how the language used in PSDs may impact the learners' WTC, OLP, and LSC by influencing their feelings, including comfort, anxiety, and confidence. A frequently highlighted theme was the importance of emotional comfort when preparing for oral tasks. The L1 groups (TG, AG, and KG) described the role of their native languages during PSDs in creating a safe space for brainstorming and organizing ideas. For example, one participant from the Arabic L1 group explained that, while the discussion did not necessarily affect his overall willingness to speak, "it helped me organize my thoughts in a way that it would be more understandable to the audience."

According to them, L1 lowered anxiety because it helped them manage the task time better, offered a 'roadmap', and let them focus more on expressing their thoughts rather than grappling with English grammar and vocabulary. One KG member acknowledged the structural differences between L1 and L2 and the difficulty of finding equivalents in both languages. However, he described feeling "less pressure" when using Kurdish, as it allowed for "clearer and more coherent thinking and higher levels of understanding" before and during their endeavors to communicate ideas in English.

Conversely, participants in EG reported higher anxiety levels, as the need to exclusively use English limited their ability to articulate complex ideas and added to the pressure of the oral tasks. One participant in this group explained that "unlike our mother tongue, it may be more challenging to explain complex ideas in English," which led to nervousness. In contrast, MG participants described how alternating between L1 and L2 during discussions enhanced their motivation to participate by creating a more comfortable environment that fostered self-confidence and readiness for subsequent tasks. Linguistic scaffolding also emerged as a central theme influencing the participants' confidence. Learners in the L1-only groups noted that using their native language provided a strong foundation for building vocabulary and clarifying their thoughts before switching to English. A TG participant likened this process to "building a roadmap for speaking," implying the critical role of L1 in facilitating cognitive organization before speaking in L2.

In contrast, MG participants emphasized how this mixed approach strengthened L1 support for brainstorming and L2 practice for refining ideas. One MG member shared, "In our group, L1 was used to discuss ideas and then L2 for practicing in English. This significantly increased our readiness to perform well in the speaking tasks."

Finally, participants reflected on how each approach impacted their perceived skill development. Despite appreciating the emotional and cognitive benefits of L1s, some TG, AG, and KG learners acknowledged that relying heavily on L1 limited their chance to practice English. An AG participant remarked that "thinking in L1 is a prerequisite for speaking in L1, and the same for English (L2). Although Arabic helped me feel confident during the discussions, I wish I had more chances to practice English before the actual task." The English-

only group interviewees reported that the English-only scenarios provided more practice opportunities and ultimately contributed to long-term improvement in their command of the language, despite higher levels of anxiety. Another participant in this group noted that "it was hard at first but useful in the long run." Meanwhile, MG participants consistently said the L1-L2 approach struck the right balance between preparation and practice, helping them feel more confident and improve their speaking through mutual collaboration. As another MG participant noted:

Using both L1 and L2 helped me not only feel confident but also learn many L2 words and idioms from my groupmates. I, for example, was so excited to see how my groupmates translated the sentences to L2 and reconstructed the sentences and learned a lot.

Across the interviews, learners in the L1-only and mixed-language groups more often described the preparation phase as relaxed and as supporting idea generation. In contrast, learners in the English-only group more often described it as effortful and anxiety-provoking. Participants in the mixed-language group were the most consistent in reporting both comfort and useful English practice. These accounts match the quantitative pattern, in which affective outcomes differentiated the groups while oral performance did not.

RQ3: Integration of quantitative and qualitative findings

Integrating the two strands, the affective differences observed quantitatively were mirrored in participants' accounts. The mixed-language group combined the significant WTC gain with descriptions of reduced pressure and easier brainstorming, and the L1-only groups combined significant LSC gains with descriptions of the first language as a planning resource. English-only participants, who showed no significant affective change, described exclusive L2 use as adding pressure during preparation while also providing valuable practice. Participants' reports also reflected the lack of measurable change in oral performance, focusing on how they felt about speaking rather than how well they spoke.

Discussion

The language mode used in PSDs was associated with differences in the two affective outcomes but not in oral performance. Between groups, WTC and LSC both showed significant effects, and post-test LSC was higher in every group whose PSDs involved the first language than in the English-only group. Within groups, LSC increased significantly in all four L1-involving groups, and WTC increased significantly in the mixed-language group. No group showed a significant change in oral performance. The pattern therefore separates L1-involving preparation from exclusively L2 preparation, rather than establishing the mixed mode as superior to L1-only preparation. Outcomes did not differ appreciably across the three L1-only implementations. Because each language was represented by a single group of six or seven learners, however, this similarity indicates only that no differential effect was detectable, not that the three languages functioned equivalently.

The affective differences associated with PSD modes involving the first language—and, for willingness to communicate, the mixed mode in particular—align with an expanding body of research highlighting the pedagogical value of translanguaging during pre-task planning. However, unlike previous studies that primarily focused on linguistic outcomes such as fluency or complexity (Aubrey & Philpott, 2025; Leeming et al., 2022; Scott & de la Fuente, 2008), the present study found that differences associated with PSD language mode were confined to affective outcomes. This suggests that translanguaging may function not merely as a linguistic resource but also as an affective scaffold that prepares learners psychologically for oral communication (Lin, 2019). The LSC gains observed in the mixed-language and L1-only groups are consistent with previous research highlighting the cognitive and affective benefits of integrating L1 and L2 in language learning activities (Jouibar et al., 2022; Scott & de la Fuente, 2008). This mixed-mode strategy can be understood as a form of pedagogical translanguaging, which has been shown to positively affect learners' performance during pre-task planning positively (Luo & Sun, 2025). Ultimately, these findings align with Clément's (1980) linguistic self-confidence framework, which emphasizes how self-confidence reduces anxiety and encourages participation in communicative tasks. They suggest that confidence gains may emerge even in relatively short interventions when learners have supportive language-learning conditions. The qualitative data further supported this interpretation, as participants in the mixed-language group consistently reported reduced anxiety, greater confidence, and enhanced motivation resulting from the complementary use of L1 and L2 during collaborative planning. These findings align with Vygotsky's (1978) perspective that learning occurs through social scaffolding and provide a basis for a broader theoretical interpretation.

Beyond supporting previous empirical findings, the present study also contributes theoretically by integrating sociocultural and affective perspectives on language learning. While Vygotsky's (1978) sociocultural theory explains how strategic use of L1 facilitates cognitive mediation and scaffolding during collaborative task preparation, Clément's (1980) linguistic self-confidence framework and MacIntyre et al.'s (1998) heuristic model of willingness to communicate explain how such cognitive support is translated into greater confidence and readiness to communicate. Taken together, the findings suggest that language choice during pre-speaking discussions functions as both a cognitive scaffold, enabling learners to organize ideas and negotiate meaning, and an affective scaffold, reducing anxiety and strengthening communicative confidence. This integrated interpretation extends previous theoretical accounts, which have typically emphasized either the cognitive benefits of translanguaging or its affective consequences in isolation, by demonstrating how these two processes interact to facilitate learners' engagement in oral communication (Fu et al., 2025; Pinho Feller, 2022).

In contrast, the absence of significant LSC change in EG may relate to anxiety reported by its members. This finding aligns with earlier studies documenting the challenges of English-

only policies in EFL contexts (Debreli, 2016; Turnbull, 2018). The qualitative interviews revealed that EG participants felt more anxious and less competent when required to use only English, supporting MacIntyre et al.'s (1998) assertion that perceived competence and anxiety are key determinants of LSC.

The significant within-group WTC increase in MG, together with a non-significant increase of similar size in TG, suggests that using L1 or a mixed-language approach can enhance learners' readiness to communicate. This echoes previous research demonstrating that L1 use in pre-task activities reduces anxiety and provides a cognitive roadmap for L2 communication (Hall & Cook, 2012; Turnbull & Dailey-O'Cain, 2009). This aligns with recent WTC models, which show that factors like lower anxiety and higher enjoyment are key contributors to a learner's motivation to speak (Sadoughi & Hejazi, 2024). This interpretation is supported by this study's own qualitative data, where participants from the L1 and MG explicitly reported greater confidence and motivation to participate in oral tasks. These outcomes resonate with findings that translanguaging in group preparation can deepen idea elaboration without compromising L2 production quality (Li & Qu, 2024; Luo & Sun, 2025) and offers a plausible explanation for the affective gains observed.

Interestingly, the English-only group (EG) revealed no significant increase in WTC. This contrasts with established theories suggesting that L2 immersion benefits long-term language acquisition (Krashen, 1982; Swain, 2000). This discrepancy may reflect the distinction between long-term language immersion and short-term task preparation. English-only instruction aims to maximize cumulative language exposure across extended periods, whereas PSDs serve primarily as a planning stage immediately preceding communication. Consequently, the psychological demands imposed by exclusive L2 use during planning may outweigh its potential linguistic benefits, particularly for learners with limited confidence. This reinforces the argument that WTC is deeply influenced by situational factors, including a classroom's emotional environment (Duong, 2026; Li et al., 2025), and aligns with Peng and Woodrow's (2010) argument that WTC is influenced by situational and enduring factors, including learners' emotional states and perceptions of their communicative competence.

The lack of significant changes in OLP across all groups is a key finding. This result supports the view that short interventions, such as PSDs, may not be sufficient to produce measurable improvements in a complex, multifaceted skill like oral proficiency. Another possible explanation concerns how oral proficiency is measured. OLP is a multidimensional construct incorporating pronunciation, fluency, grammatical accuracy, lexical sophistication, and discourse management. Consequently, improvements in planning quality, confidence, or communicative readiness may not immediately translate into statistically detectable gains on comprehensive speaking assessments. Mixed-language PSDs may first enhance learners' affective readiness and cognitive preparation, with measurable improvements in oral proficiency emerging only after sustained opportunities for language production and practice.

This interpretation also aligns with recent large-scale research, including a meta-analysis

that found task planning for oral production has complex and variable effects (Johnson & Abdi Tabari, 2022), and other work showing that significant gains in oral production depend on the development of *automatized* knowledge. This process typically requires more time than was available in the present study (Khezrlou, 2023). Therefore, while the PSD intervention clearly affected affective factors, its 10-session duration was likely too short to translate those affective gains into statistically significant improvements in OLP.

The qualitative findings shed light on the mechanisms underlying the quantitative results by highlighting two key themes: emotional comfort and linguistic scaffolding. Participants in the L1 and mixed-language groups noted that emotional comfort and reduced anxiety boosted their confidence, aligning with Clément et al.'s (2003) assertion that self-confidence is a key predictor of WTC. Conversely, learners in the English-only group expressed greater anxiety levels and difficulty articulating complex ideas, which likely hindered their WTC. Linguistic scaffolding also emerged as a key theme, with the 'roadmap' metaphor used by participants offering a practical illustration of Vygotsky's (1978) cognitive development. This interpretation is consistent with process-focused research showing that planning language influences the type and richness of preparation, with mixed-language approaches fostering both conceptual scaffolding and L2 rehearsal (Aubrey & Philpott, 2025; Leeming et al., 2022). From a mixed-methods perspective, these qualitative findings not only corroborated the quantitative results but also explained the mechanisms through which language choice influenced learners' willingness to communicate and linguistic self-confidence. By revealing how emotional comfort and linguistic scaffolding mediated learners' experiences during pre-speaking discussions, the qualitative phase added explanatory depth to the statistical findings and demonstrated the value of integrating quantitative and qualitative evidence within a sequential explanatory mixed-methods design.

Conclusion

This study explored the effects of PSDs conducted in L1, L2, and mixed-language on Iranian EFL learners' WTC, LSC, and OLP. PSD modes involving the learners' first language, including the mixed L1–L2 mode, were associated with gains in linguistic self-confidence, and the mixed mode was additionally associated with a gain in willingness to communicate. Although no significant improvements in OLP were observed, the findings highlight the significance of accounting for affective factors in language learning and provide practical recommendations for optimizing pre-speaking activities in multilingual classrooms. These findings contribute to the growing body of research on translanguaging and pre-task planning by highlighting language choice as a pedagogically meaningful resource for promoting learners' readiness to communicate in multilingual EFL classrooms.

The findings suggest that teachers in multilingual classrooms may use L1-inclusive or mixed-language pre-speaking discussions to support learners' confidence and willingness to participate, and that rigid English-only planning may not address learners' immediate affective

needs. Teacher-education programs and speaking materials could correspondingly include structured bilingual planning as a legitimate preparation strategy alongside target-language practice.

The findings should be interpreted in light of several limitations. First, the limited number of participants and the brief 10-session intervention may have constrained the generalizability of the findings. Therefore, future studies should replicate these effects with larger, more diverse samples over longer periods. Second, because the study included only male participants, the generalizability of the findings may be somewhat limited. Future studies should investigate whether gender influences the effects of PSD language modes. Third, the study focused exclusively on collaborative group discussions, which may not fully capture the complexity of authentic classroom communication. Future research could examine the effects of PSD language modes across a wider range of communicative contexts, including pair work, individual planning, online interactions, and tasks requiring different levels of cognitive and linguistic demand. Fourth, each language-mode condition was implemented in a single intact discussion group, so group membership is fully confounded with condition and, in the case of the L1-only groups, with first language. Any differences involving Turkish, Kurdish, or Arabic groups may therefore reflect group composition, peer dynamics, or pre-existing differences rather than language of preparation. Fifth, participants were assigned to conditions based on the design's linguistic requirements rather than at random; although pre-test scores were entered as covariates, covariate adjustment does not eliminate selection bias in non-randomized intact groups, and the findings should be interpreted as associational rather than causal. Finally, although the explanatory sequential mixed-methods design provided valuable insights into learners' experiences, future research could further examine the processes underlying these effects through longitudinal designs and classroom discourse analyses that capture learners' interactional behaviors during pre-speaking discussions.

Acknowledgments

We thank the editorial team of TESL Quarterly for the opportunity to submit and publish this synthesis. We also thank the anonymous reviewers for their careful, detailed reading of our manuscript and their insightful comments and suggestions. We also acknowledge all the participants who took part in this study.

Generative AI (Gemini and DeepSeek) was used solely for language editing to improve clarity, grammar, and organization. Napkin AI was used to create one schematic diagram illustrating the study procedure. All conceptualization, data analysis, interpretation, and final decisions were completed and verified by the authors, who take full responsibility for the content of the manuscript.

Declaration of conflicting interests

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for this article's research, authorship, and/or publication.

References

- Adams, R., Alwi, N. A. N. M., & Newton, J. (2015). Task complexity effects on the complexity and accuracy of writing via text chat. *Journal of Second Language Writing*, 29, 64–81. <https://doi.org/10.1016/j.jslw.2015.06.002>
- Aubrey, S., Lambert, C., & Leeming, P. (2020). The impact of first as opposed to second language pre-task planning on the content of problem-solving task performance. *Language Teaching Research*, 26(5), 867-892. <https://doi.org/10.1177/1362168820917844>
- Aubrey, S., & Philpott, A. (2025). Use of the L1 and L2 in strategic planning and rehearsal for task performances in an online classroom. *Language Teaching Research*, 29(3), 1139-1164. <https://doi.org/10.1177/13621688221077421>
- Cheng, Z., & Sun, P. (2026). Effects of translanguaging and task rehearsal on L2 speaking performance. *Applied Linguistics Review*, 17(4), 1503-1533. <https://doi.org/10.1515/applirev-2025-0057>
- Clément, R. (1980). Ethnicity, contact and communicative competence in a second language. In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychological perspectives* (pp. 147–154). Pergamon.
- Clément, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of Language and Social Psychology*, 22(2), 190–209. <https://doi.org/10.1177/0261927X03022002003>
- Cook, V. (2001). Using the first language in the classroom. *Canadian Modern Language Review*, 57(3), 402–423. <https://doi.org/10.3138/cmlr.57.3.402>
- Creswell, J. W., & Plano Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage Publications.
- Debreli, E. (2016). Perceptions of non-native EFL teachers' on L1 use in L2 classrooms: Implications for language program development. *English Language Teaching*, 9(3), 24–32. <https://doi.org/10.5539/elt.v9n3p24>
- de la Fuente, M. J., & Goldenberg, C. (2022). Understanding the role of the first language (L1) in instructed second language acquisition (ISLA): Effects of using a principled approach to L1 in the beginner foreign language classroom. *Language Teaching Research*, 26(5), 943-962. <https://doi.org/10.1177/1362168820921882>
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins Publishing Company. <https://doi.org/10.1075/llt.42>
- DiCamilla, F. J., & Antón, M. (2012). Functions of L1 in the collaborative interaction of beginning and advanced second language learners. *International Journal of Applied Linguistics*, 22(2), 160–188. <https://doi.org/10.1111/j.1473-4192.2011.00302.x>
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum Associates.
- Duong, T. N. T. (2026). Exploring students' willingness to communicate in a translanguaging classroom: A qualitative study in Vietnam. *International Journal of Linguistics and Translation Studies*, 7(2), 168–180. <https://doi.org/10.36892/ijlts.v7i2.737>

- Ebadi, S., & Azizimajd, H. (2026). Investigating the effects of speaking tasks in Clubhouse on EFL learners' speaking and foreign language anxiety. *Computer Assisted Language Learning*. Advance online publication. <https://doi.org/10.1080/09588221.2024.2371381>
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474–509. <https://doi.org/10.1093/applin/amp010>
- Ellis, R. (2021). The effects of pre-task planning on second language writing: A systematic review of experimental studies. *Chinese Journal of Applied Linguistics*. <https://doi.org/10.1515/CJAL-2021-0009>
- Ford, K. (2009). Principles and practices of L1/L2 use in the Japanese university EFL classroom. *JALT Journal*, 31(1), 63–80.
- Freire, R., & Jones, T. (2015). *Q: Skills for Success: Listening and Speaking 4* (2nd ed.). Oxford University Press.
- Fu, D., Hadjioannou, X., & Zhou, X. (2025). *Translanguaging for emergent bilinguals: Inclusive teaching in the linguistically diverse classroom*. Teachers College Press.
- Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, 20(3), 321–340. <https://doi.org/10.1177/1362168815570142>
- Galdeano Chasco, M., & García Mayo, M. P. (2024). L1 use in pre-task planning by child EFL learners. *Instructed Second Language Acquisition*, 8(1). <https://doi.org/10.1558/isla.27488>
- Gao, J., & Min, S. (2021). A comparative study of the effects of L1 and L2 prewriting discussions on L2 writing performance. *System*, 103, 102654. <https://doi.org/10.1016/j.system.2021.102654>
- García Mayo, M. P., & Hidalgo, M. Á. (2017). L1 use among young EFL mainstream and CLIL learners in task-supported interaction. *System*, 67, 132–145. <https://doi.org/10.1016/j.system.2017.05.004>
- Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language Teaching*, 45(3), 271–308. <https://doi.org/10.1017/S0261444812000067>
- Hanifa, R. (2018). Factors generating anxiety when learning EFL speaking skills. *Studies in English Language and Education*, 5(2), 230–239. <https://doi.org/10.24815/siele.v5i2.10927>
- Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching*, 43(2), 154–167. <https://doi.org/10.1017/S026144480999036X>
- Izquierdo, J., Martínez, V. G., Pulido, M. G. G., & Zúñiga, S. P. A. (2016). First and target language use in public language education for young learners: Longitudinal evidence from Mexican secondary-school classrooms. *System*, 61, 20–30. <https://doi.org/10.1016/j.system.2016.07.006>
- Johnson, M. D., & Abdi Tabari, M. (2022). Task planning and oral L2 production: A research synthesis and meta-analysis. *Applied Linguistics*, 43(6), 1143–1164. <https://doi.org/10.1093/applin/amac026>
- Jouibar, B., Hadian, B., & Vaez Dalili, M. (2022). The effect of code-switching on EFL learners' speaking accuracy, fluency, and complexity. *Journal of Language Horizons*, 6(2), 65–80. <https://doi.org/10.22051/lghor.2021.34089.1407>
- Khezrlou, S. (2023). The role of task-supported language teaching in developing automatized knowledge and learners' task engagement. *Language Awareness*, 32(3), 247–266. <https://doi.org/10.1080/09658416.2023.2177659>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Leeming, P., Aubrey, S., & Lambert, C. (2022). Collaborative pre-task planning processes and second language task performance. *REL C Journal*, 53(3), 534–550. <https://doi.org/10.1177/0033688220953225>
- Levelt, W. J. (1993). *Speaking: From intention to articulation*. MIT Press.
- Li, C., Dewaele, J.-M., Pawlak, M., & Kruk, M. (2025). Classroom environment and willingness to communicate in English: The mediating role of emotions experienced by university students in China. *Language Teaching Research*, 29(5), 2140–2160. <https://doi.org/10.1177/13621688221111623>

- Li, M., & Qu, Z. (2024). Encouraging translanguaging in collaborative talk in EFL classrooms: An epistemic network comparative study. *Linguistics and Education*, 84, 101360. <https://doi.org/10.1016/j.linged.2024.101360>
- Lin, A. M. (2019). Theories of trans/languageing and trans-semiotizing: Implications for content-based education classrooms. *International Journal of Bilingual Education and Bilingualism*, 22(1), 5-16. <https://doi.org/10.1080/13670050.2018.1515175>
- Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal*, 92(1), 71–86. <https://doi.org/10.1111/j.1540-4781.2008.00687.x>
- Luo, X., & Sun, P. P. (2025). Greater complexity and higher efficiency: Effects of translanguaging during collaborative pre-task planning on EFL learners' speaking performance. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/13621688241307029>
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in an L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545–562. <https://doi.org/10.2307/330224>
- Neumann, H., & McDonough, K. (2015). Exploring student interaction during collaborative prewriting discussions and its relationship to L2 writing. *Journal of Second Language Writing*, 27, 84–104. <https://doi.org/10.1016/j.jslw.2015.01.001>
- Oxford University Press. (2019). Oxford Test of English: Test specification (Ver. 2.0). Oxford University Press. <https://www.oxfordtestofenglish.com/>
- Peng, J.-E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning*, 60(4), 834–876. <https://doi.org/10.1111/j.1467-9922.2010.00576.x>
- Pinho Feller, N. (2022). Translanguaging and scaffolding as pedagogical strategies in a primary bilingual classroom. *Classroom Discourse*, 13(3), 312-335. <https://doi.org/10.1080/19463014.2021.1954960>
- Pyun, D. O., Kim, J. S., Cho, H. Y., & Lee, J. H. (2014). Impact of affective variables on Korean as a foreign language learners' oral achievement. *System*, 47, 53–63. <https://doi.org/10.1016/j.system.2014.09.017>
- Richards, J. C. (2008). *Teaching listening and speaking: From theory to practice*. Cambridge University Press.
- Sadoughi, M., & Hejazi, S. Y. (2024). How can L2 motivational self system enhance willingness to communicate? The contribution of foreign language enjoyment and anxiety. *Current Psychology*, 43(3), 2173–2185. <https://doi.org/10.1007/s12144-023-04479-3>
- Scott, V. M., & de la Fuente, M. J. (2008). What's the problem? L2 learners' use of the L1 during consciousness-raising, form-focused tasks. *The Modern Language Journal*, 92(1), 100–113. <https://doi.org/10.1111/j.1540-4781.2008.00689.x>
- Shabaka-Fernández, S. (2021). The effect of teacher language use in Spanish EFL classrooms. *The Language Learning Journal*, 51(2), 161–177. <https://doi.org/10.1080/09571736.2021.1970794>
- Shin, J. Y., Dixon, L. Q., & Choi, Y. (2020). An updated review on the 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>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). Routledge.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Thompson, G. L., & Harrison, K. (2014). Language use in the foreign language classroom. *Foreign Language Annals*, 47(2), 321–337. <https://doi.org/10.1111/flan.12079>

- Turnbull, B. (2018). Reframing foreign language learning as bilingual education: Epistemological changes towards the emergent bilingual. *International Journal of Bilingual Education and Bilingualism*, 21(8), 1041–1048. <https://doi.org/10.1080/13670050.2016.1238866>
- Turnbull, M., & Dailey-O'Cain, J. (2009). *First language use in second and foreign language learning*. Multilingual Matters. <https://doi.org/10.21832/9781847691972>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing*, 11(3), 225–246. [https://doi.org/10.1016/S1060-3743\(02\)00084-X](https://doi.org/10.1016/S1060-3743(02)00084-X)
- Zarrinabadi, N., & Tanbakooei, N. (2016). Willingness to communicate: Rise, development, and some future directions. *Language and Linguistics Compass*, 10(1), 30-45.
- Zohrabi, M., & Bimesl, L. (2022). Exploring EFL teachers' perceptions of strategies for promoting learners' willingness-to-communicate in online classes. *Applied Research on English Language*, 11(1), 89–110. <https://doi.org/10.22108/ARE.2024.140274.2210>
- Zohrabi, M., & Jafari, H. (2020). The role of think-pair-share interactional activity on improving Iranian EFL learners' willingness to communicate. *Teaching English Language*, 14(1), 153–182. <https://doi.org/10.22132/TEL.2020.106921>