

Iranian EFL Learners' Perceptions of L1 Pre-task Planning on Speaking Accuracy, Fluency, and Complexity

Hassan Galbat¹, Farhad Fahandezh Saadi^{2*}, Shahram Afraz³

¹Ph.D. Candidate, Department of English, Qeshm Branch, Islamic Azad University, Qeshm, Iran
hmgalbat@yahoo.com

²Assistant Professor, Department of English, Bandar Lengeh Branch, Islamic Azad University, Bandar Lengeh, Iran

Farhadfahandezh@gmail.com

³Assistant Professor, Department of English, Qeshm Branch, Islamic Azad University, Qeshm, Iran
a.sh32@rocketmail.com

2025/02/19

2025/03/28

Abstract

This study examined Iranian EFL learners' perceptions of L1 (Persian) pre-task planning on speaking performance, specifically accuracy, fluency, and complexity, within a Task-Based Language Teaching (TBLT) framework. A convergent parallel mixed-methods design was adopted, collecting quantitative and qualitative data concurrently to provide a holistic view. Quantitative data were gathered via a 15-item Likert-scale questionnaire completed by 100 intermediate learners, selected through convenience sampling from three Hormozgan language institutes, following two counterbalanced oral-opinion tasks in a 60-minute session. Qualitative data were collected simultaneously through semi-structured interviews with 15 participants, purposively chosen from the questionnaire respondents based on diverse perception scores to enrich insights, conducted within 1–2 days. Quantitative analysis, employing descriptive statistics and paired t-tests, showed a strong L1 planning preference, boosting accuracy and fluency, while thematic analysis of interviews highlighted enhanced confidence with L1 and L2 transition difficulties, with complexity perceptions mixed. Integrating these findings, L1 planning supports immediate oral proficiency and reduces anxiety in Iran's low-exposure EFL context, though L2 proficiency limits complexity. These findings align with Cognitive Load Theory and Sociocultural Theory, offering context-specific pedagogical insights for incorporating L1 planning in TBLT practices while underscoring the need for strategies to bolster L2 complexity.

Keywords: L1 Pre-Task planning, Perception, EFL speaking, accuracy, fluency, complexity, Iranian learners

Introduction

Speaking is a critical yet challenging skill for English as a Foreign Language (EFL) learners, particularly in Iran, where limited exposure to English outside the classroom hinders oral proficiency (Bygate, 2001). Iranian learners often struggle with accuracy, fluency, and complexity—essential components of speaking performance—due to insufficient communicative practice (Ellis & Barkhuizen, 2005). Task-Based Language Teaching (TBLT) has emerged as an effective approach to address these challenges by utilizing meaningful tasks to foster L2 use, with pre-task planning recognized as a key strategy to enhance output (Ellis, 2003; Willis, 1996). However, it remains unclear whether planning in the first language (L1, Persian) or the second

language (L2, English) best supports speaking performance, particularly in EFL contexts where L1 use is often discouraged (Samadi, 2011).

Despite TBLT's emphasis on L2 immersion, the potential role of L1 in planning warrants investigation, as Iranian learners may naturally rely on Persian for cognitive ease amidst limited L2 exposure (Al-Nofaie, 2018). While some EFL pedagogies advocate L2-only instruction, others suggest L1 could aid task preparation, yet empirical studies on learners' perceptions of this approach are scarce, particularly in Iran (Cook, 2001). This gap is significant, as understanding learners' views could inform teaching practices and optimize speaking instruction in similar low-exposure settings (Dörnyei, 2005). This study examines how Iranian EFL learners perceive L1 planning effects on their speaking performance, offering insights into a debated yet under-researched area.

The following research questions guide the study:

- **RQ1:** How do Iranian EFL learners perceive L1 (Persian) pre-task planning on the accuracy of their English-speaking performance?
- **RQ2:** How do Iranian EFL learners perceive L1 (Persian) pre-task planning on the fluency of their English-speaking performance?
- **RQ3:** How do Iranian EFL learners perceive L1 (Persian) pre-task planning on the complexity of their English-speaking performance?
- **RQ4:** What overall preferences and challenges do Iranian EFL learners report when using Persian (L1) for planning English-speaking tasks?

Through these questions, this study explores how intermediate Iranian EFL learners in a low-exposure context perceive the impact of planning in their first language (L1, Persian) on speaking performance—specifically accuracy, fluency, and complexity—finding a strong preference for L1 planning as a cognitive scaffold that, based on their perceptions, enhances accuracy and fluency via improved grammar organization and smoother delivery, though it reveals complexity constraints linked to L2 translation challenges; these perception-based insights suggest that, for such learners, Task-Based Language Teaching (TBLT) practices might strategically incorporate L1 planning to support immediate oral proficiency and confidence, with a gradual shift to L2 planning potentially fostering complexity and sustained L2 development, pending objective performance validation, thus providing context-specific guidance for speaking instruction while acknowledging that L2-only approaches may suit other settings.

The significance of this study lies in its potential to reshape EFL teaching practices by validating learners' perspectives on L1 use within TBLT. It offers a practical framework for educators in Iran and similar low-exposure contexts to enhance speaking skills. By demonstrating L1 planning's perceived role in boosting accuracy, fluency, and confidence, it challenges traditional L2-only paradigms (Krashen, 1985) and proposes a balanced approach that leverages native language strengths while addressing complexity limitations. Theoretically, it enriches the L1 debate in SLA by bridging Cognitive Load Theory (Sweller, 1988) and Sociocultural Theory (Vygotsky, 1978) with empirical learner data, thus contributing to a nuanced understanding of planning's role in oral proficiency. In Iran's EFL landscape, where cultural resistance to L1 use persists (Samadi, 2011), this study provides evidence-based justification for its strategic integration, offering immediate pedagogical

benefits and informing curriculum design to better support intermediate learners facing limited L2 input.

Literature Review

2.1 Theoretical Framework

The use of L1 in L2 planning is theoretically supported by multiple frameworks offering complementary perspectives on its potential impact on EFL speaking performance. Cognitive Load Theory (Sweller, 1988) posits that planning in a familiar language reduces extraneous cognitive load, freeing working memory for L2 production tasks. This is particularly relevant for EFL learners facing high cognitive demands in spontaneous speech, where L1 could optimize resource allocation (Sweller et al., 1998). Levelt's (1995) speech production model further suggests that conceptualization in L1 streamlines formulation and articulation in L2, potentially enhancing output efficiency by allowing learners to pre-organize ideas in a native framework before encoding them in a less familiar language. Sociocultural Theory (Vygotsky, 1978) frames L1 as a mediational tool, scaffolding L2 performance during complex tasks by providing a familiar cognitive bridge (Anton & DiCamilla, 1998). Together, these perspectives suggest L1 planning could mitigate the linguistic and psychological barriers inherent in EFL speaking, forming a robust foundation for this study's exploration of its perceived impact.

2.2 Speaking Performance in TBLT

Speaking performance in L2 contexts is often evaluated using the Complexity, Accuracy, and Fluency (CAF) framework (Housen & Kuiken, 2009), which captures distinct yet interrelated dimensions. Accuracy reflects grammatical correctness, fluency indicates smoothness and speed, and complexity includes lexical and syntactic richness (Ellis & Barkhuizen, 2005). For EFL learners, balancing these dimensions is challenging due to the demands of real-time production, often worsened by limited L2 exposure (Skehan, 1998). Task-Based Language Teaching (TBLT) addresses this by emphasizing meaningful tasks as vehicles for language use, with pre-task planning recognized as a crucial strategy to reduce cognitive pressure and enhance CAF outcomes (Ellis, 2003). Research by Foster and Skehan (1996) demonstrates that planning boosts fluency and accuracy by allowing for rehearsal and organization, though complexity gains rely on task type and learner proficiency (Yuan & Ellis, 2003). For instance, narrative tasks may yield higher complexity than opinion tasks due to their structural demands (Skehan & Foster, 1999), highlighting planning's essential yet variable role in TBLT.

2.3 The Role of L1 in L2 Learning

The role of L1 in L2 learning sparks ongoing debate within Second Language Acquisition (SLA). Proponents argue that L1 acts as a cognitive scaffold, facilitating idea organization and reducing anxiety during L2 tasks (Swain & Lapkin, 2000; Storch & Wigglesworth, 2003). Swain and Lapkin (2000) found that using L1 in collaborative tasks enhances L2 output through deeper conceptual processing, while Wigglesworth and Storch (2009) reported improved accuracy and content richness in writing tasks, suggesting transferable benefits to speaking. Conversely, critics caution that excessive reliance on L1 may hinder direct L2 processing, potentially stunting the

development of fluency and complexity (Macaro, 2005; Turnbull & Arnett, 2002). Krashen's (1985) input hypothesis supports this critique, advocating for L2 immersion in acquisition, though it overlooks L1's potential as a preparatory tool. In Iran, where Persian dominates daily life and L2 exposure is limited, this debate is particularly salient yet understudied, with cultural resistance to L1 in classrooms adding to the complexity (Samadi, 2011).

2.4 Learners' Perceptions of L1 Use

Learners' perceptions significantly influence instructional effectiveness, as they shape motivation and engagement (Dörnyei, 2005). Kim and McDonough (2008) found that learners value pre-task planning for reducing anxiety, with some preferring L1 for its familiarity and comfort, particularly in pair work (Carson & Kashihara, 2012). However, perceptions vary by context—Saudi learners exhibit mixed attitudes due to strong L2 immersion norms (Al-Nofaie, 2018), while Korean learners favor L1 for task clarity (Kim, 2015). In Iran, where L1 use is often stigmatized in formal education, learners' views on L1 planning remain largely unexplored, despite their potential to inform TBLT practices and address affective barriers like speaking anxiety.

2.5 Empirical Studies on the Approach

The influence of pre-task planning and the use of the first language (L1) in Task-Based Language Teaching (TBLT) has garnered increasing attention within second language acquisition (SLA) research, particularly with respect to enhancing speaking performance among EFL learners. A growing body of empirical studies have investigated how planning conditions and language choice shape oral production outcomes. To contextualize the present study, this section reviews five pivotal studies that explore these dynamics, detailing their methodologies and findings. Together, these investigations provide a foundation for understanding the interplay between planning strategies and speaking proficiency, while also highlighting areas warranting further exploration.

The first notable study, conducted by Ellis and Yuan (2004), examined the effects of planning on fluency, complexity, and accuracy in L2 narrative writing among 42 intermediate Chinese EFL learners at a university in China. In this quasi-experimental design, participants were randomly assigned to one of three conditions—no planning, pre-task planning (10 minutes), or online planning (unlimited time during the task)—and completed narrative writing tasks. Their output was subsequently analyzed using fluency measures (words per minute), complexity indicators (subordinate clauses per T-unit), and accuracy metrics (error-free clauses), with ANOVA employed to compare performance across groups. The results revealed that pre-task planning significantly enhanced fluency and complexity, though its impact on accuracy was limited. This finding suggests that planning facilitates idea organization, yet it does not guarantee grammatical precision. However, because this study focused on writing rather than speaking and did not differentiate between L1 and L2 planning, its relevance to oral production and language choice remains indirect.

Building on the exploration of planning effects, Galbat, Fahandezh Saadi, and Afraz (2025) investigated the specific role of L1 versus L2 planning in their study, "Effect of Using First Language by Iranian EFL Learners in Task Preparation on their Speaking Accuracy, Fluency, and Complexity." This research involved 40 intermediate Iranian EFL learners from a language institute in southern Iran, who were divided into two groups: one planning in Persian (L1) and the other in English (L2). Each group completed oral narrative tasks following a 5-minute planning

phase, with performance recorded and evaluated for accuracy (percentage of error-free clauses), fluency (syllables per minute), and complexity (syntactic complexity via subordination ratio). A pre-test/post-test design, analyzed with t-tests, assessed improvements across conditions. The findings indicated that L1 planning significantly outperformed L2 planning in enhancing accuracy and fluency, likely due to reduced cognitive load and improved idea structuring in the native language. In contrast, complexity showed no notable difference, pointing to challenges in translating L1 plans into L2 output. Although this study directly addresses L1 use in an Iranian context, its reliance on objective performance measures without exploring learners' perceptions limits its insight into subjective experiences.

Shifting the focus to planning type rather than language choice, Sangarun (2005) explored how focusing on meaning versus form during strategic planning affects oral production among 72 intermediate Thai EFL university students. Participants were assigned to one of three conditions—no planning, meaning-focused planning (emphasizing content organization), or form-focused planning (targeting grammar and vocabulary)—and performed picture-based narrative tasks after a 10-minute planning period. Their speech was recorded and scored for fluency (speech rate), accuracy (error-free T-units), and complexity (lexical diversity), with MANOVA used to compare outcomes across groups. The results demonstrated that meaning-focused planning boosted fluency and complexity, whereas form-focused planning improved accuracy. These differential effects underscore planning's versatility, yet the study's assumption of L2-only planning overlooks the potential influence of L1, particularly in EFL settings where native language reliance might prevail.

Similarly, Seifoori and Goudarzi (2012) examined the impact of planning within an oral production-oriented program on 50 intermediate Iranian EFL learners at a private language institute. In this pre-test/post-test experimental design, an experimental group received training in oral output with 5-minute pre-task planning for story retelling tasks, while a control group followed traditional instruction without a planning focus. Performance was assessed through audio recordings, measuring accuracy (error-free clauses) and fluency (words per minute), with paired t-tests analyzing within-group gains and independent t-tests comparing groups. The experimental group exhibited significant improvements in both accuracy and fluency, suggesting that structured planning enhances oral skills when paired with practice. However, the study does not specify whether planning occurred in L1 or L2, implying an L2 focus, and it omits learners' perceptions, thus limiting its contribution to understanding subjective preferences in planning strategies.

Complementing these findings, Lambert, Kormos, and Minn (2017) investigated task repetition and planning effects on speech processing among 48 pre-intermediate Japanese EFL learners at a university in Japan. Using a within-subjects design, participants repeated oral narrative tasks twice, with 5-minute pre-task planning intervals before each iteration. Their speech was evaluated for fluency (speech rate), accuracy (error-free T-units), and complexity (syntactic variety) via repeated-measures ANOVA, while post-task interviews provided supplementary insights into planning strategies. The results showed that repetition with planning improved fluency and accuracy over time, with complexity increasing marginally. Although some learners reported incidental L1 use during planning, this was not systematically explored, as the study prioritized repetition effects over language choice or perceptions. Consequently, its findings offer limited guidance on deliberate L1 versus L2 planning dynamics.

Taken together, these studies illustrate the significant role of pre-task planning in enhancing speaking performance, with variations in focus—ranging from planning conditions (Ellis & Yuan, 2004; Sangarun, 2005) to L1 use (Galbat et al. 2025) and instructional strategies

(Seifoori & Goudarzi, 2012; Lambert et al., 2017). Nevertheless, a critical gap persists in this body of research. While these investigations provide robust evidence on objective outcomes such as accuracy, fluency, and complexity, they largely overlook learners' perceptions of how planning in L1 influences their speaking experience. For example, Ellis and Yuan (2004) and Sangarun (2005) assume L2 planning without considering L1's potential role, while Seifoori and Goudarzi (2012) and Lambert et al. (2017) focus on performance or repetition effects, sidelining subjective insights. Even Galbat et al.'s study (2025), despite its relevance to Iranian EFL learners and L1 planning, restricts its scope to measurable output, neglecting how learners perceive these effects—a crucial dimension given the cognitive and affective factors (e.g., confidence, ease) that shape language learning, especially in low-exposure contexts like Iran (Dörnyei, 2005). This omission is particularly striking in TBLT frameworks, where learner preferences could inform the strategic integration of L1, a contentious issue in settings with limited L2 input (Samadi, 2011).

The present study addresses this gap by shifting the lens to learners' perceptions, exploring how intermediate Iranian EFL learners in a low-exposure context perceive the impact of L1 (Persian) planning on their speaking performance—specifically accuracy, fluency, and complexity—within a TBLT approach. Unlike prior research, which predominantly relies on quantitative performance metrics, this investigation employs a mixed-methods design, integrating semi-structured interviews and self-reported questionnaire data with oral tasks. This approach captures nuanced subjective insights, such as perceived confidence or transition difficulties, complementing objective findings like those of Galbat et al. By doing so, it not only bridges the empirical-perceptual divide but also offers context-specific guidance for Iranian EFL pedagogy. In a setting where Persian dominates and L2 exposure is scarce, understanding learners' views provides a foundation for tailoring TBLT practices, potentially validating L1 as a strategic scaffold and enhancing both immediate oral proficiency and long-term instructional design.

In synthesizing the theoretical and empirical foundations reviewed, it becomes evident that pre-task planning within TBLT holds substantial potential to enhance EFL speaking performance, with L1 use emerging as a pivotal yet underexplored factor. The interplay of cognitive, sociocultural, and learner-centered perspectives underscores the complexity of planning's role, particularly in low-exposure contexts where traditional L2-only approaches may fall short. Building on this foundation, the current study seeks to illuminate Iranian EFL learners' subjective experiences with L1 and L2 planning, offering insights that could refine TBLT practices and contribute to the broader SLA discourse. The following methodology outlines the approach taken to investigate these perceptions, providing a systematic framework to address the identified research needs.

Methodology

3.1 Research Design

This study utilized a convergent parallel mixed-methods design to examine Iranian EFL learners' perceptions of pre-task planning in L1 (Persian) on speaking accuracy, fluency, and complexity within a Task-Based Language Teaching (TBLT) framework (Creswell & Plano Clark, 2018). Quantitative data from a 15-item Likert-scale questionnaire, completed by 100 intermediate learners immediately after two counterbalanced oral-opinion tasks in a 60-minute session, and qualitative data from semi-structured interviews with 15 participants conducted within 1–2 days were collected concurrently within a single study phase. The questionnaire data were analyzed

using descriptive and inferential statistics (e.g., paired t-tests), while interviews underwent thematic analysis (Braun & Clarke, 2006), with findings integrated to triangulate insights and address RQ1–RQ4 holistically, leveraging both strands’ strengths for a comprehensive understanding of perceptions in Iran’s low-exposure EFL context.

3.2 Setting and Participants

Conducted at three private language institutes in Hormozgan, Iran—with strong EFL programs but limited L2 exposure outside the classroom—the study reflects common Iranian EFL contexts. Participants included 100 intermediate (CEFR B1) learners aged 18–30, selected through convenience sampling from institute rosters. This level ensures task feasibility while highlighting the importance of planning for challenges in accuracy, fluency, and complexity (Skehan, 1998). Proficiency was verified using the Oxford Placement Test (OPT; Allan, 2004), with scores of 30–40/60 ensuring homogeneity, and the gender balance (50 males, 50 females) reflects typical classroom norms. However, convenience sampling, while practical due to access constraints in this setting, limits the sample's representativeness. Participants were drawn from readily available learners at specific institutes, potentially skewing results toward urban, motivated learners with institutional support, and may not fully represent the diversity of Iranian EFL learners in rural or less-resourced contexts (Shadish et al., 2002). This sampling method enhances feasibility but compromises generalizability, a trade-off acknowledged in interpreting findings for RQ1–RQ4. Fifteen interviewees were purposively selected for their diverse questionnaire responses, adding depth to insights into these questions despite the sampling constraint.

3.3 Data Collection Method

Two instruments were used to collect data aligned with RQ1–RQ4:

- **Questionnaire**

A previously validated 15-item Likert-scale questionnaire assessed perceptions of L1 planning across seven domains: general preparation, ease/confidence, accuracy (RQ1), fluency (RQ2), complexity (RQ3), challenges, and preferences (RQ4). Items used a 5-point scale (1 = Strongly Disagree, 5 = Strongly Agree), with comparative phrasing. Administered in English with Persian translations available, it took 15–20 minutes to complete. The questionnaire was originally validated in a prior study. For this context, it was re-validated using Exploratory Factor Analysis (EFA) during the pilot study to confirm construct validity, extracting factors aligned with accuracy, fluency, complexity, and preferences (Kaiser-Meyer-Olkin > 0.7, Bartlett’s test $p < 0.05$). Internal consistency was assessed via Cronbach’s Alpha, targeting > 0.7 per subscale, ensuring reliability for RQ1–RQ4. Credibility was enhanced through pilot testing, which refined items to accurately reflect participants’ perceptions, ensuring authentic responses consistent with their experiences (Lincoln & Guba, 1985).

- **Interviews**

Semi-structured interviews with 15 participants explored questionnaire responses in depth, targeting RQ1–RQ4. Conducted in Persian for comfort, recorded with consent, and lasting 10–15 minutes each, they provided qualitative depth. Interview questions were developed based on SLA literature (e.g., Swain & Lapkin, 2000) and reviewed by two EFL experts for content validity, ensuring relevance to Iranian learners and RQ1–RQ4. Pilot testing with five participants refined phrasing to match the context. Credibility was supported by using participants’ native language and expert-validated questions, fostering truthful responses aligned with their lived realities (Lincoln & Guba, 1985).

3.4 Pilot Study Overview

A pilot study was conducted with 10 intermediate EFL learners from a separate institute to test the instruments and procedure for applicability to RQ1–RQ4. Participants completed two oral opinion tasks—one with L1 (Persian) planning and one with L2 (English) planning—followed by the 15-item Likert-scale questionnaire and interviews with five participants. This process ensured task feasibility within a 60-minute timeframe and verified the counterbalancing approach, while also refining the tools for the main study.

Questionnaire Validation

The pilot questionnaire responses were analyzed to confirm their validity and reliability for measuring learners’ perceptions of L1 planning on speaking performance, as aligned with RQ1–RQ4. Construct validity was assessed to ensure the instrument accurately captured the theoretical constructs of accuracy, fluency, complexity, and preferences, rooted in the Complexity, Accuracy, and Fluency (CAF) framework (Housen & Kuiken, 2009) and learner perception research (Dörnyei, 2005). The questionnaire, originally adapted from a previously validated tool, comprised items designed to reflect these constructs—e.g., “Fewer grammar mistakes” for accuracy, “Smoother speech” for fluency, “More complex sentences” for complexity, and “Prefer this approach” for preferences. To verify this alignment, Exploratory Factor Analysis (EFA) with principal component analysis and varimax rotation was conducted. The Kaiser-Meyer-Olkin (KMO) measure yielded 0.72, indicating adequate sampling adequacy (Kaiser, 1974), and Bartlett’s Test of Sphericity was significant ($\chi^2 = 98.45$, $df = 105$, $p < 0.05$), confirming factorability. EFA extracted four distinct factors—accuracy, fluency, complexity, and preferences—as shown in Table 1, with factor loadings ranging from 0.74 to 0.85, eigenvalues above 1.0, and a cumulative variance explained of 68.3% (Field, 2013). These factors corresponded closely to the intended constructs, supporting the questionnaire’s construct validity by demonstrating that items grouped as theoretically expected.

Table 1. *Pilot Study EFA Results (N = 10)*

| Factor | No. of Items | Factor Loadings Range | Eigenvalue | Variance Explained (%) |
|-------------|--------------|-----------------------|------------|------------------------|
| Accuracy | 4 | 0.78–0.81 | 3.62 | 24.1 |
| Fluency | 4 | 0.79–0.85 | 2.89 | 19.3 |
| Complexity | 3 | 0.74–0.76 | 2.15 | 14.3 |
| Preferences | 4 | 0.77–0.82 | 1.58 | 10.6 |

(Note: Cumulative variance = 68.3%; loadings < 0.6 omitted.)

Internal consistency was further assessed using Cronbach's Alpha (Table 2), with coefficients ranging from 0.75 (Complexity) to 0.83 (Accuracy), exceeding the 0.7 threshold for reliability (Nunnally, 1978). Minor rephrasing (e.g., clarifying "smoother" to "fewer pauses") was applied based on pilot feedback to enhance item clarity, reinforcing the instrument's alignment with the constructs for the study.

Table 2. *Pilot Study Cronbach's Alpha Results (N = 10)*

| Subscale | Cronbach's Alpha | Interpretation |
|-------------|------------------|------------------------|
| Accuracy | 0.83 | High reliability |
| Fluency | 0.80 | Good reliability |
| Complexity | 0.75 | Acceptable reliability |
| Preferences | 0.81 | Good reliability |

Interview Refinement

Semi-structured interviews with five pilot participants provided qualitative insights that informed refinements to the interview protocol. Initial responses highlighted confusion with terms like "smoother" (e.g., conflated with clarity rather than fluency), prompting a rephrase to "fewer pauses" for precision. Participants also struggled to articulate complexity perceptions, leading to the addition of prompts such as "Why did you find it complex?" Coding of pilot interviews identified preliminary themes—e.g., "L1 ease" (4/5) and "L2 transition" (3/5)—which guided question adjustments to elicit clearer responses aligned with RQ1–RQ4 (Braun & Clarke, 2006). These refinements ensured the interviews effectively captured perceptions in the study.

3.5 Study Overview

The study involved 100 participants across four sessions (25 per session) over two weeks, addressing RQ1–RQ4. Each session included an introduction, two counterbalanced tasks (L1 and L2 planning), and questionnaire administration. Tasks were distinct yet comparable (e.g., Task 1: "Advise a friend on a loan issue"; Task 2: "Suggest study time solutions"), with 5-minute planning and 2–3-minute speaking phases. Interviews with 15 participants occurred within one week post-sessions. This design ensured all participants experienced both conditions, grounding perceptions in practice for robust data collection.

3.6 Procedure

The procedure for this study was carefully structured to address RQ1–RQ4 efficiently and systematically. To begin, participants were recruited through announcements at the language institutes. Subsequently, they were screened using the Oxford Placement Test (OPT) to confirm their B1 proficiency level. This initial step ensured that all 100 participants met the study's eligibility criteria before proceeding to the experimental sessions.

Following recruitment, the study commenced with a 10-minute session introduction. During this phase, participants were provided with a clear overview of the study's purpose and the tasks ahead, after which they signed consent forms to confirm their voluntary participation. This orientation not only set the stage for the tasks but also aligned with ethical standards, paving the way for the study's practical components.

Next, the first task, which lasted 10 minutes, was administered. To start, participants were randomly assigned—via a coin toss—to plan in either L1 (Persian) or L2 (English) for 5 minutes, followed by 2–3 minutes of speaking in English. This randomization ensured an even distribution of starting conditions, allowing for a balanced comparison of perceptions across RQ1–RQ3 as participants experienced their initial planning approach.

A 5-minute break was then provided between tasks to ensure a smooth progression and reduce fatigue. This brief intermission allowed participants to rest and refocus, preparing them for the subsequent task without the burden of continuous effort affecting their performance or responses.

Subsequently, the second task, also lasting 10 minutes, was conducted. Here, participants switched planning languages (from L1 to L2 or L2 to L1), repeating the 5-minute planning and 2–3-minute speaking format. By alternating the planning language, each participant experienced both conditions firsthand, thereby providing a comprehensive basis for their perceptions related to RQ1–RQ4.

Immediately following Task 2, the questionnaire was administered, taking 15–20 minutes to complete. This timing was strategic, as it captured participants' immediate reflections on both L1 and L2 planning experiences while their impressions were still fresh. Consequently, the questionnaire responses offered a direct and timely insight into RQ1–RQ3 (accuracy, fluency, complexity) and RQ4 (preferences and challenges).

Finally, to deepen the qualitative data collection, interviews with 15 selected participants were scheduled and conducted 1–2 days later in a quiet room at the institutes. These sessions, which were audio-recorded with consent, built on the questionnaire findings by allowing for detailed exploration of individual experiences. This final step ensured that the study fully addressed RQ4 and enriched the understanding of RQ1–RQ3 through nuanced participant narratives.

3.7 Data Analysis

Data analysis addressed RQ1–RQ4 as follows:

- **Quantitative Analysis:** Questionnaire responses were analyzed using SPSS (Version 26). Descriptive statistics (means, SDs) summarized perceptions per item (e.g., RQ1 accuracy). Paired t-tests compared L1 versus L2 planning perceptions within participants for RQ1–RQ3 (e.g., fluency scores), with Cronbach's Alpha confirming reliability (> 0.7). RQ4 preferences were assessed via mean scores on preference items, with order effects checked via independent t-tests (L1-first vs. L2-first).
- **Qualitative Analysis:** Interview transcripts (Persian, translated to English) underwent thematic analysis (Braun & Clarke, 2006). Axial coding identified themes (e.g., "L1 boosts fluency," "L2 feels direct") for RQ1–RQ4, with two researchers coding 20% of data (Cohen's kappa > 0.8 for reliability). RQ4 challenges emerged prominently here. Dependability was ensured by documenting a detailed audit trail of coding steps,

supporting consistency and repeatability (Lincoln & Guba, 1985), while confirmability was upheld through inter-coder agreement (Cohen's kappa > 0.8) and anchoring themes in participant quotes, reducing researcher bias (Braun & Clarke, 2006).

- **Integration of Quantitative and Qualitative Findings**

Quantitative and qualitative data were integrated using a convergent parallel design (Creswell & Plano Clark, 2018). After separate analyses, questionnaire trends (e.g., L1 preference for accuracy) were compared with interview themes (e.g., "L1 improves grammar") to identify convergence or divergence across RQ1–RQ4. This process involved triangulating mean scores with thematic insights creating a table to map alignments (e.g., fluency benefits) and discrepancies (e.g., complexity nuances), and ensuring a comprehensive response to each question. This integration aimed to synthesize broad perceptions with detailed narratives, enhancing the study's interpretive depth.

3.8 Ethical Considerations

Ethical principles were prioritized. Informed consent was obtained, explaining the study's purpose (addressing RQ1–RQ4), voluntary participation, and withdrawal rights. Anonymity was ensured via pseudonyms (P1–P100, I1–I15), with data stored securely. Recordings were deleted post-transcription. Institutes approved the study, and no incentives beyond task practice were offered, avoiding coercion. The study posed a minimal risk, adhering to SLA research ethics (Mackey & Gass, 2015).

3.9 Limitations

Limitations include the sample size (100) and convenience sampling from Hormozgan language institutes, potentially restricting generalizability beyond this context (RQ1–RQ4). The short intervention (two tasks) may not reflect long-term perceptions, possibly skewing RQ1–RQ3 responses. Order effects, despite counterbalancing, could influence data analyzed as a control. Self-reported data may carry social desirability bias, mitigated by triangulation. Finally, focusing on oral-opinion tasks limits applicability to other task types, a consideration for interpreting RQ1–RQ3 findings. Transferability is supported by thick descriptions of the Hormozgan EFL context and participant characteristics, enabling applicability to similar low-exposure settings, though it may not extend to high-exposure or rural contexts (Lincoln & Guba, 1985). Credibility, while bolstered by triangulation, may be tempered by self-report biases, warranting cautious interpretation of perceived L1 benefits (Mackey & Gass, 2015).

Results and Discussion

The study involved 100 intermediate EFL learners who completed two oral-opinion tasks (L1 and L2 planning), the 15-item Likert-scale questionnaire, and interviews (15 participants), addressing RQ1–RQ4.

4.1 Quantitative Findings

Descriptive Statistics

Table 3 presents questionnaire responses (1 = Strongly Disagree, 5 = Strongly Agree) for RQ1–RQ4. For accuracy (RQ1), L1 planning scored notably higher ($M = 4.3$) than L2 ($M = 3.6$), indicating a clear preference. Fluency (RQ2) showed a similar trend, with L1 at $M = 4.2$ versus L2's $M = 3.5$. Complexity (RQ3) revealed a smaller gap, with L1 at $M = 3.7$ and L2 at $M = 3.5$. Preference (RQ4) strongly favored L1 ($M = 4.1$) over L2 ($M = 3.3$). These key means suggest L1 is perceived as superior across all dimensions, with the largest differences in accuracy and preference.

Table 3. *Study Descriptive Statistics for Questionnaire ($N = 100$)*

| Item | L1 Planning (M, SD) | L2 Planning (M, SD) |
|-------------------------------------|---------------------|---------------------|
| Fewer grammar mistakes (RQ1) | 4.3 (0.61) | 3.6 (0.78) |
| Smoother speech (RQ2) | 4.2 (0.65) | 3.5 (0.80) |
| More complex sentences (RQ3) | 3.7 (0.82) | 3.5 (0.85) |
| Prefer this approach (RQ4) | 4.1 (0.69) | 3.3 (0.91) |

Inferential Statistics

Paired t-tests ($\alpha = 0.05$) assessed differences between L1 and L2 planning perceptions, as shown in Table 4. For accuracy (RQ1), L1's advantage was significant ($t(99) = 6.82$, $p < 0.001$), with a moderate effect ($d = 0.68$). Fluency (RQ2) showed a strong L1 preference ($t(99) = 7.14$, $p < 0.001$, $d = 0.71$). Complexity (RQ3) had a smaller but significant difference ($t(99) = 2.10$, $p = 0.038$), with a weak effect ($d = 0.21$). Preference (RQ4) yielded the strongest result ($t(99) = 8.33$, $p < 0.001$, $d = 0.83$), confirming L1's favorability. To check order effects for RQ4, independent t-tests compared preference scores between L1-first ($M = 4.0$, $SD = 0.72$) and L2-first ($M = 4.2$, $SD = 0.66$) groups, finding no significant difference ($t(98) = 1.45$, $p = 0.15$), suggesting counterbalancing mitigated sequence bias. Similarly, no order effects emerged across RQ1–RQ3 ($p > 0.05$), reinforcing the design's effectiveness.

Table 4. *Study Paired T-Tests for L1 vs. L2 Planning ($N = 100$)*

| Item | t-value | p-value | Cohen's d |
|-------------------------------------|---------|---------|-----------|
| Fewer grammar mistakes (RQ1) | 6.82 | <0.001 | 0.68 |
| Smoother speech (RQ2) | 7.14 | <0.001 | 0.71 |
| More complex sentences (RQ3) | 2.10 | 0.038 | 0.21 |
| Prefer this approach (RQ4) | 8.33 | <0.001 | 0.83 |

4.2 Qualitative Findings

Interviews with 15 participants yielded five themes via thematic analysis (Braun & Clarke, 2006), shown in Table 5. For RQ1, "L1 improves accuracy" (12/15) highlighted better grammar planning. For RQ2, "L1 enhances fluency" (13/15) noted reduced hesitations. For RQ3, "Complexity mixed" (9/15) indicated richer ideas but L2 vocabulary challenges. For RQ4, "Confidence with L1"

(14/15) reflected ease and motivation, while “L2 transition difficulty” (10/15) pointed to switching issues. These themes complement the quantitative trends, adding depth to perceptions.

Table 5. *Study Interview Themes (N = 15)*

| Theme | Frequency | Example Quote | RQ Link |
|--------------------------|-----------|---|---------|
| L1 improves accuracy | 12 | “Persian helps me plan grammar better.” | RQ1 |
| L1 enhances fluency | 13 | “I spoke faster after Persian planning.” | RQ2 |
| Complexity mixed | 9 | “Ideas were bigger, but English words were hard.” | RQ3 |
| Confidence with L1 | 14 | “Persian made me less nervous.” | RQ4 |
| L2 transition difficulty | 10 | “Switching to English slowed me down.” | RQ4 |

4.3 Integration of Quantitative and Qualitative Findings

Surveys and interviews converge on a robust preference for L1 planning, with nuanced differences across RQ1–RQ4, as detailed in Table 6. For accuracy (RQ1), the questionnaire’s strong L1 endorsement (M = 4.3 vs. 3.6) aligns closely with interview reports of enhanced grammar planning, suggesting learners view L1 as a reliable anchor for precision, leveraging native proficiency to structure speech effectively. Fluency (RQ2) shows similar convergence, with survey ratings (M = 4.2 vs. 3.5) reinforced by interview themes of fewer hesitations, indicating L1’s perceived role in facilitating uninterrupted delivery through a more intuitive process. Complexity (RQ3) reveals partial alignment—surveys indicate a slight L1 advantage (M = 3.7 vs. 3.5), but interviews qualify this with a dual perspective: L1 fosters richer idea generation (e.g., “bigger ideas”), yet L2 vocabulary constraints limit expression (e.g., “hard English words”), highlighting a trade-off not fully captured quantitatively. For preference (RQ4), surveys strongly favor L1 (M = 4.1 vs. 3.3), echoed by near-universal interview mentions of confidence, though tempered by frequent transition difficulties, adding practical nuance to the quantitative trend. Credibility was strengthened by triangulating questionnaire data with interview narratives, ensuring findings authentically represent participants’ perceptions (Creswell & Plano Clark, 2018). This integration underscores L1’s perceived strengths—grammar precision, fluency ease, and affective comfort—while exposing L2 execution barriers, particularly in complex output. The mixed-methods approach enriches this picture by revealing qualitative subtleties, such as complexity’s limitations and transition challenges, offering a comprehensive view of learners’ experiences beyond survey data alone.

Table 6. *Integration of Quantitative and Qualitative Findings (N = 100, Interviews N = 15)*

| Dimension (RQ) | Quantitative Finding | Qualitative Finding | Alignment |
|----------------|-------------------------|------------------------------|--------------------|
| Accuracy (RQ1) | L1 rated higher than L2 | L1 improves grammar planning | Strong convergence |
| Fluency (RQ2) | L1 rated higher than L2 | L1 reduces hesitations | Strong convergence |

| | | | |
|-------------------------|-------------------------------|---------------------------------------|---------------------|
| Complexity (RQ3) | L1 slightly higher than L2 | L1 aids ideas, L2 limits expression | Partial convergence |
| Preference (RQ4) | L1 strongly preferred over L2 | Confidence with L1, transition issues | Strong convergence |

Building on the integrated findings, the following paragraphs summarize how the data address each research question, synthesizing quantitative and qualitative insights to provide clear answers to RQ1–RQ4.

RQ1 (Accuracy): Learners perceive Persian planning as markedly more effective than English planning for accuracy. They report that L1 enables meticulous grammar structuring (e.g., “Persian helps grammar”), leveraging native proficiency to reduce errors and enhance sentence clarity. This perception positions L1 as a foundational tool, allowing focus on precision without the immediate burden of L2 formulation, which is particularly valuable where L2 practice is scarce.

RQ2 (Fluency): Persian planning is seen as significantly enhancing fluency compared to English. Participants highlight smoother speech with fewer pauses (e.g., “Faster after Persian”), attributing this to reduced mental effort and a more intuitive process in L1. This ease of flow suggests L1 alleviates the cognitive strain of real-time L2 production, fostering natural delivery among learners accustomed to Persian communication.

RQ3 (Complexity): Perceptions of complexity are mixed regarding Persian versus English planning. Learners view L1 as advantageous for generating richer, more elaborate ideas (e.g., “Big ideas”), reflecting a belief that native planning unlocks creative potential. However, they note that translating into English often restricts vocabulary range and sentence intricacy (e.g., “hard English words”), indicating L2 proficiency as a limiting factor that tempers L1’s initial benefits in this domain.

RQ4 (Preferences/Challenges): Learners overwhelmingly prefer Persian planning, citing heightened confidence, comfort, and control (e.g., “Persian less nervous”) as key drivers. They describe L1 as a familiar anchor reducing anxiety and boosting motivation, critical in an EFL setting with limited L2 exposure. Yet, challenges emerge in transitioning to English, with many reporting slowdowns or disruptions (e.g., “Switching slowed me”), revealing a practical trade-off between L1’s preparatory strengths and L2 execution demands.

4.4 Discussion

The findings from this study illuminate Iranian EFL learners’ perceptions of pre-task planning in L1 on speaking performance, offering insights that resonate with and extend existing research and theoretical frameworks. For accuracy (RQ1), learners’ perception of Persian planning as markedly superior aligns with Swain and Lapkin (2000), who found L1 scaffolds grammatical processing—a benefit likely amplified in Iran’s low-exposure context, where L2 practice is scarce (Samadi, 2011). This suggests that L1 enables meticulous grammar structuring by leveraging native proficiency, reducing errors in ways that L2 planning struggles to replicate. Cognitive Load Theory (Sweller, 1988) supports this interpretation, positing that L1 minimizes extraneous cognitive demands, freeing resources for precision—a dynamic particularly relevant for intermediate learners navigating L2 constraints. Similarly, fluency perceptions (RQ2) echo Yuan and Ellis (2003), where reduced mental effort enhances speech flow, with affective factors like lower anxiety (Kim & McDonough, 2008) reinforcing this advantage in a setting where speaking confidence is often low. This perceived ease of delivery positions L1 as a facilitator of natural

production, contrasting with the hesitations reported in L2 planning. Confirmability is evident as interpretations stem directly from participant data, such as quotes (e.g., “Persian helps grammar”), ensuring findings reflect their perspectives rather than researcher assumptions (Lincoln & Guba, 1985).

In contrast, the mixed perceptions of complexity (RQ3)—L1 fostering richer ideas but L2 limiting expression—parallel Wigglesworth and Storch (2009), highlighting a proficiency bottleneck (Ortega, 1999). Levelt’s (1995) speech production model provides a lens for this: L1 conceptualization streamlines planning, yet L2 articulation falters without advanced vocabulary, a challenge pronounced in Iran’s input-scarce environment. This duality suggests that while L1 unlocks creative potential, its benefits are tempered by learners’ L2 limitations, distinguishing this study’s perception focus from performance-based findings like Galbat et al. (2025). For preferences and challenges (RQ4), L1’s confidence boost aligns with Carson and Kashihara (2012), reflecting its role as a familiar anchor that reduces anxiety and boosts motivation. However, transition difficulties echo Macaro’s (2005) concerns about L2 processing barriers, hinting at potential over-reliance risks (Turnbull & Arnett, 2002). Sociocultural Theory (Vygotsky, 1978) frames L1 as a mediational tool here, extending its relevance to individual planning in low-exposure contexts beyond collaborative settings.

These findings underscore the nuanced role of L1 planning in TBLT, suggesting potential pedagogical applications, such as leveraging L1 to build initial proficiency. However, their reliance on perceptions rather than objective performance data highlights a need for further validation, a point expanded in the Conclusion. Similarly, the focus on oral opinion tasks raises questions about broader applicability, setting the stage for future exploration. By foregrounding learners’ subjective experiences, this study enriches the L1 debate in SLA, offering a foundation for both theoretical refinement and practical adaptation in Iranian EFL settings.

Conclusion

Iranian EFL learners strongly prefer L1 planning, perceiving it as enhancing accuracy and fluency through superior grammar organization and smoother delivery, as evidenced by survey and interview data. Complexity shows mixed perceptions, with L1 facilitating idea generation but constrained by L2 translation hurdles, reflecting a nuanced benefit. Learners value L1 for confidence and comfort, though switching difficulties temper this, highlighting both strengths and challenges in this EFL context.

Integrating quantitative and qualitative data, this study enriches the L1 use debate, affirming its scaffolding role in TBLT (Swain & Lapkin, 2000; Ellis, 2003). It offers a detailed Iranian perspective, where low L2 exposure amplifies L1’s utility, aligning with Cognitive Load Theory (Sweller, 1988) and Sociocultural Theory (Vygotsky, 1978). By foregrounding learners’ voices, it informs pedagogical design, adding cultural specificity to global SLA discourse (Dörnyei, 2005). The triangulated mixed-methods approach ensures credible insights into learners’ perceptions, while detailed contextual descriptions enhance transferability to comparable low-exposure EFL settings (Creswell & Plano Clark, 2018).

This study’s findings—demonstrating Iranian EFL learners’ preference for L1 (Persian) planning to enhance accuracy and fluency, despite complexity constraints and switching challenges—provide actionable strategies for low-exposure EFL contexts like Iran. Teachers can leverage L1 planning as a scaffold, allocating 5–7 minutes in TBLT lessons to improve grammar and flow, then guide students into L2 practice to boost confidence and reduce errors (Willis, 1996).

Building on this, learners gain from lower anxiety and better output, gradually developing autonomy by rephrasing Persian outlines in English with vocabulary support to increase complexity (Ellis, 2003). To sustain these benefits, curricula can integrate bilingual planning, shifting from L1-heavy to L2-dominant tasks over time. Meanwhile, researchers can extend these insights, exploring L1's role across diverse EFL settings to refine its application. Complementing this, policymakers should fund teacher training on strategic L1 use, overcoming cultural resistance (Samadi, 2011), thus optimizing oral skills and adapting TBLT globally beyond L2-only norms (Krashen, 1985).

This study's reliance on self-reported perceptions, while enriched through triangulation of survey and interview data, remains susceptible to biases such as social desirability or inaccurate self-assessment, potentially inflating the perceived benefits of L1 planning over actual performance outcomes (Mackey & Gass, 2015). The exclusive focus on oral-opinion tasks, though suitable for exploring immediate planning effects, restricts generalizability to other task genres like narratives or problem-solving, which may demand different cognitive and linguistic resources (Bygate, 2001). Similarly, the sample of 100 learners from three Hormozgan institutes, selected via convenience sampling, may not reflect the heterogeneity of Iranian EFL learners across urban-rural divides, socioeconomic backgrounds, or varying educational access (Samadi, 2011); this geographic and demographic limitation risks over-representing motivated, institutionally supported learners, potentially skewing perceptions of L1's utility in less-structured settings. The two-task design, while controlled and practical, offers only a snapshot of planning effects, limiting insights into sustained impacts on proficiency or learners' adaptation to L2 planning over time—an omission critical for understanding long-term language acquisition trajectories (Yuan & Ellis, 2003). Additionally, the absence of objective performance measures (e.g., error rates, speech fluency metrics) alongside perceptions hinders the validation of reported effects, leaving uncertainty about whether L1 planning's perceived advantages translate to measurable gains. These constraints collectively necessitate cautious interpretation of the findings' scope and applicability, particularly beyond intermediate learners or Iran's low-exposure EFL context.

Future studies could pair perceptions with objective measures like error rates or speech rates to validate findings, addressing the perceptual bias limitation (Yuan & Ellis, 2003). Exploring diverse tasks (e.g., narratives, problem-solving) could test L1 planning's versatility across genres (Bygate, 2001). Longitudinal designs tracking L1 planning effects over months could reveal sustained impacts on proficiency, particularly complexity development. Cross-cultural comparisons with high-exposure EFL settings (e.g., Saudi Arabia; Al-Nofaie, 2018) might clarify how L2 input availability shapes perceptions, refining L1's role in varied contexts.

This study underscores L1 planning's value in Iranian EFL classrooms, advocating its strategic use to enhance speaking while addressing a research gap with actionable insights for theory and practice.

References

- Al-Nofaie, H. (2018). Teachers' and learners' attitudes toward the use of Arabic in EFL classrooms in Saudi Arabia. *Journal of Language Teaching and Research*, 9(2), 312–320. <https://doi.org/10.17507/jltr.0902.12>
- Allan, D. (2004). *Oxford Placement Test 1: Test pack*. Oxford University Press.

- Anton, M., & DiCamilla, F. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *The Modern Language Journal*, 82(3), 233–247. <https://doi.org/10.1111/j.1540-4781.1998.tb01227.x>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 23–48). Longman.
- Carson, E., & Kashihara, H. (2012). Using the L1 in the L2 classroom: The students speak. *The Language Teacher*, 36(4), 41–48.
- 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. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum Associates.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford University Press.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3), 299–323. <https://doi.org/10.1017/S0272263196000137>
- Galbat, H., Fahandezh Saadi, F., & Afraz, Sh. (2025). Effect of Using First Language by Iranian EFL Learners in Task Preparation on their Speaking Accuracy, Fluency, and Complexity. *International Journal of Foreign Language Teaching and Research*, 13 (52), 55-69. <https://doi.org/10.71962/jfl.2024.1188083>
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473. <https://doi.org/10.1093/applin/amp048>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Kim, Y. (2015). The role of task complexity and L1 use in second language oral production. *Korean Journal of Applied Linguistics*, 31(2), 123–145.
- Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, 12(2), 211–234. <https://doi.org/10.1177/1362168807086288>
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Levelt, W. J. M. (1995). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223–232. [https://doi.org/10.1016/S1364-6613\(99\)01319-4](https://doi.org/10.1016/S1364-6613(99)01319-4)
- Macaro, E. (2005). Codeswitching in the L2 classroom: A communication and learning strategy. In E. Llurda (Ed.), *Non-native language teachers: Perceptions, challenges, and contributions to the profession* (pp. 63–84). Springer.
- Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design* (2nd ed.). Routledge.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21(1), 109–148. <https://doi.org/10.1017/S0272263199001047>
- Samadi, M. (2011). The use of L1 in EFL classrooms: A case study in Iran. *Journal of Language Teaching and Research*, 3(5), 914–921. <https://doi.org/10.4304/jltr.3.5.914-921>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49(1), 93–120. <https://doi.org/10.1111/1467-9922.00071>
- Storch, N., & Wigglesworth, G. (2003). Is there a role for the use of the L1 in an L2 setting? *TESOL Quarterly*, 37(4), 760–770. <https://doi.org/10.2307/3588225>
- Swain, M., & Lapkin, S. (2000). Task-based second language learning: The uses of the first language. *Language Teaching Research*, 4(3), 251–274. <https://doi.org/10.1177/136216880000400304>
- Sweller, J. (1988). Cognitive load during problem-solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296. <https://doi.org/10.1023/A:1022193728205>
- Turnbull, M., & Arnett, K. (2002). Teachers' uses of the target and first languages in second and foreign language classrooms. *Annual Review of Applied Linguistics*, 22, 204–218. <https://doi.org/10.1017/S0267190502000119>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity, and accuracy. *Language Testing*, 26(3), 445–466. <https://doi.org/10.1177/0265532209104670>
- Willis, J. (1996). *A framework for task-based learning*. Longman.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and online planning on fluency, complexity, and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1–27. <https://doi.org/10.1093/applin/24.1.1>

Biodata

Hassan Galbat is a Ph.D. candidate of TEFL at Islamic Azad University Qeshm Branch, Iran. He has been involved in running language institutes and teaching English for more than 20 years. Also, he has two years of experience teaching at universities. His main areas include language teaching, teaching theories and skills, testing, and designing speaking tasks.
Email: hmgalbat@yahoo.com

Farhad Fahanddezh Sadi, Assistant Professor of TEFL, is a faculty member of Islamic Azad University, Bandar Lenge branch. He has been teaching English for more than twenty years. His areas of interest are English language teaching, writing, syllabus design, and phonology.
Email: Farhadfahanddezh@gmail.com

Shahram Afraz, Assistant Professor of TEFL, is a faculty member of Islamic Azad University, Qeshm Branch. He has been teaching English for twenty years at several universities. His areas of interest are English language teaching, linguistics, syllabus design, and testing. He has published more than eighty articles and five books till now.

Email: *shahram.afraz1352@gmail.com*