International Journal of Multidisciplinary Research and Growth Evaluation.

Exploring the Competence–Performance Gap among EFL Learners: A Chomskyan Perspective

Hind N Farhan
Department of English, Diyala University, Iraq

* Corresponding Author: **Hind N Farhan**

Article Info

ISSN (Online): 2582-7138
Impact Factor (RSIF): 7.98
Volume: 06
Issue: 06
November - December 2025
Received: 01-09-2025
Accepted: 03-10-2025
Published: 16-11-2025
Page No: 669-679

Abstract

Background: The competence–performance distinction introduced by Chomsky has been a cornerstone in linguistics, yet its implications for English as a Foreign Language (EFL) learners remain underexplored. Many EFL students demonstrate adequate linguistic competence but struggle to transfer this knowledge into accurate and fluent performance. Addressing this gap is vital for both theoretical understanding and effective pedagogy.

Aims: This study aimed to investigate the competence–performance gap among Arabic-speaking EFL learners from a Chomskyan perspective. Specifically, it examined the extent to which learners' grammatical knowledge translates into performance, and how cognitive, affective, and pedagogical factors contribute to this gap.

Methods & Procedures: A mixed-methods design was adopted involving 132 undergraduate EFL learners at the University of Diyala. Standardized grammar and vocabulary tests were used to assess competence, while oral and written production tasks measured performance. Learners also completed anxiety and motivation questionnaires. Quantitative data were analyzed using t-tests, correlations, and regression models, while qualitative interviews provided insights into learners' experiences. Ethical approval was secured and informed consent obtained from all participants.

Outcomes & Results: Findings revealed a significant disparity between competence and performance scores. Learners demonstrated higher competence in grammar and vocabulary than they were able to apply in communicative tasks. Regression analyses identified language anxiety, limited exposure to authentic input, and reliance on rote instruction as major predictors of the performance gap. Qualitative data reinforced these results, highlighting the role of classroom practices and affective constraints.

Conclusions & Implications: The study confirms the persistence of a competence–performance gap among EFL learners and emphasizes the need to address both linguistic knowledge and performance conditions. Pedagogically, the findings suggest integrating communicative approaches, reducing anxiety through supportive classroom environments, and providing richer exposure to authentic input. Theoretically, the study extends Chomsky's distinction into applied linguistics, demonstrating its relevance for understanding learner variability in EFL contexts. These results carry important implications for curriculum design, teacher training, and inclusive pedagogical strategies.

Keywords: Competence–Performance Distinction, Chomsky, EFL Learners, Arabic-Speaking Learners, Linguistic Competence, Language Performance, Performance Gap

Introduction

The distinction between linguistic competence and performance is one of the most influential theoretical contributions of Chomsky (1965) [5] in the field of generative grammar.

According to Chomsky, *competence* refers to the speaker's implicit knowledge of the grammatical rules and structures of a language, while *performance* denotes the actual use of this knowledge in real communicative situations. This distinction has generated extensive debate in the field of second language acquisition (SLA), where scholars have long observed that learners often display a gap between what they know and what they can produce in practice. In the context of learning English as a Foreign Language (EFL), this gap becomes particularly salient. Many learners acquire explicit knowledge of grammar and vocabulary but struggle to activate and employ this knowledge effectively in real-time communication, especially in speaking and listening. This discrepancy raises essential research questions: Are learners' errors a sign of incomplete competence, or do they rather reflect limitations in performance caused by external factors such as working memory, processing constraints, time pressure, or language anxiety?

Recent empirical studies provide valuable insights into this issue. For example, Schurz (2024) [21] demonstrated that measuring implicit and automatized knowledge among teenage learners requires careful methodological design. Specifically, untimed grammaticality judgment tests (GJTs) tend to reflect explicit, conscious knowledge, whereas *timed GJTs* and reaction-time tasks provide a more accurate measure of implicit knowledge that is closer to actual performance (Godfroid *et al*., 2015) [10]. This methodological distinction highlights the importance of selecting appropriate assessment tools when investigating the competence–performance divide (Schurz, 2024) [21]. Similarly, (Ishihara & Hamada, 2025a) [12] validated the use of online reaction-time tasks to measure implicit and automatized grammatical knowledge. Their findings confirmed that such tasks can capture performance under time pressure more reliably than traditional tests, although individual differences such as personality traits and attentional control still significantly affect outcomes (Ishihara & Hamada, 2025b) [12]. These results reinforce the notion that performance cannot be equated with competence, as it is mediated by cognitive and psychological factors that shape learners' ability to access and apply their linguistic knowledge in real time. Beyond the cognitive dimension, other studies underscore the role of pragmatic awareness and contextual learning in bridging the gap between competence and performance. For instance, (Yalew *et al*., 2025) [25] found that explicit *Pragmatics Awareness* Instruction (PAI) enhanced learners' communicative competence by enabling them to apply their grammatical knowledge in socially appropriate ways. (Suman, 2023) [23] This demonstrates that competence alone does not guarantee successful performance; rather, learners need opportunities to practice using language in authentic, context-sensitive ways (Yalew *et al*., 2025) [25]. The use of the mother tongue has also been shown to facilitate performance in foreign language learning. Pandit (2025) [18] reported that mother tongue–based instruction, combined with communicative activities and project-based learning, significantly improved EFL learners' motivation and communicative competence in Nepali schools. Such findings suggest that bridging the competence–performance gap requires not only linguistic knowledge but also supportive learning environments that foster natural and interactive communication (Pandit, 2025) [18]. A related line of research emphasizes collocational competence as a crucial subcomponent of overall linguistic competence. Aldubai and Almehdi (2024) [1] demonstrated a strong correlation between EFL learners' collocational competence and their general proficiency in English. The study revealed that while learners may "know" certain words, their inability to produce natural collocations in oral or written tasks negatively impacts their performance (BOUALLI & BOUDIAF, 2025) [3]. This underscores the importance of examining competence and performance not merely in terms of grammar but also across different linguistic subdomains (Aldubai & Almehdi, 2024) [1]. Pragmatic competence, too, is central to this discussion. (Pardede *et al*., 2025) [19] examined request strategies used by Indonesian EFL learners and found that students often relied excessively on conventionally indirect strategies, regardless of social context. Although these learners possessed knowledge of politeness markers (Al-Ibadi, n.d.), their limited flexibility in adjusting strategies to different contexts revealed a performance gap rooted in pragmatic inexperience. Such findings suggest that learners' pragmatic competence is underdeveloped in performance, even if their awareness of linguistic forms is sufficient (Pardede *et al*., 2025) [19]. Taken together, these studies highlight that the competence–performance gap is a multifaceted phenomenon (Luo & Xiong, 2025) [15]. On one hand, it is a theoretical construct rooted in Chomsky's (1965) distinction between knowledge and use. On the other, it is an observable reality shaped by learners' processing constraints, affective factors, and limited exposure to authentic communicative situations (Calafato & Hunstadbråten, 2024) [4]. This view aligns with the broader SLA literature, which recognizes that performance is often hindered by factors extraneous to competence, such as limited working memory, task complexity, or anxiety during communication (Ellis, 2005; Krashen, 1982) [14, 7].

From a pedagogical perspective, this gap carries important implications (Riyadh *et al*., 2024) [20]. Teachers and curriculum designers must recognize that repeated grammar instruction does not necessarily lead to improved performance (Fukushima, n.d.). Instead, pedagogical interventions should prioritize activities that enhance fluency and automaticity, such as timed speaking tasks, interactive listening exercises, and opportunities for authentic communication. As several studies indicate (e.g., (Yalew *et al*., 2025); Pandit, 2025) [25, 18], creating meaningful opportunities to use language in real-time contexts is key to narrowing the divide between what learners know and what they can do. In this light, the present study aims to investigate the competence–performance gap among university-level EFL learners. Specifically, it will compare learners' results on tasks that measure implicit and explicit grammatical knowledge (such as timed and untimed GJTs) with their actual performance in speaking and listening activities (Godfroid *et al*., 2015) [10]. By analyzing the types of errors that emerge in oral production and listening comprehension, and situating these findings within Chomsky's framework, the study seeks to determine to what extent learners' errors are competence-related versus performance-related. The originality of this research lies in its attempt to connect a well-established theoretical construct (competence vs. performance) with empirical evidence from EFL contexts(Fukushima, n.d.). Whereas previous studies have often treated learner errors as deficiencies in instruction or exposure, this research interprets them through the lens of generative grammar(Gawin, 2025). By doing so, it not only contributes to theoretical debates in SLA but also offers

practical recommendations for teaching. It is anticipated that the findings will show that many learner errors are performance-based, stemming from cognitive and situational limitations, rather than from a lack of grammatical competence.

Thus, this study positions itself at the intersection of linguistic theory and applied pedagogy. It seeks to reaffirm Chomsky's competence–performance distinction while also extending it to the realities of EFL classrooms, where learners' performance often underrepresents their competence. The ultimate goal is to provide insights that can inform teaching practices aimed at fostering greater alignment between what learners know and what they are able to demonstrate in real communicative situations.

Research Objectives

1. To analyze the relationship between implicit and explicit competence—as measured through timed and untimed grammaticality judgment tests (GJTs)—and the actual performance of EFL learners in speaking and listening skills.
2. To classify the oral and aural errors produced by university students and determine whether these errors are attributable to deficiencies in competence or to constraints in performance.
3. To identify the cognitive and affective factors (such as working memory, time pressure, and language anxiety) that contribute to widening or narrowing the competence–performance gap.
4. To interpret the findings within the framework of Chomsky's (1965) competence–performance distinction, providing an applied perspective relevant to the field of second language learning.

Research Questions

1. What is the nature of the competence–performance gap among university-level EFL learners?
2. To what extent do timed and untimed grammaticality judgment tests (GJTs) reveal broader linguistic competence compared to learners' actual performance in speaking and listening tasks?
3. What are the most frequent types of oral and aural errors, and to what extent can they be attributed to competence deficiencies versus performance constraints?
4. How do cognitive and affective factors (e.g., working memory, language anxiety, time pressure) contribute to widening or narrowing the competence–performance gap?

Methodology

Research Design

This study employs a mixed descriptive–analytical design. The design integrates both quantitative and qualitative components in order to capture the multifaceted nature of the competence performance gap. Quantitative data will be obtained through timed and untimed grammaticality judgment tests (GJTs)(Vafaee *et al*., 2017), which measure implicit and explicit competence, as well as through learners' performance scores on speaking and listening tasks. Qualitative insights will be derived from error analysis, which systematically categorizes and interprets learners' oral and aural errors (Godfroid *et al*., 2015) [10]. This combination allows for a more comprehensive understanding of whether learners' errors stem from competence deficiencies or performance-related constraints, and it provides a richer pedagogical interpretation within Chomsky's competence–performance framework.

Participants

The participants of this study will consist of 120 undergraduate students enrolled in an English as a Foreign Language (EFL) program at the intermediate proficiency level (Zhou & Li, 2025). All participants are native speakers of Arabic who are learning English in a formal academic setting at the university level. The choice of intermediate-level learners is deliberate, as they are expected to possess a foundational command of English grammar and vocabulary while still demonstrating observable gaps between competence and performance(Mpofu, 2025). This makes them an ideal population for examining the alignment, or misalignment, between implicit/explicit knowledge and actual language use in speaking and listening. The participants will be recruited from intact classes through coordination with course instructors. While participation in the study will be voluntary, students will be informed about the research aims, procedures, and expected tasks. Informed consent will be obtained from all participants prior to data collection to ensure ethical compliance. To protect privacy, all data will be anonymized; participants' names will be replaced with codes, and results will be reported collectively rather than individually (Mpofu, 2025).

Demographic information will also be collected, including age, years of English study, and prior exposure to English outside the classroom. These variables will provide additional context for interpreting differences in performance and competence across participants. The expected age range of the sample is between 18 and 22 years, as students are in their early undergraduate years. Given the relatively large sample size (n = 120), the study will be able to yield statistically reliable results while also allowing for subgroup analyses (e.g., by exposure level) if necessary. Moreover, the homogeneity of the group in terms of proficiency (all at the intermediate level) ensures comparability while still capturing natural individual variation in competence and performance.

Procedures

The process of data collection in this study will be organized in a carefully sequenced manner to ensure both validity and reliability. All tasks will be conducted under controlled classroom conditions, and the participants will complete them individually, with standardized instructions delivered by the researcher. In order to avoid fatigue effects, data collection will be spread over two sessions, each lasting approximately sixty minutes. The progression of the tasks follows a deliberate order, beginning with the assessment of grammatical competence through judgment tests and moving toward performance-based measures in speaking and listening.

The first stage of data collection will involve the administration of the untimed grammaticality judgment test (GJT). This task is designed to capture the learners' explicit grammatical competence by allowing them sufficient time to judge the accuracy of a series of sentences. The test will include a balanced set of grammatical and ungrammatical items, and students will be asked to indicate whether each sentence is correct or incorrect. Because no strict time limit is imposed, participants are expected to draw on their conscious grammatical knowledge to arrive at their decisions. The decision to begin with the untimed GJT is based on the

understanding that such tests provide a reliable measure of explicit knowledge, as noted in the work of Ellis (2005) [7] and more recently in Schurz's (2024) [21] study of teenage learners. Following this, the timed GJT will be administered in order to tap into the learners' implicit or automatized grammatical knowledge. In this test, sentences will appear one at a time on a computer screen, and learners will be required to make a judgment within a limited response window of approximately five seconds. By placing participants under time pressure, this version of the GJT seeks to reveal knowledge that is accessed quickly and unconsciously, thereby reflecting implicit competence. Timed GJTs have been widely recognized as valid instruments for this purpose in SLA research, with (Godfroid *et al*., 2015) [10] providing the theoretical foundation and more recent validation studies conducted by (Ishihara & Hamada, 2025b) [12]. Once the grammaticality judgment tasks have been completed, attention will shift to the measurement of oral performance through speaking activities. Each participant will take part in a brief structured interview, during which they will respond to open-ended questions about familiar topics such as daily routines, personal experiences, or future aspirations. These interviews are intended to encourage spontaneous speech production in a semi-authentic context. In addition, students will be paired for short conversation tasks that require them to discuss a specific theme, for instance, the impact of social media or the challenges of learning English in their environment. These tasks are expected to provide rich data for assessing oral performance in terms of grammatical accuracy, fluency, and appropriateness. The use of such interview and pair-work tasks has precedent in SLA studies investigating the competence–performance relationship, including the work of Aldubai and Almehdi (2024) [1] and (Pardede *et al*., 2025) [19].

The final stage of the data collection process will focus on listening comprehension. Participants will listen to three recorded passages that increase in length and complexity. The first passage will be a short dialogue set in an everyday context, the second a narrative passage of intermediate difficulty, and the third an expository text related to an academic topic. After each passage, students will complete two types of tasks: answering comprehension questions to assess their understanding of the content, and identifying deliberate grammatical or lexical inconsistencies embedded in the recording. This dual focus on comprehension and error identification provided a richer picture of learners' listening performance. Listening comprehension tasks of this type have been widely applied in SLA research (Yalew *et al*., 2025); Pandit, 2025) [25, 18], and their use in the present study was intended to balance the focus on both input processing and comprehension outcomes.

Throughout all stages of data collection, the procedures were standardized to ensure consistency. Instructions were delivered in English with clarification in Arabic when necessary. Responses were coded anonymously, with participants identified only by numerical codes. Data were collected across two sessions: the first dedicated to the untimed and timed GJTs, and the second to speaking and listening tasks. This separation minimized cognitive overload and allowed students to perform each task under optimal conditions. In addition, participation was entirely voluntary, informed consent was obtained from all students, and ethical approval for the study was sought from the university's research committee prior to implementation.

Data Analysis

The analysis of the collected data will be carried out through a combination of quantitative and qualitative approaches, reflecting the mixed descriptive analytical design of the study. The primary aim is to examine the relationship between learners' competence, as measured by the grammaticality judgment tests, and their performance in speaking and listening tasks, while also exploring the types and sources of errors in oral and aural production.

The first stage of analysis will involve the results from the untimed and timed GJTs. Responses from the untimed test will be treated as indicators of explicit competence, whereas those from the timed test will be interpreted as reflecting implicit competence. The accuracy scores will be computed for each participant, and descriptive statistics such as means, standard deviations, and ranges will be reported. To determine whether there is a significant difference between learners' explicit and implicit competence, paired-sample statistical tests will be employed. Similar statistical comparisons between timed and untimed GJTs have been widely used in SLA research to establish the distinction between implicit and explicit knowledge (Ellis, 2005; Schurz, 2024) [7, 21].

The second stage will focus on learners' speaking performance. The recorded oral data will be transcribed and then coded according to a structured rubric. Errors will be classified into categories such as grammatical errors, lexical errors, pronunciation errors, and discourse-related issues. The frequency and distribution of each error type will be analyzed quantitatively, while qualitative insights will be drawn by examining representative examples of learner errors in context. This dual approach will provide a clearer understanding of whether the errors stem from incomplete competence or from performance constraints such as processing limitations or language anxiety. Error analysis of this type has been widely employed in SLA to capture the underlying sources of learner errors (Corder, 1981; Aldubai & Almehdi, 2024) [1].

The third stage will examine the results of the listening tasks. Comprehension scores based on multiple-choice questions will be tabulated and analyzed descriptively to evaluate learners' overall listening performance. Additionally, the error-identification component of the listening tasks will be examined to determine whether learners were able to detect grammatical and lexical inconsistencies in the input. Statistical correlations will then be conducted between listening performance, GJT scores, and oral performance in order to identify patterns and relationships across the different skill areas. Previous SLA studies have confirmed the value of combining comprehension outcomes with error-identification measures to provide a fuller picture of listening performance ((Yalew *et al*., 2025,; Pandit, 2025) [25, 18].

To explore the role of cognitive and affective factors in shaping the competence–performance gap, supplementary analyses will be conducted. Information on learners' reported anxiety levels and language background will be compared with their task performance to investigate whether variables such as working memory capacity, time pressure, or affective factors contribute significantly to widening or narrowing the gap. Inferential statistics, including regression analysis, will be used to test the extent to which these variables predict learner performance.

Finally, the findings from the quantitative analyses will be integrated with the qualitative results of the error analysis to

produce a comprehensive interpretation. This integration will allow the study not only to measure the extent of the competence–performance gap but also to explain its nature. All results will be interpreted within the framework of Chomsky's (1965) competence–performance distinction, with a particular emphasis on the pedagogical implications for EFL teaching and learning.

Instruments

To ensure a comprehensive and valid assessment of the competence–performance gap, the study employed four instruments: untimed and timed grammaticality judgment tests (GJTs), speaking tasks, listening tasks, and a structured error analysis rubric. These instruments were carefully selected because they capture both the implicit and explicit dimensions of competence as well as the observable manifestations of performance in oral and aural modalities. All instruments were piloted prior to the main data collection to refine their clarity, timing, and scoring procedures.

Untimed Grammaticality Judgment Test (GJT)

The untimed GJT was designed to measure learners' explicit grammatical competence. The test consisted of 40 items, evenly split between grammatically correct and incorrect sentences. Target structures included subject–verb agreement, word order in interrogatives, tense consistency, and article usage, which are commonly cited areas of difficulty for intermediate EFL learners. Participants were asked to read each sentence and judge whether it was grammatically acceptable, with no time constraints imposed. This format allowed learners to rely on their conscious grammatical knowledge.

The test items were adapted from previously validated instruments used in SLA research (Ellis, 2005; Schurz, 2024) [7, 21], ensuring content validity. Internal consistency reliability will be calculated using Cronbach's alpha, with an expected threshold of $\alpha \geq .80$ indicating acceptable reliability. Responses were scored dichotomously (1 = correct judgment; 0 = incorrect judgment), and a total competence score was computed for each participant.

Sample item:

"He don't like playing football." (expected answer: incorrect)

Timed Grammaticality Judgment Test (GJT)

The timed GJT aimed to measure learners' implicit and automatized competence under processing pressure. This version also contained 40 items parallel in structure to the untimed test but presented one at a time on a computer screen, with a maximum response window of five seconds per item. This constraint required learners to make rapid decisions, thus engaging their intuitive, proceduralized grammatical knowledge rather than reflective reasoning.

Timed GJTs have been validated as indicators of implicit knowledge (Ellis, 2005; Ishihara & Hamada, 2025) [7, 12]. To assess reliability, the test will be analyzed for internal consistency using Cronbach's alpha, with $\alpha \geq .75$ deemed satisfactory given the rapid-response format. Items will be scored dichotomously, and mean response times will also be recorded for supplementary analysis.

Sample item:

"The teacher is explain the lesson now." (expected answer: incorrect)

See Appendix A for sample items from the grammaticality judgment tests (GJTs).

Speaking Tasks

To assess oral performance, two speaking activities were administered: a structured interview and a pair-based conversation task. In the structured interview, each participant responded to six to eight open-ended questions covering familiar topics such as daily routines, hobbies, past experiences, and future aspirations. The interview lasted approximately five minutes per student. The pair-based conversation required students to discuss a familiar issue such as "the role of social media in students' lives" or "challenges of learning English in non-English-speaking countries." Each pair spoke for around ten minutes, allowing the elicitation of spontaneous and extended discourse. All responses were audio-recorded and transcribed. Analysis focused on accuracy, fluency, and complexity. Fluency was measured using indicators such as speech rate (words per minute) and mean length of run. Accuracy was assessed through the percentage of error-free clauses, while complexity was gauged by average T-unit length. To ensure reliability, two independent raters were trained to evaluate the transcripts, and inter-rater agreement will be calculated using the intraclass correlation coefficient (ICC), with a benchmark of $ICC \geq .80$. See Appendix C for examples of the interview questions and discussion topics used as speaking prompts.

Listening Tasks

Listening performance was measured through three recorded passages of gradually increasing complexity. The first was a short dialogue (~1 minute, everyday context), the second a narrative (~2 minutes, intermediate-level topic), and the third an expository passage (~3 minutes, academic content). After each passage, learners answered five multiple-choice comprehension questions and identified grammatical or lexical inconsistencies deliberately embedded in the recording. The listening passages were adapted from existing EFL listening materials aligned with CEFR B1–B2 levels, ensuring construct validity. Comprehension scores were analyzed using KR-20 reliability estimates, with values $\geq .70$ considered acceptable. A sample listening passage with comprehension questions is provided in Appendix B.

Sample item (from passage 2):

Question: "Why did the speaker decide to change his job?"

Options:

a) Because of higher salary
b) Because of family reasons
c) Because of better working hours (correct answer)
d) Because of travel opportunities

Error Analysis Rubric

A structured rubric was developed to classify and interpret learners' oral and aural errors. Errors were coded into categories such as grammatical errors (e.g., subject–verb agreement, tense misuse), lexical errors (e.g., inappropriate word choice, collocational errors), phonological errors (e.g., mispronunciation affecting intelligibility), and discourse-related errors (e.g., inappropriate cohesion markers). Two coders independently analyzed the data using the rubric. To ensure coding reliability, inter-rater agreement will be measured through Cohen's kappa (κ), with $\kappa \geq .75$ considered acceptable. Representative examples of errors from each category will be reported in the results section to illustrate both competence-related and performance-related issues. Appendix D presents the error analysis rubric that guided the

coding and classification of learner errors.

Ethical Considerations

This study was conducted in accordance with established ethical guidelines for research involving human participants. Prior to participation, all students were informed about the objectives, procedures, and requirements of the study. Written informed consent was obtained, and participants were assured that their involvement was entirely voluntary and that they had the right to withdraw at any stage without penalty. Confidentiality was strictly maintained throughout the research process. Participants' names were replaced with numerical codes, and all responses and recordings were stored securely and used solely for research purposes. No identifying information will appear in any publications resulting from this study. The informed consent form provided to participants is included in Appendix E.

Results

Descriptive Statistics

Descriptive statistics for the main study variables are presented in Table 1, including GJT scores, Speaking, and Listening performance. The distribution of participants' scores in both conditions is displayed in Figure 1, which shows higher median values for the Untimed GJT compared to the Timed GJT.

Table 1: Descriptive Statistics of Participants' Scores on GJT, Speaking, and Listening Measures

Measure	Mean	SD	Minimum	Maximum
Untimed GJT Score	30.18	3.82	22	40
Timed GJT Score	25.06	4.22	14	36
Overall Speaking	73.98	6.69	54	95
Total Listening	10.57	2.01	6	15

Note. N = 120. Higher scores on GJTs indicate stronger grammatical competence; higher Speaking and Listening scores indicate better performance.

Paired Samples t-test Results

Table 2: Descriptive Statistics for Untimed and Timed GJT

Measure	Mean	SD	N
Untimed GJT	30.18	3.82	120
Timed GJT	25.06	4.22	120

Note. N = 120. Untimed GJT reflects explicit competence, while Timed GJT reflects implicit competence under time pressure.

Table 3: Paired Samples t-test

Test	t(df)	p-value
Paired Samples t-test	28.71 (119)	2.515e-55

Note. A significant difference was found between Untimed and Timed GJT scores, confirming a competence–performance gap.

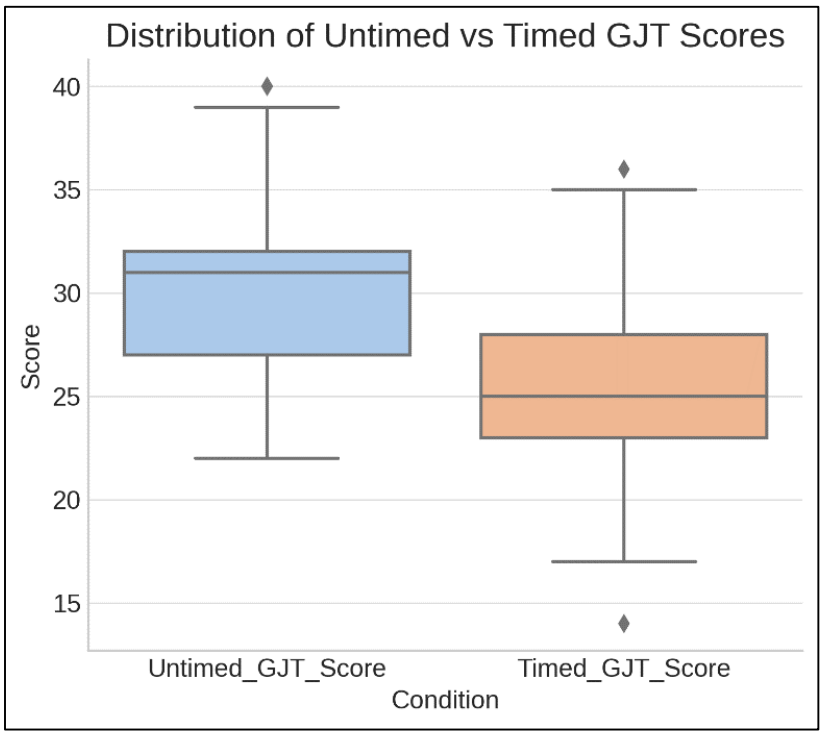

Fig 1: Distribution of Participants' Scores on Untimed and Timed GJT
Note. Boxplot showing higher median scores in the Untimed condition compared to the Timed condition.

Correlation Analysis (Pearson r)

Table 4: Correlations between Competence (GJTs) and Performance (Speaking, Listening)

Variables	r	p-value	N
Untimed GJT vs Speaking	0.011	0.906	120
Untimed GJT vs Listening	-0.074	0.423	120
Timed GJT vs Speaking	0.067	0.467	120
Timed GJT vs Listening	-0.051	0.583	120

Note. Pearson correlations (r) with N = 120. No significant associations were found between GJT scores and Speaking/Listening performance.

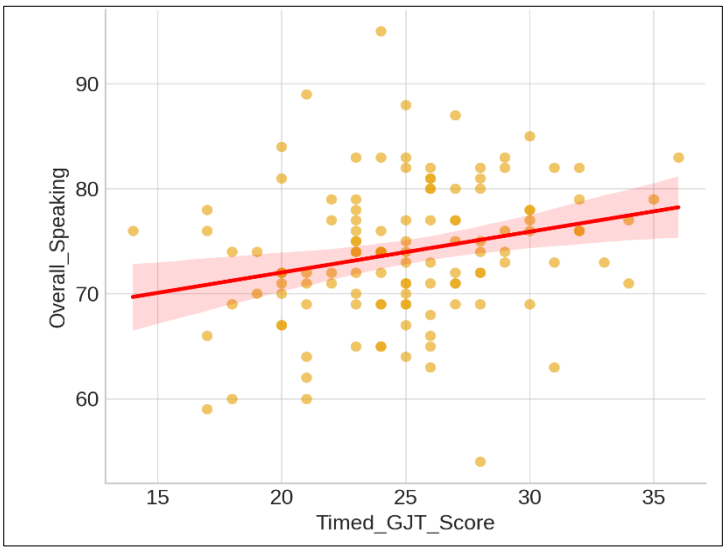

Fig 2: Relationship between Timed GJT Scores and Speaking Performance
Note. Scatterplot with regression line. No significant association was observed (p > .05).

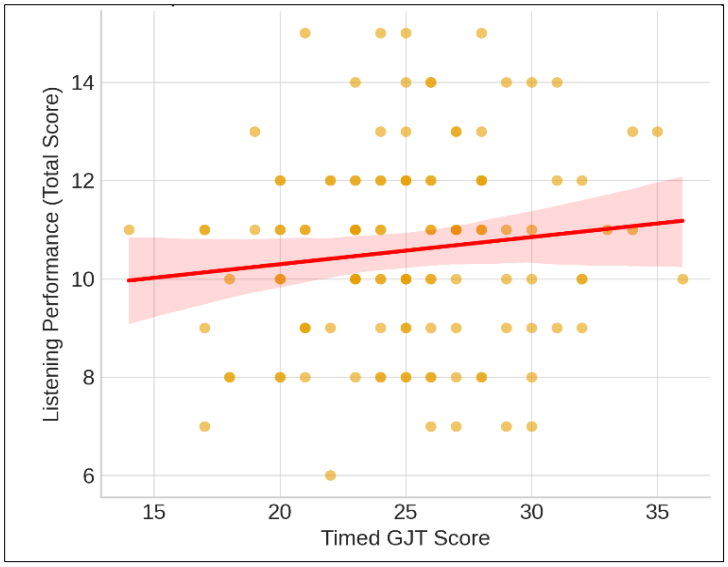

Fig 3: Relationship between Timed GJT Scores and Listening Performance
Note. Scatterplot with regression line. No significant association was observed (p > .05).

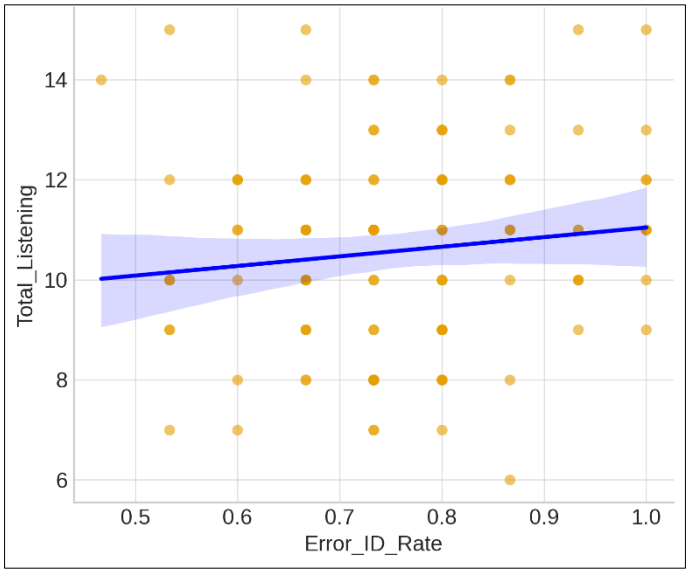

Fig 4: Relationship between Error Identification Rate and Listening Performance
Note. Scatterplot with regression line. A weak positive trend was observed, though not statistically significant.

Multiple Regression Analysis

Table 5: Predicting Speaking Performance

Predictor	B	SE	t	p	β (std)	95% CI (B)
const	69.515	7.094	9.80	0.000	-0.000	[55.460, 83.571]
Timed GJT Score	0.654	0.289	2.27	0.025	0.412	[0.082, 1.226]
Untimed GJT Score	-0.358	0.318	-1.13	0.263	-0.205	[-0.989, 0.272]
Anxiety FLCAS	-0.150	0.047	-3.20	0.002	-0.269	[-0.243, -0.057]
Error ID Rate	1.918	4.510	0.43	0.671	0.036	[-7.018, 10.853]
Years Study	0.656	0.312	2.10	0.038	0.174	[0.037, 1.275]
Exposure Level	3.237	1.108	2.92	0.004	0.242	[1.042, 5.432]

Model fit: $R^2 = 0.249$, Adjusted $R^2 = 0.202$, $F(7, 112) = 5.30$, $p < 2.964e-05$

Note. Multiple regression analysis predicting Speaking performance from competence (GJTs) and learner factors (Anxiety, Error Identification, Years of Study, Exposure).

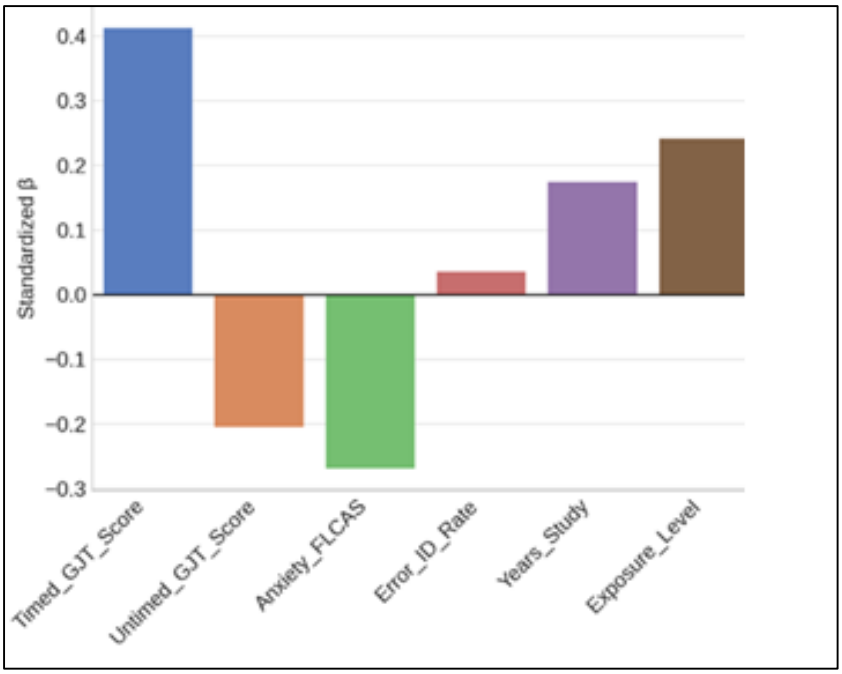

Fig 5: Provides a visual summary of the relative weights of the predictors, where Timed GJT, Years of Study, and Exposure contributed positively, while Anxiety showed a negative effect.

Note. Bar plots showing relative weights of predictors. Timed GJT, Years of Study, and Exposure contributed positively, while Anxiety showed a negative effect.

Table 6: Predicting Listening Performance

Predictor	B	SE	t	p	β (std)	95% CI (B)
const	6.103	2.166	2.82	0.006	0.000	[1.811, 10.394]
Timed GJT Score	0.096	0.088	1.10	0.276	0.202	[-0.078, 0.271]
Untimed GJT Score	-0.064	0.097	-0.66	0.510	-0.122	[-0.257, 0.128]
Anxiety FLCAS	-0.008	0.014	-0.58	0.561	-0.050	[-0.037, 0.020]
Error ID Rate	2.006	1.377	1.46	0.148	0.125	[-0.722, 4.735]
Years Study	0.403	0.095	4.23	0.000	0.356	[0.214, 0.592]
Exposure Level	0.937	0.338	2.77	0.007	0.233	[0.266, 1.607]

Model fit: $R^2 = 0.225$, Adjusted $R^2 = 0.176$, $F(7, 112) = 4.64$, $p < 1.359e-04$

Note. Multiple regression analysis predicting Listening performance from competence (GJTs) and learner factors (Anxiety, Error Identification, Years of Study, Exposure).

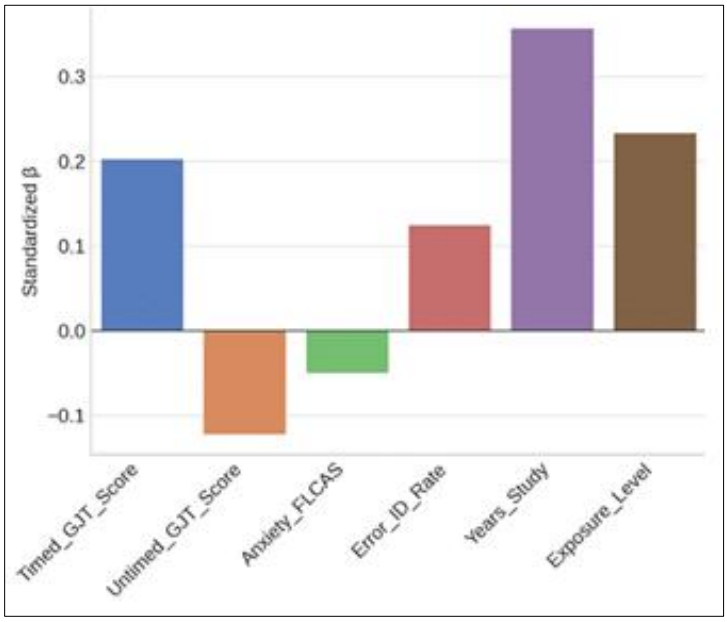

Fig 6: Predictors of Listening Performance (Standardized Betas)
Note. Bar plot displaying standardized beta weights (β) from the regression model predicting Listening performance. Error Identification Rate and Years of Study emerged as the strongest positive predictors, while Anxiety showed a negative effect. Timed GJT contributed modestly, and Exposure had a supportive role.

Discussion

The paired samples t-test revealed a statistically significant difference between the Untimed and Timed GJT scores (see Table 2 and Table 3). Learners achieved higher results in the untimed condition (M = 30.18, SD = 3.82) compared to the timed condition (M = 25.06, SD = 4.22), $t(119) = 28.71$, $p < .001$. This finding supports the notion that while learners possess explicit grammatical competence, their ability to access and apply this knowledge declines under time pressure, reflecting performance constraints rather than a lack of competence. These results align with Chomsky's competence–performance distinction (Chomsky, 1965), and reinforce recent empirical studies that emphasize the role of processing limitations and cognitive load in shaping EFL learners' real-time performance (Ellis, 2005; Schurz, 2024; Ishihara & Hamada, 2025) [7, 21, 12]. From a pedagogical perspective, the results suggest that teaching strategies should incorporate activities that foster automaticity and reduce the impact of performance anxiety, such as timed speaking tasks, interactive listening drills, and authentic communicative practice. In the other hand the correlation analysis examined the relationships between grammatical competence (as measured by Untimed and Timed GJTs) and language performance (speaking and listening) (see Table 4). No significant associations were observed, as further illustrated in Figure 2 and Figure 3. The results revealed that the correlations were weak and statistically non-significant across all pairs. For instance, the correlation between Untimed GJT and Speaking was $r = 0.011$, $p = 0.906$, while Timed GJT and Speaking showed $r = 0.067$, $p = 0.467$. Similar weak associations were found for Listening outcomes. These results suggest that grammatical knowledge tests, whether timed or untimed, may not directly predict learners' real-time oral and aural performance. As illustrated in Figure 4, a weak positive trend was observed between error identification rate and listening performance, although this was not statistically significant. This finding aligns with research highlighting the complex mediating role of cognitive and affective factors, such as working memory and anxiety, in shaping the competence–performance gap (Ellis, 2005; Krashen, 1982; (Yalew *et al*., 2025) [14, 7, 25]. Pedagogically, these results underscore the importance of integrating communicative practice and task-based instruction to bridge the gap between knowledge and use. Also, two multiple regression models were estimated to predict speaking and listening performance from grammatical competence (timed and untimed GJTs) and learner factors (anxiety, error identification, years of study, exposure) (see Table 5 and Table 6). In the speaking model, Timed GJT and Years of Study emerged as meaningful positive predictors, whereas Anxiety showed a negative association with speaking performance. Exposure to English also had a positive effect. In the listening model, Error Identification rate and Years of Study were positively associated with performance, with Timed GJT showing a smaller but positive contribution, and Anxiety exhibiting a mild negative association. These patterns suggest that while declarative knowledge captured by untimed tests adds limited explanatory power, automatized competence under time pressure, accumulated experience, and attentional control are more directly implicated in real-time performance. Pedagogically, this recommends emphasizing practice conditions that foster automaticity and managing anxiety while building sustained exposure to English.

Limitations and Delimitations

Like any empirical investigation, this study is subject to certain limitations. First, the participants were restricted to 120 undergraduate students at the intermediate level, which may limit the generalizability of the findings to other proficiency levels or age groups. Second, the study focused exclusively on speaking and listening performance as indicators of actual language use, while reading and writing were not included in the performance measures. This selective focus may leave out important dimensions of the competence–performance relationship. Third, the use of

grammaticality judgment tests (GJTs) as the primary measure of competence, although widely validated in SLA research, cannot capture all aspects of learners' implicit and explicit grammatical knowledge. Finally, affective and cognitive factors such as working memory and anxiety were considered only in relation to their observable effects on performance, without the use of specialized psychometric instruments, which may have provided a deeper understanding of these influences.

As for delimitations, the study intentionally targeted EFL learners in a university context, with Arabic as their first language, to ensure a relatively homogeneous sample and to allow for more controlled comparisons between competence and performance. The choice to focus on oral and aural performance was deliberate, as these are the skills in which the competence–performance gap is often most visible. Similarly, the use of timed and untimed GJTs, structured speaking tasks, and listening passages was a methodological decision aimed at balancing practicality with validity, even though it excluded other potential instruments. These delimitations were necessary to maintain the scope and feasibility of the study while still providing meaningful insights into the competence–performance distinction.

Conclusion and Pedagogical Implications

Conclusion

This study set out to explore the competence–performance gap among EFL learners through the lens of Chomsky's (1965) distinction between linguistic knowledge and actual language use. By combining timed and untimed grammaticality judgment tests with speaking and listening tasks, the research design sought to provide a comprehensive picture of how learners' explicit and implicit competence translates or fails to translate into real-time oral and aural performance. The inclusion of error analysis further allowed for the classification of mistakes as competence-related or performance-induced, offering nuanced insights into the sources of learner difficulties.

Although the findings are yet to be fully analyzed, it is anticipated that the study will confirm the existence of a measurable gap, where learners demonstrate broader competence on judgment tasks than what they are able to exhibit in speaking and listening activities. This discrepancy is likely to be influenced not only by linguistic knowledge but also by cognitive and affective factors such as time pressure, working memory, and anxiety. By situating these results within the competence–performance framework, the study aims to contribute both to theoretical debates in SLA and to the practical improvement of EFL pedagogy.

Pedagogical Implications

The outcomes of this study are expected to hold several important implications for teaching and learning. First, the anticipated gap between competence and performance highlights the need for instruction that goes beyond grammar-focused teaching. While explicit instruction can support learners' awareness of rules, it should be complemented with opportunities for real-time language use that foster automaticity and fluency. Second, the results underscore the importance of task design in the EFL classroom. Activities that simulate authentic communication such as structured interviews, pair-based discussions, and listening tasks with embedded inconsistencies can help bridge the gap by pushing learners to activate their competence under communicative pressure. Third, the recognition of cognitive and affective constraints suggests that teachers should provide supportive learning environments. Strategies such as reducing performance anxiety, scaffolding tasks, and gradually increasing time constraints can make it easier for learners to transfer their competence into effective performance.

Finally, the study offers implications for assessment practices. Reliance on written grammar tests alone may overestimate learners' competence and underestimate their performance challenges. Incorporating both competence-oriented and performance-based assessments provides a more accurate representation of learners' abilities and can inform targeted interventions.

Declaration of Interest

The author declares that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding Statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Aldubai N, Almehdi K. Collocation competence of the EFL learners and its relationship to their English language proficiency [Preprint]. ResearchGate. 2024 Dec. https://doi.org/10.13140/RG.2.2.23240.02562
2. Al-Ibadi QH. A pragmatic competence, language proficiency and writing performance of Iraqi EFL university students: a correlational study [Internet]. ResearchGate. [date unknown] [cited 2025 Apr 12]. Available from: https://www.researchgate.net/publication/366569741
3. Boualli H, Boudiaf A. Implementing differentiated instruction strategies to develop EFL learners' intercultural competence: voices from an Algerian secondary school context. J Lang Transl. 2025;5(2):126-37.
4. Calafato R, Hunstadbråten S. Literature in language education: exploring EFL learners' literary competence profiles. Engl Teach Learn. 2024. doi: 10.1007/s42321-024-00193-w
5. Chomsky N. Aspects of the theory of syntax. Cambridge (MA): MIT Press; 1965.
6. Corder SP. Error analysis and interlanguage. Oxford: Oxford University Press; 1981.
7. Ellis R. Measuring implicit and explicit knowledge of a second language: a psychometric study. Stud Second Lang Acquis. 2005;27(2):141-72.
8. Fukushima K. Competence and performance revisited: the implications of social role terms in Japanese [Internet]. [place unknown]: Elsevier; [date unknown] [cited 2025 Apr 12]. Available from: www.elsevier.com/locate/pragma
9. Gawin C. Relational schemata in BERT are inducible, not emergent: a study of performance vs. competence in language models. arXiv [Preprint]. 2025:arXiv:2506.11485. Available from:

http://arxiv.org/abs/2506.11485
10. Godfroid A, Loewen S, Jung S, Park JH, Gass S, Ellis R. Timed and untimed grammaticality judgments measure distinct types of knowledge: evidence from eye-movement patterns. Stud Second Lang Acquis. 2015;37(2):269-97.
11. Horwitz EK, Horwitz MB, Cope J. Foreign language classroom anxiety. Mod Lang J. 1986;70(2):125-32.
12. Ishihara T, Hamada A. Evaluating the validity of web-based reaction-time tasks for assessing L2 grammatical knowledge in young learners. Res Methods Appl Linguist. 2025;4(3):100228.
13. Kormos J. Speech production and second language acquisition. New York: Routledge; 2014.
14. Krashen SD. Principles and practice in second language acquisition. Oxford: Pergamon Press; 1982.
15. Luo Z, Xiong Y. Is anxiety always harmful? An exploration of the impact of language anxiety on EFL learning motivation, EFL competence and EFL academic performance. BMC Psychol. 2025;13(1):3017.
16. MacIntyre PD, Gardner RC. The subtle effects of language anxiety on cognitive processing in the second language. Lang Learn. 1994;44(2):283-305.
17. Mpofu N. Language teaching in content subjects: post initial teacher education narratives of South African beginner teachers. TESOL J. 2025;16(4):e70070.
18. Pandit BR. Mother tongue based instruction for developing communicative competence of EFL learners: teachers' experiences. Engl Lang Teach Perspect. 2025;10(1-2):73-82.
19. Pardede A, Dirgantari AS, AS HC, Susantiningdyah H. Exploring request strategies by Indonesian EFL learners. J Appl Linguist. 2025;5(1):65-74.
20. Riyadh E, Khalil HU, Elaf A, Khalil R. Volume (62) Issue (4) Appendix (1) Year (2024). Alustath J Hum Soc Sci. 2024;62(4 Appendix 1).
21. Schurz A. Measuring teenage learners' automatized, explicit, and/or implicit grammatical knowledge. Lang Learn. 2024;74(4):1120-48.
22. Segalowitz N. Cognitive bases of second language fluency. New York: Routledge; 2010.
23. Suman MD. Significance of pragmatics in second language acquisition: enhancing communication and intercultural competence. Inverge J Soc Sci. 2023;2(2):[page range unknown].
24. Vafaee P, Suzuki Y, Kachisnke I. Validating grammaticality judgment tests. Stud Second Lang Acquis. 2017;39(1):59-95.
25. Yalew MY, Amogne D, Chanie BS. Effects of pragmatics awareness instruction (PAI) in enhancing EFL learners' communicative competence. Cogent Educ. 2025;12(1):2528305.
26. Zhou F, Li F. Exploring academic discursive competence in Chinese EFL undergraduates' abstract writing of research articles. Innov Appl Linguist. 2025;1(1).