

Research Article:

Linguistic and Affective Barriers to Speaking among Yemeni EFL Second-Year Students: Evidence from a Conflict-Affected University Context

Ameen Ali Mohammed Algamal^{*}, Yosra Alkhajah¹, and Amatalateef Alsaeedi¹

¹Department of English Language, Faculty of Education, Thamar University - Yemen

^{*}Corresponding author: ameenalgalamal@tu.edu.ye

Abstract

Speaking is widely regarded as one of the most cognitively and socially demanding skills for English as a Foreign Language (EFL) learners, yet the linguistic and affective mechanisms constraining it remain unevenly documented, particularly in Arabic-speaking, conflict-affected higher education contexts. This study examined speaking difficulties reported by second-year students in the Department of English, Faculty of Education, Thamar University, Yemen, using an integrated framework of Communicative Competence, Willingness to Communicate, and Affective Filter/Foreign Language Anxiety. A 15-item, five-point Likert questionnaire was administered to 18 female second-year students aged 21–25 through a quantitative descriptive survey design. Data were analyzed descriptively using SPSS. Limited vocabulary, grammatical difficulties, and inaccurate pronunciation emerged as the principal linguistic constraints. An affective barrier was also evident, particularly in spontaneous, unplanned oral responses, despite uniformly high self-reported motivation. The findings suggest that linguistic and affective barriers may operate as partially independent mechanisms rather than as a single underlying deficit. The study contributes context-specific evidence from a war-affected, resource-constrained Arabic-speaking setting, complementing research dominated by East and Southeast Asian contexts. The findings support combining targeted linguistic instruction with structured, low-stakes opportunities for spontaneous speaking practice. Given the small, single-sex, single-institution sample, the findings should be interpreted as theory-generating rather than statistically generalizable.

ARTICLE HISTORY

Submitted: 9 December 2024

Revised: 5 May 2025

Accepted: 21 March 2026

KEYWORDS

Communicative Competence; Foreign Language Anxiety; Speaking Difficulties; Willingness to Communicate; Yemeni EFL Learners

HOW TO CITE:

Algamal, A. A. M., Alkhajah, Y., & Alsaeedi, A. (2026). Linguistic and affective barriers to speaking among Yemeni EFL second-year students: Evidence from a conflict-affected university context. *ETERNAL: English, Teaching, Learning, and Research*, 12(1), 20–35.

<https://journal.uin-alauddin.ac.id/index.php/Eternal/article/view/53134>



INTRODUCTION

English now functions as the primary language of cross-border scholarship, commerce, and digital communication, and speaking it competently is widely treated as a gatekeeping skill for participation in global academic and professional life (Almusharraf, 2024). This demand has reshaped what English language education is expected to deliver. Grammatical accuracy or written competence are no longer

treated as sufficient outcomes on their own; contemporary curricula worldwide increasingly foreground oral communicative competence, placing speaking instruction at the centre of what counts as effective English language education, a shift reflected in the growing volume of technology-mediated pronunciation and speaking interventions being trialled across EFL settings (Tseng, Chen, Wang, Cheng, Yang, & Gao, 2022). Speaking is also the skill EFL learners across very different regions consistently report as the hardest to acquire, since it demands the simultaneous, real-time coordination of lexical retrieval, grammatical encoding, phonological production, and socio-affective self-regulation, all performed under the direct scrutiny of listeners (Horwitz, Horwitz, & Cope, 1986).

This general difficulty is compounded, at the regional level, for learners across the Arabic-speaking Middle East and North Africa (MENA). The phonological and morphosyntactic distance between Arabic and English, combined with historically grammar-translation-oriented curricula, has been repeatedly linked to persistent gaps in oral proficiency across the region (Almusharraf, 2024). Recent large-sample research on Saudi university EFL instruction has documented a systematic gap between how important instructors say pronunciation instruction is and what they are actually trained and resourced to deliver in the classroom (Almusharraf, 2024). Broader research on willingness to communicate and foreign language anxiety, in both Arabic- and Asian EFL contexts, shows that learners' oral participation is shaped as much by affective state as by linguistic knowledge (Bai, 2023; MacIntyre, Clément, Dörnyei, & Noels, 1998).

Yemen presents a particularly acute version of these regional pressures. Its higher education sector has operated for over a decade under armed conflict, which has driven a substantial brain drain of university faculty, degraded institutional infrastructure, and disrupted the continuity of teacher training and curriculum development (Bin-Hady & Al-Tamimi, 2021). English instruction in many Yemeni universities still relies on teacher-centred, grammar-focused methods even where curricular documents nominally endorse communicative approaches. Learners have responded by turning to informal, technology-mediated strategies, including social media, entertainment media, and self-directed digital practice, to compensate for limited classroom speaking opportunities (Bin-Hady & Al-Tamimi, 2021). This is itself a small but telling instance of the broader transformation ETERNAL is concerned with. As formal institutional capacity for spoken-English instruction comes under strain, learners and, informally, teachers improvise alternative pathways to communicative competence. Designing interventions that meaningfully address the underlying difficulties, technological or pedagogical, first requires knowing where those difficulties actually sit.

At the local level, no published study has examined how the linguistic components of speaking (vocabulary, grammar, pronunciation) and the affective components (anxiety, willingness to communicate, motivation) jointly manifest among university-level EFL learners at Tamar University's Faculty of Education, a context combining the general Arabic-English typological distance documented across the MENA region with the specific structural strain of a decade-long conflict. That a study has not previously been conducted in this location is not, by itself, a sufficient scholarly rationale, and it is not the contribution sought here (see also the Conclusion). The genuinely open question is whether the componential structure of speaking difficulty established mainly in other English-medium and foreign-language contexts, where vocabulary, grammar, and pronunciation typically travel together with broader

linguistic challenges and constrain academic language performance jointly (Aizawa, Rose, Thompson, & Curle, 2023; Rehman, Silpachai, Levis, Zhao, & Gutierrez-Osuna, 2022), also holds for adult learners in a typologically distant, conflict-affected Arabic-speaking university setting. As this study will show, it does not hold cleanly: high motivation here coexists with persistent linguistic gaps instead of resolving them.

This study therefore addressed one research question: what are the principal linguistic and affective sources of speaking difficulty reported by Yemeni EFL second-year students at Tamar University, and how do these linguistic and affective components relate to one another? Weighing these components against one another within a single sample, and interpreting the pattern through an integrated theoretical framework, allows the study to pursue three aims: extending the evidence base on EFL speaking difficulty to an under-researched, conflict-affected Arabic-speaking context; testing whether linguistic and affective barriers to speaking function as independent or overlapping constraints in this population; and generating evidence-based, transferable guidance for speaking pedagogy in comparably resource-constrained EFL settings undergoing similar institutional change.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Across the international literature, vocabulary, grammar, and pronunciation recur as the three linguistic pillars of EFL speaking difficulty, though their relative weighting varies by population. Aizawa et al. (2023), studying Japanese university students in an English-medium instruction programme, found that vocabulary, grammar, and pronunciation difficulties did not resolve into a single dominant barrier but instead combined to constrain students' academic language performance, with different components becoming salient depending on the demands of the task at hand. A parallel, more fine-grained picture comes from work on Arabic-speaking learners specifically: Rehman, Silpachai, Levis, Zhao, and Gutierrez-Osuna (2022) used a data-driven analysis of Arabic learners' spoken English to show that segmental pronunciation errors cluster in predictable, L1-specific patterns rather than occurring randomly across sounds, which is a useful corrective to any assumption that pronunciation difficulty is simply diffuse or unstructured. Taken together, this work reinforces the view that vocabulary, grammar, and pronunciation are interlocking, not separable, since a deficit in any one constrains performance in the others under the real-time demands of speech production. Comparable patterns turn up outside this context too: Koizumi's often-cited work on the vocabulary–speaking relationship, and subsequent replications, indicate that productive vocabulary knowledge is a strong statistical predictor of EFL oral performance across typologically diverse learner populations. That makes it unlikely that the vocabulary-speaking link identified in this study's own data is simply a context-specific artefact, though it does not by itself tell us whether Yemeni learners' errors follow the same L1-specific clustering Rehman et al. documented for other Arabic speakers, a question this study's instrument was not designed to answer.

Pronunciation research has undergone a specific theoretical shift that matters here. The dominant contemporary goal is no longer native-like accuracy but intelligibility and comprehensibility, a reframing visible in a recent meta-analysis of computer-assisted pronunciation training, which found that training effectiveness depended heavily on educational level and delivery setting rather than on any single technology working uniformly well across contexts (Almusharraf, N., 2024). This complicates how difficulty reports should be read: a learner may rate their own

pronunciation as inadequate against a native-speaker standard they were never realistically expected to meet. At the same time, large-sample research on Arabic-speaking instructors' own pronunciation practices in Saudi Arabia found that most instructors, despite recognizing pronunciation's importance, lacked specific training in teaching it and devoted little classroom time to segmental or suprasegmental instruction (Almusharraf, A., 2024). Related experimental work on Saudi EFL learners' production of English diphthongs found that variability in phonetic training modality affected how reliably learners produced these sounds, underscoring that pronunciation difficulty is shaped as much by the input and training learners receive as by any fixed first-language constraint (Almusharraf, Aljasser, Mahdi, Al-Nofaie, & Ghobain, 2024). This supply-side explanation for learner-reported pronunciation difficulty has rarely been considered in the Yemeni literature.

Grammar difficulties are also frequently linked to instructional context, not just aptitude. A recent structural equation modelling study of Indonesian university EFL students found that a supportive academic environment, including institutional resources and classroom practice opportunities, significantly predicted grammar mastery, which in turn was the strongest predictor of English-speaking proficiency among the factors tested (Rahmawati, Herlina, & Lustyantje, 2025). Yemeni English instruction has historically emphasized grammar memorization for examinations over communicative use, a pattern that fits reports that Yemeni classrooms remain teacher-centred and grammar-focused even where curricula nominally prescribe communicative methods (Bin-Hady & Al-Tamimi, 2021). Two limitations in this linguistic-factors literature deserve flagging. It is overwhelmingly drawn from Indonesian contexts, leaving open whether the same componential structure holds for adult Arabic-speaking university learners. And it rarely specifies whether reported difficulty reflects a genuine competence gap or a performance constraint under the specific pressure of spontaneous, real-time speech, a distinction this study's results (reported below) suggest is analytically important.

Beyond these linguistic factors, a substantial and theoretically mature literature treats EFL speaking difficulty as partly, and sometimes primarily, an affective phenomenon. Horwitz, Horwitz, and Cope's (1986) foundational construct of foreign language classroom anxiety established that a learner's linguistic knowledge and their capacity to deploy it under communicative pressure are empirically and theoretically separable. Krashen (1982) later formalized this distinction in the Affective Filter Hypothesis, which proposes that anxiety, low motivation, and low self-confidence raise a psychological barrier to language production independent of underlying competence. Subsequent research operationalized the behavioural consequence of this filter through the construct of willingness to communicate (WTC). MacIntyre, Clément, Dörnyei, and Noels (1998) modelled L2 WTC as a situational, rapidly fluctuating readiness to initiate discourse at a particular moment in time, an idea MacIntyre (2007) later extended into an account of communication as a volitional process shaped jointly by trait-level anxiety and situational confidence.

Recent large-sample quantitative studies back up the practical relevance of this distinction. Bai (2023), studying 662 Chinese EFL learners, found that foreign language anxiety and foreign language boredom each independently and significantly predicted lower willingness to communicate, net of learners' underlying language ability. Almusharraf and Bailey (2023) similarly showed, across a large Saudi sample, that self-reported proficiency, standardized test scores, and foreign language anxiety were only

partially overlapping predictors of learners’ communicative confidence: anxiety is not simply a stand-in for low competence. Yemeni-specific research on this affective strand remains comparatively thin. Existing studies have documented Yemeni learners’ identity-related ambivalence about native-like pronunciation and its motivational correlates, but none has directly modelled how anxiety and willingness to communicate relate to the specific linguistic components (vocabulary, grammar, pronunciation) that learners also report as difficult.

Two points of theoretical tension matter for interpreting this study’s results. First, a strong reading of the Affective Filter Hypothesis would predict that high motivation should lower anxiety and, in turn, raise willingness to communicate; but Bai’s (2023) finding that anxiety and boredom independently suppress WTC even among motivated learners suggests the relationship is looser than that. Second, most prior studies treat WTC as a single global disposition, while MacIntyre’s (2007) volitional-process account implies that willingness can vary sharply by communicative situation: planned classroom participation is not the same thing as unplanned spontaneous response. The present instrument was able to probe this distinction directly, and it proves central to interpreting the findings in the Discussion, below.

Bringing these linguistic and affective strands together, this study integrates three complementary frameworks into a single analytical lens, summarized in Table 1 alongside the questionnaire items each component informed. Communicative Competence theory (Canale & Swain, 1980; Hymes, 1972) supplies the grammatical and phonological competence dimensions against which vocabulary, grammar, and pronunciation difficulties are interpreted. The Willingness to Communicate model (MacIntyre et al., 1998; MacIntyre, 2007) supplies a situationally sensitive account of when learners choose, or fail to choose, to speak, letting the study distinguish difficulty in planned classroom participation from difficulty in spontaneous, unplanned response. The Affective Filter Hypothesis and its empirical descendant, foreign language anxiety (Krashen, 1982; Horwitz et al., 1986), supplies the mechanism by which fear of negative evaluation and low confidence are theorized to suppress production independently of competence.

Table 1. *Integrated theoretical framework mapped to questionnaire items*

Framework component	Core construct measured	Corresponding questionnaire items
Communicative competence – grammatical competence (Canale & Swain, 1980)	Vocabulary range and grammatical accuracy in spoken production	Items 1, 3
Communicative competence – phonological control	Pronunciation accuracy and intelligibility	Item 2
Willingness to communicate (MacIntyre et al., 1998; MacIntyre, 2007)	Readiness to initiate spontaneous, unplanned oral communication	Items 6, 7, 8
Foreign language anxiety / affective filter (Krashen, 1982; Horwitz et al., 1986)	Fear of negative evaluation, shame, and confidence while speaking	Items 4, 5, 9
Motivational orientation (integrative/instrumental)	Desire to learn and independently practise spoken English	Items 10, 11, 12, 13, 14, 15

Used together, these frameworks generate a testable expectation that sets this study apart from most prior descriptive surveys in the Yemeni and Indonesian

literatures. If linguistic competence and affective state are genuinely separable mechanisms, as Horwitz et al. (1986) and MacIntyre (2007) theorize, then a cohort could plausibly report high motivation and low general shame (an affectively favourable profile) while still reporting substantial vocabulary, grammar, and pronunciation difficulty (a competence-based constraint), and while separately reporting acute difficulty specifically with spontaneous, unplanned response (a situational WTC constraint distinct from general classroom anxiety). The findings reported below test for exactly this three-way pattern, not a single undifferentiated “speaking is hard” finding.

Three distinct gaps in the existing literature motivate this study, and they deserve to be stated plainly rather than left implicit in the review above. The first is empirical and geographic, though geography alone is not sufficient justification (noted above): the componential structure of speaking difficulty, meaning which linguistic and affective factors dominate and how they relate to one another, has been established mainly among English-medium-instruction contexts elsewhere in Asia and among Arabic speakers studied outside the Gulf and Levant (Aizawa et al., 2023; Rehman et al., 2022) and, more recently, large Chinese and Saudi university samples (Almusharraf & Bailey, 2023; Bai, 2023). Its applicability to adult, Arabic-speaking, conflict-affected university learners has not been tested. The second gap is theoretical: the reviewed literature rarely tests whether willingness to communicate operates as a single global disposition or, as MacIntyre’s (2007) volitional-process account implies, varies by communicative situation independently of general anxiety, a distinction this study’s instrument was designed to probe directly. The third gap is explanatory: prior Yemeni-specific research has documented technology-mediated coping strategies (Bin-Hady & Al-Tamimi, 2021), without consistently connecting learner-reported difficulty to the institutional and instructional causes documented in adjacent regional literature, such as conflict-driven faculty attrition and limited instructor pronunciation training (Almusharraf, A., 2024). This study addresses the first gap empirically, the second theoretically, and the third interpretively, while treating all three, as elaborated in the Discussion, as hypotheses for future confirmatory research and not as settled conclusions.

RESEARCH METHOD

This study employed a quantitative, descriptive survey design to document the prevalence and patterning of theoretically pre-specified categories of speaking difficulty (linguistic and affective, per Table 1) within a defined population. A descriptive design fit the study’s aim, which was confirmatory and structural: testing whether an existing theoretical distinction, linguistic competence versus affective/situational willingness, holds in a new population. Exploratory theory-generation would have needed qualitative or mixed-methods data instead. No interview, observation, or open-response data were collected; the analysis is accordingly quantitative and descriptive, and the study does not claim mixed-methods status.

The study was conducted in the Department of English, Faculty of Education, Tamar University, Yemen, during the 2023/2024 academic year. The target population comprised second-year students in the Department. A convenience sample of 18 female students, aged 21-25, was recruited from the available second-year cohort. This sampling approach and sample size are modest next to recent comparable survey research in the region: Almusharraf’s (2024) study of Saudi EFL instructors sampled 163

participants, and a comparable Indonesian speaking-anxiety survey sampled 110 students. That comparison is used deliberately here, as a benchmark for reading the present study’s limitations (discussed further in the Conclusion), not as a claim of equivalence.

Data were collected using a structured questionnaire of 15 close-ended statements rated on a five-point Likert scale (strongly agree, agree, neutral, disagree, strongly disagree). Matching the framework in Table 1, three items addressed grammatical/phonological competence (vocabulary, grammar, pronunciation), three addressed situational willingness to communicate (classroom enjoyment, first-language use, spontaneous response), three addressed foreign-language anxiety and confidence (public-speaking confidence, fear of ridicule, shame), and six addressed motivational orientation and independent practice. Recent instrument-development practice in this field sets a higher bar than this study met: Almusharraf (2024), for instance, reports content validity indices from expert review and Cronbach’s alpha reliability coefficients ($\alpha = .74-.75$) for comparable constructs. The present instrument went through no equivalent pilot test, expert content-validity review, or reliability analysis before this administration. This is acknowledged here as a methodological limitation, not glossed over, and it limits how confidently the item groupings in Table 1 can be treated as validated subscales, as opposed to theoretically motivated groupings of individual items.

The questionnaire was administered in paper form to the full sample during regular class time. Students were given sufficient time to respond, and the researchers stayed present to clarify procedural questions and ensure complete responses. Participation was voluntary, and students were informed of the study’s purpose beforehand. Completed questionnaires were coded and analyzed using SPSS. Responses to each item were summarized as percentage frequency distributions across the five response categories. Given the descriptive design and small sample size, no inferential statistical tests were performed. Results are reported and interpreted as descriptive frequencies, organized according to the theoretical groupings in Table 1 instead of raw item order, to keep the presentation analytically focused. On ethical procedures: participation was voluntary, and students were informed of the study’s purpose before data collection. Responses were treated confidentially and reported only in aggregate form (see the Ethical Statement And Informed Consent section).

FINDINGS

Table 2 presents the percentage distribution of responses across all 15 items. The narrative below is organized by theoretical construct (Table 1), not by item order, and reports the patterns most relevant to the research question without repeating every cell of Table 2 in prose.

Table 2. *Second-year students’ responses to the speaking-difficulties questionnaire*

No.	Statement	SA	A	N	D	SD	Total
1	I master a lot of vocabulary in English.	5.6%	50.0%	27.8%	16.7%	0.0%	100%
2	I can pronounce words in English well.	5.6%	38.9%	38.9%	16.7%	0.0%	100%
3	When speaking English, I am capable of constructing sentences correctly.	0.0%	33.3%	33.3%	33.3%	0.0%	100%

4	Speaking the language in front of others in the auditorium gives me a sense of confidence.	16.7%	38.9%	27.8%	11.1%	5.6%	100%
5	I fear that people will laugh at me if I make mistakes while speaking English.	16.7%	33.3%	0.0%	38.9%	11.1%	100%
6	I enjoy speaking English and participating in the classroom.	44.4%	5.6%	38.9%	11.1%	0.0%	100%
7	When I am in class, I prefer to use Arabic rather than English.	16.7%	11.1%	22.2%	44.4%	5.6%	100%
8	I am capable of responding spontaneously to the teacher's questions in English.	11.1%	11.1%	50.0%	27.8%	0.0%	100%
9	I am ashamed of my ability to speak English.	11.1%	11.1%	5.6%	33.3%	38.9%	100%
10	My motivation is to learn how to speak English.	72.2%	27.8%	0.0%	0.0%	0.0%	100%
11	I enjoy watching English movies because I want to speak English.	55.6%	16.7%	27.8%	0.0%	0.0%	100%
12	I need to be able to speak English so I can continue my studies abroad.	22.2%	33.3%	22.2%	11.1%	11.1%	100%
13	I enjoy singing English songs because I want to improve my English speaking skills.	33.3%	22.2%	33.3%	5.6%	5.6%	100%
14	I have a habit of speaking English at home.	16.7%	16.7%	11.1%	27.8%	27.8%	100%
15	I practice speaking English while I am alone.	44.4%	16.7%	22.2%	11.1%	5.6%	100%

Note. SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree.

Turning first to grammatical and phonological competence, vocabulary was the domain in which students reported the greatest self-assessed adequacy: 55.6% agreed or strongly agreed that they command sufficient English vocabulary (Item 1), though 27.8% remained neutral, so this area is comparatively favourable but not fully resolved. Pronunciation (Item 2) showed a less favourable and more polarized pattern, with agreement and neutrality tied at 38.9% each and 16.7% disagreeing. Grammar and sentence-level accuracy (Item 3) was the most evenly, and unfavourably, distributed of the three competence items: responses split into exact thirds (33.3% agree, 33.3% neutral, 33.3% disagree), with no student choosing either extreme. Grammatical accuracy in spoken production is a genuinely unresolved difficulty for this cohort as a whole, not a concern confined to a struggling minority.

On willingness to communicate, responses diverged sharply depending on whether the communicative act was planned or spontaneous. For planned classroom participation, students reported a broadly favourable profile: 50.0% enjoyed speaking and participating in class (Item 6), and a plurality (44.4%) disagreed with preferring Arabic over English in class (Item 7), pointing to a general orientation toward English

use during structured activities. For spontaneous, unplanned response, the profile reversed: only 22.2% felt capable of responding spontaneously to a teacher's question in English (Item 8), half (50.0%) were neutral, and 27.8% disagreed. This planned/spontaneous divergence is the clearest structural finding in the affective data, and the Discussion examines it further.

On foreign language anxiety and confidence, confidence in front of an audience (Item 4) was net favourable (55.6% combined agree/strongly agree versus 16.7% disagree/strongly disagree), but fear of ridicule for mistakes (Item 5) was more evenly split, with exactly one-third (33.3%) agreeing they hold this fear against half (50.0%) who disagreed or strongly disagreed. General shame about one's speaking ability (Item 9) was the affective item with the clearest favourable skew: 72.2% disagreed or strongly disagreed with feeling ashamed. Together, these three items suggest that generalized shame is not a dominant barrier for this cohort, but a specific, situational fear, being laughed at for an error, affects a non-trivial third of respondents.

Finally, on motivational orientation and independent practice, motivation-related items were the most uniform in the dataset. All students (100% combined) reported high motivation to learn to speak English (Item 10), and large majorities reported watching English-language films (72.3% combined, Item 11) and practising speaking alone (61.1% combined, Item 15) as self-directed strategies. This fits the technology-mediated informal learning strategies documented among Yemeni undergraduates more broadly (Bin-Hady & Al-Tamimi, 2021). Speaking English at home (Item 14) was the exception, with 55.6% (combined disagree/strongly disagree) reporting they do not do so, which matches reports that spoken English is rarely used domestically in Yemeni households regardless of proficiency.

Overall, the cohort shows uniformly high motivation and generally low generalized shame, alongside three distinguishable constraints: an unresolved grammatical-accuracy gap affecting the cohort as a whole, a pronunciation gap affecting a substantial minority, and a specific, situational willingness-to-communicate deficit concentrated in spontaneous, not planned, speech. This pattern matches the three-way dissociation the integrated framework outlined earlier predicted, rather than showing a single undifferentiated difficulty.

DISCUSSION

The clearest theoretical contribution of these findings concerns the relationship between motivation, competence, and situational willingness to communicate, and it qualifies a simplified reading of Krashen's (1982) Affective Filter Hypothesis directly. A strong version of that hypothesis would predict that this cohort's uniformly high motivation and low generalized shame should translate into correspondingly high willingness to communicate across all speaking situations. The data do not support that prediction. Motivation and low shame coexisted with a sharp, specific deficit in spontaneous response (Item 8) that was not evident in planned classroom participation (Item 6). This fits MacIntyre's (2007) argument that willingness to communicate is a situationally volatile, moment-to-moment volitional process, not a stable trait, and it extends that argument empirically by showing the planned/spontaneous distinction can hold even where global anxiety indicators such as shame and general confidence are favourable. This pattern has not, to this study's knowledge, been previously documented in the Yemeni or wider Arabic-speaking literature. Bai's (2023) finding that anxiety and boredom independently suppress willingness to communicate net of ability helps explain it: what looks limited here is not a general affective disposition but a

situation-specific readiness that ordinary classroom-confidence measures do not capture.

This theoretical claim should be advanced cautiously, not asserted as established. A sample of 18 respondents on an unpiloted, non-validated instrument cannot demonstrate a psychological dissociation on its own; it can only show that responses to differently worded items did not move together. The interpretation offered here, that this reflects a genuine planned/spontaneous distinction in willingness to communicate, is the most theoretically motivated reading of the pattern, but it is not the only one, and the study's design cannot adjudicate between it and at least two rival explanations considered next. Describing the dissociation as a finding that qualifies Krashen (1982) is more accurate if it is read as a hypothesis the data are consistent with, not a conclusion the data establish.

The first rival explanation is a competence confound, not an affective one. Item 8 asked whether students can respond spontaneously to a teacher's question in English, which plausibly draws on exactly the grammatical encoding and lexical retrieval skills that Items 1–3 already show to be unresolved for this cohort, especially under real-time pressure. If so, the low spontaneous-response score may simply reflect the same grammar and vocabulary gap documented elsewhere in the data, expressed under a harder performance condition, and not a distinct affective or volitional barrier operating independently of competence. The present data cannot separate a competence-based explanation from an affect-based one for this item, since no independent, competence-only measure of spontaneous production was collected. This is a genuine limitation of a single-instrument descriptive design, not a minor caveat, and it means the three-way dissociation proposed earlier is better described as consistent with the data than confirmed by them.

The second rival explanation concerns the validity of self-report for the competence items themselves. Item 1 asks students to judge their own vocabulary adequacy, and self-assessed vocabulary knowledge is a notoriously unreliable proxy for measured vocabulary size or productive use, particularly among learners with limited exposure to calibrating feedback. That concern is amplified here, since teacher-centred instruction, as noted earlier, affords few opportunities to test spoken vocabulary against real communicative demand. The comparatively favourable vocabulary result (55.6% agreement) may therefore partly reflect miscalibrated self-confidence and not genuine lexical adequacy, a possibility the descriptive design cannot rule out. A performance-based vocabulary measure in future research (discussed further in the Conclusion) could resolve it.

The linguistic findings refine prior descriptive work rather than simply repeating it, though again with an important caveat about the direction of inference. The near-even three-way split on grammatical accuracy (Item 3) fits Aizawa et al.'s (2023) finding that grammar is rarely the single dominant barrier to language performance on its own, and Bin-Hady and Al-Tamimi's (2021) observation that Yemeni English instruction remains teacher-centred and grammar-focused for examination purposes offers a plausible explanation: this cohort may hold adequate declarative grammatical knowledge that has not been proceduralized for real-time spoken use. This competence-versus-performance account is attractive precisely because it points to something instructionally actionable, but that attractiveness is also a reason for caution: it should be labelled for what it is, an inference borrowed from studies of Yemeni classroom practice in general, not a claim this study's own data can directly test,

since no measure of declarative grammatical knowledge, such as a written grammar test, was collected alongside the spoken-difficulty item for comparison.

The pronunciation finding (Item 2) invites a similar, appropriately hedged, supply-side reading. Almusharraf's (2024) large-sample study of Saudi university EFL instructors found that most instructors, despite valuing pronunciation instruction, lacked specific training in teaching it and devoted comparatively little classroom time to segmental and suprasegmental instruction. Extending this explanation to Tamar University is plausible but unverified: Almusharraf's (2024) evidence concerns instructors in a different country, a different institutional tier, and a different sample entirely, and no equivalent data on Yemeni instructors' own pronunciation training were collected here or, as far as this study can establish, exist elsewhere in the published literature on this specific institution. The claim is kept in the discussion because it generates a testable and instructionally consequential hypothesis (discussed further in the Conclusion), not because the present data confirm it. A purely learner-internal account, consistent with Almusharraf et al.'s (2024) finding that training modality and variability shape learners' production of difficult English sounds, fits the same result equally well and cannot be ruled out.

A further limitation bears directly on the affective findings and has not been adequately addressed yet: the sample was entirely female, and gender is not a neutral variable where foreign language speaking anxiety is concerned. Survey research elsewhere has reported measurable gender differences in university students' speaking anxiety (Alamri & Qasem, 2024). The anxiety and willingness-to-communicate profile documented here may therefore be specific to female Yemeni second-year students, not representative of the department's second-year cohort as a whole, which also includes male students. Because no male comparison group was sampled, this study cannot say whether the planned or spontaneous asymmetry central to its theoretical argument would replicate, weaken, or reverse in a mixed-sex sample, and the generalized claims about "Yemeni EFL second-year students" in the Conclusion should be read with this constraint in mind.

The Affective Filter Hypothesis itself deserves scrutiny too, not uncritical use as an explanatory backdrop. Krashen's (1982) original formulation has been criticized in the wider second language acquisition literature for being hard to operationalize and, in its strongest form, effectively unfalsifiable, since almost any suppressed production can be redescribed after the fact as evidence of a raised filter. A recent meta-analysis and systematic review of foreign language anxiety interventions found that even well-designed interventions produced inconsistent reductions in anxiety across studies, with effects varying considerably by intervention type and learner population (Xiong, Zhang, Zhao, Liu, Guan, Sui, Feng, & Lee, 2024), which is a fairly blunt way of saying that the field still does not have a reliable recipe for lowering the filter even when it wants to. Set against that backdrop, this study's affective interpretation of the spontaneous-response gap is best treated as one plausible mechanism among several, sitting alongside the competence-confound explanation above, and not as the demonstrated cause of the pattern.

These findings are relevant to the broader transformation of English language education, although the connection is inferential and should be stated precisely rather than asserted loosely. The transformation visible in this context is not yet happening in formal instruction; grammar-focused, teacher-centred practice persists. It is instead happening informally, as students substitute technology-mediated self-directed practice for classroom speaking opportunities that the conflict-strained institutional

system has been unable to reliably provide (Bin-Hady & Al-Tamimi, 2021). To be explicit: this claim rests on a different published sample of Yemeni undergraduates surveyed elsewhere, not on data collected from the present cohort, and it functions here as contextual corroboration, not direct evidence about these 18 students' own informal learning behaviour beyond what Items 11 and 15 report. With that caveat noted, the present findings show what informal, self-directed adaptation has apparently not resolved: high motivation and independent practice have not closed the grammatical-accuracy and spontaneous-response gaps documented here, plausibly because self-directed exposure to English media and solitary rehearsal do not supply corrective feedback or low-stakes opportunities for unplanned, evaluated speech the way structured classroom interaction could. A genuinely transformative response in this and comparable contexts would need to bring feedback-rich, spontaneous-speech-focused instruction closer to the informal, technology-mediated practice learners are already undertaking on their own, instead of treating institutional reform and informal learner adaptation as separate, unconnected tracks.

CONCLUSION

This study provides a clearer account of speaking difficulties among EFL students by showing that these difficulties involve more than a lack of linguistic knowledge. Students' experiences of speaking English were shaped by the relationship between their language resources, their readiness to communicate in particular situations, and the opportunities available for oral practice. This suggests that speaking difficulty should not be treated as a single or uniform problem. In particular, being motivated to learn English or having generally positive attitudes toward speaking does not necessarily mean that students will feel equally prepared to respond spontaneously in English.

One important implication of the study concerns the distinction between planned and spontaneous speaking. These two forms of communication place different demands on learners. When students have time to prepare, they can organize their ideas, select vocabulary, monitor grammar, and anticipate possible difficulties. Spontaneous speaking provides much less time for these processes. Students must draw on their linguistic knowledge while also responding to the demands of an ongoing interaction. For this reason, willingness to communicate may be better understood in relation to particular speaking situations rather than as a fixed characteristic of an individual learner. This distinction may help explain why students who are motivated and willing to learn English can nevertheless hesitate or experience difficulty during unplanned oral interaction.

The study also indicates that students' speaking difficulties need to be considered within their educational setting. Learners do not develop speaking ability independently of the instruction they receive. Regular opportunities to speak, useful feedback, exposure to spoken English, and teachers' ability to support oral communication all contribute to the development of speaking skills. In contexts where such opportunities are limited, difficulties that appear to be individual learner problems may also reflect limitations in the learning environment. This does not remove the importance of learners' own linguistic development, but it suggests that improving speaking ability requires attention to both learners and the conditions in which they learn.

These observations have several implications for English teaching at the university

level. Speaking instruction should provide students with regular opportunities to produce language without extensive preparation, while still giving them sufficient support to develop accuracy and pronunciation. Short impromptu responses, pair and small-group interaction, role plays, discussion tasks, and other forms of low-stakes speaking practice can give students repeated experience in responding in real time. Such activities should not be used simply to make students speak more often; they should also provide opportunities for feedback that helps students recognize and improve specific linguistic problems.

Teacher development is also important. If students have limited opportunities to practise spoken English, increasing the quality of classroom interaction becomes particularly important. Teachers need practical strategies for integrating speaking into regular instruction and for providing feedback on grammar, pronunciation, vocabulary, and fluency without discouraging participation. Professional development in communicative teaching and pronunciation instruction may therefore complement learner-focused interventions. Similarly, technology can provide additional opportunities for speaking practice through self-recording, guided speaking activities, and other forms of structured practice; mobile-assisted language learning applications, for instance, have been shown to enhance EFL learners' speaking skills alongside foreign language enjoyment and language-specific grit (Hwang, Rahimi, & Fathi, 2024). However, such activities are likely to be more useful when they are connected to clear learning objectives and followed by appropriate feedback rather than used simply as additional language exposure.

The conclusions of this study should be interpreted in light of its methodological limitations. The study involved a small sample from one institution and one cohort, and the participants were all from the same sex. The findings therefore cannot be generalized to Yemeni EFL students more broadly. The study also relied on self-reported responses rather than direct assessment of students' speaking performance. Consequently, the reported difficulties represent students' perceptions and experiences and should not be taken as direct evidence of their actual level of speaking competence. In addition, the instrument was not subjected to extensive pilot testing and reliability analysis, which limits the strength of conclusions that can be drawn from the measured categories. Finally, the cross-sectional descriptive design does not allow causal relationships among linguistic, affective, and situational factors to be established.

These limitations point to several directions for future research. Studies involving students from several Yemeni universities and more diverse participant groups would provide a stronger basis for examining whether similar patterns occur across institutions and populations. Future research should also combine self-report measures with direct speaking assessments. Comparing students' perceptions with their performance in planned and spontaneous speaking tasks could clarify whether reported grammatical and pronunciation difficulties correspond to observable performance problems. Qualitative data from students and teachers could further explain how classroom practices, institutional conditions, and opportunities for oral interaction influence speaking development. Longitudinal studies would also be useful for examining how changes in instruction and opportunities for practice affect students' willingness and ability to communicate over time.

Overall, the study suggests that understanding EFL speaking difficulty requires attention to both the learner and the learning environment. Students may have the motivation to communicate but still need greater linguistic support, more

opportunities for spontaneous interaction, and instructional conditions that allow them to use and develop their English regularly. A more careful distinction between linguistic difficulty and situational difficulty can therefore help teachers and institutions design speaking instruction that responds to students' actual needs. For EFL contexts where opportunities for oral communication are limited, strengthening both students' linguistic resources and the conditions in which they use those resources remains an important direction for teaching and research.

AI USE AND DISCLOSURE

The authors used ChatGPT (OpenAI) during the preparation of this manuscript to assist with language editing, grammar correction, and improvement of clarity and readability. ChatGPT was not used to generate or fabricate research data, conduct statistical analyses, interpret the results, or make substantive research decisions. All AI-assisted text was reviewed, revised, and verified by the authors. The authors remain fully responsible for the accuracy, originality, scholarly interpretation, and integrity of the manuscript and its references.

ETHICAL STATEMENT AND INFORMED CONSENT

The present study was conducted after the approval of the Department of English in the Faculty of Education at Tamar University, Yemen. This educational institution supervised, reviewed, and approved all the research procedures. Also, the study was approved by the Ethical Committee of the Faculty of Education at Tamar University, Yemen (dated 28 July 2024) and in accordance with the Declaration of Helsinki principles for research involving human participants. An informed, oral consent was also obtained from participants prior to data collection. All participants were informed about the purpose of the study and study procedures. They were also informed that the participation in the study is completely voluntary, and they can withdraw at any stage of the study. They were informed that their provided data will remain confidential and will be merely used for research purposes.

CRedit AUTHORSHIP CONTRIBUTION STATEMENT

- **Author 1:** Conceptualization, Project administration, Visualization, Writing – review & final draft, Supervision, Resources, Validation.
- **Author 2 :** Formal analysis, Data curation, Investigation, Formal analysis, Writing – original draft.
- **Author 3:** Resources, Visualization, Writing – original draft, Funding acquisition.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this article.

DATA AVAILABILITY

The data collected in order to support the findings of this study were collected through a questionnaire from 18 Yemeni EFL learners from the Department of English in the Faculty of Education at Tamar University, Yemen. Due to ethical and privacy constraints, these data are not publicly available. However, they may be made available by the corresponding author upon reasonable request and subject to appropriate ethical approval.

REFERENCES

- Aizawa, I., Rose, H., Thompson, G., & Curle, S. (2023). Beyond the threshold: Exploring English language proficiency, linguistic challenges, and academic language skills of Japanese students in an English medium instruction programme. *Language Teaching Research*, 27(4), 837–861. <https://doi.org/10.1177/1362168820965510>
- Alamri, W., & Qasem, F. (2024). Foreign language speaking anxiety in relation to gender: students and teachers' perceptions. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2423440>
- Almusharraf, A. (2024). Pronunciation instruction in the context of world English: Exploring university EFL instructors' perceptions and practices. *Humanities and Social Sciences Communications*, 11, 847. <https://doi.org/10.1057/s41599-024-03365-y>
- Almusharraf, A., Aljasser, A., Mahdi, H. S., Al-Nofaie, H., & Ghobain, E. (2024). Exploring the effects of modality and variability on EFL learners' pronunciation of English diphthongs. *Humanities and Social Sciences Communications*, 11(1), 1–11. <https://doi.org/10.1057/s41599-024-02632-2>
- Almusharraf, N. (2024). The effect of settings, educational level and tools on computer-assisted pronunciation training: A meta-analysis. *Journal of Computer Assisted Learning*, 40(5), 1934–1953. <https://doi.org/10.1111/jcal.12974>
- Almusharraf, N., & Bailey, D. R. (2023). Students know best: Modeling the influence of self-reported proficiency, TOEIC scores, gender, and study experience on foreign language anxiety. *SAGE Open*, 13(3). <https://doi.org/10.1177/21582440231179929>
- Bai, S. (2023). The predictive effects of foreign language anxiety and boredom on willingness to communicate among Chinese struggling EFL learners. *Heliyon*, 9(9), e19610. <https://doi.org/10.1016/j.heliyon.2023.e19610>
- Bin-Hady, W. R. A., & Al-Tamimi, N. O. M. (2021). The use of technology in informal English language learning: Evidence from Yemeni undergraduate students. *Learning and Teaching in Higher Education: Gulf Perspectives*, 3(2), 115–125. <https://doi.org/10.1108/LTHE-09-2020-0037>
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/1.1.1>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.2307/327317>
- Hwang, G.-J., Rahimi, M., & Fathi, J. (2024). Enhancing EFL learners' speaking skills, foreign language enjoyment, and language-specific grit utilising the affordances of a MALL app: A microgenetic perspective. *Computers & Education*, 214, Article 105015. <https://doi.org/10.1016/j.compedu.2024.105015>
- Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

- MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The Modern Language Journal*, 91(4), 564–576. <https://doi.org/10.1111/j.1540-4781.2007.00623.x>
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545–562. <https://doi.org/10.1111/j.1540-4781.1998.tb05543.x>
- Rahmawati, E., Herlina, H., & Lustyantje, N. (2025). Unleashing English-speaking proficiency: The role of academic environment, technology, and linguistic adaptability evidence from Indonesia. *Social Sciences & Humanities Open*, 12, Article 101821. <https://doi.org/10.1016/j.ssaho.2025.101821>
- Rehman, I., Silpachai, A., Levis, J., Zhao, G., & Gutierrez-Osuna, R. (2022). The English pronunciation of Arabic speakers: A data-driven approach to segmental error identification. *Language Teaching Research*, 26(6), 1055–1081. <https://doi.org/10.1177/1362168820931888>
- Tseng, W.-T., Chen, S., Wang, S.-P., Cheng, H.-F., Yang, P.-S., & Gao, X. A. (2022). The effects of MALL on L2 pronunciation learning: A meta-analysis. *Journal of Educational Computing Research*, 60(5), 1220–1252. <https://doi.org/10.1177/07356331211058662>
- Xiong, Y., Zhang, Q., Zhao, L., Liu, S., Guan, H., Sui, Y., Feng, J., & Lee, K. M. R. (2024). A meta-analysis and systematic review of foreign language anxiety interventions among students. *Journal of Language and Social Psychology*. <https://doi.org/10.1177/0261927X241291258>