

Stalled at A2 Low: Speaking Attainment and Teacher-Perceived Barriers in the CEFR-Aligned Malaysian Primary ESL Classroom

Mahaletchumy Ponnusamy, Khairunnisa Sharom*, Nur Salwa Abdul Wahid, Evarina Amiron

Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.1026EDU0547>

Received: 16 August 2026; Accepted: 21 August 2026; Published: 02 September 2026

ABSTRACT

Malaysia's CEFR-aligned curriculum benchmarks Year 5 English to level A2 Low, whose defining achievement in speaking is the move from the isolated utterance to connected, extended talk. Whether pupils reach it, and what prevents those who do not, has been examined only indirectly: Malaysian research has documented speaking barriers extensively and scrutinised CEFR implementation closely, but the two literatures have not been brought into contact, and neither reports curriculum-referenced attainment. This study addresses that gap. Speaking Band records for 245 Year 5 pupils were analysed descriptively alongside a thematic analysis, following Braun and Clarke's six-phase procedure, of open-ended responses from the 19 primary ESL teachers responsible for that cohort. The attainment distribution proved sharply compressed: 69.8% of pupils were placed at Bands 3 and 4, 20.4% at Bands 1 and 2, and only 9.8% at Bands 5 and 6, with three pupils (1.2%) reaching the highest level. Thematic analysis yielded five themes, dominated by affective barriers (47.4% of respondents) and limited lexical resource (42.1%). Integration through the curriculum's own descriptors proved analytically decisive: the Band 4 to 5 boundary is defined not by accuracy, range or pronunciation but by whether a pupil sustains communication at times or most of the time. The cohort has therefore stalled precisely where volitional extended speech becomes the criterion, the capacity that teacher-reported anxiety most directly suppresses. The study proposes that this plateau is maintained by a self-reinforcing cycle linking limited vocabulary, fear of peer evaluation and avoidance of extended talk, and identifies a diagnosis-remedy gap: teachers named affective barriers most often (47.4%) but affective remedies least often (5.3%). While 63.2% conceptualised technology as a source of authentic input, only one identified its capacity to remove the peer audience, an underexploited affordance in the dimension teachers identify as most limiting.

Keywords: CEFR; speaking proficiency; primary ESL; foreign language anxiety; classroom-based assessment

INTRODUCTION

Among the four language skills, speaking occupies a paradoxical position in Malaysian English language education. It is simultaneously the skill most explicitly valued in policy discourse, most closely associated with employability and social mobility, and most reliably reported as underdeveloped in the classroom. The Malaysia Education Blueprint 2013-2025 set the target of producing operationally bilingual school leavers, and the English Language Education Reform Roadmap 2015-2025 restructured the entire English curriculum around internationally benchmarked communicative outcomes (Ministry of Education Malaysia, 2013, 2015). A decade into that reform, teachers continue to report classrooms in which pupils understand a good deal of English but will not, or cannot, speak it.

This gap between receptive competence and productive performance is not merely an instructional inconvenience. Speaking is the skill through which language learning becomes socially visible, and therefore the skill through which learners form judgements about their own competence. A pupil who consistently fails to produce oral English before peers accumulates evidence, however unwarranted, that they are not a language learner. The affective consequences of that accumulation extend well beyond the English lesson.

The adoption of the CEFR gives this problem an unusually precise formulation. Year 5 is benchmarked to A2 Low (Ministry of Education Malaysia, 2020), and what distinguishes A2 from A1 in spoken production is the shift from "simple phrases and sentences" to "a series of phrases and sentences" (Council of Europe, 2001). The Year 5 target is thus, in essence, connected talk. This specificity makes it possible to ask not merely whether pupils are performing poorly but exactly which capacity they have not acquired, and the national Band descriptors, which state in explicit language what separates each level from the next, provide the instrument for doing so.

The Malaysian empirical literature is nonetheless structurally unable to answer that question, for reasons developed in the Literature Review. One body of work documents learner-reported and teacher-reported speaking obstacles without reporting attainment; another scrutinises the fidelity of CEFR implementation without reporting pupil outcomes. Neither reads attainment against perceived barriers, and neither exploits the descriptor scales as an interpretive resource. The consequence is that the field can list many plausible obstacles to speaking without establishing which of them constrains outcomes, and therefore without any principled basis for prioritising intervention.

The present study responds to that gap. It reports the speaking attainment distribution of 245 Year 5 pupils across the six national Bands, alongside a thematic analysis of open-ended responses from the 19 primary ESL teachers who taught and assessed that cohort, concerning the challenges they encounter, the remedies they propose, and the role they envisage for technology. Because the two strands derive from the same schools, they can be integrated more strongly than parallel datasets would permit, though this linkage also introduces a methodological consideration addressed in the Methodology.

Research Objectives

The study is guided by four objectives:

1. To describe the distribution of speaking attainment among Year 5 pupils across the six national Bands and to locate that distribution against the CEFR A2 Low target set for the year group.
2. To identify the principal challenges that primary ESL teachers perceive in the teaching of speaking, together with the remedial strategies they propose.
3. To examine how teachers conceptualise the contribution of technology to speaking development.
4. To integrate the two strands through the curriculum's Band descriptors in order to identify which specific capacity constrains movement into the upper bands.

Research Questions

RQ1. How is speaking attainment distributed across the six Bands among Year 5 pupils, and how does this distribution relate to the CEFR A2 Low target?

RQ2. What challenges do primary ESL teachers perceive in the teaching of speaking, and what solutions do they propose?

RQ3. What affordances do teachers attribute to technology in supporting pupils' speaking performance?

RQ4. What does integrating the two strands through the Band descriptors reveal about the constraint on Year 5 speaking attainment?

Significance of the Study

The contribution is threefold. Empirically, the study supplies a Band attainment profile for a sizeable Year 5 cohort, a granularity largely absent from Malaysian speaking research, which more commonly reports aggregate pass rates or Likert-scale self-assessments. Methodologically, it demonstrates the analytic value of using criterion-referenced descriptors as the integration device in mixed-methods educational research: because the descriptors state what distinguishes each band from the next, a distribution concentrated at a particular band is a statement about a particular capacity, not merely about a quantity of achievement. Practically, it converts a

diffuse teacher wish list for technology into design requirements that can be evaluated in subsequent intervention research.

THE CEFR CONTEXT OF MALAYSIAN PRIMARY ENGLISH

Adoption of the CEFR in Malaysia

The Malaysia Education Blueprint 2013-2025 committed the national system to producing school leavers proficient in both Bahasa Malaysia and English, and identified English language education as a priority area for structural reform (Ministry of Education Malaysia, 2013). The operational response was developed by the English Language Standards and Quality Council and published as English Language Education Reform in Malaysia: The Roadmap 2015-2025 (Ministry of Education Malaysia, 2015). The Roadmap's central instrument was the adoption of the Common European Framework of Reference for Languages (Council of Europe, 2001) as the organising standard for curriculum, materials, teacher education and assessment across the whole system.

The choice was consequential in three respects. First, it replaced a locally defined and largely norm-referenced conception of achievement with an internationally benchmarked, criterion-referenced one, so that a Malaysian pupil's attainment became expressible in the same terms as a learner anywhere else. Second, because the CEFR is organised around functional can-do statements rather than around grammatical inventories, it shifted the object of assessment from knowledge about English towards demonstrable performance in it. Third, and most relevant here, the CEFR treats spoken interaction and spoken production as distinct competences with separate descriptor scales, which obliges a CEFR-aligned curriculum to specify and assess speaking explicitly rather than to fold it into general language achievement.

The Roadmap set stage-level targets, with primary education as a whole benchmarked to A2 and secondary education to B1 and B2 (Ministry of Education Malaysia, 2015). The finer year-by-year calibration was subsequently specified in the revised primary curriculum documents. Implementation proceeded through the Standard Curriculum for Primary Schools (Revised 2017) and its accompanying Curriculum and Assessment Standard Documents, and was supported by CEFR-aligned commercial textbook series licensed from international publishers.

Adoption has not been uncontested. A substantial critical literature questions the transferability of a framework developed for adult learners in European multilingual settings to a Malaysian primary population learning English in a predominantly non-English-speaking environment, and questions whether the imported materials that accompany the framework carry cultural content appropriate to Malaysian classrooms (Jamain & Jamaludin, 2023). These reservations form part of the context in which the attainment data reported below should be read.

The Year 5 Speaking Standard

The Curriculum and Assessment Standard Document for Year 5 English states that the curriculum standards for Year 5 are aligned to CEFR proficiency level A2 Low (Ministry of Education Malaysia, 2020). The document specifies a graduated progression across the primary years, reproduced in Table I.

TABLE I. CEFR Curriculum Target Levels, Malaysian Primary English

Year	CEFR target	Year	CEFR target
Year 1	Working towards A1	Year 4	A1 High
Year 2	A1 Low	Year 5	A2 Low
Year 3	A1 Mid	Year 6	A2 Mid

Note. Adapted from Table 1, Curriculum Target Levels based on the CEFR (Ministry of Education Malaysia, 2020, p. 2).

Year 5 therefore represents the transition point at which pupils are expected to cross from the A1 band into A2.

At A2, the CEFR self-assessment grid describes spoken interaction as the capacity to "communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities" and to "handle very short social exchanges," and spoken production as the capacity to "use a series of phrases and sentences to describe in simple terms" familiar subject matter (Council of Europe, 2001, Table 2). The revised descriptor scales of the Companion Volume retain this general profile while elaborating mediation and online interaction (Council of Europe, 2020).

The distinction the CEFR draws between A1 and A2 spoken production is important for the present analysis. At A1 the learner produces "simple phrases and sentences"; at A2 the learner produces "a series of phrases and sentences" (Council of Europe, 2001, Table 2). The A1-to-A2 transition in speaking is thus, in essence, a transition from the isolated utterance to connected, extended talk. This is precisely the transition that Year 5 is designed to accomplish.

The Year 5 curriculum operationalises this through three Content Standards for speaking, divided between spoken interaction and spoken production, as set out in Table II.

TABLE II. Year 5 Speaking Content and Learning Standards

Domain	Content Standard	Learning Standards
Spoken interaction	2.1 Communicate simple information intelligibly	2.1.1 Give detailed information about themselves; 2.1.2 Find out about and describe experiences up to now; 2.1.3 Ask for, give and respond to simple advice; 2.1.4 Ask about and describe future plans; 2.1.5 Describe people, places and objects using suitable statements
Spoken interaction	2.2 Use appropriate communication strategies	2.2.1 Keep interaction going in short exchanges by asking suitable questions; 2.2.2 Agree a set of basic steps needed to complete short classroom tasks
Spoken production	2.3 Communicate appropriately to a small or large group	2.3.1 Narrate short basic stories and events

Note. Adapted from Ministry of Education Malaysia (2020), Speaking Skills, pp. 25-28.

Two features of this standard warrant emphasis. First, Learning Standard 2.2.1 requires pupils to keep interaction going by asking suitable questions, which places responsibility for sustaining the exchange on the pupil rather than on the teacher. Second, Learning Standard 2.3.1 requires narration, the most extended and least scaffolded of the speaking tasks in the Year 5 syllabus, since a narrative must be organised and delivered without the turn-by-turn support that dialogue supplies. Both standards presuppose a learner willing to hold the floor.

Band Descriptors for Speaking

Attainment against these standards is reported through classroom-based assessment (Pentaksiran Bilik Darjah) on a six-point scale. The curriculum document labels the six points Performance Levels, rendered in Malay as Tahap Penguasaan; in school reporting practice, and throughout this paper, they are referred to as Bands 1 to 6. The descriptors are criterion-referenced and holistic, reflecting the teacher's cumulative judgement of a pupil's oral performance across a reporting period rather than performance on a single timed test. Table III reproduces the Year 5 speaking descriptors.

TABLE III. Year 5 Band Descriptors for Speaking Skills (Bands 1-6)

Level	Descriptors
Band 1	Can communicate simple information about themselves using fixed phrases with a lot of support. Can manage interaction and classroom task by providing short and simple responses with a lot of support. Can describe people, places and objects using fixed phrases with a lot of support. Can narrate short basic stories and events with difficulty even with a lot of support.

Level	Descriptors
Band 2	Can communicate simple information about themselves with a lot of support. Can manage interaction and classroom task appropriately with a lot of support. Can describe people, places and objects using basic statements with a lot of support. Can narrate short basic stories and events with a lot of support.
Band 3	Can communicate simple information about themselves clearly. Can manage interaction and classroom task appropriately. Can describe people, places and things using suitable statements adequately. Can narrate short basic stories and events adequately.
Band 4	Can communicate simple information about themselves clearly by providing some relevant details. Can manage interaction and classroom task appropriately by sustaining communication at times. Can describe people, places and things clearly using suitable statements with very few relevant details. Can narrate short basic stories and events clearly at an appropriate pace.
Band 5	Can communicate simple information about themselves clearly by providing a lot of relevant details. Can manage interaction and classroom task appropriately by sustaining communication most of the time. Can describe people, places and things clearly using suitable statements with some relevant details. Can narrate short basic stories and events with clear diction and articulation.
Band 6	Can communicate simple information about themselves with a lot of relevant details clearly and confidently. Can manage interaction and classroom task appropriately by sustaining communication naturally. Can describe people, places and things creatively using suitable statements. Can narrate short basic stories and events creatively with clear diction and articulation. Can display exemplary model of language use and guide others.

Note. Reproduced from Ministry of Education Malaysia (2020), Performance Level Descriptors for Speaking Skills. The source document labels these Performance Levels; they are referred to here as Bands.

Read as a progression, the scale is organised around two thresholds rather than six evenly spaced steps. The first threshold falls between Band 2 and Band 3 and concerns support: every Band 1 and Band 2 descriptor is qualified by the phrase "with a lot of support," and every Band 3 descriptor drops it. Crossing into Band 3 therefore means performing the same functions without scaffolding.

The second threshold falls between Band 4 and Band 5 and concerns sustained talk. The Band 4 descriptor permits a pupil to sustain communication "at times"; Band 5 requires sustaining it "most of the time"; Band 6 requires sustaining it "naturally" and additionally requires the pupil to "display exemplary model of language use and guide others." What separates the intermediate from the excellent bands is therefore not accuracy, range or pronunciation, none of which distinguishes Band 4 from Band 5, but duration and self-direction: the capacity and willingness to keep talking. The descriptor scale, in other words, makes volitional extended speech the criterion for the upper bands. This observation becomes central to the interpretation offered in the Discussion.

The attainment records analysed in this study are Band classifications from this Pentaksiran Bilik Darjah system, assigned against the Year 5 speaking descriptors reproduced above. The terms weak, intermediate and excellent used in those records correspond to the aggregation of Bands 1-2, 3-4 and 5-6 respectively, following common school reporting practice; the Band scale itself is not divided into named categories in the curriculum document.

LITERATURE REVIEW

Theoretical Perspectives on Speaking Development

Theoretical accounts of second language acquisition distinguish between the conditions necessary for comprehension and those necessary for production. Krashen's (1985) input hypothesis holds that acquisition proceeds when learners encounter comprehensible input slightly beyond their current level. Swain's (1985)

output hypothesis, developed partly in response to the observed limits of Canadian immersion programmes in which learners achieved high comprehension but persistently non-target-like production, argues that comprehensible input alone is insufficient. Learners must be pushed to produce language in order to notice gaps in their interlanguage, test hypotheses about form, and shift from the semantic processing that suffices for comprehension to the syntactic processing that production demands. A classroom that supplies input generously while providing few occasions for pushed output will predictably produce learners who understand more than they can say.

From a sociocultural perspective, Vygotsky's (1978) zone of proximal development frames oral competence as emerging through assisted performance, in which a more capable interlocutor supplies scaffolding that is progressively withdrawn. This framing bears directly on the Band scale described above, whose first threshold is defined precisely by the removal of support. It also implies that the quantity of speaking opportunity matters less than its structure: an unscaffolded invitation to speak may simply expose the gap it was meant to close.

On the lexical side, Nation (2013) distinguishes receptive from productive vocabulary knowledge and argues that the latter requires substantially more encounters and, critically, retrieval under production conditions. A learner may recognise a word reliably while remaining unable to retrieve it under the time pressure of live conversation. This distinction anticipates one of the findings reported below, in which teachers describe pupils who understand the lesson but cannot assemble an utterance quickly enough to participate in it.

Affective Dimensions of Oral Performance

Horwitz et al. (1986) established foreign language classroom anxiety as a situation-specific construct distinct from general trait anxiety, comprising communication apprehension, test anxiety and fear of negative evaluation. The third component is particularly consequential in primary settings, where the peer audience is permanent, socially significant and not chosen by the learner.

MacIntyre et al. (1998) subsequently modelled willingness to communicate as the immediate antecedent of actual second language use, positioned at the apex of a pyramid of enduring and situational influences including self-confidence, intergroup attitudes and communicative competence. The model's central implication is that competence and use are separable: raising linguistic competence without addressing the affective layer may leave observable oral behaviour unchanged. Since classroom-based assessment records observable behaviour rather than latent competence, this implication bears directly on how a Band distribution should be interpreted.

The Malaysian evidence on this point is now substantial. Thandavaraj and Mohd-Said (2025) administered an age-adapted Foreign Language Classroom Anxiety Scale to 132 pupils across Years 4, 5 and 6 in Malaysian public primary schools, supplemented by classroom observation and group interviews. They report moderate-to-high anxiety at every year level, with mean scores above 100 and no statistically significant difference between year groups, and identify fear of negative judgement, social comparison and perceived proficiency as the principal contributors. The absence of year-level variation is notable: anxiety does not appear to diminish as pupils progress through the primary years and accumulate instruction.

Empirical Studies of Speaking Difficulties in Malaysia

A considerable body of Malaysian research documents the obstacles learners face in speaking English. Kashinathan and Aziz (2021) synthesised fifteen studies published between 2014 and 2021 and concluded that learners encounter both internal challenges, principally anxiety and low self-confidence, and external ones arising from classroom environment, teaching methodology and limited opportunity for authentic practice. The stability of this inventory across a decade of independent studies, conducted in different states and at different educational levels, suggests a structural rather than incidental condition.

Such studies share a methodological profile that also constitutes their limitation. They typically elicit perceptions, whether from learners through questionnaires or from teachers through interviews, and report those perceptions thematically. They rarely place the reported obstacles alongside any measure of what pupils

actually attain. The result is a well-developed account of what participants believe impedes speaking, unaccompanied by evidence about the shape of the attainment such impediments produce.

Studies of CEFR Implementation and Classroom-Based Assessment

A second and largely separate research strand examines how the CEFR-aligned curriculum has been enacted. Sidhu et al. (2018) investigated CEFR-aligned school-based assessment in Malaysian primary ESL classrooms through surveys, interviews and document analysis across five schools, and found that although teachers viewed school-based assessment favourably, they lacked full understanding of the method and possessed limited knowledge of the revised CEFR-aligned curriculum. Assessment practice fell short of its intended formative purpose, with little constructive feedback and minimal peer or self-assessment, constrained by time, large classes, workload and insufficient training.

Mohamad Uri (2023) surveyed 331 secondary English teachers and interviewed a member of the English Language Standards and Quality Council, reporting that some 41% of respondents found implementation procedures unclear and over 61% acknowledged that inadequate understanding of the framework would impede classroom application, with the burden of designing CEFR-appropriate activities and adapting imported materials cited as major practical obstacles. Jamain and Jamaludin (2023), through document analysis, identified limited teacher awareness, insufficient professional development and a shortage of contextually appropriate CEFR-aligned resources as the framework's principal implementation problems.

This strand has produced a detailed picture of teacher readiness and assessment literacy. Its object of inquiry, however, is the fidelity of implementation rather than its outcome. These studies establish that teachers find the framework difficult to enact; they do not report what pupil attainment against the framework's own descriptors looks like, nor whether the resulting distribution is consistent with the barriers teachers describe.

Technology and Speaking Development

Research on technology-enhanced language learning offers two rationales for its use in speaking instruction. The first is input-based: multimedia resources supply authentic, contextualised models of spoken English, including prosody, register and paralinguistic features that print cannot convey, and Mayer's (2009) work on multimedia learning provides principled grounds for expecting dual-channel presentation to aid retention.

The second rationale is affective. Huang and Liu (2025), synthesising 99 empirical studies published between 2004 and 2024, characterise technology as a double-edged instrument in relation to foreign language anxiety: it reduces anxiety through psychological safety and personalised feedback while potentially increasing it through cognitive overload and diminished human interaction. Asynchronous and non-judgemental practice environments emerged as the most reliably anxiety-reducing applications. The authors note, however, that research has concentrated heavily on higher education, leaving the primary sector comparatively unexamined. This is a substantive gap rather than a bibliographic one, since the mechanism by which technology might assist an eleven-year-old who fears peer ridicule in a fixed classroom group is not self-evidently the mechanism operating for an anxious undergraduate who selected into a language course.

Malaysian primary evidence on this point is, however, beginning to accumulate. Sharom and Kew (2022) reported that a flipped classroom design, in which instructional content was delivered through digital materials before the lesson so that class time could be redirected to production and feedback, significantly improved the English writing performance of Malaysian primary pupils relative to conventional instruction. This finding matters here for two reasons. First, it establishes that a technology-mediated redistribution of instructional time is workable in precisely this population and setting, rather than only in the higher education contexts that dominate the review literature. Second, the mechanism it exploits, moving transmission out of the lesson so that the lesson itself becomes practice time, addresses directly the time constraint that teachers report below as a barrier to speaking instruction. Whether the same design transfers from writing, which is produced privately, to speaking, which is produced before an audience, is an open question, and the affective findings reported below suggest it should not be assumed.

The Research Gap

Table IV summarises the four strands reviewed above against the question each leaves open.

TABLE IV. Research Strands and the Gap Addressed

Strand	Representative studies	What is established	What is not addressed
Speaking difficulties	Kashinathan & Aziz (2021)	A stable inventory of learner and teacher perceived obstacles	Whether these obstacles correspond to any measured attainment pattern
Affective measurement	Thandavaraj & Mohd-Said (2025); Horwitz et al. (1986)	Anxiety is moderate to high and does not decline across primary years	Whether anxiety maps onto curriculum-referenced attainment outcomes
CEFR implementation	Sidhu et al. (2018); Mohamad Uri (2023); Jamain & Jamaludin (2023)	Teachers find the framework difficult to understand and enact	What pupil attainment against CEFR-aligned descriptors actually looks like
Technology and anxiety	Huang & Liu (2025); Mayer (2009)	Technology can reduce anxiety through psychological safety	Whether this holds for primary learners, and whether teachers perceive it

Across these strands, a specific gap emerges. Malaysian research has documented the barriers to speaking extensively and has scrutinised the implementation of the CEFR-aligned curriculum closely, but the two literatures have not been brought into contact. No study located for this review reports a Band distribution for speaking in a defined primary cohort and reads that distribution against the barriers identified by the very teachers who assigned those bands, using the curriculum's own descriptors as the interpretive frame.

The consequence of that separation is not merely a missing publication. Without the attainment distribution, the barrier literature cannot establish which of its many documented obstacles actually constrains outcomes, and consequently offers no basis for prioritising among them. Without the barrier data, the implementation literature can report that teachers struggle to enact the framework but cannot say whether pupil attainment is limited by that difficulty or by conditions upstream of it. And because the CEFR descriptors specify in explicit terms what distinguishes one band of speaking from the next, they offer an interpretive resource that neither literature has exploited: the boundary between two Bands is a statement about exactly which capacity a pupil has not yet acquired.

The present study addresses this gap by reporting the speaking Band distribution of a Year 5 cohort, analysing the barriers perceived by the teaching team responsible for that cohort, and integrating the two through the descriptor language of the curriculum itself.

METHODOLOGY

Research Design

The study adopted a convergent mixed-methods design with linked samples (Creswell & Plano Clark, 2018). Quantitative and qualitative data were collected within the same schools, analysed separately using procedures appropriate to each, and integrated at the interpretation stage against the curriculum's Band descriptors. The 19 teachers who provided the qualitative data were the English teachers responsible for the 245 pupils whose attainment records constitute the quantitative dataset.

This linkage is the design's principal strength and also the source of its principal caution. The strength is that the two strands describe the same instructional setting rather than two settings assumed to be comparable: when these teachers report that pupils will not speak in front of peers, they are reporting on the pupils whose Bands are analysed here, so the integration reported in the Findings is substantive rather than merely thematic. The caution is that the teachers who described the barriers are also the assessors who assigned the bands, which

introduces the possibility of common-rater consistency: a teacher who believes pupils lack confidence may, without any lapse in professional judgement, apply the descriptors in a manner congruent with that belief. The two strands are therefore not fully independent sources of evidence, and the analysis below is framed as a coherent single-source account of one instructional setting rather than as triangulation between independent observers. This consideration is developed further in the Limitations.

The design is described as convergent rather than explanatory sequential because the qualitative instrument was not constructed in response to the attainment distribution; both datasets pertain to the same academic period. Where a study collects attainment data first and then designs the qualitative strand specifically to explain it, the explanatory sequential label is the appropriate one.

Participants and Data Sources

Quantitative strand

The quantitative dataset comprised speaking attainment records for 245 Year 5 pupils (aged approximately 11 years), each classified into one of the six Bands of the Pentaksiran Bilik Darjah classroom-based assessment system on the basis of teacher assessment of oral performance against the Year 5 speaking standards reproduced in Table III. For analytical purposes the six bands were collapsed into the three categories used in school reporting practice: weak (Bands 1 and 2), intermediate (Bands 3 and 4) and excellent (Bands 5 and 6). Records were complete for all 245 pupils, with no missing cases.

Qualitative strand

The qualitative dataset comprised responses from the 19 primary ESL teachers of the cohort to an online questionnaire administered through Google Forms. The instrument collected one categorical item on teaching experience and three open-ended items: the challenges encountered in teaching English speaking classes; suggested solutions to speaking problems among Level 2 primary pupils; and views on how technology could help pupils achieve better speaking performance. Participation was voluntary and responses were anonymous apart from the timestamp, so individual teachers cannot be matched to individual classes; the linkage between strands therefore holds at cohort level rather than at class level. The experience profile of the sample is reported in Table V.

TABLE V. Teaching Experience Profile of Respondents (N = 19)

Teaching experience	n	%
More than 10 years	12	63.2
6-10 years	4	21.1
1-5 years	3	15.8
Total	19	100.0

The sample was heavily weighted towards experienced practitioners, with 84.2% having taught for more than five years. This composition is a strength for the purpose at hand: the accounts reported below reflect sustained observation of successive cohorts rather than early-career impressions, and experienced teachers are better positioned to distinguish recurrent structural barriers from the difficulties of a particular class.

Data Analysis

Quantitative data were analysed descriptively. Frequencies and percentages were computed for each Band and for each aggregated category, with percentages calculated against the total cohort of 245 and rounded to one decimal place.

Qualitative data were analysed using Braun and Clarke's (2006) six-phase thematic analysis. The three open-ended items were coded separately, since they address analytically distinct questions, and the resulting theme sets are reported separately. Phase one involved repeated reading of the full response set to establish familiarity. Phase two generated initial codes inductively at the semantic level, staying close to participants' own formulations. Phase three collated codes into candidate themes; phase four reviewed these against the coded extracts and the full dataset; phase five defined and named the final themes; phase six produced the analytic narrative reported below.

Because individual responses frequently referenced more than one difficulty, coding was not mutually exclusive: a single response could contribute to several themes. Reported frequencies therefore represent the number of respondents who raised a theme, not a partition of the sample, and theme percentages do not sum to 100.

Trustworthiness and Ethical Considerations

Several measures were taken to support the credibility of the qualitative findings, following established criteria for trustworthiness in naturalistic inquiry (Lincoln & Guba, 1985). Verbatim extracts are supplied for each theme so that readers may evaluate the fit between data and interpretation. Theme frequencies are reported transparently, including those with low support, rather than only those that fit the emerging argument, and divergent cases are retained and discussed rather than set aside. Original spelling and punctuation are preserved in all quoted extracts, with clarifications supplied in square brackets where required for intelligibility. Responses are identified by anonymised codes (R1 to R19).

Participation in the survey was voluntary and no personally identifying information was collected. Pupil attainment records were supplied in aggregate form, classified only by Band, with no individual identifiers, and were used solely for the analysis reported here.

FINDINGS AND ANALYSIS

Distribution of Speaking Attainment (RQ1)

Table VI presents the distribution of the 245 Year 5 pupils across the six Bands.

TABLE VI. Speaking Performance by Band (N = 245)

Level	Category	n	%
Band 1	Weak	1	0.4
Band 2	Weak	49	20.0
Band 3	Intermediate	101	41.2
Band 4	Intermediate	70	28.6
Band 5	Excellent	21	8.6
Band 6	Excellent	3	1.2
Total		245	100.0

The distribution is markedly asymmetric. Band 3 is the modal category, accounting for 101 pupils or 41.2% of the cohort, with Band 4 the second largest at 70 pupils (28.6%). Together these adjacent bands account for 171 pupils, almost seven in ten of the entire year group. At the lower extreme, Band 2 accounts for 49 pupils (20.0%) while Band 1 contains a single pupil (0.4%). At the upper extreme, Band 5 accounts for 21 pupils (8.6%) and Band 6 for only three (1.2%).

Table VII presents the same data aggregated into the three reporting categories.

TABLE VII. Speaking Performance by Aggregated Category

Category	Levels	n	%
Weak	1-2	50	20.4
Intermediate	3-4	171	69.8
Excellent	5-6	24	9.8
Total	1-6	245	100.0

Figure 1 presents the same distribution graphically, with the boundary between Bands 4 and 5 marked. The visual effect of the compression is more immediate than the tabulated frequencies convey: the cohort forms a single mass to the left of the threshold and thins to almost nothing on its right.

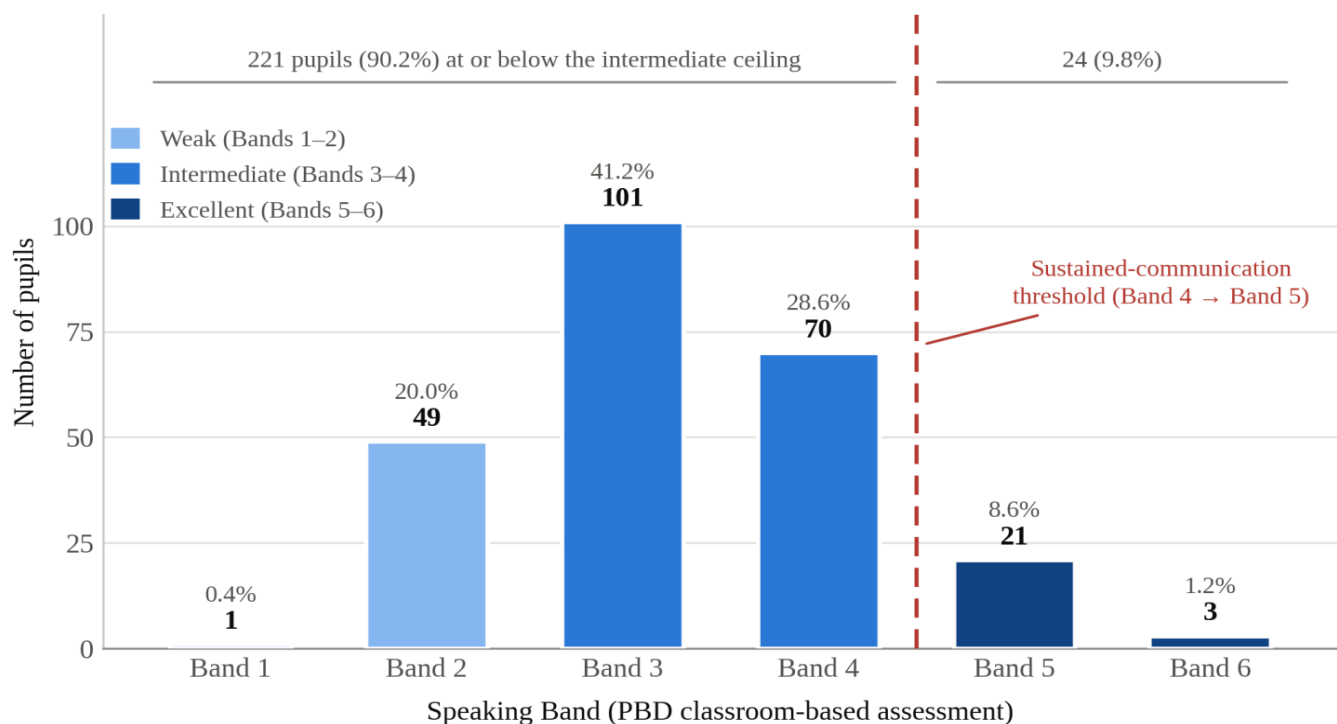


Figure 1. Distribution of Year 5 speaking attainment by Band (N = 245), with the sustained-communication threshold between Band 4 and Band 5 marked. Bar shading indicates the three aggregated reporting categories.

Three features warrant emphasis. First, the cohort is compressed at the intermediate level: 69.8% of pupils occupy the two middle bands. Second, the upper end of the scale is close to empty. Fewer than one pupil in ten reached the excellent category, and Band 6 was attained by three pupils in a cohort of 245. Third, and by contrast, genuine failure is also rare: only one pupil was placed at Band 1. Taken together, 90.2% of pupils sit at or below the intermediate ceiling.

The shape of this distribution is analytically informative. It is not the profile of a cohort that has failed to learn English, which would concentrate at Bands 1 and 2. Nor is it the profile of a cohort progressing evenly through a criterion-referenced scale, which would show fuller representation in the upper bands by the penultimate year of primary schooling. It is instead the profile of a cohort that has reliably acquired a functional floor and then stalled: pupils have accumulated enough oral English to be assessed as satisfactory, and very few have moved beyond it. This pattern is described here as the intermediate plateau.

The plateau read against the descriptors

Because the Bands are criterion-referenced, the location of the plateau can be interpreted directly. The 20.4% of pupils at Bands 1 and 2 are, by descriptor, pupils who require "a lot of support" to perform any speaking function. The 69.8% at Bands 3 and 4 perform those functions without support but, per the Band 4 descriptor,

sustain communication only "at times." The 9.8% at Bands 5 and 6 sustain communication "most of the time" or "naturally."

The plateau therefore falls exactly at the second threshold identified in the CEFR Context section: the boundary between Bands 4 and 5, which is defined by sustained rather than intermittent talk. Nearly seven in ten Year 5 pupils can speak when required, unassisted, but do not keep speaking. Against the CEFR A2 Low target for the year group, which requires "a series of phrases and sentences" rather than isolated ones (Council of Europe, 2001), this is a substantively precise finding: the cohort has largely secured the A1 achievement of the intelligible independent utterance and has largely not secured the A2 achievement of connected, extended talk.

Perceived Challenges in Teaching Speaking (RQ2)

Thematic analysis of teacher responses to the first open-ended item yielded five themes, summarised in Table VIII.

TABLE VIII. Themes in Teacher-Perceived Challenges (N = 19)

Theme	n	%	Sources
Affective barriers: low confidence, reticence, fear of ridicule	9	47.4	R6, R7, R8, R10, R12, R13, R15, R16, R18
Limited lexical resource	8	42.1	R1, R3, R4, R7, R9, R11, R13, R19
Weak foundational proficiency and L1 dependence	3	15.8	R5, R15, R17
Structural and resource constraints	3	15.8	R3, R14, R15
Restricted exposure beyond the classroom	2	10.5	R1, R2

Affective barriers

The most frequently raised theme, cited by nine of nineteen respondents (47.4%), concerned pupils' unwillingness rather than inability to speak. Responses described a spectrum running from hesitancy to outright refusal. At the milder end, R6 observed simply that pupils "need a lot of prompting," while R11 described a processing delay: pupils "take sometime to talk/ think about the words they need to use." At the more severe end, R8 reported that "Students refuse to try conversing in English even simple ones," attributing this to "Lack of confidence," and R16 reduced the problem to three words: "Pupils dont speak."

Two respondents specified the social mechanism at work. R13 listed "Peer mocking" as a distinct challenge alongside low confidence and limited vocabulary, and R18 described pupils as "so timid and afraid n become a laughing stock." These accounts locate the barrier not in the individual learner's disposition but in the classroom as an evaluative arena. What is feared is not the language task but the audience. The finding aligns closely with Thandavaraj and Mohd-Said's (2025) identification of fear of negative judgement and social comparison as leading contributors to speech anxiety in this population.

R15, the most detailed response in the dataset, listed self-confidence alongside "Vocal to speak," a formulation that captures the distinction between having something to say and being willing to voice it. R12 similarly paired the absence of confidence with a lack of "communicative skill to speak on their own," pointing to a dependence on teacher scaffolding that has not been successfully withdrawn.

"The pupils are not confident and lack of communicative skill to speak on their own." (R12, more than 10 years' experience)

Limited lexical resource

Eight respondents (42.1%) identified vocabulary as a primary constraint, making this the second dominant theme and the one most often stated without elaboration, as though self-evident. R9 and R19 each answered the question in three words: "Low vocab" and "Lack of vocab" respectively. R1 paired vocabulary with exposure,

citing "pupils' lack of vocabulary and exposure to english," and R7 paired it with affect, noting that "Students have limited vocabularies and less confidence."

Two responses framed the lexical problem from the instructional rather than the learner side. R4 identified the challenge as "Using vocabularies that suit students' level," locating the difficulty in the teacher's calibration of input, while R11 described the retrieval cost that a thin lexicon imposes in real time. This distinction is consequential and corresponds directly to Nation's (2013) separation of receptive from productive vocabulary knowledge: the first framing treats vocabulary as a stock to be increased, the second as an access problem under production conditions. Both are supported by the data, and they imply different instructional responses.

Weak foundational proficiency and L1 dependence

Three respondents (15.8%) located the difficulty below the level of the English curriculum itself. R5 was the most direct, reporting "Very poor profieciency of the pupils. Quite a number of level 1 pupils can't even read L1 so L2 is worst." This observation reframes the speaking problem as partly a literacy problem, and one not confined to English. R17 noted simply that the "Student doesn't understood," while R15 identified "Obstacles in thinking process (based on mother tongue)."

"Very poor profieciency of the pupils. Quite a number of level 1 pupils can't even read L1 so L2 is worst." (R5, 6-10 years' experience)

Although raised by a minority, this theme deserves weight disproportionate to its frequency, because it identifies a group for whom interventions pitched at the Year 5 speaking standards are unlikely to be effective.

Structural and resource constraints

Three respondents (15.8%) pointed to conditions outside the teacher's control. R14 cited "Time constraints" without elaboration; R15 listed both time constraint and "Disturbance in surrounding" and proposed reducing class size; R3 reported "no ICT facilities in the class." These echo the constraints Sidhu et al. (2018) found limiting CEFR-aligned classroom-based assessment, namely time, large classes and workload, suggesting that the same conditions constrain both the assessment and the teaching of speaking. The infrastructure point is significant for RQ3: a teacher enthusiastic about technology's potential may nevertheless be unable to enact it.

Restricted exposure beyond the classroom

Two respondents (10.5%) explicitly named out-of-school exposure in response to the challenges item. R2 wrote that "PUPILS DIDN'T UNDERSTAND THE LANGUAGE DUE TO LACK OF EXPOSURE AT HOME," and R1 cited lack of exposure alongside vocabulary. The low frequency here is misleading if read in isolation, because exposure emerged as the single most prominent theme in responses to the solutions item. Teachers appear to treat limited exposure less as a classroom challenge than as the condition their remedies must address.

Proposed Solutions (RQ2 continued)

Analysis of the second open-ended item produced the theme distribution shown in Table IX.

TABLE IX. Themes in Teacher-Proposed Solutions (N = 19)

Theme	n	%	Sources
Expanding exposure to English input	6	31.6	R1, R2, R3, R9, R13, R15
Structured and scaffolded oral practice	5	26.3	R4, R10, R11, R12, R18
Policy and structural change	4	21.1	R8, R14, R15, R16
Technology-mediated practice	3	15.8	R3, R5, R19
Increasing opportunity to use English	3	15.8	R6, R9, R15
Affective and motivational support	1	5.3	R7

The most frequently proposed remedy, offered by six respondents (31.6%), was to widen pupils' exposure to English. Proposals ranged from early intervention, with R1 recommending "introduction to english picture story books (physical or digital) in kindergarten to give early exposure of english to the pupils," to institutional partnership, with R2 suggesting teachers "Work hand-in-hand with school's library; book talk activity where pupils can share about the book they've read." R13 proposed a three-part programme: "Read more English materials. Listen to English broadcast. English time - only English in class."

A second cluster of five respondents (26.3%) proposed structured oral practice with explicit scaffolding, and these responses are notable for their pedagogical specificity. R18 described a graduated sequence: "1-Drilling with sample answers, 2-fill in blanks with their own answers to the fixed phrase, drill again, 3.small group intervention." This is a recognisable progression from full support through partial support to independent production in a low-exposure setting, and it corresponds closely to the withdrawal of scaffolding that the Band 2 to 3 threshold formalises. R11 proposed to "Provide guidance like simple dialogue to assist them," while R4 and R10 both suggested performance-based framing, through mimicking film characters and "acting session" respectively, which have the additional property of licensing pupils to speak as someone other than themselves.

A third cluster of four respondents (21.1%) proposed changes beyond the individual teacher's remit. R15 recommended reducing class size and instituting a school-wide speaking programme; R14 proposed a maximum of two English classes per teacher; R8 and R16 proposed assessment leverage, through a speaking test and by making English a compulsory pass subject respectively. The last is worth noting as a divergent case: it proposes to solve a problem the same respondents attribute to anxiety by raising the stakes of performance, a remedy that the affective literature would predict to be counterproductive.

Only one respondent (5.3%) proposed directly addressing the affective dimension, with R7 suggesting teachers "Motivate and provide them vocabularies." This is a striking asymmetry given that affective barriers were the most frequently identified challenge.

Perceived Affordances of Technology (RQ3)

The third open-ended item produced the most convergent responses in the dataset. Themes are summarised in Table X.

TABLE X. Themes in Perceived Technological Affordances (N = 19)

Theme	n	%	Sources
Authentic audiovisual input and pronunciation modelling	12	63.2	R1-R6, R10, R11, R13, R14, R17, R18
Interactive and responsive practice applications	4	21.1	R8, R12, R16, R19
Vocabulary development through media	4	21.1	R1, R3, R9, R13
Anxiety reduction through asynchronous self-recording	1	5.3	R7
Qualified or conditional endorsement	2	10.5	R8, R15

Twelve respondents (63.2%) identified authentic audiovisual material as technology's principal contribution, and their accounts were unusually specific about mechanism. R5 explained that technology "Helps to provide audio n visual media to help students to improve their pronunciation as they can take note of the words (spelling & sound of letters)," identifying the orthography-phonology mapping that print alone cannot supply. R11 pointed to non-verbal dimensions: "Videos can show how to use the language inc facial expressions etc." R6 emphasised contextual variety, noting that pupils "can watch and listen to more English program to give pupils opportunity to see English used in different real life situation."

R18 offered the most fully specified pedagogical design in the dataset, integrating authentic input with structured production:

"By giving a real life situation eg video of conversation in a cafe etc, pupils will have to replace some info with their own preferred answers (listed on the slides)." (R18, more than 10 years' experience)

This proposal is essentially a substitution drill built on an authentic model, combining the input affordance that most respondents identified with the scaffolded production that the solutions theme emphasised.

Four respondents (21.1%) looked beyond input towards interaction. R19 asked for "An interactive and user friendly app that they can access and practice by themselves," and specified the requirement further as an application that would "Response to their conversation." This is a request for conversational feedback rather than playback, and it identifies a functional gap that current classroom technology in many Malaysian primary schools does not fill.

A single respondent identified the affective affordance, but the observation is disproportionately significant:

"Students feel comfortable recording themselves rather than speaking in front of others." (R7, more than 10 years' experience)

R7 here identifies precisely the mechanism that Huang and Liu (2025) isolate as technology's principal anxiety-reducing property: the removal of the live evaluative audience. Given that nine respondents identified fear of that audience as the central obstacle, the fact that only one connected technology to its removal represents a notable gap between problem diagnosis and solution design.

Two respondents qualified their endorsement. R8 answered "A lot if properly manage," and R15's response, "Personally put them in the English setter," suggests scepticism that mediated exposure substitutes for immersion. These qualifications are consistent with the double-edged characterisation in the review literature and are retained here as legitimate counterweights rather than treated as non-responses.

Integration of Strands (RQ4)

Table XI sets the two strands side by side, using the Band descriptors as the integrating frame. Because the respondents are the assessors of this cohort, each row pairs an attainment fact with an account of the same pupils by the same professionals.

TABLE XI. Joint Display of Quantitative and Qualitative Findings

Attainment finding	Governing descriptor	Corresponding theme	Inference
69.8% at Band 3-4	Sustains communication "at times"	Affective barriers (47.4%)	Extended talk is attempted only intermittently, not absent
9.8% at Band 5-6; 1.2% at Band 6	Sustains "most of the time" / "naturally"; guides others	Fear of peer evaluation (R13, R18)	Upper bands require volitional public performance that anxiety forecloses
0.4% at Band 1	Requires "a lot of support" for all functions	Pupils comprehend but will not produce (R6, R8, R16)	Deficit is productive and volitional, not receptive
20.4% at Band 1-2	All functions need heavy scaffolding	Weak foundational literacy including L1 (15.8%)	A subgroup needs literacy-level rather than speaking-level intervention

The two strands converge on a coherent account. The near-absence of pupils at Band 1 indicates that the cohort has not failed to acquire English; the near-absence at Band 6 indicates that few have moved beyond functional adequacy. The descriptors specify what stands between the two, and it is not accuracy, range or pronunciation, none of which distinguishes Band 4 from Band 5. It is sustained, self-directed talk. Teacher accounts identify precisely what suppresses sustained, self-directed talk: a lexical resource sufficient for reactive exchange but not for extension, and an affective climate in which holding the floor invites ridicule.

One point of divergence deserves note. Teachers identified affective barriers as the leading challenge but proposed affective remedies in only one instance, directing their solutions instead towards input and structural

change. The most direct implication of the integrated analysis is therefore that current remedial thinking is aimed at a different target from the one teachers themselves identify as primary.

DISCUSSION

The Plateau Sits at the Descriptor Boundary

The central finding of this study is not that most pupils perform at an intermediate level, which would be unremarkable, but that the plateau coincides exactly with the descriptor boundary defined by sustained talk. This coincidence is what converts a distribution into an explanation.

Had the cohort clustered at Bands 1 and 2, the constraint would be scaffolding: pupils cannot perform without support. Had it spread evenly, the constraint would be differential rates of acquisition. Instead, 69.8% sit at the two bands whose descriptors record functions performed independently but intermittently, and the population thins abruptly at the point where the descriptor requires those same functions to be sustained. Nothing in the Band 4 to Band 5 transition demands new linguistic knowledge; it demands more of the same behaviour, produced for longer and without prompting.

Read against the CEFR target for the year group, the finding is sharper still. A2 Low requires "a series of phrases and sentences" where A1 required "simple phrases and sentences" (Council of Europe, 2001). The Year 5 curriculum target is, in substance, connected talk. The cohort has achieved the utterance and not the series. The plateau is therefore not a general shortfall against the CEFR benchmark but a specific and locatable one, and it falls on the dimension that the affective literature identifies as most susceptible to anxiety, since anxiety curtails duration and self-initiation long before it degrades accuracy.

The Plateau as a Self-Reinforcing Cycle

The integrated evidence suggests the plateau is produced by a cycle rather than by any single deficit. A limited productive lexicon means that attempting extended speech carries a high risk of visible failure, since the pupil is likely to run out of words mid-utterance; Nation's (2013) distinction between receptive and productive knowledge explains why a pupil who follows the lesson may still be unable to sustain a turn. In a classroom where peers mock such failure, as R13 and R18 both reported, the rational response is to minimise exposure by speaking only when required and only in the shortest adequate form. But it is precisely extended, pushed output that consolidates vocabulary into retrievable productive knowledge (Swain, 1985). Avoidance therefore protects the pupil in the short term while preserving the lexical limitation that made avoidance necessary.

Each element sustains the others, and the pupil settles at the level of oral performance that can be produced without risk, which is the Band 3 and Band 4 descriptor almost exactly. This account explains a feature of the data that a single-deficit explanation cannot. If limited vocabulary alone were responsible, one would expect a broader spread as pupils acquired vocabulary at different rates. If anxiety alone were responsible, one would expect linguistically stronger pupils to escape it and populate the upper bands. The concentration of nearly seven in ten pupils at adjacent bands is more consistent with a stable equilibrium than with a distribution of individual deficits, and it is consistent with MacIntyre et al.'s (1998) argument that willingness to communicate mediates between competence and actual use.

The cycle also offers an account of why anxiety does not decline across the primary years in Thandavaraj and Mohd-Said's (2025) data. If avoidance preserves the very limitation that motivates it, then additional years of instruction will not by themselves dissolve the anxiety, because the pupil's evidence about their own likely performance does not improve.

The Diagnosis-Remedy Gap

The asymmetry between the challenges teachers identify and the solutions they propose is the study's second substantive finding. Affective barriers were named by 47.4% of respondents, affective remedies by 5.3%. Solutions concentrated instead on expanding input and on structural or policy change.

This gap is unlikely to reflect any failure of insight, given that the sample was composed overwhelmingly of experienced teachers who described the affective mechanism with precision. A more plausible reading is that the available repertoire of professional action is better stocked for input problems than for affective ones. A teacher can obtain more reading material, play more video, or institute an English-only period; the classroom techniques for reducing fear of peer evaluation are less codified and are rarely addressed in curriculum documents or in-service training. This reading is consistent with the CEFR implementation literature, which repeatedly finds teachers under-supported in enacting the framework's requirements (Sidhu et al., 2018; Mohamad Uri, 2023; Jamain & Jamaludin, 2023); the present finding suggests that the support deficit extends to the affective conditions the descriptors implicitly presuppose, since a descriptor requiring a pupil to sustain communication naturally is as much a statement about willingness as about competence.

Two responses point in a more troubling direction. R8's proposal of a speaking test and R16's proposal to make English a compulsory pass subject would, if the cycle described above is accurate, increase the evaluative stakes of exactly the performances pupils are already avoiding. The practical implication is that professional development should treat anxiety reduction as a teachable instructional technique with specific classroom procedures, rather than as a matter of general encouragement.

Technology as an Affective Instrument

Teachers overwhelmingly conceptualised technology as an input device: 63.2% identified authentic audiovisual material as its principal affordance, against 5.3% who identified reduced-anxiety practice. Read against the challenge data and against the descriptor analysis, this allocation is inverted relative to need. If the constraint on movement into the upper bands is sustained volitional talk, and if what suppresses it is the peer audience, then technology's most valuable property in this setting is not the quality of the models it delivers but the audience it removes.

R7's observation that pupils "feel comfortable recording themselves rather than speaking in front of others" describes an asynchronous, low-stakes production channel in which a pupil may attempt an extended turn, hear it, discard it, and try again without a single peer witnessing the failed attempts. This preserves the pushed output that development requires while suspending the evaluative conditions that suppress it, and it corresponds directly to the psychological-safety mechanism Huang and Liu (2025) identify as technology's clearest anxiety-reducing pathway. It is also, notably, a format in which the Band 5 criterion of sustaining communication can be practised at all, since a pupil rehearsing alone faces no interlocutor whose patience might expire. R19's request for an application that would "Response to their conversation" extends the same logic towards interaction with non-judgemental feedback.

The caution registered by R3, that some classrooms have "no ICT facilities," is essential to any recommendation drawn from this analysis. An intervention requiring reliable classroom devices will not reach the settings that may need it most, which argues for designs operating on a small number of shared devices or on pupils' home devices rather than on one-to-one provision.

Implications for Practice and Policy

Five implications follow. First, interventions should target the lexical and affective dimensions jointly rather than sequentially, since the cycle will regenerate if either element is left in place: vocabulary instruction that does not create safe conditions for using the new vocabulary is unlikely to move pupils off the plateau.

Second, instruction and assessment should target duration explicitly. Because the Band 4 to 5 boundary is defined by sustained rather than intermittent communication, tasks that elicit a single correct utterance cannot generate evidence of Band 5 attainment however well performed. Speaking tasks in Year 5 should require pupils to hold the floor for a specified stretch, and Learning Standard 2.3.1, narration, is the natural vehicle.

Third, asynchronous self-recording merits systematic trial as a routine practice format rather than an occasional activity. It is inexpensive, requires only modest infrastructure, and addresses the barrier teachers most frequently identify while providing a setting in which extended turns can be rehearsed without an audience. That such technology-mediated redesign is feasible in Malaysian primary classrooms is already established for

writing (Sharom & Kew, 2022); the present analysis indicates why the case for speaking may be stronger still, since the affective barrier that self-recording removes has no equivalent in written production.

LIMITATIONS

Three limitations bound the claims made here. First, the study is confined to a single cohort and its teaching team, and the qualitative sample of 19 teachers is small. The findings describe one instructional setting in depth rather than supporting generalisation to Malaysian primary ESL provision as a whole. The sample's concentration of experienced teachers, while analytically useful, further limits transferability to settings staffed by early-career practitioners.

Secondly, and relatedly, Band placement rests on teacher professional judgement and may be subject to central-tendency bias, whereby assessors avoid extreme categories. The near-absence of pupils at Bands 1 and 6 is consistent with a genuine plateau but also, in principle, with conservative band assignment. The CEFR implementation literature's finding that teachers possess limited assessment literacy in relation to the framework (Sidhu et al., 2018; Mohamad Uri, 2023) makes this a live rather than a formal concern.

Thirdly, the study is cross-sectional and descriptive. It cannot establish whether the observed distribution represents deterioration, improvement or stability relative to earlier cohorts, and no causal claim is advanced.

DIRECTIONS FOR FURTHER RESEARCH

Four lines of inquiry follow. Most urgently, a study in which speaking performance is rated by assessors independent of the teachers surveyed would establish whether the congruence reported here reflects the instructional reality or the shared perspective of a single professional group. A design-based intervention study trialling asynchronous self-recording against conventional oral practice, with willingness to communicate and Band movement as outcomes, would test the central practical claim advanced here. A longitudinal study tracking band movement from Year 4 through Year 6 would establish whether the plateau is a stable feature of the primary years or an artefact of the Year 5 stage, and would be particularly informative given that Year 4 targets A1 High and Year 6 targets A2 Mid. Finally, an observational study of how teachers actually apply the speaking descriptors would test the interpretive assumption identified in the fourth limitation above.

CONCLUSION

This study set out to describe the speaking attainment of Year 5 Malaysian ESL pupils against the CEFR-aligned curriculum standard and to read that distribution against the barriers their teachers perceive. The attainment data revealed a cohort compressed at the intermediate levels, with 69.8% at Bands 3 and 4, fewer than one pupil in ten reaching the excellent category, and three of 245 attaining the highest level. The teacher data identified affective barriers and limited lexical resource as the dominant obstacles, with weak foundational literacy, restricted exposure and structural constraints raised by smaller numbers.

Integrating the strands through the curriculum's own descriptors produced the study's principal finding. The plateau falls exactly at the boundary where the descriptors stop asking whether a pupil can speak and start asking whether a pupil keeps speaking. Nothing in that transition requires new linguistic knowledge; it requires the willingness to hold the floor. Against a year-group target of CEFR A2 Low, whose defining achievement in spoken production is the series of sentences rather than the single sentence, the cohort has secured the utterance and not the series.

The analysis further identified a gap between diagnosis and remedy: teachers named affective barriers most often but proposed affective solutions least often, and conceptualised technology overwhelmingly as a source of input rather than as an instrument for removing the audience whose presence they identified as the core problem.

The most actionable conclusion is that speaking interventions in Malaysian primary classrooms are unlikely to succeed if they address vocabulary and confidence separately, and that assessment tasks eliciting single utterances cannot generate evidence of the sustained talk the upper bands require. The single teacher

observation that pupils are more comfortable recording themselves than speaking in front of others points towards a class of intervention that addresses both dimensions at once, and that observation deserves systematic trial rather than the incidental status it currently holds.

DECLARATIONS

Ethical Approval

Ethical approval for this study was granted by the Educational Planning and Research Division (EPRD), Ministry of Education Malaysia [KPM.600-3/2/3-eras(12635)]. Further written permission to conduct the research was obtained from the relevant District Education Office and from the administration of participating school prior to any data collection. All teacher participants took part voluntarily, were informed of the purpose and scope of the study, and provided informed consent before completing the questionnaire; no individual teacher is identifiable in the data reported here. Pupil speaking attainment records were released by the participating schools in aggregate form, classified only by Band and containing no individual identifiers.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this article. No funding was received that could have influenced the design, conduct, analysis or reporting of this research.

Data Availability Statement

The anonymised qualitative dataset comprising the 19 teacher questionnaire responses, and the aggregate speaking attainment frequencies reported in Tables VI and VII, are available from the corresponding author on reasonable request. The pupil attainment records are not publicly deposited because they were released by the participating schools for the purposes of this study only, and their release beyond the research team would exceed the terms under which access was granted.

REFERENCES

1. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
2. Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
3. Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing.
4. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
5. Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>
6. Huang, L., & Liu, M. (2025). The impact of technology on foreign language anxiety: A systematic review of empirical studies from 2004 to 2024. *Humanities and Social Sciences Communications*, 12, Article 1665. <https://doi.org/10.1057/s41599-025-05921-6>
7. Jamain, S., & Jamaludin, K. (2023). Issues and challenges in the implementation of the Common European Framework of Reference (CEFR) in Malaysian primary school. *International Journal of Academic Research in Business and Social Sciences*, 13(12), 5863–5877. <https://doi.org/10.6007/IJARBSS/v13-i12/20394>
8. Kashinathan, S., & Aziz, A. A. (2021). ESL learners' challenges in speaking English in Malaysian classroom. *International Journal of Academic Research in Progressive Education and Development*, 10(2), 983–991. <https://doi.org/10.6007/IJARPED/v10-i2/10355>
9. Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
10. Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.
11. 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>

12. Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
13. Ministry of Education Malaysia. (2013). *Malaysia education blueprint 2013–2025: Preschool to post-secondary education*.
14. Ministry of Education Malaysia. (2015). *English language education reform in Malaysia: The roadmap 2015–2025*. English Language Standards and Quality Council.
15. Ministry of Education Malaysia. (2020). *Dokumen standard kurikulum dan pentaksiran: Bahasa Inggeris tahun 5, sekolah kebangsaan* [Curriculum and assessment standard document: English language Year 5, national primary school]. Curriculum Development Division.
16. Mohamad Uri, N. F. (2023). Challenges in CEFR adoption: Teachers' understanding and classroom practice. *International Journal of Modern Languages and Applied Linguistics*, 7(1), 49–62. <https://doi.org/10.17576/ijmlal-2023-0701-04>
17. Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
18. Sharom, K., & Kew, S. N. (2022). The effects of the flipped classroom on primary school students' writing performance in English. *International Journal of Computer-Assisted Language Learning and Teaching*, 12(1), 1–22. <https://doi.org/10.4018/IJCALLT.301195>
19. Sidhu, G. K., Kaur, S., & Chi, L. J. (2018). CEFR-aligned school-based assessment in the Malaysian primary ESL classroom. *Indonesian Journal of Applied Linguistics*, 8(2), 452–463. <https://doi.org/10.17509/ijal.v8i2.13311>
20. Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
21. Thandevaraj, E. J., & Mohd-Said, N.-E. (2025). A mixed-methods study of speech anxiety among Malaysian primary ESL pupils. *International Journal of Academic Research in Progressive Education and Development*, 14(4), 13–30. <https://doi.org/10.6007/IJARPED/v14-i4/26478>
22. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.