# Visual and Discourse-Structured Scaffolding for Impromptu Speaking: A Classroom Action Research Study

**Suzila Mohd Shukor*, Ajuntha Kuppan, Agnes Shanti Joseph**

**School of Languages, Literacies and Translation, Universiti Sains Malaysia**

***Corresponding Author**

**DOI: https://doi.org/10.47772/IJRISS.2026.101000811**

**Received: 24 October 2026; Accepted: 29 October 2026; Published: 18 November 2026**

## ABSTRACT

Impromptu speaking is a demanding test of B2-level proficiency because learners must find something to say, organise it, and keep talking with almost no planning time. Many learners respond with short, loosely connected utterances and frequent hesitation. This classroom action research study examined whether pairing visual prompts with an explicit discourse framework helped 20 B2-level engineering undergraduates in an English for Specific Purposes (ESP) General English II course at a Malaysian public university to speak more readily and with clearer organisation. Over an eight-week intervention, learners practised the course's 2+2-minute assessment task: two minutes describing a picture in the simple present tense, followed by two minutes of extended discussion organised around four moves (problem, consequences, suggested actions, and hope). Data came from pre- and post-intervention speaking assessments scored on a five-subskill discourse management rubric, a 20-item self-rating questionnaire, semi-structured interviews, and reflection journals. Descriptive results show that rubric scores rose on all five subskills (idea initiation, idea development, logical organisation, discourse markers, and turn sustaining), with the largest gains for discourse markers and idea initiation. Learners' self-ratings rose by a similar amount, although learners underestimated their progress in using discourse markers. Thematic analysis produced four themes: greater confidence, better organisation of ideas, reduced cognitive load at the start of the task, and a shift from description to idea development. Learners tended to credit the picture with helping them start and the framework with helping them continue. The findings suggest that the two forms of support do different jobs and work best together, and they offer ESP teachers a practical routine for tasks that require organised, unplanned speech.

**Keywords**: Discourse Management, Engineering Students, Impromptu Speaking, Speaking Fluency, Visual Scaffolding

## INTRODUCTION

Speaking is one of the core components of second language proficiency, particularly at the B2 level of the Common European Framework of Reference (CEFR), where learners are expected to produce clear, detailed, and organised speech across a range of communicative contexts (Council of Europe, 2020). Within this domain, impromptu speaking is a constant source of challenge because it requires learners to generate ideas, retrieve appropriate vocabulary, and focus on fluency as well as organisation simultaneously, with little or no time to prepare.

In English for Specific Purposes (ESP) contexts and among undergraduate engineering students at a Malaysian public research university, English grammar knowledge, but not impromptu speaking skills, is widespread. General English courses with this title are considered ESP courses, and the course syllabus for speaking includes topics that may be relevant to engineering graduates, such as problem reporting and resolution, as well as assessments that include an impromptu speaking task similar to the one described here. Impromptu speaking assessment tasks are challenging for B2-level students in the course, who tend to deliver brief, loosely structured speeches. Findings are consistent with predictions from cognitive load theory (Sweller, 1988; Sweller et al., 2019) that tasks requiring students to speak and plan and monitor their own speech will overload working memory and reduce student performance. Deficits in other pragmatic language skills, such

as turn-taking and topic-maintenance skills, may also affect student performance and increase anxiety about impromptu speaking tasks.

To address this gap, the present study investigates the combined use of visual scaffolding and an explicit discourse framework to facilitate speaking tasks conducted in an impromptu manner. Visual prompts are added

with the intention of easing the cognitive demands of idea generation by supplying concrete stimuli, while the discourse framework serves to guide learners in organising those ideas into a more extended and coherent response. As the Literature Review section develops further into detail, previous studies have typically investigated these two types of support independently; while the present study considers them as complementary components of a single pedagogical approach.

An action research design was adopted because a pedagogical intervention could be designed, implemented, and evaluated by the teacher-researcher within an authentic classroom setting through iterative cycles of planning, action, observation, and reflection (Kemmis et al., 2014). The study was conducted over eight weeks with B2-level learners in the ESP General English II course described above, in which spontaneous speaking is a required component of the assessed curriculum. Students in this course are expected to deliver a four-minute impromptu presentation based on a picture prompt, following a 2+2-minute structure that was already in institutional use for assessment purposes but was not, prior to this study, supported by any visual scaffolding or explicit discourse guidance during teaching. The present intervention introduced both forms of support into the teaching of this existing task.

The visual prompts used in the intervention were drawn from real-world, socially relevant issues (for example, traffic congestion, food security, and public health), a design choice made on the assumption that topics learners already have background knowledge of would lower the burden of content generation and allow classroom time to focus on discourse organisation rather than topic comprehension; this rationale, and the specific criteria used to select and sequence the prompts, is detailed in the Methodology section. By examining the effects that visual scaffolding and discourse structuring have on learners' ability to arrange and extend their ideas in impromptu speaking, this study seeks to provide practical, classroom-tested guidance for ESP practitioners who need learners to produce on the spot, organised speech under assessment conditions.

The research questions guiding this study are:

RQ1: To what extent does the combined use of visual scaffolding and a structured discourse framework improve learners' discourse management in impromptu speaking?

RQ2: How do learners perceive the use of visual scaffolding and structured support in improving their impromptu speaking?

## LITERATURE REVIEW

This section reviews research on scaffolding in second language speaking, drawing mainly on studies published from 2021 onwards together with the foundational work on which they build. Scaffolding theory serves as the organising framework. The review first sets out the origins of the construct, then considers visual scaffolding and discourse-structure scaffolding as two ways of putting it into practice and then turns to cognitive load theory as an account of what kind of support should make impromptu speech easier. It closes by identifying what the literature has not yet examined, namely the two forms of support used together.

### Scaffolding as the Guiding Theoretical Framework

The concept of scaffolding has its roots in Vygotsky's (1978) sociocultural theory and, in particular, in the Zone of Proximal Development (ZPD): the distance between what a learner can do alone and what the same learner can do with help from a more capable person or a mediating tool. For Vygotsky, learning is most productive in this zone. Support pitched far above the learner's current ability cannot be used, and a task the learner can already manage alone offers little room for development. The term “scaffolding” itself comes from Wood et al. (1976), who used it to describe how a tutor makes the ZPD workable in practice. The tutor reduces

the degrees of freedom in a task so that the learner can concentrate on the part that lies just beyond their current competence, and then withdraws that support, a process usually termed fading, as competence grows.

Three points from this definition shape both the review and the design of the present study. First, a scaffold is temporary; it is meant to be removed once the learner can perform without it. Second, a scaffold narrows a particular gap rather than removing the difficulty of the whole task, so different scaffolds can be expected to act on different parts of a complex task. Third, impromptu speaking is not one skill but several working at once: finding content, organising it, and delivering it under time pressure. It follows that a single scaffold is unlikely to cover all these demands. Visual scaffolding and discourse-structure scaffolding, the two strands reviewed below, can therefore be read as acting on two different parts of the same ZPD, an argument we return to in the synthesis.

**Visual Scaffolding: Narrowing the Gap in Content Generation**

Visual input is one of the most common ways of giving learners something to talk about. Picture-based tasks are widely used in speaking assessment partly because they hold content relatively constant across speakers. In a study of 377 Japanese learners of English performing picture-narration tasks, Koizumi and In'nami (2024) found that speech rate was the only measure of complexity, accuracy and fluency that significantly predicted how well learners achieved the communicative purpose of the task, and it was also the measure that behaved most consistently across different picture prompts. For the present study, this suggests that the immediate value of a picture lies in helping learners start and keep talking, rather than in shaping how their ideas are organised. A related line of evidence concerns familiarity with content. Bui and Huang (2018) found that both content familiarity and pre-task planning improved L2 fluency, with planning having the stronger effect and partly compensating for unfamiliar content. Impromptu tasks remove planning time almost entirely, which arguably makes familiar content, and a concrete visual cue to that content, more important than it would be in a prepared presentation.

Studies of technology-enhanced visual and situated scaffolding point in a similar direction. Huang (2024) found that virtual reality (VR) scaffolding significantly reduced speaking anxiety among 34 Taiwanese medical students, although learners also reported new anxieties linked to the technology itself and to being assessed in the VR setting. Yan et al. (2024) reported significant gains in fluency, vocabulary, pronunciation and grammar among 16 first-year English majors in China who took part in VR role-play classes built on a situated learning approach, although the study had no control group. More broadly, Pitura and Chang (2023), working with university students in Poland and Australia, mapped scaffolding features onto instructor-, peer- and technology-based tools for L2 speaking and found that students in both settings rated these tools as helpful, with instructor feedback rated more highly than peer feedback. Taken together, the picture-based and VR studies locate the main contribution of visual and situated support at the start of a speaking task, where a concrete context narrows the gap between what learners can produce from nothing and what the task demands. None of them, however, examine whether such support also helps learners organise an extended response once they have begun. This is the first gap addressed by the present study.

**Discourse-Structure Scaffolding: Narrowing the Gap in Organisation**

A second form of scaffolding gives learners an explicit structure for what to say and in what order. Saienko and Nazarenko (2021) used “speaking frames” with 41 third-year engineering students in an ESP course in Ukraine. At sentence level, the frames organised topic vocabulary and grammar around a core situation; at text level, students analysed a model monologue and built a frame that captured its structure. After the course, 70% of the experimental group reached a sufficient speaking level, compared with 43% of the control group, and hesitations, repetitions, false starts and pauses all fell, while continuous utterances more than doubled. Because the participants were engineering undergraduates in an ESP setting, this study is particularly close to the present one. Nguyen and AlSaqqaf (2025) tested a module built on the DMGA model (Diagnosing, Modelling, Sharing, Guiding and Applying) with 25 ESP undergraduates in Vietnam. Breakdown fluency and speech rate improved significantly, but repair fluency and vocabulary measures did not. The model's sequence, which moves from teacher modelling to guided and then independent application, is a practical form of fading, and

the uneven results are a useful reminder that structural support does not improve every aspect of speech equally within a short intervention.

Structure can also be supplied by peers. Homayouni (2022) combined peer assessment with scaffolding and group work for lower-intermediate EFL learners at a language institute in Iran and found that the experimental group outperformed a control group on an oral narrative post-test. Across these studies, explicit structure, whether supplied through frames, a staged teaching model or peer feedback, is associated with longer, more continuous and better-organised output. These studies suggest that a discourse-structure scaffold, in effect a map of what to say at each stage of a response, makes it less likely that learners will stop after one or two short utterances. What this strand of research seldom examines is how learners get started in the first place. In the studies reviewed, the scaffold mainly shapes how learners express and sequence content; the moment of facing an unfamiliar prompt with no preparation time is not their focus. This is the second gap the present study addresses.

### Cognitive Load as the Mechanism behind Scaffolding's Effectiveness

Cognitive load theory (Sweller, 1988; Sweller et al., 2019) offers an account of why scaffolding of either kind should make impromptu speaking easier. Working memory has limited capacity. When a task requires learners to generate ideas, retrieve vocabulary and monitor delivery at the same time, the combined load can exceed that capacity, leaving fewer resources for language production itself. In this view, a scaffold does not remove the inherent difficulty of the task but redistributes the load. A picture moves part of the content-generation demand onto an external representation, while a discourse framework moves part of the organisational demand onto a sequence the learner has already learned, and each frees working memory for the parts of the task it does not cover. Sweller (2023) describes how the theory has expanded by incorporating ideas from neighbouring fields, and its application here, to real-time L2 speech, should be read in that spirit: as a working explanation rather than a direct test. It does, however, explain why the two scaffolds should not be expected to substitute for each other. Each relieves a different source of load, so neither should fully make up for the absence of the other.

### Synthesis and Research Gap

Three observations follow from this review. First, research on visual and situated scaffolding (Bui & Huang, 2018; Huang, 2024; Koizumi & In'nami, 2024; Yan et al., 2024) indicates that concrete content and context help learners begin and sustain speech and can reduce anxiety, but these studies do not measure whether learners organise extended responses more effectively. Second, research on discourse-structure scaffolding (Homayouni, 2022; Nguyen & AlSaqqaf, 2025; Saienko & Nazarenko, 2021) shows gains in continuity, fluency and organisation, but generally takes the initial generation of content for granted. Third, we found no study that combines both forms of scaffolding in a single intervention with engineering ESP undergraduates, using a classroom action research design in which discourse management is assessed as five separate rubric-scored subskills (idea initiation, idea development, logical organisation, discourse markers, and turn sustaining).

The present study addresses this gap from within scaffolding theory. Rather than treating visual support and discourse structure as competing interventions, it treats them as two scaffolds acting on different parts of the same ZPD. Both were used within a single 2+2-minute impromptu speaking task, and teacher support was faded across the eight-week intervention, as scaffolding theory prescribes (Wood et al., 1976). The design asks whether scaffolding, both content generation and content organisation produces gains across all five discourse management subskills, rather than the partial gains that the single-scaffold studies reviewed above would lead us to expect.

## METHODOLOGY

### Research Design

Following the Plan–Act–Observe–Reflect cycle (Kemmis et al., 2014), a classroom-based action research

design was adopted for this study. Rather than utilizing a controlled experimental design, action research was deemed more appropriate for the current study as it sought to develop, implement, and refine a pedagogical solution to the issues that were directly identified by the teacher-researcher in her own classroom of English for Specific Purposes (ESP). Within this eight-week cycle, a mixed-methods research design was adopted to address both research questions that were posed for this study. Quantitative measures were collected and used

to address the first research question of this study, whilst qualitative measures were used to address the second research question. These mixed methods were adopted to allow for the explanation of the findings from the quantitative data collected for this study with the findings from the qualitative data that were collected.

During the planning stage, the teacher-researcher identified a recurring difficulty among students in previous cohorts, who struggled to give coherent responses during the 2+2-minute impromptu speaking exercise. To address this, a structured intervention combining visual aids with a four-step discourse structure was implemented over eight weeks with the current cohort. Students' spontaneous speaking abilities were assessed using the established 2+2-minute format, with data collected through the instruments described in the following sections. In the final week, the teacher-researcher analysed the data and provided feedback to students. Minor adjustments were also made to the intervention, including revisions to the discourse structure and the addition of sentence-level language support, in response to observed difficulties in using appropriate language to discuss the visual aids. Such adjustments based on students' needs are a key feature of action research, distinguishing it from experimental research (Kemmis et al., 2014).

**Research Context and Participants**

The study was conducted at a Malaysian public research university, in an ESP General English II course for engineering undergraduates (referred to here without its institutional course code to preserve participant privacy). Although administratively categorised as a general English course, its content and assessment are ESP-oriented: the speaking component specifically targets the kind of spontaneous, structured oral communication engineering graduates are expected to produce in professional contexts, such as explaining a problem, justifying a course of action, and proposing a solution under time pressure.

Twenty B2-level engineering undergraduates participated in the study. All participants had completed the prerequisite General English I course and had their CEFR B2 proficiency verified by course teachers through prior examination results before enrolling in the course in which this study took place. Participants were not selected through random sampling; they constituted an intact class taught by the teacher-researcher, which is consistent with the classroom-embedded nature of action research but is also acknowledged as a limitation in the Discussion and Conclusion section below.

**Intervention Design**

The intervention ran for eight weeks and was built around the course's existing 2+2-minute impromptu speaking task, in which learners are required to produce a four-minute impromptu presentation based on a visual prompt: the first two minutes require a description of the prompt in the simple present tense, and the remaining two minutes require extended discussion. Prior to this study, learners completed this task without any visual scaffolding or explicit discourse guidance, using only the assessment rubric as a reference. The intervention added three forms of support to this existing task structure.

First, the visual prompts used across the intervention were purposively selected to depict real-world social issues that engineering undergraduates could reasonably be expected to hold background knowledge and opinions about, for example, urban traffic congestion and everyday environmental waste. This selection criterion was adopted on the rationale that familiar, opinion-generating topics should lower the cognitive cost of content generation, allowing more of learners' working memory to be directed toward organisation and language use rather than toward comprehending an unfamiliar topic.

Second, learners were provided with express language scaffolding that included useful expressions about guessing, expressing opinions, and indicating logical relationships between ideas (see Appendix 4). As the third step, learners went through the four-step discourse framework, which consists of problem definition,

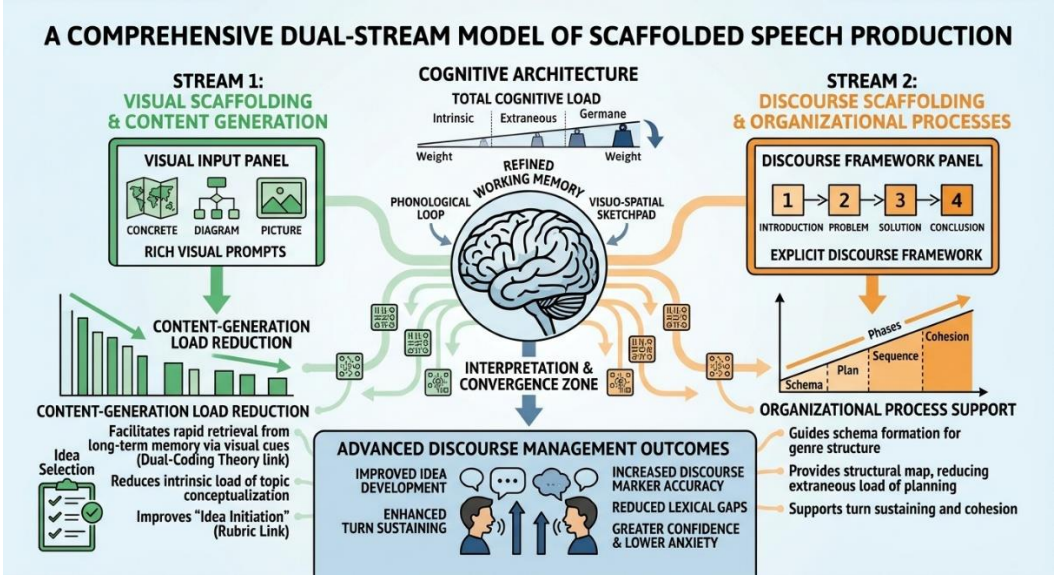
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue X October 2026

results, solutions, and expressing hope. These four steps were linked to the longer discussion phase of the 2+2-minute task, providing students with a robot-like sequence of actions in their speech instead of allowing them to just develop their ideas without having a specific structure.

The teacher clearly explained and demonstrated the discourse framework and the language scaffolding. Model

answers were used to clarify what the teacher meant (see Appendix 3). This was done at the beginning of the training process and continued in every following session, and the support given by the teacher gradually decreased as students became more familiar with the framework, which was in line with the withdrawal scaffolding principle discussed in the Literature Review. After each session, the students got feedback concerning the way they applied discourse structure, generated ideas, and used fluency.

**The framework in a diagram:**




**Instruments**

**Pre- and Post-Intervention Speaking Assessment**

Learners' impromptu speaking performance was assessed in Week 1 (pre-test, prior to any scaffolding instruction) and again in Week 8 (post-test, after the full intervention), using the same 2+2-minute task format and the discourse management rubric reproduced in Appendix 6. The rubric scores five subskills (idea initiation, idea development, logical organisation, use of discourse markers, and turn sustaining), each on a four-band scale (1 = Emerging to 4 = High), aligned with CEFR B2 descriptors for extended, coherent spoken discourse. This instrument produced the actual-performance data reported under RQ1 in the Results and Findings section.

**Perception Questionnaire**

A 20-item questionnaire (Appendix 1) was developed directly from the descriptors used in the speaking assessment rubric, with four items representing each of the five subskills, so that learners' self-perceived performance could be compared against their actual rubric-assessed performance. Each item was rated on the same four-band scale used in the rubric (Band 1 = Emerging to Band 4 = High). The questionnaire was administered twice, immediately before the pre-test and immediately after the post-test and produced the self-perceived performance data reported alongside the rubric data under RQ1.

**Semi-Structured Interviews**

Individual semi-structured interviews were conducted with participants after the post-test in Week 8, using two guiding questions (Appendix 2): a primary question asking learners to describe how the visual scaffolding and

structured framework had helped their speaking, and a follow-up question asking what further support they felt they needed. The interviews were 10–15 minutes long, and they were conducted in a combination of English and Malay to facilitate appropriate expression. Audio recordings were made with agreement by participants, who were aware that they were to be transcribed for analysis purposes. The semi-structured format allowed the teacher-researcher to clarify learner responses with examples, while keeping the interviews focused on the same two main questions allowing improvements in comparability.

**Learner Reflection Journals**

After each of the three intervention sessions, learners completed a short-written reflection (Table 1) with the help of sentence starters from an organized template (Appendix 7; e.g. "I learned that...", "I modified... to..."). This produced up to three reflection entries per learner across the intervention, capturing changes in learners' experience and self-assessed difficulty over time in a way that a single post-intervention interview could not. Journals were collected after each session and read alongside the interview transcripts during analysis.

**Data Collection Procedure**

Table 1 summarises the timeline over which the four instruments above were administered across the eight-week intervention.

Table 1 Timeline of Intervention and Data Collection

| **Week** | **Activity** |
|---|---|
| Week 1 & 2 | Pre-test speaking assessment; pre-intervention questionnaire; introduction and modelling of visual |
| | scaffolding and the four-step discourse structure |
| Week 3 & 4 | Intervention Session 1: guided practice with visual prompts and the discourse framework; 2+2-minute speaking task; teacher feedback; reflection journal entry |
| Week 5 & 6 | Intervention Session 2: continued practice with visual scaffolding; reinforcement of discourse |
| | structure; speaking task; teacher feedback; reflection journal entry |
| Week 7 & 8 | Intervention Session 3: independent application of the framework; speaking task; teacher feedback; reflection journal entry; post-test speaking assessment; post-intervention questionnaire; semi-structured interviews |

**Data Analysis**

Because RQ1 and RQ2 called for different kinds of evidence, quantitative and qualitative data were analysed separately before being brought together in the Discussion and Conclusion section.

**Quantitative Analysis (RQ1)**

Speaking assessment scores and questionnaire responses were analysed descriptively, in keeping with the classroom action research design and the size of the group. For each of the five discourse management subskills, mean pre-test and post-test scores were calculated for (a) the rubric-scored speaking assessment (actual performance) and (b) the four corresponding questionnaire items (self-perceived performance), and the change between the two points was computed for both. Standard deviations and percentage change were also reported for the questionnaire data. Placing the two data sources side by side makes it possible to see whether learners' sense of their own progress matched the progress recorded by the assessor or departed from it.

**Qualitative Analysis (RQ2)**

Interview transcripts and reflection journal entries were analysed together following the six phases of thematic analysis set out by Braun and Clarke (2006) and their later reflexive account of the method (Braun & Clarke, 2019). In Phase 1 (familiarisation), the teacher-researcher read all interview transcripts and journal entries

repeatedly to develop an overall sense of the dataset. In Phase 2 (generating initial codes), each transcript and journal entry was coded line by line for recurring ideas, using an inductive approach so that codes were grounded in what learners said rather than imposed from the theoretical framework in advance. In Phase 3 (searching for themes), related codes occurring across both the interview and journal data were grouped into candidate themes.

In Phase 4 (reviewing themes), each candidate theme was checked against the full dataset to confirm that it was supported by evidence from multiple participants and, where possible, from both data sources rather than from a single instrument alone. In Phase 5 (defining and naming themes), four themes were finalised and named. In Phase 6 (report production), the most illustrative quotations from the data obtained and the source of each quote were indicated in the Results and Findings section of the paper.

## RESULTS AND FINDINGS

Findings are reported in two parts, corresponding to the two research questions and the two data types described in the Data Analysis section: quantitative results addressing RQ1 (reported separately for actual, rubric-assessed performance and self-perceived, questionnaire-based performance), followed by qualitative results addressing RQ2 (reported by theme, with each illustrative extract labelled by its data source).

RQ1: Discourse Management — Actual Performance (Speaking Assessment)

RQ1: To what extent does the combined use of visual scaffolding and a structured discourse framework improve learners' discourse management in impromptu speaking?

Table 2 Below are the descriptive metrics (Frequencies, Percentages, Means, and Band Distributions for Pre & Post Test.

20 Respondents rated themselves via the self-assessment scale ranges from 1 to 4:

*Rating Scale: Band 1(B1): Emerging (1.00), Band 2(B2): Developing (2.00), Band 3(B3): Competent (3.00), Band 4(B4): High (4.00)*

Table 2: Descriptive Metrics (Pre vs. Post)

| No. | Questionnaire Item | Pre-Test Dist. f (%) | Pre M | Post-Test Dist. f (%) | Post M | Δ M |
|---|---|---|---|---|---|---|
| **A. Idea Initiation (Opening & Topic Engagement)** | | | **1.55** | | **2.79** | **1.24** |
| 1 | I can start speaking immediately when I am given a topic or visual. | B1: 8 (40%)<br>B2: 12 (60%)<br>B3: 0 (0%)<br>B4: 0 (0%) | 1.6 | B1: 0 (0%)<br>B2: 1 (5%)<br>B3: 17 (85%)<br>B4: 2 (10%) | 3.05 | 1.45 |
| 2 | I can begin my response without long pauses. | B1: 7 (35%)<br>B2: 13 (65%)<br>B3: 0 (0%)<br>B4: 0 (0%) | 1.65 | B1: 0 (0%)<br>B2: 5 (25%)<br>B3: 15 (75%)<br>B4: 0 (0%) | 2.75 | 1.1 |
| 3 | I can understand the prompt and start with a relevant idea. | B1: 11 (55%)<br>B2: 8 (40%)<br>B3: 1 (5%)<br>B4: 0 (0%) | 1.5 | B1: 0 (0%)<br>B2: 6 (30%)<br>B3: 13 (65%)<br>B4: 1 (5%) | 2.75 | 1.25 |
| 4 | I can confidently introduce my topic at the | B1: 11 | 1.45 | B1: 0 (0%) | 2.6 | 1.15 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | beginning of my speech. | (55%)<br>B2: 9 (45%)<br>B3: 0 (0%)<br>B4: 0 (0%) | | <br>B2: 10 (50%)<br>B3: 8 (40%)<br>B4: 2 (10%) | | |
| **B. Idea Development (Elaboration & Expansion)** | | | **1.71** | | **2.58** | **0.86** |
| 5 | I can explain my ideas clearly when I speak. | B1: 9 (45%)<br>B2: 8 (40%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.7 | B1: 0 (0%)<br>B2: 6 (30%)<br>B3: 12 (60%)<br>B4: 2 (10%) | 2.8 | 1.1 |
| 6 | I can give examples or reasons to support my ideas. | B1: 8 (40%)<br>B2: 9 (45%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.75 | B1: 0 (0%)<br>B2: 12 (60%)<br>B3: 4 (20%)<br>B4: 4 (20%) | 2.6 | 0.85 |
| 7 | I can go beyond describing and add more meaningful information. | B1: 9 (45%)<br>B2: 8 (40%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.7 | B1: 0 (0%)<br>B2: 13 (65%)<br>B3: 6 (30%)<br>B4: 1 (5%) | 2.4 | 0.7 |
| 8 | I can expand my ideas into longer responses. | B1: 9 (45%)<br>B2: 8 (40%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.7 | B1: 0 (0%)<br>B2: 12 (60%)<br>B3: 6 (30%)<br>B4: 2 (10%) | 2.5 | 0.8 |
| **C. Logical Organisation (Discourse Structuring)** | | | **1.46** | | **2.63** | **1.16** |
| 9 | I can organise my ideas in a clear sequence. | B1: 13 (65%)<br>B2: 5 (25%)<br>B3: 2 (10%)<br>B4: 0 (0%) | 1.45 | B1: 0 (0%)<br>B2: 10 (50%)<br>B3: 8 (40%)<br>B4: 2 (10%) | 2.6 | 1.15 |
| 10 | I can present my ideas step by step (e.g., problem → solution). | B1: 11 (55%)<br>B2: 7 (35%)<br>B3: 2 (10%)<br>B4: 0 (0%) | 1.55 | B1: 0 (0%)<br>B2: 10 (50%)<br>B3: 8 (40%)<br>B4: 2 (10%) | 2.6 | 1.05 |
| 11 | My speech has a clear beginning, middle, and ending. | B1: 13 (65%)<br>B2: 6 (30%)<br>B3: 1 (5%)<br>B4: 0 (0%) | 1.4 | B1: 0 (0%)<br>B2: 11 (55%)<br>B3: 7 (35%)<br>B4: 2 (10%) | 2.55 | 1.15 |
| 12 | My ideas are easy for others to follow. | B1: 12 (60%)<br>B2: 7 (35%)<br>B3: 1 (5%)<br>B4: 0 (0%) | 1.45 | B1: 0 (0%)<br>B2: 7 (35%)<br>B3: 11 (55%)<br>B4: 2 (10%) | 2.75 | 1.3 |
| **D. Use of Discourse Markers (Linking & Cohesion)** | | | **1.68** | | **2.19** | **0.51** |
| 13 | I use linking words (e.g., because, so, however) when speaking. | B1: 11 (55%)<br>B2: 5 (25%)<br>B3: 4 (20%) | 1.65 | B1: 2 (10%)<br>B2: 12 (60%)<br>B3: 5 (25%) | 2.25 | 0.6 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | B4: 0 (0%) | | B4: 1 (5%) | | |
| 14 | I can connect my ideas smoothly using appropriate expressions. | B1: 13 (65%)<br>B2: 3 (15%)<br>B3: 4 (20%)<br>B4: 0 (0%) | 1.55 | B1: 1 (5%)<br>B2: 15 (75%)<br>B3: 3 (15%)<br>B4: 1 (5%) | 2.2 | 0.65 |
| 15 | I can show relationships between ideas (e.g., cause, contrast). | B1: 7 (35%)<br>B2: 12 (60%)<br>B3: 1 (5%)<br>B4: 0 (0%) | 1.7 | B1: 1 (5%)<br>B2: 14 (70%)<br>B3: 4 (20%)<br>B4: 1 (5%) | 2.25 | 0.55 |
| 16 | My speech sounds connected rather than separate sentences. | B1: 6 (30%)<br>B2: 12 (60%)<br>B3: 2 (10%)<br>B4: 0 (0%) | 1.8 | B1: 3 (15%)<br>B2: 13 (65%)<br>B3: 4 (20%)<br>B4: 0 (0%) | 2.05 | 0.25 |
| **E. Turn Sustaining (Maintaining Speech)** | | | **1.71** | | **1.85** | **0.14** |
| 17 | I can speak continuously without long pauses. | B1: 8 (40%)<br>B2: 10 (50%)<br>B3: 2 (10%)<br>B4: 0 (0%) | 1.7 | B1: 6 (30%)<br>B2: 12 (60%)<br>B3: 1 (5%)<br>B4: 1 (5%) | 1.85 | 0.15 |
| 18 | I can keep speaking even when I need time to think. | B1: 9 (45%)<br>B2: 11 (55%)<br>B3: 0 (0%)<br>B4: 0 (0%) | 1.55 | B1: 5 (25%)<br>B2: 12 (60%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.9 | 0.35 |
| 19 | I can maintain my speech without stopping frequently. | B1: 6 (30%)<br>B2: 11 (55%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.85 | B1: 5 (25%)<br>B2: 12 (60%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.9 | 0.05 |
| 20 | I can complete my response without losing my flow. | B1: 8 (40%)<br>B2: 9 (45%)<br>B3: 3 (15%)<br>B4: 0 (0%) | 1.75 | B1: 7 (35%)<br>B2: 11 (55%)<br>B3: 2 (10%)<br>B4: 0 (0%) | 1.75 | 0 |

Below is the descriptive narrative analysis for Table 2.

**Dimension A: Idea Initiation (Opening & Topic Engagement)**

In the pre-intervention, students exhibited high initial hesitation and anxiety. Category average mean was M = 1.55 (between Emerging and Developing). Over half of respondents (55%) rated themselves at Band 1 for understanding prompts (f = 11) and introducing topics confidently (f = 11).

Meanwhile, in post-intervention, this dimension experienced the largest positive shift (ΔM = +1.24). The post-test category means rose to M = 2.79. Notably, 0% of students remained in Band 1 for all items. Item 1 (Immediate start) saw 85% of students (f = 17) advance to Band 3 (Competent), demonstrating that visual prompts effectively eliminated content-generation bottlenecks.

**Dimension B: Idea Development (Elaboration & Expansion)**

In the pre-intervention, elaboration was severely restricted (M = 1.71). Between 40% and 45% of learners clustered in Band 1 across all elaboration items.

However, in post-intervention, the category mean increased to M = 2.58 (ΔM = +0.86). For Item 5 (Explaining ideas clearly), 60% of learners (f = 12) achieved Band 3. Additionally, Item 6 (Providing supporting examples) saw 20% (f = 4) reach Band 4 (High). This supports Cognitive Load Theory, as relieving content load via visual cues freed working memory for detailed elaboration.

**Dimension C: Logical Organisation (Discourse Structuring)**

As evident in pre-intervention, structural organisation recorded the lowest overall baseline mean (M = 1.46). Up to 65% of respondents placed themselves in Band 1 for sequential ordering (f = 13) and structural completeness (f = 13).

In contrast, post-intervention recorded an increase. Following explicit instruction on the 4-step framework, the category mean surged to M = 2.63 (ΔM = +1.16). Band 1 counts completely disappeared (0%). For Item 12 (Ideas easy to follow), 55% of students (f = 11) reached Band 3, with 10% (f = 2) advancing to Band 4 across all structural items.

**Dimension D: Use of Discourse Markers (Linking & Cohesion)**

In the pre-intervention, cohesion was fragmented (M = 1.68). Item 14 (Smooth connections) showed 65% of students (f = 13) at Band 1.

On the contrary, during post-intervention, the overall mean reached M = 2.19 (ΔM = +0.51). The primary movement was a strong migration into Band 2 (Developing): 75% (f = 15) for smooth connections and 70% (f = 14) for showing causal/contrast relationships. This indicates that phrase banks provided essential structural scaffolds for lower-proficiency learners.

**Dimension E: Turn Sustaining (Maintaining Speech)**

During pre-intervention, real-time speech maintenance had a baseline mean of M = 1.71, with responses divided between Bands 1 and 2.

As per post-intervention, turn sustaining showed modest, incremental growth (M = 1.85, ΔM = +0.14). While Band 1 frequencies dropped across all items (reducing from 8-9 students down to 5-7), most learners consolidated in Band 2 (55%-60%). This reflects the high cognitive demands of real-time spontaneous production, where fluency gains require longer intervention periods.

It can be observed that student experiences varied based on individual baseline proficiencies.

Thematic analysis of the interview transcripts and reflection journal entries, following the six-phase process described in the Data Analysis section, identified four themes: (1) increased confidence in impromptu presentation, (2) improved ability to organise ideas, (3) visual scaffolding reduces cognitive load, and (4) transition from description to idea development. Each theme is presented below with illustrative extracts; the Source column indicates whether an extract was drawn from the post-intervention interview or from a weekly reflection journal entry.

Theme 1: Increased Confidence in Impromptu Presentation

| Example Response | Source |
|---|---|
| "Before this activity, I was very nervous when I had to speak suddenly. Now I feel more prepared even if I don't know the topic." | Interview |
| "I don't panic anymore when the teacher asks me to speak. I can at least say | Interview |

| | |
|---|---|
| something and continue." | |
| "The structure gives me confidence because I know I won't get stuck halfway." | Journal |
| "Previously I always paused a lot, but now I can speak more smoothly without thinking too much." | Journal |
| "I feel more comfortable sharing my ideas, even if my grammar is not perfect." | Interview |
| "Now I am more willing to try speaking because I have a guide to follow." | Journal |

Learners reported reduced anxiety and greater willingness to participate in impromptu speaking. The structured framework appeared to function as a form of psychological, as well as linguistic, support, giving learners a sense that they would not run out of things to say partway through the task. This idea has a lot of things in common with the turn sustaining category from Table 2. Turn sustaining did make progress, but it was the least advanced of all five subskills. In other words, the students started feeling more comfortable before their speech started getting smoother.

Theme 2: Improved Ability to Organise Ideas

| **Example Response** | **Source** |
|---|---|
| "Before this my ideas were all mixed. Now I can arrange them step by step." | Journal |
| "The framework helps me connect my ideas better, not just say random sentences." | Interview |
| "I know how to expand my answer instead of stopping after one point." | Interview |
| "The steps make my speaking clearer because I follow a sequence." | Interview |
| "I can explain my ideas more completely, not just give short answers." | Journal |
| "It helps me think deeper about the topic, not only to describe what I see." | Journal |

The students found the reason for their improved ability to organize their thoughts in terms of the four-step structure rather than the visual prompts. This finding from the students' report supports the conclusion mentioned later in the Discussion and Conclusion that visual scaffolding and the structure of discourse both played their own role in improving the students' performance, rather than a single impact.

Theme 3: Visual Scaffolding Reduces Cognitive Load

| **Example Response** | **Source** |
|---|---|
| "The picture helps me generate ideas quickly, so I didn't waste time thinking about what to say." | Interview |
| "Without the visual, I feel blank. With the picture, I can start immediately." | Interview |
| "The image gives me something to talk about, so I don't feel lost." | Interview |
| "I don't need to translate too much in my head because I can see the situation." | Journal |
| "The visual makes it easier to focus because I already have a starting point." | Journal |
| "It reduces the pressure because I don't need to imagine everything by myself." | Journal |

This theme provides a qualitative counterpart to Cognitive Load Theory (Sweller, 1988): learners explicitly described the visual prompt as reducing the mental effort otherwise required to generate a starting idea, which corresponds to the largest quantitative gain being recorded for idea initiation in Table 2.

Theme 4: Transition from Description to Idea Development

| **Example Response** | **Source** |
|---|---|
| "At first I only describe the picture, but later I can give opinions and suggestions." | Interview |
| "The second part helps me think beyond what I see." | Journal |
| "I learn how to continue speaking, not just stop after describing." | Interview |

This theme reflects the specific progression the 2+2-minute structure was designed to produce: from simple

description in the first two minutes to opinion, speculation, and suggestion in the second. Learners' own descriptions of this shift correspond to the quantitative gains recorded for idea development and turn sustaining in Table 2 and suggest that it was the second half of the task, rather than the picture itself, that moved learners beyond description, the level the rubric places at Band 2 (“mostly descriptive with minimal elaboration”).

## DISCUSSION AND CONCLUSION

This study used a mixed-methods action research design to examine the effect of combining visual scaffolding with an explicit discourse framework on the impromptu speaking of B2-level engineering undergraduates, operationalising discourse management as five measurable subskills: idea initiation, idea development, logical organisation, discourse markers, and turn sustaining.

Viewed through the scaffolding framework set out in the Literature Review, the results support the study's central premise: visual scaffolding and discourse-structure scaffolding act on different parts of learners' ZPD and using them together was followed by gains across all five discourse management subskills. The two elements also seem to have made separate contributions. The start of ideas has been noted to pull in one of the two highest additions. Moreover, learners have mentioned the visual as something that gives a person the starting point instantly (Theme 3). This corresponds with the belief found in various studies of picture-originated activities (Koizumi & In'nami, 2024) as well as studies on familiarity with the content (Bui & Huang, 2018) that says that effective means of communication has a lot to do with the ability to start a conversation with respect to some tangible, known topic. The changes seen in producing ideas and in sequential organization, along with the fact that learners believe that the efficient structure is what helped them to learn not a picture (Theme 2), suggests that the visual alone was not enough to have the learners improved in the organization. It follows the improvements of organization and fluency when speaking tasks are used (Saienko & Nazarenko, 2021) as well as staged scaffolding (Nguyen & AlSaqqaf, 2025). Cognitive load theory (Sweller, 1988; Sweller et al., 2019) can be used as an explanation here since each kind of structure appeared to have facilitated a different kind of load imposed on working memory.

The increases in discourse markers, which is reflected in the highest score presented in Table 2, implies that the explicit language scaffolding (Appendix 4) enabled learners to have a ready list of linking devices available when needed instead of having to produce the devices on their own. The improvement observed for the turn sustaining variable but of a smaller level could suggest that fluency is less influenced by the scaffolding in terms of structure and rather requires practice, the time of which was not sufficient within the scope of eight-week instruction period.

Qualitative data presents an effective dimension that is not represented by the rubric scores solely. The participants in the focus group as well as in reflective journals expressed that there was a reduction in anxiety and the readiness to do impromptu speaking (Theme 1), even though they admitted having problems with grammar. It means that the gains from the intervention may occur due to confidence and unwillingness to face the task instead of the effect on the measurements of performance through the rubric. However, it is worth mentioning that there was no direct measure of confidence in this study, so this interpretation serves solely as an assumption for future studies.

Taken together, the findings support treating visual scaffolding and discourse-structure scaffolding as complementary rather than interchangeable. Visual scaffolding appears best suited to narrowing the gap in content generation, while discourse-structure scaffolding appears necessary to turn that initial output into extended, coherent discourse. Teacher modelling was heaviest in Weeks 1–2 and minimal by Weeks 7–8, in line with the principle that support should be withdrawn as competence develops (Wood et al., 1976). The research didn’t explore whether student’s acquisitions could be preserved without scaffolding yet. Future studies therefore may study this question. This brings an important implication for ESP teachers planning speaking tasks: for discourse management to take place in the context of the task, it’s fundamental to use both types of scaffolding in combination, as one type of it might not be enough on its own.

The study has several limitations. The sample (20 learners) and the eight-week duration are small in relation to what would be needed for generalizable conclusions. Without a comparison group, the gains cannot be

attributed to the intervention with the confidence a controlled design would allow; they are also consistent with ordinary practice effects from repeated attempts at the same task format. All quantitative results were reported descriptively, in keeping with the classroom action research design, so no claims can be made about the statistical significance of the gains. The questionnaire data, moreover, reflects learners' perceptions rather than measured performance. The intervention bundles three forms of support (visual scaffolds, discourse frameworks, and language-expression scaffolds), making it difficult to isolate the individual contribution of each component. Finally, as an intact-class study in a single institution, the findings are best treated as context-specific evidence rather than as generalisable to other ESP settings without further testing. Note that the intervention was designed as a holistic pedagogical package rather than a factorial experiment.

Future research could extend this design with a longer intervention period, with a controlled group, establishing a comparison group receiving one form of scaffolding only (to isolate the separate contributions of visual and discourse-structure support suggested by the qualitative data), and inferential statistical analysis with a larger sample. Further research could also examine whether the same combined scaffolding design transfers to other proficiency levels, other ESP disciplines, or speaking tasks with a different rhetorical structure than the problem–consequence–action–hope sequence used here.

In conclusion, this study provides classroom-based evidence that combining visual scaffolding with an explicit discourse framework supported B2-level engineering undergraduates in moving from brief, fragmented impromptu responses toward more organised, extended, and confidently delivered speech. Subsequent studies may employ multi-arm comparison groups (e.g., isolating visual-only vs. discourse-framework-only groups) to parse out individual effects. By addressing both the cognitive demand of generating content and the organisational demand of structuring it, this two-part approach offers ESP practitioners a practical, classroom-tested strategy for teaching impromptu speaking as a discourse management skill rather than as an unstructured test of spontaneous fluency alone.

## 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. Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806
3. Bui, G., & Huang, Z. (2018). L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy. Language Teaching Research, 22(1), 94– 114. https://doi.org/10.1177/1362168816656650
4. Council of Europe. (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing.
5. Homayouni, M. (2022). Peer assessment in group-oriented classroom contexts: On the effectiveness of peer assessment coupled with scaffolding and group work on speaking skills and vocabulary learning. Language Testing in Asia, 12, Article 61. https://doi.org/10.1186/s40468-022-00211-3
6. Huang, H. L. (2024). Exploring the impact of VR scaffolding on EFL teaching and learning: Anxiety reduction, perceptions, and influencing factors. Multimodal Technologies and Interaction, 8(10), Article 85. https://doi.org/10.3390/mti8100085
7. Kemmis, S., McTaggart, R., & Nixon, R. (2014). The action research planner: Doing critical participatory action research. Springer. https://doi.org/10.1007/978-981-4560-67-2
8. Koizumi, R., & In'nami, Y. (2024). Predicting functional adequacy from complexity, accuracy, and fluency of second-language picture-prompted speaking. System, 120, Article 103208. https://doi.org/10.1016/j.system.2023.103208
9. Nguyen, T. U. N., & AlSaqqaf, A. (2025). Improving ESP Vietnamese learners' EFL speaking fluency and vocabulary with DMGA scaffolding: A modular approach. International Journal of Evaluation and Research in Education, 14(2), 1528–1536. https://doi.org/10.11591/ijere.v14i2.30468
10. Pitura, J., & Chang, H. (2023). Tools for scaffolding the development of L2 speaking in English-medium higher education: Lessons from Poland and Australia. Theory and
11. Practice of Second Language Acquisition, 9(2), 1–35. https://doi.org/10.31261/TAPSLA.12816

12. Saienko, N., & Nazarenko, I. (2021). Using speaking frames as scaffolding tools to teach university students to speak in ESP. International Journal of Learning, Teaching and Educational Research, 20(4), 99–115. https://doi.org/10.26803/ijlter.20.4.6
13. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
14. Sweller, J. (2023). The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. Educational Psychology Review, 35, Article 95. https://doi.org/10.1007/s10648-023-09817-2
15. Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, 31(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5
16. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
17. Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17(2), 89–100. https://doi.org/10.1111/j.1469-7610. 1976.tb00381.x
18. Yan, W., Lowell, V. L., & Yang, L. (2024). Developing English language learners' speaking skills through applying a situated learning approach in VR-enhanced learning experiences. Virtual Reality, 28, Article 167. https://doi.org/10.1007/s10055-024-01061-5

# APPENDICES

## Appendix 1: 20-Item Questionnaire (Aligned with the Rubric)

Scale: Band 1 – Emerging; Band 2 – Developing; Band 3 – Competent; Band 4 – High

A. Idea Initiation (Opening & Topic Engagement)

- I can start speaking immediately when I am given a topic or visual.
- I can begin my response without long pauses.
- I can understand the prompt and start with a relevant idea.
- I can confidently introduce my topic at the beginning of my speech.

B. Idea Development (Elaboration & Expansion)

- I can explain my ideas clearly when I speak.
- I can give examples or reasons to support my ideas.
- I can go beyond describing and add more meaningful information.
- I can expand my ideas into longer responses.

C. Logical Organisation (Discourse Structuring)

- I can organise my ideas in a clear sequence.
- I can present my ideas step by step (e.g., problem → solution).
- My speech has a clear beginning, middle, and ending.
- My ideas are easy for others to follow.

D. Use of Discourse Markers (Linking & Cohesion)

- I use linking words (e.g., because, so, however) when speaking.
- I can connect my ideas smoothly using appropriate expressions.
- I can show relationships between ideas (e.g., cause, contrast).
- My speech sounds connected rather than separate sentences.

E. Turn Sustaining (Maintaining Speech)

- I can speak continuously without long pauses.
- I can keep speaking even when I need time to think.
- I can maintain my speech without stopping frequently.
- I can complete my response without losing my flow.

**Appendix 2: Interview Questions**

- 1. How did the visual scaffolding and structured framework help you in your speaking?
- 2. What other support do you need to expand your ideas?

**Appendix 3: Samples of Scaffolded Materials**

Visual Prompt 1

[Original image retained from source file — not reproduced in this excerpt]

Q1: Look at the picture and make up your version of what is happening. (2 minutes)

Sample answer: "As I observe, this is a trader who earns a living selling poultry at a local wet market. His day starts as early as 6 a.m. daily. His tasks include handling the produce, preparing orders, and trimming items into portions as customers request. Once done, he weighs and packs them for collection. He enjoys interacting socially with his customers as part of his daily routine and uses his expertise to suggest a suitable portion for a named recipe, helping customers choose according to their needs and budget."

Grammatical structure for sample answer: describe states/conditions in the picture, including location and occupation, using the simple present tense; describe feelings and emotions in the picture; describe actions using the present continuous tense.

Q2: What would you do if you were in a similar situation? (2 minutes)

Tips: communicating about sustainability effectively involves being informed, positive, and relatable — highlight the benefits of sustainable action (health, cost savings, quality of life), frame the response around shared values, share a personal or practical example, and emphasise solutions rather than only problems.

Structure of sample answer: state the problem clearly; explain the consequences relating to the area of focus (e.g., environment); make a call to action or express hope.

Sample answer: "I would love to see a reduction in single-use plastics among traders and consumers. Single-use plastics pose real concerns for small traders and for the environment more broadly — the main issue is the environmental impact of plastic waste, including pollution and harm to ecosystems.

The consequences of this problem are serious: plastic waste increases pollution and can affect public health over time, and it also represents a lost economic opportunity as reusable alternatives become more available.

To improve this, I believe traders and customers need better access to affordable reusable packaging, and clearer public information about its benefits. I hope that in the future, local markets will rely far less on single-use plastic, for the benefit of both traders and the environment."

Visual Prompt 2 (example)

Q1: Look at the picture and make up your version of what is happening. (2 minutes)

Sample answer: "This picture shows heavy traffic congestion during peak hours in a city area. Many vehicles are queued along the road, and drivers appear to be waiting patiently as the traffic barely moves. Some pedestrians can be seen trying to cross carefully despite the congestion."

Q2: What would you do if you were in a similar situation? (2 minutes)

Sample answer: "In a situation like this, facing a traffic jam, I will surely try to remain calm and act according to the rules of the road. Traffic congestion is a big problem that has an impact on how people live in cities. There are serious problems connected with this phenomenon: it causes air pollution since cars are still moving

while not going anywhere; it utilizes gasoline and time and creates a lot of stress. I believe creating an efficient public transportation system is a good way to solve this problem, as it will make it possible for people to travel without using their vehicles."

**Appendix 4: Useful Phrases for B2 Speaking Tasks**

1. Start your answer

- From what I can see in the picture...
- It seems like...
- I think this situation happened because...
- In my opinion...

2. Speculating

- It could be that...
- Perhaps this is due to...
- One possible reason is...

3. Linking ideas

- First, ...
- On the one hand... on the other hand...
- As a result...
- This leads to...

4. Expressing consequences

- This can cause serious problems such as...
- One consequence of this is...
- This might affect...

5. Making a call to action

- We need to...
- It's important to...
- I believe we should...
- My hope is that...

**Appendix 5: Tips for Structuring the Problem–Consequence–Action–Hope Response**

- Identify the problem or issue shown in the picture or question (e.g., traffic congestion, pollution, road safety).

- State the problem clearly in your answer (e.g., "This shows a problem with traffic congestion, which affects daily life in urban areas.").

- Explain the impact or consequence (e.g., "Traffic jams increase air pollution and waste fuel, which harms the environment and people's health.").

- Give a solution or call to action (e.g., "Governments should improve public transport and encourage carpooling to reduce congestion.").

- Express hope or a positive outlook (e.g., "I hope that in the future, cities will have better transport systems and cleaner air for everyone.").

**Appendix 6: Assessment Rubric — Discourse Management in Impromptu Speaking**

| **Subskill** | **4 – High** | **3 – Competent** | **2 – Developing** | **1 – Emerging** |
|---|---|---|---|---|
| Idea Initiation (Opening & Topic Engagement) | Initiates promptly with no noticeable hesitation; clearly engages with the visual prompt and establishes a relevant, focused starting point. | Initiates with minor hesitation; generally relevant to the prompt. | Noticeable hesitation before starting; limited or partially relevant response to prompt. | Unable to initiate effectively; long pauses or irrelevant response. |
| Idea Development (Elaboration & Expansion) | Ideas are well-developed with clear explanations, examples, or reasoning; moves beyond description to meaningful elaboration. | Ideas are developed with some supporting details; attempts to go beyond description. | Limited development; mostly descriptive with minimal elaboration. | Very limited or no development; ideas remain fragmented or repetitive. |
| Logical Organisation (Discourse Structuring) | Ideas are clearly organised in a logical sequence (e.g., problem → consequence → solution → hope); smooth progression of ideas. | Generally organised with a clear sequence, though minor lapses may occur. | Some attempt at organisation, but ideas may be loosely connected or inconsistent. | Lacks clear organisation; ideas are disjointed or randomly presented. |
| Use of Discourse Markers (Linking & Cohesion) | Uses a range of appropriate discourse markers effectively (e.g., because, therefore, however); ideas are well-connected and cohesive. | Uses some discourse markers appropriately; connections between ideas are mostly clear. | Limited or repetitive use of basic connectors; weak connections between ideas. | Rare or no use of discourse markers; ideas lack cohesion. |
| Turn Sustaining (Maintaining Speech) | Sustains speech smoothly with minimal hesitation; able to continue speaking without breakdown. | Sustains speech with occasional pauses but maintains overall flow. | Frequent pauses and hesitation; difficulty maintaining speech. | Unable to sustain speech; frequent breakdowns and long pauses. |

The rubric was designed to align with CEFR B2 descriptors, particularly in relation to learners' ability to produce extended, coherent, and well-structured spoken discourse (Council of Europe, 2020).

Appendix 7: Reflection Journal Template

- I realised that...

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue X October 2026

- I adjusted the... as to...
- It can be seen that...
- The learners were able to...