JEELS
(Journal of English Education and Linguistics Studies)
P-ISSN: 2407-2575 E-ISSN: 2503-2194
https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/jeels

ASSESSING COMMUNICATIVE PERFORMANCE THROUGH A MULTIMODAL VIDEO PROJECT IN ENGLISH FOR TOURISM AND HOSPITALITY

***Alan Jaelani[1]; Ninuk Lustyantie[2]; Fathiaty Murtadho[3]**
[1]English Department, Universitas Ibn Khaldun Bogor, Indonesia
[2,3]Applied Linguistics, Postgraduate, Universitas Negeri Jakarta, Indonesia
**alan.jaelani@uika-bogor.ac.id; ninuklustyantie@unj.ac.id; fathiaty.murtadho@unj.ac.id*
(*) Corresponding Author

Abstract: In contemporary tourism and hospitality education, communicative competence is increasingly multimodal and digitally mediated. However, conventional assessments often fail to capture this complexity. Thus, this study examines the use of a multimodal video project to assess communicative language performance at a private tourism and hospitality school in Indonesia. The participants were 30 fourth-semester undergraduate students who worked collaboratively to produce short tourism promotion videos integrating spoken language, visual imagery, and audio elements. Data were collected through classroom

[1] **Citation in APA style:**
Jaelani, A., Lustyantie, N., & Murtadho, F. (2026). Assessing communicative performance through a multimodal video project in English for Tourism and Hospitality. *JEELS (Journal of English Education and Linguistics Studies, 13(1),* 187-212.
DOI: 10.30762/jeels.v13i1.5668

Submission: June 2025, Revision: January 2026, Publication: April 2026

observations, rubric-based performance assessments, analysis of student-produced video artefacts, and semi-structured interviews with 18 purposively selected students. The assessment focused on key components of communicative performance, including fluency, grammatical accuracy, pronunciation, vocabulary use, coherence, and multimodal integration. Findings revealed that the project provided an authentic assessment context, capturing both linguistic and paralinguistic features, while students highlighted motivation and engagement as critical to their participation. Although perceived as a fair assessment method, challenges in time management, technical skills, linguistic demands, and group coordination emerged. The study advocates for a shift in ESP assessment practices towards incorporating multimodal, performance-based tasks that reflect real-world communication and digital literacy skills.

Keywords: *Multimodal assessment, communicative language performance, video project, digital literacy, project-based learning*

INTRODUCTION

Tourism and hospitality industry operates within a communication landscape that is increasingly multimodal, digitally mediated, and performance oriented. Professionals in this sector are required to establish grammatical accuracy and fluency in English and to integrate speech, visuals, gestures, and digital media to communicate persuasively with diverse audiences. As English functions as the primary lingua franca of global tourism, the ability to perform communicative tasks authentically has become a core competency for students in tourism and hospitality education (Zou et al., 2019). This shift has considerable implications both for language instruction and more critically for how communicative

language performance is assessed in vocational English for Specific Purposes (ESP) contexts.

Traditional language assessment practices in ESP programs have largely relied on monomodal formats such as written tests, scripted oral examinations, or isolated speaking tasks. While these instruments can capture discrete linguistic knowledge, they often fail to reflect the complex, situated, and multimodal nature of professional communication in tourism settings (Jaiswal, 2019). In real workplace contexts, communication rarely occurs through spoken or written language alone; instead, meaning is constructed through the interaction of verbal language with visual design, spatial organization, prosody, gesture, and digital affordances (Kress & van Leeuwen, 2021; O'Halloran, 2008; T. Xiong et al., 2022). Consequently, there is growing concern that conventional assessment approaches may inadequately represent students' actual communicative competence and professional readiness.

In response to these limitations, multimodal assessment has emerged as a promising alternative that aligns assessment practices more closely with real-world communication demands. Multimodal assessment recognizes that communicative performance involves the transposition of multiple semiotic resources and therefore evaluates language use as it is enacted in authentic, meaning-making contexts (Ross et al., 2020; W. Xiong & Lai, 2024). Rather than isolating linguistic forms, this approach consents learners to determine how language interacts with visual, auditory, and embodied modes to achieve communicative goals. In tourism and hospitality education, where promotional discourse, service encounters, and digital marketing are central professional genres, multimodal assessment holds relevance (Bonsignori & Cappelli, 2022).

Existing scholarship has established strong theoretical foundations for multimodality in language education. Research on multimodal literacy emphasizes learners' capacity to interpret and

produce meaning across modes which highlight its importance for contemporary communication practices Liaw & Chen (2023) & Unsworth et al. (2019). Similarly, studies on digital multimodal composing have verified positive effects on learner engagement, creativity, and language development (Hafner, 2020; Pham & Li, 2023). However, much of this work has focused on multimodal learning and pedagogy rather than on assessment. As a result, assessment is often treated as a secondary or implicit outcome rather than as a primary object of inquiry.

Research examining multimodal tasks frequently positions video projects, digital storytelling, or vlog production as instructional tools that enhance motivation and learning pham processes (Huang, 2021; Ramírez Mera & Tur, 2021). While these studies provide valuable insights into learner engagement and skill development, they tend to foreground learning outcomes rather than interrogating how such tasks function as systematic assessment instruments. Questions concerning what aspects of communicative language performance are assessed, how performance criteria are operationalized, and how multimodal products are evaluated remain insufficiently addressed in ESP contexts with clear professional orientations.

Moreover, although Ross et al. (2020) propose a comprehensive multimodal assessment framework that accounts for both process and product, empirical studies applying this framework to assess communicative language performance in vocational settings are still limited. In tourism and hospitality education, where assessment must capture both linguistic competence and professional communicative effectiveness, the lack of context-specific multimodal assessment research constitutes a significant gap. This gap is evident in the Indonesian context, where tourism is a strategic economic sector and English proficiency is a key employability requirement, yet assessment practices remain predominantly traditional and exam oriented.

Despite a growing interest in multimodal pedagogy, there is a clear lack of empirical research that theorizes and examines multimodal video projects as assessment tools for communicative language performance in tourism and hospitality ESP contexts. Existing studies tend to emphasize learning and engagement rather than assessment validity, focus on general EFL or higher education settings rather than vocational tourism programs, and under-theorize how multimodal resources contribute to the evaluation of communicative competence. Accordingly, there is insufficient understanding of how multimodal assessment can be systematically designed, implemented, and interpreted to assess professional communicative performance in tourism education.

Addressing this gap requires shifting analytical attention from multimodal learning outcomes to multimodal assessment practices. In tourism communication, effectiveness is not determined solely by grammatical accuracy or lexical range, but by how language is deployed strategically alongside visuals, sound, and non-verbal cues to inform, persuade, and engage audiences. From this perspective, communicative language performance must be assessed as a situated, multimodal act. Video-based assessment tasks, when carefully designed and evaluated, offer a viable means of capturing this complexity, as they allow learners to reveal communicative competence in contextually authentic scenarios (Bonsignori & Cappelli, 2022; Ross et al., 2020).

This study investigates how a multimodal video project can be used to assess key dimensions of communicative language performance, including fluency, accuracy, pronunciation, vocabulary use, coherence, and multimodal integration. It also explores how students perceive this form of assessment in comparison to traditional testing formats and identifies the challenges and affordances associated with implementing multimodal assessment in a vocational ESP context. By focusing explicitly on assessment, this study contributes conceptual and empirical insights into the role of multimodal assessment in evaluating professional communication.

METHOD

Research Design

This study employed a qualitative case study approach to explore the use of multimodal video projects as assessment tools for evaluating students' communicative language performance in a private hospitality and tourism school in Bogor, Indonesia. The approach was chosen because it enables an in-depth, holistic investigation of complex educational practices within their real-life contexts (Johnson & Stake, 1996).

Participants

The participants in this study were 30 students enrolled in a fourth-semester course titled Professional English II. From this course, eighteen students were selected for in-depth interviews using purposive sampling. Selection was based on ensuring diverse representation in terms of English proficiency levels and roles assumed within their respective project groups, including scriptwriters, presenters, and editors. Observational and artefact data involved all students working in collaborative groups to produce the final video projects.

Research Instruments

Data were collected using three primary instruments: classroom observation notes, student-created video projects, and semi-structured interviews. Classroom observations focused on documenting students' verbal and non-verbal communication, collaborative interactions, and task engagement throughout the four-week project implementation. An observation checklist was used to guide data collection, adapted from communicative competence and multimodal learning frameworks of Ross et al. (2020)**.** The checklist included indicators such as frequency and clarity of spoken English use during group work, use of gestures, facial expressions, and eye contact during rehearsals and recordings, degree of peer interaction and collaborative decision-making, engagement in task-related discussions (e.g., scripting,

revising, rehearsing), and responsiveness to feedback from peers and the instructor. These indicators were selected to capture both linguistic performance and multimodal communicative behavior within the ESP tourism project context. These observations provided contextual data on how the multimodal project facilitated authentic opportunities for communicative language use.

The student video artefacts were collected at the end of the project as summative performance tasks and assessed using a structured analytic rubric specifically designed to evaluate students' communicative language performance in an ESP context for tourism and hospitality. The rubric assessed students' ability to use spoken English effectively within a multimodal video project that simulated professional tourism communication tasks. The employed six assessment indicators were: fluency, grammatical accuracy, pronunciation, vocabulary use, organization and coherence, and multimodal integration. These indicators reflect core components of communicative performance while acknowledging the role of visual, auditory, and non-verbal semiotic resources in meaning-making. Each indicator was rated on a four-point scale ranging from poor to excellent for detailed evaluation of both linguistic and multimodal aspects of student performance within an authentic ESP tourism context.

Semi-structured interviews were conducted with selected students to gain insights into their perceptions of the project as both a learning and assessment tool, addressing aspects such as motivation, perceived language development, technical challenges, and collaborative dynamics.

Learning Instruction

Students were asked to produce a short promotional video for a chosen tourist destination in Indonesia. The task involved four main stages: conducting research and selecting a site, scriptwriting in English, filming on location, and editing the final product. Each stage was supported by the instructor who provided feedback on language use, organizational structure, and technical production aspects. This

instructional design aligned with a task-based learning approach that integrates language use with authentic performance, and encourages collaboration, creativity, and digital literacy.

Data Collection and Analysis

The data collected from observations, video artefacts, and interviews were analysed using thematic analysis. This process involved transcribing all interview recordings and systematically coding them for recurring patterns and themes. Deductively, the coding was informed by the research questions and key theoretical constructs from multimodal learning and communicative competence. Inductively, new themes were allowed to emerge from the data without imposing pre-existing categories.

Classroom observation notes and video content were similarly coded and analyzed using the same coding framework. To ensure coding reliability, a two-stage coding process was implemented. First, the primary researcher conducted initial coding based on predefined categories derived from the research questions and theoretical constructs of communicative competence and multimodal discourse. Second, a subset of the data (approximately 25%) was independently reviewed by a peer researcher with expertise in qualitative analysis. Any discrepancies in coding were discussed until consensus was reached, thereby enhancing the dependability of the analytic process.

For the video artefacts, a multimodal discourse analysis framework (Kress & van Leeuwen, 2006) was applied, focusing on how verbal narration, visual representation, and non-verbal behavior (such as gestures and facial expressions) worked together to create coherent and persuasive communicative performances.

Findings from all three data sources were triangulated to enhance the trustworthiness of the study addressing credibility, dependability, and confirmability. Credibility was established through data triangulation across observations, video artefacts, and interviews to ensure consistency of findings. Dependability was strengthened using systematic coding procedures and peer review of the coding

process. Confirmability was ensured by maintaining an audit trail consisting of observation notes, coded transcripts, analytic memos, and video artefacts, allowing findings to be traced back to the original data sources.

FINDINGS

Facilitation of Communicative Language Performance Assessment

The first research question focused on how multimodal video projects facilitated the assessment of students' communicative language performance in a professional English. This inquiry was grounded in the need for authentic and contextually rich language assessment that goes beyond traditional written or oral tests. By integrating audiovisual components, the multimodal video project aimed to capture a fuller spectrum of students' communicative competencies both linguistic and paralinguistic within a meaningful task-based setting.

To evaluate this, performance data were collected using teacher observation notes, structured rubrics, and peer evaluations. These instruments assessed six core components of communicative language performance: fluency, accuracy, pronunciation, vocabulary use, organization and coherence, and multimodal integration. Each component was rated on a 4-point scale: excellent, good, fair, poor. The assessment focused on students' verbal production and their use of multimodal tools to enhance the clarity, persuasiveness, and engagement of their video presentations.

Table 1.
The Structured Rubric for Multimodal Video Project

Category	Excellent (16–20)	Good (11–15)	Fair (6–10)	Poor (1–5)
Fluency	Smooth, natural flow; confident and spontaneous delivery	Mostly smooth with minor hesitations	Noticeable pauses, some disfluency	Frequent hesitations, robotic or heavily scripted

Accuracy (Grammar)	Grammatically accurate with varied structures	Mostly correct grammar with few errors	Frequent grammar errors; limits clarity	Persistent basic errors; affects understanding
Pronunciation	Clear, intelligible, appropriate stress and intonation	Mostly clear; minor pronunciation issues	Occasionally unclear; mispronunciations noticeable	Often unclear; affects comprehension
Vocabulary Use	Rich, precise, and domain-appropriate vocabulary	Good range; mostly relevant and effective	Limited vocabulary; some awkward or vague expressions	Basic or incorrect vocabulary; lacks relevance
Organization & Coherence	Clear structure; effective transitions; logical flow	Mostly clear organization with minor issues	Some disorganization; unclear connections	Lacks structure; confusing or random sequencing
Multimodal Integration	Excellent use of visuals/audio to enhance message	Good use of multimodal elements; mostly relevant	Basic use; some mismatches or distractions	Little or no integration; poor or irrelevant visuals/ audio
				Total Score: __/120

Although the video project was produced collaboratively in groups, performance assessment was conducted at the individual level using a structured rubric. Each student was rated separately across six criteria. Group scores reported in Table 2 represent the aggregation of individual scores within each group which allows the analysis to reflect both individual communicative performance and overall group pattern. The students were organized into six groups (Groups A–F), with each group consisting of five members. Each group was tasked with creating a 3–5 minute-promotional video of a tourism site in Indonesia, integrating narration, dialogues, background music, captions, transition effects, and relevant imagery. The task required

students to collaboratively script their content, rehearse and record narration or live speech, edit the video, and submit both the final product and a reflective commentary on the process.

The following table provides a breakdown of the performance assessment across the groups based on the cumulative individual scores for each indicator.

Table 2.
Group-Based Communicative Language Performance Ratings (Individual Scores Aggregated)

Group	F	A	P	V	O&C	MI	Total Score (Max 120)
A	18	15	17	19	18	20	107/120
B	17	14	16	18	17	18	100/120
C	14	13	15	17	15	17	91/120
D	12	12	13	15	14	16	82/120
E	16	14	16	18	16	19	99/120
F	13	12	14	16	14	17	86/120

F=Fluency, A=Accuracy, P=Pronunciation, V=Vocabulary Use, O&C= Organization & Coherence, MI=Multimodal Integration

In table 2, each group's score for a component represents the sum of individual scores (5 students × 4 max points = 20). This method was to reflect both group collaboration and individual achievement. Group A demonstrated the highest fluency as its members delivered their scripts with natural pacing and rhythm, avoiding excessive hesitations or repetition. This group adopted a host-tourist conversational format, which allowed for turn-taking, spontaneous interaction, and a dynamic flow. Group D and Group F, on the other hand, showed less fluency due to heavy reliance on reading from scripts, which resulted in robotic delivery and uneven speech patterns. The multimodal format, however, seemed to reduce speaking anxiety in some students, as the option to pre-record and edit segments created a low-stress environment for language production.

General grammatical accuracy was moderate with common issues including incorrect verb tenses, article misuse, and subject-verb agreement particularly among Groups D and F. However, groups that invested more time in scriptwriting and peer-editing, e.g., Groups A

and E displayed better structural control in their sentence construction. The nature of the task encouraged repeated revisions, which appeared to help students' self-correct errors before the final recording. While the multimodal setting did not directly eliminate grammar errors, it fostered a process-oriented workflow that encouraged careful language planning.

Most groups performed well in pronunciation, with Groups A, B, and E earning high marks. Their speech was generally intelligible, with appropriate stress, intonation, and articulation. Students who rehearsed with peer feedback or used voiceover functions were able to achieve more polished pronunciation. In contrast, live recordings posed pronunciation challenges for Groups D and F especially when recorded outdoors where background noise affected clarity. The multimodal format supported pronunciation development by allowing students to listen to their own recordings, reflect, and re-record as needed.

Vocabulary use was one of the strongest aspects across all groups. Most students were able to employ topic-specific and descriptive terms accurately, such as *"eco-tourism," "heritage site," "culinary experience,"* and *"hospitality services."* Group A and E excelled in this area by incorporating persuasive and context-appropriate expressions suited for tourism promotion.

High scores in organization and coherence mirrored students' ability to sequence ideas logically and maintain cohesion throughout the video. Groups A and E structured their content into clear sections e.g., introduction, attractions, local food, conclusion and used transition markers both verbally and visually. Lower-performing groups, such as D and F, often had disjointed segments or abrupt transitions, indicating less planning and coordination during the editing process. Nonetheless, even in these cases, the need to synchronize speech with visuals appeared to push students toward more mindful content organization.

Multimodal integration was executed very well by most groups. Students creatively combined narration, music, imagery, subtitles, and

camera work to reinforce their spoken content. Group A's use of background music and aerial footage added a professional polish, while Group E's inclusion of local interviews provided authentic contextual flavor. This multimodal enhancement did not distract from language performance but rather reinforced key messages and made student communication more engaging and effective.

Student Perceptions of the Multimodal Video Project

Qualitative data were gathered from semi-structured interviews and reflective journals of 18 student participants. Thematic coding and frequency analysis were applied to explore recurring sentiments regarding their emotional response, cognitive engagement, skill development, and perceived fairness of the project. Interview responses were transcribed, coded, and grouped into five overarching themes and the Likert-scale interpretation was applied as shown in table 3 below. Although this study adopts a qualitative case study design, Likert-scale summaries are included for descriptive purposes only to provide an overview of response tendencies and to support the thematic interpretation of interview data rather than to make inferential or statistical claims.

Table 3.
Distribution of Student Perceptions

Perception Theme	Indicator	SA	A	N	D	SD
Affective Attitude	Enjoyed the video project	11	6	1	0	0
	Felt proud of final product	10	7	1	0	0
Perceived Learning Effectiveness	Learned more through project work	10	6	2	0	0
	Improved speaking and writing simultaneously	9	7	2	0	0
Engagement and Motivation	Felt more motivated than in traditional tests	9	7	2	0	0
	Practiced English more often during project	8	8	2	0	0
Real-World Communication Relevance	Project resembled workplace communication	8	7	3	0	0

	Helped develop confidence for speaking with tourists	9	6	3	0	0
Fairness and Preference	Felt the project fairly assessed language ability	9	6	2	1	0
	Preferred video projects over oral tests	12	5	1	0	0

SA=Strongly Agree, A=Agree, N=Neutral, D=Disagree, SD=Strongly Disagree

Most students expressed a high level of enjoyment in participating in the multimodal video project. 11 out of 18 respondents "strongly agreed" that the project was enjoyable, and an additional six "agreed." One respondent commented, "*It didn't feel like a test. It felt more like a creative job that I could be proud of" (R5).* Another said, *"I watched my group's video with my family. They also liked it." (R8).* The sense of ownership and pride contributed to heightened emotional engagement with the language task. Many students admit that filming themselves created a sense of accomplishment not normally associated with standard classroom tests.

Students overwhelmingly believed that the project allowed them to learn more effectively than through conventional methods. Several students pointed to the integrative nature of the task, combining speaking, writing, vocabulary application, and grammar accuracy. For instance, a student said, "*When we write the script, we also learn grammar. Then, when we say it in the video, we check our pronunciation." (R6).* This feedback confirms that the process-oriented nature of the task allowed students to engage in repeated practice, revision, and reflection.

Another key finding was the project's positive impact on learner motivation. Over 88% of participants reported feeling more motivated compared to traditional testing formats. Students emphasized the appeal of using technology, working in groups, and creating a tangible end product. One student noted, *"I practiced English almost every day to prepare. I even used apps to check my pronunciation before recording." (R13).* The notion of performance for a real audience (peers and instructors) also pushed students to prepare more thoroughly, indicating increased intrinsic and extrinsic motivation.

Students reported that the multimodal task felt more aligned with the kind of English they expect to use in their future careers in tourism and hospitality. Several mentioned how the task of promoting a destination resembled what a tour guide or hotel receptionist might do. One student reflected, *"I think this is the first project where we speak like we're working, not like we're studying." (R3).* The inclusion of visual elements such as maps, images, and video clips anchored the task in real-life communicative practice.

Lastly, students viewed the project as a fairer assessment of their communicative abilities. Many commented on the ability to rehearse and edit their videos before final submission, which reduced anxiety and increased performance quality. One participant explained, *"In an oral test, I'm nervous and forget my words. But in a video, I can prepare and speak well." (R11).* This finding raises important considerations for assessing communicative language performance in ways that reduce test-related stress and allow for greater learner agency.

Teacher notes and observations reinforced the student-reported experiences. Teachers noted that student participation and effort were higher than during traditional speaking tests. Moreover, students who were typically quiet in class seemed more confident on camera. This was particularly noticeable among students who assumed roles behind the camera but still engaged in scripting and narration.

Peer feedback collected during post-project presentations supported the idea of improved learner autonomy and communicative competence. Peers recognized each other's strengths and noted improvement in areas such as pronunciation, natural speech, and clarity of message.

Challenges and Benefits Encountered in the Project

Drawing from student interviews, reflective journals, and researcher observations, the data were thematically coded into two broad domains: perceived challenges and benefits. A nuanced analysis reveals the multifaceted nature of learner experiences within the

context of integrating language learning with digital media production.

Perceived Challenges

Students reported encountering a diverse range of challenges that can be classified into four major categories: time management, technical challenges, linguistic challenges, collaborative issues. Table 4 summarizes the specific issues and the frequency with which they were mentioned across the participant pool.

Table 4.
Challenges Encountered by Students

Challenge Domain	Specific Challenge	Freq.	Participant's Comment
Time Management	Aligning schedules for group work	12	*"It was hard to find time when everyone was free to shoot the video together." (R18)*
Technical Challenges	Editing video content	10	*"I didn't know how to add background music or text. I had to learn from YouTube." (R2)*
	File formatting and uploading issues	5	*"The file was too big to send. I had to compress it several times." (R4)*
Linguistic Challenges	Writing scripts in English	9	*"We spent a long time checking grammar in our script before filming." (R15)*
	Pronunciation and delivery	6	*"I had to practice several times because I was afraid of mispronouncing words." (R16)*
Collaborative Issues	Unequal group participation	4	*"Some members did more than others. Not everyone contributed equally." (R9)*
	Decision-making and role negotiation	8	*"We argued a bit about who should be the host and who should do the editing."(R12)*
Performance Anxiety	Nervousness about being recorded	6	*"I was very shy at first, especially knowing my face would be seen by others."(R11)*

A key challenge was time management, especially in synchronizing schedules for group filming and editing sessions. Students found it difficult to coordinate in-person meetings. Although mobile devices made filming easier, several students still struggled with video editing, including adding captions, adjusting audio, and ensuring smooth transitions. While some had prior experience with apps like CapCut or Canva, others relied on peers or online tutorials to gain the required technical proficiency.

Linguistically, students faced difficulties with scriptwriting. Unlike spontaneous speaking tasks, the project required them to carefully craft and deliver coherent grammatically correct scripts. Several students noted that even though they had time to revise and practice, ensuring natural pronunciation and appropriate intonation remained a hurdle. Collaboratively, though generally seen as beneficial, sometimes introduced friction. Students cited issues such as unequal participation, miscommunication over task roles, and difficulties in decision-making. A few groups struggled to reach consensus on creative direction or presentation format, leading to extended project timelines. Interestingly, performance-related anxiety was another emerging theme. Despite not having to perform live, some students still experienced stress about being filmed. However, this challenge often transformed into **a** learning opportunity as they became more comfortable through repetition.

Perceived Benefits

students reported numerous personal, linguistic, and professional benefits from completing the multimodal video project. The benefits were classified into six categories: language skill improvement, digital literacy, professional relevance, increased confidence, collaborative growth, and motivational impact.

Table 5.
Students' Perceived Benefits

Benefit Domain	Specific Benefit	Freq.	Participants' Comment
Language Skills	Improved speaking fluency	14	*"We did several takes, so my speaking became more natural each time." (R2)*
	Expanded tourism-related vocabulary	13	*"Now I can describe hotels and places using more professional words." (R8)*
Digital Literacy	Learned to use editing tools and apps	10	*"I didn't know how to edit before. Now I can make short tourism videos." (R1)*
Professional Relevance	Better understanding of tourism communication	12	*"It felt like real work in the hotel industry, explaining services and tours." (R4)*
Confidence Building	Gained confidence in using English on camera	12	*"Before I was shy. Now I feel braver to speak English in public." (R6)*
Collaboration Skills	Learned teamwork and time management	9	*"We had to share ideas and solve problems together, like in a real job." (R13)*
Motivation	More motivated to learn and use English	9	*"It was more fun than normal tests. I wanted to do my best." (R7)*

Students felt that the project improved their oral communication skills in fluency and pronunciation. They appreciated the opportunity for multiple takes and peer feedback, which allowed them to self-monitor and revise. Students perceived a noticeable expansion in their vocabulary, particularly terminology relevant to tourism and hospitality. A surprising benefit was the enhancement of digital literacy. Students who had never used editing tools learned to integrate images, voice-overs, subtitles, and transitions. These skills were relevant to the project and considered valuable for future employment in the digital age of tourism promotion.

Several students connected the project directly with real-world communication practices. They reported that producing a tourism video simulated the kind of promotional content expected in hotel

marketing or travel services, which made the task more meaningful and engaging. This relevance contributed to their confidence in applying English in future workplace settings. The project also fostered teamwork and collaboration**.** Although challenges existed, many students highlighted the satisfaction of completing a group task and acknowledged the development of soft skills such as negotiation, planning, and shared responsibility.

DISCUSSION

The findings clearly indicated that the multimodal video project was effective in fostering students' communicative competence across fluency, grammar accuracy, pronunciation, vocabulary use, coherence, and multimodal integration. Most students performed at moderate to high levels in these areas, with especially strong outcomes in vocabulary use and multimodal integration. These results align with research by Kayumova & Sadykova (2019); Pham & Li (2023); Xu (2023) who found that integrating digital multimodal tasks into EFL classrooms enhances students' engagement with authentic communication and improves language performance.

A key observation in this study was that students achieved fluency and coherence more easily when assuming the role of video narrators or hosts. This finding supports the concept of situated learning Kayumova & Sadykova, 2019; Morgenstern, 2023; Owoyele et al., 2022) which emphasize the importance of context-rich, socially embedded learning experiences. By simulating real-world communication scenarios, students moved beyond rote learning and engaged in more spontaneous, purposeful language use (Goyal & Rose, 2020; Makhmudov, 2023; Twiner et al., 2022).

Notably, the integration of multimodal features amplified the communicative intent of the spoken language. This reflects the principles of multimodal discourse framework (Kress & van Leeuwen, 2021) which posits that meaning is co-constructed through multiple semiotic modes. The successful combination of verbal and visual modes in the project revealed students' increasing semiotic awareness,

an essential skill in professional hospitality contexts where visual storytelling is often central.

However, grammar and pronunciation accuracy remained at moderate levels for the majority. While this suggests that students prioritized communication over correctness, it also points to an area requiring additional instructional focus. Although multimodal projects encourage expressive freedom, scaffolding in linguistic accuracy remains necessary (Boche & Henning, 2015; Nfor, 2020).

Student perceptions were positive, particularly in terms of engagement, perceived learning, and motivation. This aligns with (Muntaha et al., 2023; Senaratne et al., 2022) who underline the affective benefits of multimodal assessment, including reduced anxiety and enhanced learner agency. Students in this study described the project as more reflective of their actual abilities than traditional tests. The iterative nature of video production from scripting, rehearsing, and editing allowed for self-correction and deeper reflection, leading to higher satisfaction with the final performance.

As expected, time management and technical difficulties emerged as key obstacles. Coordinating group schedules and navigating editing tools proved difficult for several students, confirming previous observations by (Hafner, 2020) about the digital demands of multimodal tasks. This suggests that while students are often considered "digital natives," their technological proficiency varies, particularly when required to produce professional-quality multimedia content.

Despite this challenge, the benefits reported were substantial. Students felt their speaking skills improved due to repeated practice, echoing the concept of "productive rehearsal" in spoken language learning (Cappellazzo et al., 2023). Vocabulary gains were also notable, particularly in the domain-specific lexicon of tourism. This lends support to content and language integrated learning (CLIL) approaches which posit that language learning is enhanced when embedded in meaningful content contexts (Banegas et al., 2020; Ljalikova et al., 2021; Yang, 2020; Zhyrun, 2016).

Pedagogical implications

The findings of this study carry important pedagogical implications for English language instruction, especially within the context of hospitality and tourism education. The implementation of multimodal video projects enabled students to practice and demonstrate communicative language performance in realistic, job-relevant scenarios. This task-based, context-rich approach supported the development of fluency, vocabulary use, and coherence, while also engaging students in meaningful and purposeful communication. Teachers in ESP settings can employ similar video-based projects to simulate professional genres such as promotional videos or service interactions, thereby helping learners bridge the gap between classroom practice and workplace expectations. Assessment rubrics should reflect both individual contributions and group efforts while maintaining clear performance criteria.

CONCLUSION

Within a task that represented real-world communication scenarios relevant to the tourism industry, the multimodal video project successfully facilitated the assessment of communicative language performance. Students generally perceived the project positively, reporting increased motivation, enjoyment, and confidence, along with a sense of fairness and relevance in the assessment process. Despite encountering challenges such as time constraints, video editing difficulties, and linguistic struggles during scriptwriting, students also reported significant benefits in speaking skills, tourism vocabulary acquisition, teamwork, and digital literacy. However, the study has limitations, including a small sample size and a single institutional setting, which may limit the generalizability of the results. Future research could explore similar multimodal tasks across different ESP contexts and educational levels, incorporate longitudinal designs to assess lasting impacts on language development, and investigate the

influence of technological training on student performance and engagement in multimodal assessments.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES

During the preparation of this manuscript, the authors utilized several AI-assisted technologies to support the research process. CapCut was employed to transcribe the interview recordings, ChatGPT was used to extract and refine the interview scripts by removing the timestamp, and Atlas.ti was applied to classify and code the qualitative data. The authors have carefully reviewed and edited all content generated or assisted by these tools and take full responsibility for the accuracy and integrity of the publication.

REFERENCES

Banegas, D. L., Corrales, K., & Poole, P. (2020). Can engaging L2 teachers as material designers contribute to their professional development? findings from Colombia. *System*, *91*. https://doi.org/10.1016/j.system.2020.102265

Boche, B., & Henning, M. (2015). Multimodal scaffolding in the secondary english classroom curriculum. *Journal of Adolescent and Adult Literacy*, *58*(7). https://doi.org/10.1002/jaal.406

Bonsignori, V., & Cappelli, G. (2022). Developing strategies for conceptual accessibility through multimodal literacy in the English for tourism classroom. *Multimodal Communication*, *11*(1). https://doi.org/10.1515/mc-2020-0026

Cappellazzo, U., Falavigna, D., & Brutti, A. (2023). An investigation of the combination of rehearsal and knowledge distillation in continual learning for spoken language understanding. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2023-August*. https://doi.org/10.21437/Interspeech.2023-198

Goyal, R., & Rose, H. (2020). Stilettoed Damsels in Distress: the (un)changing depictions of gender in a business English textbook. *Linguistics and Education*, *58*. https://doi.org/10.1016/j.linged.2020.100820

Hafner, C. A. (2020). Digital multimodal composing: How to address multimodal communication forms in ELT. *English Teaching(South Korea)*, *75*(3). https://doi.org/10.15858/engtea.75.3.202009.133

Huang, H. W. (2021). Effects of smartphone-based collaborative vlog projects on EFL learners' speaking performance and learning engagement. *Australasian Journal of Educational Technology*, *37*(6). https://doi.org/10.14742/ajet.6623

Jaiswal, P. (2019). Using learner-centered instructional approach to foster students' performances. *Theory and Practice in Language Studies*, *9*(9). https://doi.org/10.17507/tpls.0909.02

Johnson, K. E., & Stake, R. E. (1996). The art of case study research. *The Modern Language Journal*, *80*(4). https://doi.org/10.2307/329758

Kayumova, A., & Sadykova, G. (2019). Implementing a multimodal online program for very young learners of Russian: Educators' perspective. *Indonesian Journal of Applied Linguistics*, *9*(1). https://doi.org/10.17509/ijal.v9i1.13323

Kress, G., & van Leeuwen, T. (2006). *Reading images: the grammar of visual design*. Routledge.

Kress, G., & van Leeuwen, T. (2021). Reading Images: The Grammar of Visual Design. In *Reading Images* (3rd ed.). Routledge. https://doi.org/10.4324/9781003099857

Laadem, M., & Mallahi, H. (2020). Multimodal Pedagogies in Teaching English for Specific Purposes in Higher Education: Perceptions,

Challenges and Strategies. *International Journal on Studies in Education*, *1*(1). https://doi.org/10.46328/ijonse.3

Liaw, M. L., & Chen, H. I. (2023). Guest editorial: Contextualized multimodal language learning. *Educational Technology and Society*, *26*(3). https://doi.org/10.30191/ETS.202307_26(3).0001

Ljalikova, A., Meristo, M., Alas, E., & Jung, M. (2021). Narrative analysis as a means of investigating CLIL teachers' meaningful experiences. *Qualitative Research in Education*, *10*(3). https://doi.org/10.17583/qre.7511

Makhmudov, K. (2023). Bridging cultures through English language education: A comprehensive model for intercultural communication competence development. *Theoretical & Applied Science*, *119*(03). https://doi.org/10.15863/tas.2023.03.119.27

Morgenstern, A. (2023). Children's multimodal language development from an interactional, usage-based, and cognitive perspective. In *Wiley Interdisciplinary Reviews: Cognitive Science* (Vol. 14, Issue 2). https://doi.org/10.1002/wcs.1631

Muntaha, M., Chen, J., & Dobinson, T. (2023). Exploring students' experiences of using multimodal CMC tasks for English communication: A case with Instagram. *Educational Technology and Society*, *26*(3). https://doi.org/10.30191/ETS.202307_26(3).0006

Nfor, S. (2020). Multimodal scaffolding teaching: role-taking or role-creating in an English as a Foreign Language (EFL) oral communication class in Japan. *Scenario: A Journal of Performative Teaching, Learning, Research*, *XIV*(1). https://doi.org/10.33178/scenario.14.1.3

O'Halloran, K. L. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language

and visual imagery. *Visual Communication*, *7*(4). https://doi.org/10.1177/1470357208096210

Owoyele, B., Trujillo, J., de Melo, G., & Pouw, W. (2022). Masked-Piper: Masking personal identities in visual recordings while preserving multimodal information. *SoftwareX*, *20*. https://doi.org/10.1016/j.softx.2022.101236

Pham, Q. N., & Li, M. (2023). Digital multimodal composing using Visme: EFL students' perspectives. *Asia-Pacific Education Researcher*, *32*(5). https://doi.org/10.1007/s40299-022-00687-w

Ramírez Mera, U. N., & Tur, G. (2021). Supporting students' control and ownership of learning in the multimodal learning from a PLE perspective. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3486011.3486502

Ross, J., Curwood, J. S., & Bell, A. (2020). A multimodal assessment framework for higher education. *E-Learning and Digital Media*, *17*(4). https://doi.org/10.1177/2042753020927201

Senaratne, H., Oviatt, S., Ellis, K., & Melvin, G. (2022). A critical review of multimodal-multisensor analytics for anxiety assessment. In *ACM Transactions on Computing for Healthcare* (Vol. 3, Issue 4). https://doi.org/10.1145/3556980

Twiner, A., Major, L., & Wegerif, R. (2022). School-based simulated internships to support dialogic collaboration and authentic links with the world of work: A design-based research study. *Irish Educational Studies*, *41*(1). https://doi.org/10.1080/03323315.2021.2022515

Unsworth, L., Cope, J., & Nicholls, L. (2019). Multimodal literacy and large-scale literacy tests: Curriculum relevance and responsibility. *Australian Journal of Language and Literacy*, *42*(2). https://doi.org/10.1007/BF03652032

Xiong, T., Feng, D., & Hu, G. (2022). Researching cultural knowledge and values in English language teaching textbooks: Representation, multimodality, and stakeholders. In *Cultural Knowledge and Values in English Language Teaching Materials: (Multimodal) Representations and Stakeholders*. https://doi.org/10.1007/978-981-19-1935-0_1

Xu, Y. (2023). Investigating the effects of digital multimodal composing on Chinese EFL learners' writing performance: A quasi-experimental study. *Computer Assisted Language Learning*, *36*(4). https://doi.org/10.1080/09588221.2021.1945635

Yang, W. (2020). The development, adoption and evaluation of the integration of an ESP and CLIL textbook: Perspectives from the CLIL learners. *ESP Today*, *8*(1). https://doi.org/10.18485/esptoday.2020.8.1.4

Zhyrun, I. (2016). Culture through comparison: creating audio-visual listening materials for a CLIL course. *Latin American Journal of Content & Language Integrated Learning*, *9*(2). https://doi.org/10.5294/laclil.2016.9.2.5

Zou, J. H., Ramirez, S. I. M., Erazo, M. A. C., & Encalada, M. A. R. (2019). The significance of English language development for future Asian tourism and hotel management professionals. *Theory and Practice in Language Studies*, *9*(1). https://doi.org/10.17507/tpls.0901.02