Al-Ishlah: Jurnal Pendidikan
Vol.18, 1 (March, 2026), pp. 566-580
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v18i1.9460

Integrating YouTube into EAP Pedagogy to Foster Students' Academic Speaking Competence in EFL Classrooms

Ida Maulida

Universitas Subang, Subang, Indonesia; idamaulida@unsub.ac.id

ARTICLE INFO

Keywords:

YouTube;
English for Academic Purposes (EAP);
academic speaking;
CEFR;
multimodal learning

Article history:

Received 2026-01-22
Revised 2026-02-23
Accepted 2026-03-31

ABSTRACT

The increasing demand for English for Academic Purposes (EAP) proficiency requires EFL learners to develop advanced academic speaking skills. However, limited exposure to authentic academic discourse remains a major challenge. Digital platforms such as YouTube may provide multimodal input to support speaking development. This study examines whether integrating YouTube into EAP pedagogy improves students' academic speaking competence. A quasi-experimental pre-test–post-test non-equivalent control group design was used with N = 14 university EFL students (n = 7 per group) over 8 sessions (one semester). Data were collected through CEFR-based speaking tests, audio/video recordings, and rater scores, with an additional Likert-scale questionnaire for the experimental group. Quantitative data were analyzed using paired- and independent-samples t-tests, and N-gain scores; qualitative support came from learner perception data. Both groups improved significantly; however, the experimental group showed larger gains (M_pre = 61.14, SD = 2.41; M_post = 81.57, SD = 2.64) than the control group (M_pre = 61.00, SD = 2.16; M_post = 68.57, SD = 1.72). Paired t-tests indicated significant improvement in the experimental group, $t(6) = -55.38$, $p < .001$, with a large effect. N-gain scores were moderate for the experimental group (52.75%) and low for the control group (19.41%). Questionnaire results indicated positive perceptions (M ≈ 4.43–4.71). Findings suggest that YouTube integration may enhance academic speaking through authentic input and multimodal learning. Pedagogically, structured video-based tasks are recommended in EAP instruction. However, results should be interpreted cautiously due to the small sample size and non-random assignment.

Corresponding Author:
Ida Maulida
Universitas Subang, Subang, Indonesia; idamaulida@unsub.ac.id

1. INTRODUCTION

In higher education, English for Academic Purposes (EAP) plays an important role in preparing students to participate effectively in academic communication. EAP is not only concerned with general English proficiency, but also with the ability to use language appropriately in academic contexts, including presenting ideas, developing arguments, and communicating knowledge in a clear and

organized way (Hyland, 2006; Jordan, 2020). Among the language skills required in EAP, academic speaking is particularly significant because students are often expected to deliver presentations, participate in discussions, and explain ideas using formal and discipline-appropriate language.

However, developing academic speaking competence remains challenging for many learners in English as a Foreign Language (EFL) contexts. In many classrooms, students have limited opportunities to interact with authentic academic discourse and to observe how effective academic speaking is performed in real settings. As a result, they may struggle with fluency, organization, academic vocabulary, and the use of appropriate discourse markers in oral communication (Hamp-Lyons, 2007; Hyland, 2006). These difficulties become more apparent when students are required to speak in a structured academic format rather than in everyday conversational English.

The increasing use of digital technology in education has opened new possibilities for addressing these challenges. Multimedia resources can provide learners with richer language input by combining visual, auditory, and textual information. According to Mayer's (2009) cognitive theory of multimedia learning, such multimodal input can support comprehension and retention because learners process information through multiple channels. In language learning, this is particularly useful because students can observe not only what is said, but also how it is delivered through pronunciation, intonation, gesture, and overall discourse organization.

One digital platform that has attracted considerable attention in language education is YouTube. As a widely accessible video-sharing platform, YouTube offers abundant authentic materials, including academic lectures, presentations, interviews, and conference talks. These resources may help learners gain greater exposure to real language use and observe models of effective oral communication in academic settings. Previous studies have suggested that YouTube can support language learning by increasing learner motivation, improving comprehension, and providing models for speaking practice (Alwehaibi, 2015; Saed et al., 2021; Truong & Le, 2022). In addition, students often perceive YouTube as engaging and useful for improving their speaking ability because it allows repeated viewing and flexible access to learning materials (Budiyanti, 2023; Musdayanti et al., 2024).

From a theoretical perspective, the integration of YouTube into EAP instruction can be explained through several perspectives in second language acquisition. Krashen's (1985) Input Hypothesis emphasizes the importance of comprehensible input in language development, while Schmidt's (1990) Noticing Hypothesis highlights the role of conscious attention to linguistic features in acquisition. Through YouTube videos, learners can be exposed to meaningful and contextualized academic input and may notice important elements such as academic expressions, discourse markers, and presentation strategies. Furthermore, Swain's (1985) Output Hypothesis argues that language development is strengthened when learners are encouraged to produce language. When YouTube is integrated into structured classroom activities, students are not only exposed to input but are also guided to respond through speaking tasks, presentations, and reflection.

The use of YouTube in EAP classrooms is also closely related to task-based language teaching (TBLT), which emphasizes meaningful communication and active learner engagement through authentic tasks (Ellis, 2003; Nunan, 1999). In this context, learners can watch academic videos, analyze their structure and language features, summarize key ideas, and produce their own oral presentations based on the models they observe. Such activities may help bridge the gap between receptive exposure and productive performance, especially in developing academic speaking competence.

Although many studies have explored the use of YouTube in improving speaking skills, research specifically examining its role in EAP pedagogy remains limited. Much of the existing literature focuses on general speaking development, learner perceptions, or online learning contexts, rather than on academic speaking competence within a structured EAP course (Ilyas & Putri, 2020; Rohmah, 2022; Zulaefa & Rizal, 2023). Therefore, more empirical research is needed to investigate whether the integration of YouTube into EAP instruction can contribute meaningfully to students' academic speaking development.

This study aims to address that gap by examining the effect of YouTube-integrated instruction on students' academic speaking competence in an EFL classroom. Specifically, the study compares students who receive YouTube-based EAP instruction with those who experience conventional teaching. In addition, it explores students' perceptions of the use of YouTube as a learning medium. By doing so, this study is expected to contribute to the growing body of research on technology-enhanced language learning and to provide practical implications for EAP teachers seeking more engaging and authentic ways to support students' academic speaking development.

2. METHODS

2.1 Research Design

This study employed a quasi-experimental mixed-method design to investigate the effect of YouTube-integrated instruction on students' academic speaking competence in an English for Academic Purposes (EAP) course. Specifically, a pre-test–post-test non-equivalent control group design was used, as intact classes could not be randomly assigned due to institutional constraints.

The integration of quantitative and qualitative data was intended to provide a more comprehensive understanding of both learning outcomes and learner experiences. The quantitative component measured changes in speaking performance, while the qualitative component (questionnaire and classroom observations) explored students' perceptions of the instructional intervention.

Given the exploratory nature of the study and the limited number of participants, this research is positioned as a pilot study, aiming to identify preliminary trends rather than to produce generalizable conclusions.

2.2 Participants and Sampling

The study involved 14 first-year undergraduate students enrolled in an EAP course at a university in Indonesia, divided into:

a. Experimental group (n = 7)
b. Control group (n = 7)

Participants were selected using purposive cluster sampling, based on intact classroom grouping. Random assignment was not feasible due to administrative constraints.

All participants had relatively comparable English proficiency levels (B1–B2 CEFR), as indicated by the institutional placement test. None had received prior formal training in academic speaking.

Although the sample size is small, it is considered acceptable for exploratory classroom-based research. However, this limitation is acknowledged and addressed in the interpretation of results, particularly regarding statistical power and generalizability.

2.3 Instructional Treatment

The intervention was conducted over 8 instructional sessions within one academic semester.

2.3.1 Experimental Group (YouTube-Integrated EAP Instruction)

Students in the experimental group engaged in structured YouTube-based learning activities designed according to multimodal learning and task-based language teaching principles. The instructional procedure included:

1. Video Exposure
   Students watched selected academic videos (e.g., lectures, presentations, TED Talks) aligned with EAP topics.
2. Guided Analysis
   Students analyzed:

a. discourse structure (introduction–body–conclusion)
b. academic vocabulary and expressions
c. use of discourse markers
d. delivery features (intonation, fluency, gestures)

3. Language Noticing and Modeling
   Students identified useful phrases and practiced imitating academic language patterns.
4. Production Tasks
   Students performed structured speaking tasks, including:
   a. mini-presentations
   b. recorded academic talks
   c. peer feedback sessions
5. Reflection and Feedback
   Students received feedback from peers and the instructor to improve performance.

2.3.2 Control Group (Conventional Instruction)

The control group received traditional EAP instruction without digital media integration. Activities included:

a. textbook-based instruction
b. teacher explanation
c. limited speaking practice without video modeling

2.4 Instruments

2.4.1 Academic Speaking Test

Students completed a pre-test and post-test in the form of a short academic presentation (2–3 minutes) on comparable topics.

Speaking performances were assessed using an analytical rubric adapted from:

a. CEFR Oral Production Descriptors (B2–C1)
b. EAP Oral Presentation Criteria (Hamp-Lyons & Hyland, 2002)

The rubric included five components:

1. Fluency
2. Grammatical accuracy
3. Lexical sophistication
4. Organization and coherence
5. Use of academic discourse markers

Each component was scored on a 5-point scale.

2.4.2 Inter-Rater Reliability

To ensure the reliability of the speaking assessment, inter-rater reliability was examined using the Intraclass Correlation Coefficient (ICC). The analysis yielded an ICC value of 0.80, indicating a good level of agreement between the two raters. According to commonly accepted benchmarks, ICC values above 0.75 are considered acceptable, while values above 0.80 reflect good reliability. This result suggests that both raters applied the assessment rubric consistently across all speaking performances. Given the subjective nature of speaking evaluation and the relatively small sample size of this study, achieving an ICC of 0.80 provides sufficient evidence that the scoring process was reliable. Therefore, the speaking test results can be considered dependable and suitable for subsequent statistical analysis.

2.4.3 Student Perception Questionnaire

A Likert-scale questionnaire (1–5) was administered to the experimental group to examine:

a. learning motivation

b. comprehension of academic input
c. perceived usefulness of YouTube

The instrument was adapted from previous validated studies and reviewed by two EAP experts to ensure content validity.

2.5 Data Collection Procedure

The study followed these steps:

1. Pre-test
   Both groups performed an academic speaking task.
2. Intervention (8 sessions)
   a. Experimental group: YouTube-integrated instruction
   b. Control group: conventional instruction
3. Post-test
   Both groups completed a parallel speaking task.
4. Questionnaire Administration
   Conducted only in the experimental group.
5. Classroom Observation
   Observational notes were recorded to capture student engagement and interaction.

2.6 Data Analysis

2.6.1 Quantitative Analysis

Data were analyzed using SPSS with the following procedures:

1. Descriptive Statistics
   Mean, standard deviation, minimum, and maximum scores were calculated.
2. Assumption Testing
   a. Normality: Shapiro–Wilk test
   b. Homogeneity: Levene's test
3. Inferential Statistics
   a. Paired-samples t-test: within-group comparison
   b. Independent-samples t-test: between-group comparison
4. Effect Size
   To complement significance testing, Cohen's d was calculated to determine the magnitude of the treatment effect.
5. Normalized Gain (N-Gain)
   Used to evaluate learning improvement relative to initial performance.

2.6.2 Qualitative Analysis

Questionnaire data were analyzed descriptively using mean scores to capture students' overall perceptions of the instructional intervention. In addition, observational data collected during classroom activities were used to complement and support the interpretation of the quantitative results. To enhance the validity and depth of the analysis, data from multiple sources—including speaking test results, questionnaire responses, and classroom observations—were triangulated, allowing for a more comprehensive and robust understanding of the research findings.

3. FINDINGS AND DISCUSSION

3.1 Findings

This section presents the findings of the study on the effect of YouTube-integrated instruction on students' academic speaking competence in an EAP classroom. The findings are organized into four parts: the implementation of the instructional treatment, students' speaking performance in the pre-test and post-test, the statistical results, and students' perceptions of the use of YouTube in EAP learning.

3.1.1 Implementation of the YouTube-Integrated Instruction

The instructional treatment in the experimental group was designed to strengthen students' academic speaking competence through multimodal learning by integrating YouTube-based materials into EAP pedagogy. Across eight sessions, students were exposed to authentic academic videos such as lectures, presentation models, and public speaking materials. These videos were used to help students observe the structure of academic discourse, identify academic vocabulary and discourse markers, and notice delivery features such as intonation, fluency, and gesture. In addition, students completed structured speaking tasks, including mini-presentations, recorded academic talks, peer feedback, and reflection activities.

Table 1. Meeting Material/Topic

Meet	Topics	Learning Objectives	Activities and Assignments	Resources / YouTube Links
1.	Introduction to EAP: Characteristics of Academic Language & Purpose of EAP	Students understand the concept of EAP, the characteristics of academic language, and the purpose of developing academic skills	• Watch the introductory video of EAP • Class discussion: characteristics of academic language • Analyze examples of simple academic texts	https://youtu.be/pH0wFxgcvoI
2.	Basics of Academic Presentation	Students understand the structure of academic presentations: opening, body, closing, and presentation planning	• Watch the Presentation Skills 1 video • Identify the presentation structure in the video • Practice making a short presentation outline	https://youtu.be/gmfc0rH9Lpg
3.	Listening & Analyzing Real Academic	Lectures Students are able to recognize academic language features and information delivery strategies in real lectures	• Watch real MIT/Open Course Ware lectures - Discussion: academic expressions, cohesive devices, intonation • Note down important academic vocabulary and phrases	https://youtu.be/Unzc731iCUY
4.	Public Speaking & Delivery	Students develop delivery skills: intonation, clarity, audience engagement	• Watch public speaking videos - Practice pronunciation, intonation, gestures • Mini presentation simulation 1-2 minutes	https://youtu.be/82TU1epxJLU
5.	Creating an Effective Academic Presentation	Students are able to create a structured and interesting academic	• Watch Presenting Effectively • Practice creating slides and presentation scripts	https://youtu.be/qFLL-XB56UU

			• Peer review of presentation outlines	
6.	Exposure to Scientific Lectures & Higher Academic Register	Students recognize scientific academic register, formal vocabulary, and technical phrases	• Watch scientific lectures • Identify academic phrases and scientific register • Practice imitating the use of language in presentations	https://youtu.be/lcdbQSAkTHk
7.	Academic Presentation Production	Students are able to create and record mini academic presentations	• Students choose relevant topics • Compile presentations based on video models • Record their own presentations for evaluation	using reference videos from Sessions 2, 3, 4, and 6 as models
8.	Academic Presentation Evaluation & Reflection	Students receive feedback, reflection, and presentation improvement strategies	• Peer feedback on presentations • Watch EAP playlist for review • Discussion: difficulties, delivery improvement strategies, and structure	https://youtu.be/EszM4P1dMCg

As shown in Table 1, the instructional sequence was systematically designed to scaffold students' academic speaking development across eight sessions. The progression begins with foundational knowledge of EAP and gradually moves toward more complex skills, such as analyzing authentic lectures, developing presentation techniques, and producing independent academic presentations. The integration of YouTube videos in each session provided consistent exposure to authentic language use and multimodal input, which supported students' ability to notice discourse patterns, imitate academic expressions, and refine their delivery skills. Furthermore, the inclusion of reflective and peer-feedback activities allowed students to consolidate their learning and improve their performance iteratively. Overall, this structured and video-supported approach was intended to create a coherent learning experience that aligns with the objectives of EAP instruction.

3.1.2 Pre-Test and Post-Test Results

The speaking test used in both the pre-test and post-test required students to perform a continuous monologue on the topic of environmental challenges in their country. The performances were assessed using a CEFR-based scoring rubric.

Table 2 shows the score distribution of the experimental group. In the pre-test, all seven students in the experimental group were at the B1 level. After the treatment, their speaking performance improved: four students reached the C1 level, while three students achieved the B2–C1 level in the post-test. This result indicates substantial progress in the experimental group after receiving YouTube-integrated EAP instruction.

Table 2. Score Level Pre and Post test Experiment Test

Students	Pre-Test	CEFR Pre	Post-Test	CEFR Post
1	62	B1	82	C1
2	58	B1	79	B2–C1
3	65	B1	85	C1
4	60	B1	80	B2–C1
5	63	B1	84	C1
6	59	B1	78	B2–C1
7	61	B1	83	C1

Table 3 presents the score distribution of the control group. Similar to the experimental group, all seven students in the control group were at the B1 level in the pre-test. In the post-test, all seven students improved to the B2–C1 level. This finding shows that the control group also experienced improvement, although the degree of progress was less pronounced than that of the experimental group.

Table 3. Score Level Pre and Post test Control Test

Students	Pre-Test	CEFR Pre	Post-Test	CEFR Post
1	63	B1	70	B2–C1
2	60	B1	68	B2–C1
3	62	B1	69	B2–C1
4	59	B1	67	B2–C1
5	64	B1	71	B2–C1
6	61	B1	69	B2–C1
7	58	B1	66	B2–C1

3.1.3 Descriptive Statistics

The descriptive statistics in Table 4 show that the initial speaking ability of the two groups was relatively comparable. The experimental group obtained a mean pre-test score of 61.14 (SD = 2.41), while the control group obtained a mean pre-test score of 61.00 (SD = 2.16). These results suggest that both groups started from a similar level of speaking competence before the intervention.

Table 4. Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Pre_Test_Experiment	7	58	65	61.14	2.41
Post_Test_Experiment	7	78	85	81.57	2.64
Pre_Test_Control	7	58	64	61.00	2.16
Post_Test_Control	7	66	71	68.57	1.72
Valid N (listwise)	7				

After the intervention, the experimental group showed a marked increase in performance, with a mean post-test score of 81.57 (SD = 2.64). This indicates an average gain of 20.43 points from pre-test to post-test. In contrast, the control group reached a mean post-test score of 68.57 (SD = 1.72), representing an average gain of 7.57 points. These results suggest that although both groups improved, the experimental group demonstrated a considerably greater increase in academic speaking performance than the control group.

The relatively small standard deviations in both groups indicate that the scores were not widely dispersed, suggesting a fairly consistent pattern of performance within each group. However, given the small sample size, these results should still be interpreted cautiously.

3.1.4 Test of Normality

Before conducting inferential statistics, a normality test was performed using the Shapiro–Wilk test, which is appropriate for small sample sizes. As shown in Table 5, the significance values for all four datasets were above 0.05: 0.976, 0.949, 0.744, and 0.958. These results indicate that the data were normally distributed and that the assumption of normality was met. Therefore, parametric statistical tests, including the paired-samples t-test and independent-samples t-test, were considered appropriate for further analysis.

Tabel 5. Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Class	Statistic	df	Sig.	Statistic	df	Sig.
Score	1	.111	7	.200*	.984	7	.976
	2	.108	7	.200*	.978	7	.949
	3	.153	7	.200*	.952	7	.744
	4	.170	7	.200*	.980	7	.958

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

The normality test was used to determine whether the data in each group (Class 1 to Class 4) was normally distributed. This normality is important for determining whether researchers can use parametric statistical tests, such as the t-test or ANOVA. This study focused on the Shapiro–Wilk (S–W) test. This is because it is more accurate for small samples (N < 50). In addition, in this study, the sample size was only 7, making the Shapiro–Wilk test more accurate. Based on the results of the normality tests using the Kolmogorov–Smirnov and Shapiro–Wilk tests, all data in the four groups showed a significance value (p-value) above 0.05. In the Shapiro–Wilk test, which is more relevant due to the small sample size (N = 7), the p-values for Class 1 to Class 4 were 0.976, 0.949, 0.744, and 0.958, respectively. All these values are greater than the significance limit of 0.05, so it can be concluded that there is no significant deviation from the normal distribution across all groups. In other words, the data in Class 1, Class 2, Class 3, and Class 4 are all normally distributed. This finding confirms that the assumption of normality has been met, so researchers can appropriately use parametric statistical tests, such as the paired sample t-test to test the difference between the pre-test and post-test within the same group, and the independent sample t-test to compare the results between the experimental and control groups. The met normality also strengthens the validity of the inferential analysis conducted in this study. Based on the output above, the significance value (Sig.) for all data in both the Kolmogorov-Smirnov test and the Shapiro-Wilk test is > 0.05, so it can be concluded that the research data is normally distributed. The research data is normally distributed, so we can use parametric statistics (paired sample t-test and independent sample t-test) to analyse the research data.

3.1.5 Paired-Samples t-Test

The paired-samples t-test results in Table 6 indicate that both groups showed statistically significant improvement from pre-test to post-test. In the experimental group, the mean difference between the pre-test and post-test scores was –20.429, with $t(6) = –55.384$, $p < .001$. In the control group, the mean difference was –7.571, with $t(6) = –37.477$, $p < .001$. These findings indicate that both instructional approaches contributed to improvement in students' speaking performance.

However, the magnitude of improvement in the experimental group was substantially greater than that in the control group. This pattern is also visible in the mean score gains reported in Table 4. Therefore, although conventional instruction was associated with some improvement, the YouTube-integrated treatment appeared to support greater development in students' academic speaking competence. Because of the very small sample size and the limited score variability, the unusually large t-values should be interpreted with caution.

Table 6. Paired Samples Test

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Pre_Test_Experiment - Post_Test_Experiment	-20.429	.976	.369	-21.331	-19.526	-55.384	6	.000
Pair 2	Pre_Test_Control - Post_Test_Control	-7.571	.535	.202	-8.066	-7.077	-37.477	6	.000

The results of the paired-samples t-test indicate that both the experimental and control groups experienced statistically significant improvements in their academic speaking performance ($p < 0.05$). In the experimental group, the mean score increased by 20.43 points from pre-test to post-test. The relatively low standard deviation (SD = 0.976) and standard error (SE = 0.369) suggest that the improvement was consistent across participants. The 95% confidence interval (–21.33 to –19.53) further confirms that the observed gain was statistically meaningful.

Similarly, the control group also demonstrated a significant improvement, with a mean increase of 7.57 points. The consistency of this improvement is reflected in the low standard deviation (SD = 0.535) and standard error (SE = 0.202), with a 95% confidence interval ranging from –8.07 to –7.08. However, the magnitude of improvement in the control group was substantially smaller compared to that of the experimental group.

Although both groups showed statistically significant gains, the greater increase observed in the experimental group suggests that the YouTube-integrated instruction may have contributed more effectively to the development of students' academic speaking skills. Nevertheless, the unusually large t-values obtained in this study should be interpreted with caution, as they may be influenced by the small sample size and limited variability in the data.

Overall, these findings indicate that while conventional instruction can improve students' academic speaking ability, the integration of YouTube appears to provide additional benefits, particularly in enhancing fluency, academic language use, and confidence in oral performance.

3.1.6 Homogeneity of Variance

Table 7 shows the results of Levene's test for homogeneity of variance. The significance value based on the mean was 0.145, which is greater than 0.05. The other values based on the median, adjusted median, and trimmed mean were also above 0.05. These findings indicate that the variance between the groups was homogeneous. Therefore, the assumption of homogeneity of variance was met, and the use of parametric comparison procedures was acceptable.

Table 7. Test of Homogeneity of Variance

		Levene Statistic	**df1**	**df2**	**Sig.**
Result_Students	Based on Mean	2.425	1	12	.145
	Based on Median	1.687	1	12	.218
	Based on Median and with adjusted df	1.687	1	11.592	.219
	Based on trimmed mean	2.414	1	12	.146

Based on the results of the Test of Homogeneity of Variance (Levene's Test), the significance values obtained for all calculation methods are above 0.05, namely Based on Mean of 0.145, Based on Median of 0.218, Based on Median with adjusted df of 0.219, and Based on Trimmed Mean of 0.146. Since all Sig. values are > 0.05, it can be concluded that the variance between the experimental group and the control group is homogeneous or not significantly different. This condition indicates that the data meet the assumption of homogeneity of variance, which is the main requirement in the use of parametric tests. With the fulfillment of this assumption, the Independent Samples T-Test analysis can be continued using the Equal Variances Assumed model, because the variance of the two groups is considered the same by the statistical system. Based on the output above, the Significance (Sig.) Based on the mean value, which is 0.145> 0.05, it can be concluded that the variance of the Post-test data of the experimental class and the Post-test data of the control class are the same or homogeneous. Thus, one of the (not absolute) requirements of the independent sample t-test has been met.

3.1.7 N-Gain Analysis

The N-gain analysis was used to examine the relative improvement in students' learning outcomes. As shown in Table 9, the experimental group obtained an average N-gain score of 52.75%, which falls into the medium category. Meanwhile, the control group obtained an average N-gain score of 19.41%, which is categorized as low.

This result indicates that the improvement in the experimental group was more meaningful than that in the control group. Although the gain in the experimental group was not classified as high, it still suggests that the YouTube-integrated EAP instruction was more effective than conventional instruction in improving students' academic speaking performance.

Table 8. N-Gain Category

N-Gain Range (g)	Category	Meaning Learning Effectiveness
g ≥ 0,70	High	Very effective learning improves learning outcomes.
0,30 ≤ g < 0,70	Medium	Moderately effective learning improves learning outcomes.
g < 0,30	Low	Learning is less or ineffective.

Table 9. N-Gain Score Test calculation results

No	Class Experiment	Class Control
	N-Gain Score (%)	N-Gain Score (%)
1.	53	19
2.	50	20
3.	57	18
4.	50	20
5.	57	19
6.	46	21
7.	56	19
Average	52.75	19.41
Minimum	46	18
Maximum	57	21

The N-gain score test results show that the average N-gain score for the experimental class (learning using YouTube-based EAP Pedagogy) was 52.75%, which falls into the medium gain category. The N-gain score in the experimental class was a minimum of 46 and a maximum of 57, indicating a consistent and significant improvement in students' academic speaking skills in this class. Meanwhile, the average N-gain score for the control class (conventional learning without YouTube integration) was 19.41%, which falls into the low-gain category. The N-gain score in the control class was a minimum of 18 and a maximum of 21, indicating that the improvement in students' speaking skills in this group was very limited. Thus, it can be concluded that the use of YouTube as a medium in English for Academic Purposes (EAP) pedagogy has proven more effective in improving students'

academic speaking skills compared to conventional learning methods. Although the improvement in the experimental class was in the moderate category, these results still show that the integration of digital videos through YouTube makes a greater contribution to the development of academic speaking skills than traditional learning that is not based on digital media.

3.1.8 Students' Perceptions of YouTube Integration

The questionnaire results from the experimental group are presented in Table 9. Overall, students responded positively to the use of YouTube in EAP learning. The mean scores for the five questionnaire items ranged from 4.43 to 4.71, indicating that most students selected Agree or Strongly Agree. Specifically, students reported that YouTube helped them understand academic speaking materials, increased their motivation to practice speaking, improved their confidence in speaking in class, provided relevant academic content, and effectively supported the improvement of their academic speaking skills.

These results suggest that students perceived YouTube not only as an engaging learning medium but also as a useful resource for developing academic speaking competence in the EAP classroom.

Table 10. Questionnaire Result

No	Questionnaire Statement	R1	R2	R3	R4	R5	R6	R7	Mean
1	Using YouTube in EAP learning helps me understand academic speaking materials.	4	5	4	5	4	4	5	4.43
2	YouTube videos motivate me to practice speaking English more actively	4	4	5	4	5	4	5	4.43
3	The integration of YouTube makes me more confident when speaking in class	4	5	4	4	5	4	5	4.43
4	The content of the YouTube videos used is relevant to the academic topics being studied	5	4	4	5	4	5	4	4.43
5	YouTube-based learning is effective in improving my academic speaking skills.	5	5	4	5	4	5	5	4.71

Based on the table above, the majority of responses fell on a scale of 4 (Agree) and 5 (Strongly Agree). The average score for each statement ranged from 4.43 to 4.71, indicating that respondents agreed to strongly agreed that YouTube integration helps Understand academic speaking material, increase motivation for speaking practice, increase confidence when speaking, provide relevant content, Effective in improving academic speaking skills, in conclusion, the questionnaire results indicate positive student support for the use of YouTube in EAP learning to improve academic speaking competence.

Overall, the findings show that both the experimental and control groups improved from pre-test to post-test. However, the experimental group, which received YouTube-integrated EAP instruction, demonstrated greater improvement than the control group in terms of mean score gain, post-test performance, and N-gain category. In addition, students in the experimental group reported positive perceptions of the use of YouTube as a learning tool. Taken together, these findings suggest that integrating YouTube into EAP pedagogy may provide meaningful support for the development of students' academic speaking competence. Nevertheless, the findings should be interpreted carefully due to the small sample size and the non-randomized design of the study.

3.2 Discussion

The findings of this study indicate that both instructional approaches contributed to improvements in students' academic speaking competence; however, the YouTube-integrated instruction resulted in substantially greater gains. This result can be interpreted through several theoretical perspectives in second language acquisition and EAP pedagogy.

First, the improvement observed in the experimental group supports the principles of multimodal learning theory, which emphasizes that learning is enhanced when information is presented through multiple modes, such as visual, auditory, and textual input. The use of YouTube videos provided students with rich multimodal input, allowing them to simultaneously process linguistic forms, discourse structures, and paralinguistic features such as intonation and gesture. This is consistent with Mayer's (2009) cognitive theory of multimedia learning, which posits that dual-channel processing enhances understanding and retention of information.

Second, the findings align with input and noticing hypotheses in second language acquisition. Authentic video materials expose learners to real academic discourse, enabling them to notice salient linguistic features such as academic vocabulary and discourse markers. According to Schmidt (1990), conscious noticing is a crucial condition for language acquisition, while Krashen (1985) emphasizes the importance of comprehensible input in facilitating language development. The availability of authentic input through YouTube may therefore explain the greater improvement observed in the experimental group.

Furthermore, the results can be interpreted through a task-based language teaching (TBLT) perspective. Task-based learning emphasizes meaningful communication and active learner engagement in completing tasks (Ellis, 2003; Nunan, 2004). In this study, students engaged in activities such as analyzing video content, practicing presentations, and receiving feedback, which required them to actively process input and produce output. This aligns with Swain's (1985) output hypothesis, which highlights the importance of language production in developing communicative competence.

In addition, the findings highlight the role of affective factors, particularly motivation and confidence, in language learning. The positive responses from the questionnaire suggest that YouTube-based learning increased students' engagement and willingness to practice speaking. This supports Dörnyei's (2001) theory of motivation, which emphasizes that engaging and relevant learning environments can enhance learners' performance. The use of authentic and accessible video content may also reduce anxiety and create a more supportive learning atmosphere.

Despite these promising findings, the results should be interpreted with caution. The small sample size and limited variability in the data may have influenced the statistical outcomes, particularly the magnitude of effect sizes. In addition, the non-randomized design limits the ability to establish causal relationships. As noted by Creswell (2014), quasi-experimental studies provide useful insights but require careful interpretation due to potential threats to internal validity.

Overall, this study contributes to the growing body of research on technology-enhanced language learning (TELL) by providing empirical support for the integration of YouTube in EAP instruction. The findings suggest that combining authentic multimodal input with structured speaking tasks can create a more effective learning environment for developing academic speaking skills. Future research is recommended to involve larger samples, more rigorous experimental designs, and additional variables such as learner autonomy and interaction patterns to further validate these findings.

4. CONCLUSION

This study suggests that the integration of YouTube into English for Academic Purposes (EAP) instruction is associated with greater improvement in students' academic speaking competence compared to conventional teaching methods, as evidenced by higher mean score gains and moderate N-gain results in the experimental group. The use of multimodal and authentic video input appears to support students' development in areas such as fluency, academic vocabulary, and presentation

confidence. However, these findings should be interpreted with caution due to several limitations, including the small sample size, limited variability in scores, and the use of a non-randomized quasi-experimental design, which may affect the generalizability and robustness of the results. Therefore, future research is recommended to involve larger and more diverse samples, apply more rigorous experimental designs, and explore additional variables such as learner autonomy, interaction patterns, and long-term retention in order to provide more comprehensive evidence on the effectiveness of YouTube-based EAP instruction.

REFERENCES

Abidin, Z. (2021). *The effect of YouTube on speaking ability of senior high school students*, 49(2).

Alwehaibi, H. O. (2015). The impact of using YouTube in EFL classroom on enhancing EFL students' content learning. *Journal of College Teaching & Learning*, *12*(2), 121–126. https://doi.org/10.19030/tlc.v12i2.9182

Amal, M. I., Yamin, M., & Rosalina, E. (2025). The effectiveness of using YouTube as a medium for teaching EFL in improving students' speaking skills: A systematic review, *8*(2).

Budiyanti, K. (2023). EFL students' perception on the use of YouTube videos to encourage their speaking skill. *Indonesian Journal of Integrated English Language Teaching*, *8*(2), 40. https://doi.org/10.24014/ijielt.v8i2.21093

Hamp-Lyons, L. (2007). The impact of EAP oral assessment criteria on teaching and learning. *Journal of English for Academic Purposes*, *6*(3), 152–168.

Hamp-Lyons, L., & Hyland, K. (2002). Assessing second language writing. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 145–170). Cambridge University Press.

Heigham, J., & Croker, R. A. (Eds.). (2009). *Qualitative research in applied linguistics: A practical introduction*. Palgrave Macmillan.

Hismanoglu, M. (2012). The impact of globalization and information technology on language education policy in Turkey. *Procedia - Social and Behavioral Sciences*, *31*, 629–633. https://doi.org/10.1016/j.sbspro.2011.12.115

Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.

Ilyas, M., & Putri, M. E. (2020). YouTube channel: An alternative social media to enhance EFL students' speaking skill. *J-SHMIC: Journal of English for Academic*, *7*(1), 77–87. https://doi.org/10.25299/jshmic.2020.vol7(1).4141

Jalaluddin, M. (2016). *Using YouTube to enhance speaking skills in ESL classroom*.

Jordan, R. R. (2020). *Academic speaking: A handbook for international students*. Routledge.

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer.

Khoiroh, S. A. (2021). Using YouTube for speaking in online learning: EFL students' perception and difficulties, *9*(2).

MacIntyre, P. D., Baker, S. C., Clément, R., & Donovan, L. A. (2017). Fluency in L2 speech production. *Journal of Second Language Speaking Research*, *3*(1), 45–68.

Maulida, I. (2021). Implementation of English learning using speech recognition technology. *MBI Journal*. https://binapatria.id/index.php/MBI/article/view/423

Maulida, I. (2024). The need analysis of material design in English teaching on ESP program, *18*(11).

Maulida, I., Aminah, M., & Hermawan, V. S. (2023). The effect of students' motivation in the ESP classroom through learning English achievement on students' computer science.

Musdayanti, Sudewi, P. W., & Nurhaeni. (2024). YouTube for EFL students: Perceptions on building speaking confidence. *GLENS: Global English Insights Journal*, *1*(2), 89–97. https://doi.org/10.61220/glens.v1i2.350

Muslem, A. (2021). Students' perception toward oral corrective feedback in speaking classes: A case at English education department students. *International Journal of Language Education*, *5*(4), 244–259.

https://doi.org/10.26858/ijole.v5i4.19010

Nunan, D. (1999). *Second language teaching & learning*. Heinle & Heinle.

Rastari, K. L. (2023). YouTube as instructional media in promoting EFL Indonesian students' speaking skill, *11*(1).

Rohmah, A. A. (2022). The use of TED Talks videos for learning public speaking by English department students, *10*(1).

Rossman, G. B., & Rallis, S. F. (2003). *Learning in the field: An introduction to qualitative research*. SAGE Publications.

Saed, H. A., Haider, A. S., Al-Salman, S., & Hussein, R. F. (2021). The use of YouTube in developing the speaking skills of Jordanian EFL university students. *Heliyon*, *7*(7), e07543. https://doi.org/10.1016/j.heliyon.2021.e07543

Sudrajad, W., Jauhari, I., & Pigai, Z. K. (2022). Enhancing students' speaking skill through YouTube video in pandemic era, *2*.

Truong, N. K. V., & Le, Q. T. (2022). Utilizing YouTube to enhance English speaking skill: EFL tertiary students' practices and perceptions. *AsiaCALL Online Journal*, *13*(4), 7–31. https://doi.org/10.54855/acoj.221342

Zulaefa, R. F., & Rizal, D. (2023). The use of YouTube as a supplementary resource in students' speaking skills. *Ahmad Dahlan Journal of English Studies*, *10*(1). https://doi.org/10.26555/adjes.v10i1.302