



How Effective Are Computer-Assisted Pronunciation and YouTube-Based Instructions in Improving Oral English Proficiency Among Secondary School Students?

Olufunke O. Olatoye*✉ and **Oluwakemi Olurinola**✉

To cite this article. O. O. Olatoye and O. Olurinola, “How Effective Are Computer-Assisted Pronunciation and YouTube-Based Instructions in Improving Oral English Proficiency Among Secondary School Students?,” *Lang. Technol. Soc. Media*, vol. 4, no. 1, pp. 1–18, 2025.

DOI: <https://doi.org/10.70211/ltsm.3026-7196.73>

To link to this article:



Published online: 1 November 2025



Submit an article to this journal



View crossmark data



Watch the video on YouTube



How Effective Are Computer-Assisted Pronunciation and YouTube-Based Instructions in Improving Oral English Proficiency Among Secondary School Students?

Olufunke O. Olatoye* and Oluwakemi Olurinola

Received: 12 August 2025

Revised: 17 August 2025

Accepted: 29 October 2025

Online: 1 November 2025

Abstract

This study investigates the effectiveness of Computer-Assisted Pronunciation Instruction (CAPI) and YouTube-based instruction in enhancing oral English proficiency among Nigerian senior secondary school students. A quasi-experimental design was employed, involving three groups: the CAPI group, the YouTube group, and a control group. The study utilized an Oral English Language Proficiency Test (OELPT), administered before and after the treatments, to assess students' pronunciation improvements. Data were analyzed using Analysis of Covariance (ANCOVA) to determine the main and interaction effects of the treatment, gender, and language spoken at home. The results revealed a significant effect of the treatment on students' oral proficiency ($F = 14.530, p < 0.05$), with both CAPI and YouTube-based instruction outperforming the control group. However, gender and language spoken at home did not significantly influence the outcomes ($p > 0.05$). There were also no significant interaction effects between treatment and these demographic factors, suggesting that the treatments were equally effective for both male and female students and regardless of the language spoken at home. These findings highlight the effectiveness of CAPI and YouTube as technology-based tools for improving oral English proficiency, with implications for language instruction in diverse educational contexts. This study contributes to the growing body of research on technology-enhanced language learning, providing evidence of the efficacy of CAPI and YouTube-based instruction in secondary school settings, particularly in resource-constrained environments. The findings suggest that integrating digital tools into language education can offer a scalable, accessible solution to improving students' oral proficiency, regardless of their socio-demographic backgrounds.

Keywords: CAPI; YouTube-based instruction; Oral English proficiency; Language learning technology; Pronunciation improvement; Secondary education; Speech recognition

Publisher's Note:

WISE Pendidikan Indonesia stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright:

©

2025 by the author(s).

License WISE Pendidikan Indonesia, Bandar Lampung, Indonesia.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license

(<https://creativecommons.org/licenses/by/4.0/>).



Language,
Technology,
and Social Media



INTRODUCTION

English has become a global language, widely recognized as a universal medium for communication across various sectors, including education, business, and international diplomacy [1]. In many countries, including Nigeria, English is not only taught as a subject but also serves as the primary medium of instruction across various disciplines [2], [3]. As a result, mastering both written and spoken aspects of English has become crucial for academic success and effective participation in the global community. Among the different aspects of language acquisition, oral proficiency is particularly important as it allows students to communicate clearly and confidently in both academic and social contexts. Mastering spoken English is thus a vital skill that enables individuals to articulate their ideas effectively, participate in discussions, and express themselves in a variety of settings.

In the Nigerian education system, oral proficiency is emphasized as a key component of the English language curriculum [1]. It aims to equip students with the skills needed to speak English fluently and accurately, which are essential for their success in higher education and in the workforce. However, despite its importance, many students face significant challenges in acquiring oral English proficiency. One of the primary reasons for these challenges is the linguistic and cultural diversity in Nigeria. With over 500 indigenous languages spoken across the country, students often struggle with pronunciation due to the interference of their first languages or local accents [2], [3], [4]. This phenomenon leads to mispronunciations that hinder effective communication and reduce students' confidence in using English.

Studies have shown that learning a foreign language, especially oral English, presents substantial challenges for learners in Nigeria, where the influence of various local languages and accents can make the acquisition of correct pronunciation difficult. According to Alsuhailani et al. [1], effective pronunciation teaching requires learners to engage with accurate models of the language's sounds and rhythms. However, traditional teaching methods in Nigeria often fail to provide the necessary tools for overcoming pronunciation difficulties. Teachers frequently lack access to the appropriate technology and often do not use proper intonation, further exacerbating the challenges faced by students in mastering oral English [2], [3], [4]. Moreover, research by Al-Kresheh [5] highlighted that learners of English as a second language commonly struggle with pronunciation, either due to the interference of their first language or the lack of exposure to native English speakers.

In response to these challenges, technology has emerged as a potential solution to enhance the teaching and learning of oral English. The integration of Computer-Assisted Language Learning (CALL), particularly CAPI, offers several advantages. CAPI utilizes Automatic Speech Recognition (ASR) technology, which provides learners with immediate feedback on their pronunciation. This feedback helps students identify and correct pronunciation errors, thereby improving their ability to produce intelligible speech. Studies have demonstrated that CAPI, by offering fast feedback and incorporating visual and auditory illustrations, encourages autonomous learning and has a significant impact on improving pronunciation skills [6], [7]. Furthermore, ASR systems in CAPI provide native-speaker pronunciation models, which help learners replicate correct sounds and enhance their speech intelligibility [8], [9], [10].

In addition to CAPI, YouTube has proven to be a valuable educational tool in language learning. The platform offers a wide array of videos featuring native speakers, providing learners

with authentic pronunciation models and real-life contexts. Apoko and Waluyo [11] found that YouTube significantly improved vocabulary acquisition and pronunciation among students, as it allowed them to listen to and mimic native speakers. Moreover, YouTube not only enhances learners' vocabulary but also motivates them to engage more actively in their language learning process. The entertaining and accessible nature of YouTube videos increases student engagement, making it an effective tool for improving oral English proficiency [12], [13], [14].

Despite the promising results from studies on CAPI and YouTube in other contexts, their application in Nigerian secondary schools remains underexplored. Most studies on the effectiveness of CAPI and YouTube have been conducted in developed countries or higher education settings, where technological infrastructure is more readily available. In Nigeria, however, many schools still lack the necessary resources, which presents a unique challenge in integrating these tools effectively. Getie [15] pointed out that many Nigerian students are exposed to limited opportunities for practicing oral English with native speakers, making it harder for them to develop the correct pronunciation. This research gap highlights the need for studies that explore the effectiveness of CAPI and YouTube in enhancing oral proficiency in Nigerian secondary schools, particularly in resource-constrained settings.

The purpose of this study is to examine the effects of CAPI and YouTube-based instruction on improving oral English proficiency among senior secondary school students in Nigeria. It will also examine how gender and the language spoken at home moderate the effectiveness of these technological interventions. Specifically, the study seeks to: 1) Determine the main effect of CAPI and YouTube treatments on students' oral English proficiency; 2) Assess the main effect of gender on students' oral proficiency; 3) Examine the main effect of language spoken at home on students' oral proficiency; 4) Investigate the interaction effects of treatment and gender, as well as treatment and language spoken at home. Based on these objectives, the following hypotheses were formulated:

Table 1. Hypotheses of the Study

Hypothesis Code	Hypothesis Statement
H ₀₁	There is no significant main effect of treatment (CAPI and YouTube) on students' proficiency in oral English.
H ₀₂	There is no significant main effect of gender on students' proficiency in oral English.
H ₀₃	There is no significant main effect of language spoken at home on students' proficiency in oral English.
H ₀₄	There is no significant interaction effect of treatment and gender on students' proficiency in oral English.
H ₀₅	There is no significant interaction effect of treatment and language spoken at home on students' proficiency in oral English.
H ₀₆	There is no significant interaction effect of gender and language spoken at home on students' proficiency in oral English.
H ₀₇	There is no significant interaction effect of treatment, gender, and language spoken at home on students' proficiency in oral English.

Given the increasing role of technology in education, this study aims to contribute valuable insights into how digital tools like CAPI and YouTube can be effectively integrated into the

Nigerian educational system to enhance students' oral English proficiency. The findings of this study may have broader implications for the use of technology in language education, particularly in resource-limited settings.

LITERATURE REVIEW

The Importance of Oral Proficiency in English Language Learning

Oral proficiency in English is a critical component of language education, particularly in non-native English-speaking countries. In Nigeria, where English serves as both a subject and a medium of instruction, mastering spoken English is essential for academic success and effective communication. However, students often face challenges due to the influence of local languages and accents, which can impede clear communication. Studies have shown that mispronunciations resulting from these influences can hinder students' academic performance and confidence in using English [1], [2], [16], [17]. Research indicates that traditional methods of teaching oral English, which rely heavily on face-to-face interaction and textbook-based approaches, are often insufficient to address these challenges. These methods may not provide the individualized feedback necessary for students to improve their pronunciation skills effectively [3], [4]. Therefore, there is a need for innovative approaches that can offer personalized and accessible pronunciation instruction to students.

The Role of Technology in Enhancing Oral English Proficiency

The integration of technology into language education has been shown to offer effective solutions for enhancing oral proficiency. CAPI, which utilizes ASR technology, provides real-time feedback on students' pronunciation, allowing them to identify and correct errors independently. Studies have demonstrated that ASR-based CAPI systems can significantly improve learners' pronunciation skills by offering immediate corrective feedback and opportunities for repeated practice [10], [18], [19]. Furthermore, the self-directed nature of CAPI allows students to engage with pronunciation practice at their own pace, fostering autonomy and motivation. This aligns with constructivist learning theories, which emphasize the importance of active learner engagement in the construction of knowledge [8]. The use of CAPI has been particularly beneficial in contexts where access to native speakers or language instructors is limited, as it provides learners with consistent and personalized pronunciation practice.

The Influence of YouTube in Language Learning

In addition to CAPI, YouTube has emerged as a valuable tool for language learning, offering a vast array of authentic content that can enhance oral proficiency. Research has shown that exposure to native speaker videos on YouTube can improve students' pronunciation by providing models of correct speech patterns in various contexts [20], [21], [22]. The multimedia nature of YouTube videos, which combine auditory and visual elements, supports multimodal learning and caters to diverse learning styles. Studies have also highlighted the positive impact of YouTube on student engagement and motivation. The interactive features of YouTube, such as the ability to pause, rewind, and replay videos, allow learners to focus on specific pronunciation aspects and practice at their own pace. This flexibility makes YouTube an effective supplementary resource for pronunciation instruction, particularly in resource-constrained settings [12], [23].

Socio-Cultural Factors Influencing Language Learning

While technological interventions can enhance oral proficiency, socio-cultural factors such as gender and home language play significant roles in language learning outcomes. Research has indicated that gender differences can influence language learning strategies and achievements, with female learners often outperforming male learners in language tasks due to differences in learning styles and motivation [24], [25], [26]. Understanding these differences is crucial for designing effective language instruction that caters to the needs of all learners. Additionally, the language spoken at home can impact students' proficiency in English. Studies have shown that learners who are exposed to English or related languages at home tend to have higher levels of oral proficiency compared to those from homes where entirely different language systems are used [27], [28], [29]. This underscores the importance of considering students' linguistic backgrounds when evaluating the effectiveness of language learning interventions.

METHODS

Research Design

This study employed a quasi-experimental factorial design with a $3 \times 2 \times 2$ factorial pre-test, post-test structure, incorporating two experimental groups and one control group. The design was selected for its ability to manipulate the independent variables (types of treatment (CAPI, YouTube-based instruction, and traditional teaching), gender, and language spoken at home) to examine their effects on students' oral English proficiency. The quasi-experimental approach was appropriate as it involved intact groups, where participants were not randomly assigned, but their baseline characteristics were considered, ensuring the robustness of the findings. This design allows for the analysis of both main effects and interaction effects between treatment types and socio-demographic variables, providing a comprehensive understanding of the factors influencing oral English acquisition.

Population of the Study

The target population for this study consisted of all Senior Secondary School 2 (SSS II) students in public schools within Oyo State, Nigeria. These students were selected because they are at an intermediate level of English language proficiency, making them an ideal cohort to examine the impact of various instructional interventions on oral language skills. The study focused on students from diverse linguistic backgrounds, as Nigeria is home to over 500 indigenous languages, which influence students' acquisition of standard English pronunciation.

Sample and Sampling Technique

A purposive sampling technique was used to select a total of 108 students from two secondary schools in Akinyele and Ibadan North Local Government Areas of Ibadan, Oyo State. These schools were specifically chosen based on their availability of resources, such as access to computer systems for CAPI and internet connectivity for YouTube-based instruction. The sample was stratified into three groups: the CAPI group, the YouTube group, and the control group. Each group consisted of students from both genders and different home language backgrounds, ensuring a balanced representation across the variables of interest. The students were divided into the following groups:

- CAPI group (38 students)
- YouTube group (33 students)
- Control group (37 students)

The schools were selected to be geographically separated to minimize the risk of cross-group contamination. In each school, students were assigned to groups based on availability, ensuring that no student had prior exposure to the instructional methods being tested.

Table 2. Layout of Pre-Test, Post-Test 3×2×2 Factorial Design

Treatment	Gender	Language Spoken at Home	Total
CAPI	Male	English	5
		Mother Tongue	14
	Female	English	6
		Mother Tongue	13
YouTube	Male	English	2
		Mother Tongue	8
	Female	English	13
		Mother Tongue	10
Control	Male	English	9
		Mother Tongue	8
	Female	English	6
		Mother Tongue	14
Total			108

Key: CAPI - Computer-Assisted Pronunciation Instruction; YouTube - YouTube Instruction Group; LSH - Language Spoken at Home

Instrumentation

The primary instrument for data collection was the Oral English Language Proficiency Test (OELPT), designed to assess students' pronunciation accuracy and overall oral English proficiency. The OELPT was administered before and after the treatment to capture the pre-test and post-test scores, providing a measure of improvement in oral proficiency.

OELPT

The OELPT utilized Pronunciation Coach Software, an accent training tool that offers a standardized and objective rating of students' pronunciation. One challenge in assessing oral proficiency, especially in non-native environments, is the potential bias introduced by non-native speakers as evaluators. To overcome this, the Pronunciation Coach Software provides an ASR system that rates students' speech based on comparison with a native speaker model. The software categorizes pronunciation accuracy into four levels: Excellent, Good, Fair, and Poor.

Test items were shuffled for both pre-test and post-test to avoid any order effects. Each student's pronunciation was recorded by the software and rated according to their performance. The OELPT scores from the pre-test were used as baseline data, and the same test items were used in the post-test to measure any improvement in pronunciation following the interventions.

Table 3. Rating of Proficiency Test

Intelligibility Description	Intelligibility Score
Poor	0% - 24%
Fair	25% - 49%
Good	50% - 74%
Excellent	75% - 100%

Pronunciation Coach Software Functions

The Pronunciation Coach Software served dual purposes in this study:

1. CAPI instruction: It was used as the primary instructional tool for the CAPI group.
2. Assessment: It was also used to assess students' oral proficiency both before and after the treatment (pre- and post-test assessments).

The Pronunciation Coach software includes a Speech Intelligibility Scorer, which uses speech recognition to assess students' pronunciation accuracy. The system evaluates the intelligibility of the students' speech, with recognized words highlighted in green and unrecognized words in red. The final intelligibility score reflects the proportion of words correctly pronounced, providing valuable insights into students' progress in oral English.

Treatment Groups

CAPI Group

The CAPI group received instruction via the Pronunciation Coach Software installed on the school's computer systems. Students interacted with the software by typing words, vowels, or consonants that they needed to learn for each lesson. They listened to the correct pronunciation provided by the software and repeated it aloud. After completing each lesson, students participated in an evaluation session where the software graded their pronunciation performance.

YouTube Group

The YouTube group was exposed to video-based instruction that featured native English speakers demonstrating correct pronunciation. The videos were downloaded and shown to the students in class. After watching the videos, students were encouraged to repeat the words and sentences aloud, mimicking the native speaker models in the videos. This approach allowed students to practice pronunciation through exposure to real-world language use.

Control Group

The control group received traditional classroom instruction on oral English pronunciation. The teaching was conducted by the regular teacher using standard lesson plans and textbooks, with no use of technological tools such as CAPI or YouTube videos. This group served as a comparison to assess the effectiveness of technology-based instruction methods.

Data Collection

Data were collected in three stages: pre-test, treatment, and post-test. The pre-test and post-test used the same set of items, but these were arranged in different orders to prevent test familiarity effects. The study was conducted over a period of six weeks, with instructional sessions conducted twice a

week. During this period, students were taught oral English topics, and the OELPT was administered at both the beginning and the end of the study.

Data Analysis

The data were analyzed using ANCOVA to compare the mean scores of the three treatment groups. ANCOVA was chosen because it controls for potential covariates, such as the pre-test scores, and provides a more accurate estimate of the treatment effects. Additionally, Scheffe post-hoc analysis was conducted to identify the most effective treatment strategy. All data analyses were carried out using the Statistical Package for the Social Sciences (SPSS) version 24 at a significance level of 0.05.

RESULTS AND DISCUSSION

ANCOVA

The results of the ANCOVA are presented in Table 4. The analysis shows that the overall model was statistically significant ($F(12, 95) = 4.561, p < 0.05$), explaining approximately 36.6% of the variance in students' oral English proficiency ($R^2 = 0.366$; Adjusted $R^2 = 0.285$). After controlling for pre-test scores, there was a significant main effect of treatment on students' proficiency ($F(2, 108) = 14.530, p < 0.05$). However, the effects of gender ($F(1, 108) = 1.452, p > 0.05$) and language spoken at home ($F(1, 108) = 0.878, p > 0.05$) were not significant. Furthermore, no significant two-way or three-way interactions were found among treatment, gender, and language spoken at home.

Table 4. ANCOVA Results for Treatment, Gender, and Language Spoken at Home

Source	Type III SS	df	Mean Square	F	Sig.
Corrected Model	8213.626a	12	684.469	4.561	.000
Intercept	25297.417	1	25297.417	168.554	.000
Pretest	2213.764	1	2213.764	14.750	.000
Treatment	4361.550	2	2180.775	14.530	.000*
Gender	217.911	1	217.911	1.452	.231
Language Spoken at Home	131.820	1	131.820	0.878	.351
Treatment $\times$ Gender	257.301	2	128.651	0.857	.428
Treatment $\times$ Language Spoken at Home	32.367	2	16.183	0.108	.898
Gender $\times$ Language Spoken at Home	0.093	1	0.093	0.001	.980
Treatment $\times$ Gender $\times$ Language Spoken	123.238	2	61.619	0.411	.664
Error	14258.029	95	150.085		
Total	665300.885	108			
Corrected Total	22471.655	107			

$R^2 = 0.366$; Adjusted $R^2 = 0.285$. $p < 0.05$

Hypotheses Testing

The following hypotheses were tested to examine the impact of treatment, gender, and language spoken at home on students' oral English proficiency. The Analysis of Covariance (ANCOVA)

results provide a detailed account of these effects. Below is a detailed explanation for each hypothesis, supported by the corresponding [table 5](#).

Table 5. Summary of Hypothesis Testing Results

Hypothesis	Effect	F-value	p-value	Decision
H01	Treatment Effect	14.530	0.000	Rejected
H02	Gender Effect	1.452	0.231	Retained
H03	Language Spoken at Home Effect	0.878	0.351	Retained
H04	Interaction Effect (Treatment $\times$ Gender)	0.857	0.428	Retained
H05	Interaction Effect (Treatment $\times$ Language Spoken at Home)	0.108	0.898	Retained
H06	Interaction Effect (Gender $\times$ Language Spoken at Home)	0.001	0.980	Retained
H07	Three-Way Interaction (Treatment $\times$ Gender $\times$ Language Spoken at Home)	0.144	0.664	Retained

H₀₁: Treatment Effect

H₀₁ posited that there would be no significant effect of the treatment on students' proficiency in oral English. The ANCOVA results revealed a significant main effect of treatment ($F(2, 108) = 14.530, p < 0.05$), leading to the rejection of the null hypothesis. This indicates that the treatment types (CAPI, YouTube, and traditional) significantly influenced students' oral proficiency, with CAPI and YouTube being more effective than traditional methods. The p-value of 0.000, being less than the alpha level of 0.05, strongly supports the conclusion that both CAPI and YouTube are effective interventions for enhancing students' oral English proficiency.

H₀₂: Gender Effect

H₀₂ tested whether gender had a significant impact on students' oral English proficiency. The analysis revealed no significant effect of gender ($F(1, 108) = 1.452, p > 0.05$), meaning that male and female students showed similar improvements in their proficiency levels, regardless of the treatment they received. With a p-value of 0.231, which is greater than the 0.05 threshold, the null hypothesis is retained. This finding suggests that the treatment interventions (CAPI and YouTube) are equally effective for both genders.

H₀₃: Language Spoken at Home Effect

H₀₃ hypothesized that the language spoken at home would have a significant effect on students' oral English proficiency. The results indicated that language spoken at home did not significantly impact students' proficiency ($F(1, 108) = 0.878, p > 0.05$), meaning that students' progress in oral English was not affected by whether they spoke English or their mother tongue at home. The p-value of 0.351 further supports the retention of the null hypothesis. This implies that the treatments provided, CAPI and YouTube, were equally effective for students from different linguistic backgrounds, highlighting the potential of technology to overcome barriers posed by home language.

H₀₄: Interaction Effect of Treatment and Gender

H₀₄ examined the interaction effect of treatment and gender on students' oral English proficiency. The analysis showed that the interaction effect was not significant ($F(1, 108) = 0.857, p > 0.05$), meaning that the effectiveness of the treatment did not differ between male and female students. With a p-value of 0.428, the null hypothesis is retained. This result suggests that the CAPI and YouTube interventions were equally effective for both male and female students, further supporting the conclusion that the treatments are gender-neutral.

H₀₅: Interaction Effect of Treatment and Language Spoken at Home

H₀₅ tested whether there was an interaction effect between treatment and language spoken at home on students' oral proficiency. The results revealed no significant interaction effect ($F(2, 108) = 0.108, p > 0.05$), indicating that the language spoken at home did not influence the effectiveness of the treatments. The p-value of 0.898 confirms the retention of the null hypothesis. This suggests that the effectiveness of both CAPI and YouTube-based instruction was not influenced by the home language, highlighting the adaptability of these technologies for students from diverse linguistic backgrounds.

H₀₆: Interaction Effect of Gender and Language Spoken at Home

H₀₆ examined whether there was an interaction effect between gender and language spoken at home on students' proficiency in oral English. The results showed no significant interaction effect ($F(1, 108) = 0.001, p > 0.05$), indicating that the combination of these two variables did not affect students' performance. With a p-value of 0.980, the null hypothesis is retained. This result implies that the gender and language spoken at home of students did not interact to influence the effectiveness of the treatments, reinforcing the consistency of the treatment effects across demographic factors.

H₀₇: Three-Way Interaction Effect of Treatment, Gender, and Language Spoken at Home

H₀₇ tested the three-way interaction effect of treatment, gender, and language spoken at home on students' oral proficiency. The results showed no significant three-way interaction ($F(2, 108) = 0.144, p > 0.05$), indicating that the effectiveness of the treatments was independent of these three factors. The p-value of 0.664 confirms the retention of the null hypothesis. This suggests that the CAPI and YouTube-based instruction treatments were equally effective for all students, regardless of gender and language spoken at home.

Multiple Classification Analysis (MCA)

The MCA results are shown in Table 6. The analysis indicates that treatment exerted the strongest influence on students' proficiency (Beta = -0.287), followed by gender (Beta = -0.246), and finally language spoken at home (Beta = -0.104). The effect sizes demonstrate that the instructional approach was more critical than demographic variables in shaping oral English proficiency.

Table 6. MCA of Treatment, Gender, and Language Spoken at Home

Variable + Category	N	Unadjusted Variation	Eta	Adjusted for Factors	Beta
Treatment Group			0.423		-0.287

Variable + Category	N	Unadjusted Variation	Eta	Adjusted for Factors	Beta
CAPI	38	1.75		3.22	
YouTube	33	1.65		2.13	
Control	37	1.89		2.09	
Gender			0.230		-0.246
Male	46	0.72		1.57	
Female	62	0.41		1.28	
Language Spoken			0.119		-0.104
English	41	1.21		0.19	
Mother Tongue	67	1.92		1.21	

Grand Mean = 15.43; $R^2 = 0.366$ (Adjusted $R^2 = 0.285$)

Univariate and Post Hoc Analysis

The univariate analysis in Table 7 indicates significant mean differences among the three groups ($F(2, 105) = 11.432$, $p < 0.05$). Post-hoc comparisons further confirmed that both CAPI and YouTube groups significantly outperformed the control group, though no significant difference was observed between the CAPI and YouTube groups.

Table 7. Univariate Tests of Mean Proficiency Across Groups

Source	SS	df	MS	F	Sig.	Remark
Contrast	1607336.364	2	803668.182	11.432	.000	Sig.
Error	7381325.636	105	70298.339			

Group Performance

As shown in Table 8, students in the YouTube group achieved the highest post-test mean score ($M = 83.82$), followed by the CAPI group ($M = 79.19$), while the control group obtained the lowest mean score ($M = 69.11$). This trend is illustrated in Figure 1, confirming the superiority of technology-enhanced instruction compared to conventional teaching methods.

Table 8. Post-Test Mean Scores for Experimental and Control Groups

Group	N	Mean	SD	Min	Max
CAPI	38	79.19	14.16	46.20	96.65
YouTube	33	83.82	4.83	72.15	90.70
Control	37	69.11	16.91	30.00	93.15
Total	108	77.15	14.49	30.00	96.65

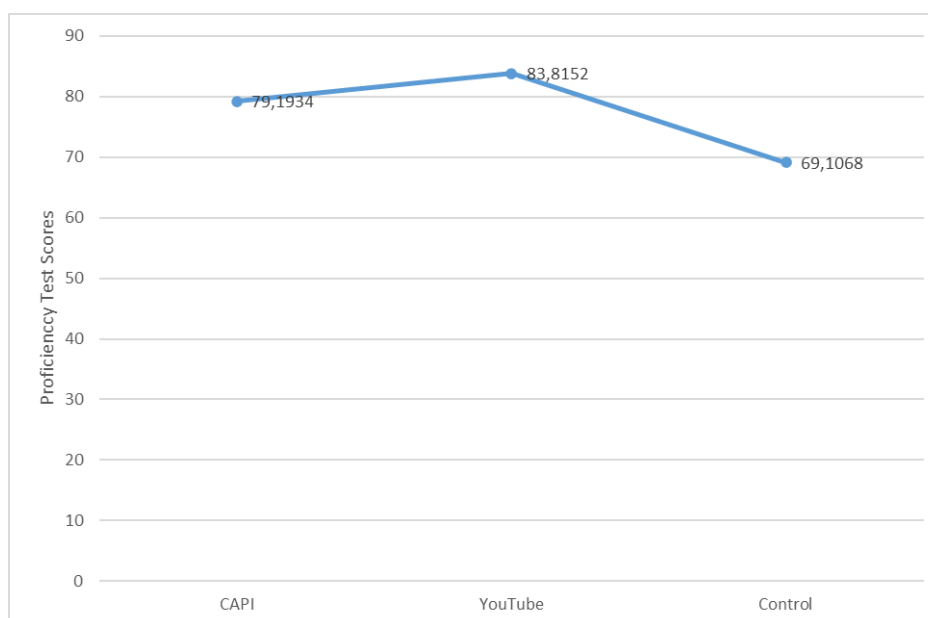


Figure 1. Mean Proficiency Scores of students in the three groups (CAPI, YouTube and Control)

Pairwise Comparisons

Pairwise comparisons (Table 9) confirmed that both CAPI and YouTube significantly outperformed the control group. However, there was no significant difference between CAPI and YouTube, suggesting that both technology-based interventions are equally effective in enhancing oral English proficiency.

Table 9. Pairwise Comparison among the Experimental and Control Groups

(I) Factor	(J) Factor	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CAPI	Video	-92.43461	63.08890	.146	-217.5282	32.6590
	Control	201.73329*	61.23652	.001	80.3126	323.1540
Video	CAPI	92.43461	63.08890	.146	-32.6590	217.5282
	Control	294.16790*	63.48392	.000	168.2910	420.0448
Control	CAPI	-201.73329*	61.23652	.001	-323.1540	-80.3126
	Video	-294.16790*	63.48392	.000	-420.0448	-168.2910

*. The mean difference is significant at the 0.05 level.

Discussion

The findings of this study demonstrate that both CAPI and YouTube-based instruction significantly improved students' oral English proficiency compared to traditional teaching methods. This outcome aligns with previous research that has emphasized the effectiveness of technology in language learning, particularly in enhancing pronunciation skills. For instance, Liu et al. [30] and Tejedor-García et al. [31] showed that the use of ASR in CAPI facilitates real-time feedback, allowing learners to refine their pronunciation and become more intelligible to others. This study confirms those findings, as the CAPI group demonstrated significant improvements in their oral proficiency, highlighting the efficacy of immediate, automated feedback in language learning.

The comparison between CAPI and YouTube-based instruction revealed no significant differences in students' proficiency, suggesting that both technological interventions are equally effective in improving oral English skills. This finding is consistent with Apoko and Waluyo [11], who suggested that YouTube is a powerful tool for language learning due to its exposure to authentic, native-level language use. In this study, the YouTube group performed at a comparable level to the CAPI group, which further supports the idea that video-based instruction, when integrated effectively into the curriculum, can be just as beneficial as more structured, software-based programs. The advantage of YouTube lies in its accessibility and variety of content, which offers learners diverse linguistic models, as highlighted by Al-khresheh [32] and Pellas [33].

However, this study found no significant effects of gender or language spoken at home on students' oral English proficiency. These results challenge some of the previous studies that identified gender and home language as significant factors influencing language learning outcomes. For example, Oga-Baldwin and Fryer [34] noted that female learners often outperform males in language-related tasks due to differences in learning strategies and motivation. Similarly, Getie [15] emphasized the role of home language in shaping students' ability to acquire English pronunciation. In contrast, the findings from this study suggest that the type of instructional treatment whether CAPI, YouTube, or traditional methods was the most significant factor affecting students' performance, rather than demographic variables such as gender or home language. This result may reflect the unique context of this study, where the technological interventions provided a level playing field, minimizing the impact of external factors that typically influence learning outcomes in more traditional settings.

One of the notable contributions of this study is its focus on the integration of both CAPI and YouTube as complementary technologies in a secondary school setting. While many previous studies have examined these tools in isolation, this study provides a comparative analysis of their effectiveness. The results suggest that both tools can significantly enhance students' oral English proficiency, making them valuable resources for language educators, particularly in resource-limited settings where access to native speakers or advanced instructors is often constrained. Furthermore, this study addresses the gap in the literature regarding the application of these technologies in Nigerian secondary schools, where limited access to language resources and technological tools often impedes effective pronunciation instruction.

The study also adds to the body of knowledge on the use of YouTube in language education. While Malakuk and Park [35] and Sumi Kim and Kim [36] highlighted YouTube's role in improving vocabulary and comprehension, this study extends these findings by demonstrating its effectiveness in directly enhancing oral proficiency. By incorporating YouTube videos featuring native English speakers, students were exposed to various pronunciation models, which supported their ability to replicate accurate speech patterns. This practical application of YouTube, combined with structured feedback, provides a new perspective on how social media platforms can be used in formal education settings to improve language skills.

A key strength of this study lies in its contextual relevance. The findings suggest that both CAPI and YouTube can be effective tools for enhancing English language proficiency in Nigerian secondary schools, where there are often significant challenges in pronunciation due to the country's linguistic diversity. As Al-khresheh [32] pointed out, students from different linguistic backgrounds face unique challenges in acquiring the correct pronunciation, which can impact their academic

performance. This study offers a potential solution to this issue by integrating technology that provides immediate, tailored feedback, regardless of the students' home language or accent.

Despite the promising results, several limitations must be acknowledged. First, the study was conducted in a small sample of students from two schools in Ibadan, which may limit the generalizability of the findings to other regions or populations. Future research should expand the sample size and include students from diverse geographical areas to examine whether the effects of CAPI and YouTube-based instruction are consistent across different cultural and educational contexts. Second, while the study focused on oral proficiency, future research could explore the long-term effects of technology-based pronunciation instruction on other aspects of language learning, such as listening comprehension, writing, and overall language fluency.

This study contributes to the growing body of literature on the use of technology in language education, particularly in improving oral proficiency. The findings underscore the potential of CAPI and YouTube as effective tools for enhancing students' pronunciation skills in secondary schools. By comparing the two methods, this study provides valuable insights into their relative effectiveness, offering a basis for educators to choose the most suitable technology for their teaching context. The lack of significant effects from gender and language spoken at home further highlights the importance of instructional methods in shaping students' language outcomes, suggesting that with the right tools, students from diverse backgrounds can achieve similar levels of proficiency.

CONCLUSION

This study examined the impact of CAPI and YouTube-based instruction on enhancing oral English proficiency among Nigerian senior secondary school students. The results demonstrated that both CAPI and YouTube-based instruction significantly improved students' pronunciation skills when compared to traditional teaching methods. Importantly, no significant difference was observed between the two technology-driven interventions, suggesting that both approaches are equally effective in supporting language learners. These findings contribute to the growing body of evidence that highlights the potential of digital tools in addressing the challenges of teaching oral English, particularly in contexts where students face linguistic diversity and limited access to native speakers. The study also found that gender and language spoken at home did not significantly influence the effectiveness of the treatments, emphasizing that instructional methods, rather than demographic factors, played a critical role in improving oral proficiency. This research adds to the existing literature by providing empirical evidence of the effectiveness of CAPI and YouTube in secondary school settings in Nigeria, offering practical implications for educators. The findings suggest that technology-enhanced pronunciation instruction can offer a scalable, flexible solution for improving language skills, particularly in resource-constrained environments, and warrant further exploration into its long-term impact on overall language proficiency.

LIMITATIONS

Despite the valuable insights provided by this study, several limitations must be acknowledged. First, the sample size was relatively small, comprising only 108 students from two schools in Ibadan, Oyo State. This limited sample size restricts the generalizability of the findings to a broader population, especially to schools in other regions of Nigeria or different countries with varying

educational contexts. Future research could benefit from a larger, more diverse sample to enhance the external validity of the results. Second, while the study focused on short-term improvements in students' oral English proficiency, it did not assess the long-term effects of CAPI and YouTube-based instruction. The lack of follow-up data means that the sustainability of the improvements observed in pronunciation remains unclear. Longitudinal studies are needed to examine whether the gains in oral proficiency persist over time and how these technologies influence other aspects of language learning, such as fluency and listening comprehension. Finally, the study was limited to the evaluation of CAPI and YouTube as instructional tools in isolation, without considering the combined use of multiple technologies or other pedagogical strategies that could enhance learning outcomes. Future research could explore the potential benefits of integrating CAPI, YouTube, and other digital tools within a more comprehensive instructional framework.

AUTHOR INFORMATION

Corresponding Author

Olufunke O. Olatoye – Department of Science and Technology, Faculty of Education, Olabisi Onabanjo University, Ago-Iwoye, Ogun State (Nigeria);

 orcid.org/0009-0005-4709-4723

Email: olufunkeolatoye.oo@gmail.com

Author

Olufunke O. Olatoye – Department of Science and Technology, Faculty of Education, Olabisi Onabanjo University, Ago-Iwoye, Ogun State (Nigeria);

 orcid.org/0009-0005-4709-4723

Oluwakemi Olurinola – Department of Science and Technology, Faculty of Education, Olabisi Onabanjo University, Ago-Iwoye, Ogun State (Nigeria);

 orcid.org/0000-0002-7502-8142

AUTHOR CONTRIBUTION

O.O.O. and O.O. worked collaboratively throughout the development of this manuscript. O.O.O. took the lead in conceptualizing the study, formulating the research objectives, and designing the methodological framework. She also handled the statistical analysis, interpretation of results, and integration of the findings into the discussion. O.O. was responsible for conducting the literature review, coordinating the data collection procedures, and drafting the initial version of the manuscript. Both authors contributed to revising and refining the content, ensuring the coherence of the argumentation and alignment with the journal's standards. They jointly reviewed and approved the final version of the manuscript before submission.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DECLARATION OF USE OF AI IN SCIENTIFIC WRITING

The authors declares that ChatGPT 5 Plus was used to assist in refining the language of this manuscript, particularly to enhance clarity, coherence, and academic tone. The tool was not employed in any part of the research design, data analysis, interpretation, or visualization process. All core ideas, arguments, and conclusions were independently developed by the authors, who take full responsibility for the originality, accuracy, and integrity of the work.

REFERENCES

- [1] Y. Alsuhaibani, H. S. Mahdi, A. Al Khateeb, H. A. Al Fadda, and H. Alkadi, "Web-based pronunciation training and learning consonant clusters among EFL learners," *Acta Psychol. (Amst.)*, vol. 249, p. 104459, Sep. 2024. <https://doi.org/10.1016/j.actpsy.2024.104459>
- [2] X. Van Ha, L. T. Nguyen, and B. P. Hung, "Oral corrective feedback in English as a foreign language classrooms: A teaching and learning perspective," *Heliyon*, vol. 7, no. 7, p. e07550, Jul. 2021. <https://doi.org/10.1016/j.heliyon.2021.e07550>
- [3] O. Sheredekina, I. Karpovich, L. Voronova, and T. Krepkaia, "Case Technology in Teaching Professional Foreign Communication to Law Students: Comparative Analysis of Distance and Face-to-Face Learning," *Educ. Sci.*, vol. 12, no. 10, p. 645, Sep. 2022. <https://doi.org/10.3390/educsci12100645>
- [4] Y. Li, "Current problems with the prerequisites for flipped classroom teaching—a case study in a university in Northwest China," *Smart Learn. Environ.*, vol. 5, no. 1, p. 2, Dec. 2018. <https://doi.org/10.1186/s40561-018-0051-4>
- [5] M. H. Al-khresheh, "Phonetic challenges in English: the impact of mispronunciation of the bilabial plosive /p/ on communication among Saudi EFL learners," *Cogent Arts Humanit.*, vol. 11, no. 1, Dec. 2024. <https://doi.org/10.1080/23311983.2024.2390777>
- [6] L. Yang, "Student Engagement With Teacher Feedback in Pronunciation Training Supported by a Mobile Multimedia Application," *Sage Open*, vol. 12, no. 2, Apr. 2022. <https://doi.org/10.1177/21582440221094604>
- [7] C. J. McCartan, J. Roberts, and J. Jordan, "Centre-based early education interventions for improving school readiness: A systematic review," *Campbell Syst. Rev.*, vol. 19, no. 4, Dec. 2023. <https://doi.org/10.1002/cl2.1363>
- [8] M. Farrús, "Automatic Speech Recognition in L2 Learning: A Review Based on PRISMA Methodology," *Languages*, vol. 8, no. 4, p. 242, Oct. 2023. <https://doi.org/10.3390/languages8040242>
- [9] S. Inceoglu, W.-H. Chen, and H. Lim, "Monitoring student behavior in autonomous automatic speech recognition-based pronunciation practice," *System*, vol. 124, p. 103387, Aug. 2024. <https://doi.org/10.1016/j.system.2024.103387>
- [10] W. Sun, "The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: a mixed methods investigation," *Front. Psychol.*, vol. 14, Aug. 2023. <https://doi.org/10.3389/fpsyg.2023.1210187>
- [11] T. W. Apoko and B. Waluyo, "Social media for English language acquisition in Indonesian higher education: Constructivism and connectivism frameworks," *Soc. Sci. Humanit. Open*, vol. 11, p. 101382, 2025. <https://doi.org/10.1016/j.ssaho.2025.101382>
- [12] S. Jiang and H. Zhao, "Learning English vocabulary via Instagram or YouTube: Surveying the impacts on motivation, growth mindfulness, willingness to communicate, and enjoyment from the lens of self-determination theory," *Learn. Motiv.*, vol. 89, p. 102089, Feb. 2025. <https://doi.org/10.1016/j.lmot.2024.102089>
- [13] Ü. Çoban Sural and Z. Yaşar Sağlık, "Next-generation learning experience: The effect of technology-assisted and printed context-based vocabulary activities on vocabulary knowledge

- and learning motivation,” *Educ. Inf. Technol.*, vol. 29, no. 16, pp. 20823–20853, Nov. 2024. <https://doi.org/10.1007/s10639-024-12688-x>
- [14] X. Zhou and Y. S. Goh, “What shapes tertiary learners' experiences and challenges in mobile-assisted seamless vocabulary learning? A case of Chinese as a foreign language,” *Acta Psychol. (Amst.)*, vol. 257, p. 105123, Jul. 2025. <https://doi.org/10.1016/j.actpsy.2025.105123>
- [15] A. S. Getie, “Factors affecting the attitudes of students towards learning English as a foreign language,” *Cogent Educ.*, vol. 7, no. 1, Jan. 2020. <https://doi.org/10.1080/2331186X.2020.1738184>
- [16] A. M. Vieriu and G. Petrea, “The Impact of Artificial Intelligence (AI) on Students' Academic Development,” *Educ. Sci.*, vol. 15, no. 3, p. 343, Mar. 2025. <https://doi.org/10.3390/educsci15030343>
- [17] A. Jatmiko, N. Armita, Irwandani, T. Saputro, and M. Aridan, “Development of Science Learning Videos with the Canva Application on Socioscientific Issues Content,” *E3S Web Conf.*, vol. 482, p. 05004, Jan. 2024. <https://doi.org/10.1051/e3sconf/202448205004>
- [18] M. Bashori, R. van Hout, H. Strik, and C. Cucchiari, “I Can Speak: improving English pronunciation through automatic speech recognition-based language learning systems,” *Innov. Lang. Learn. Teach.*, vol. 18, no. 5, pp. 443–461, Oct. 2024. <https://doi.org/10.1080/17501229.2024.2315101>
- [19] T. Dillon and D. Wells, “Effects of Pronunciation Training Using Automatic Speech Recognition on Pronunciation Accuracy of Korean English Language Learners,” *English Teach.*, vol. 78, no. 1, pp. 3–23, Mar. 2023. <https://doi.org/10.15858/engtea.78.1.202303.3>
- [20] A. Leis, “How speech-to-text technology affects pronunciation gains and self-confidence in EFL learners,” *Comput. Assist. Lang. Learn.*, pp. 1–24, Aug. 2025. <https://doi.org/10.1080/09588221.2025.2534498>
- [21] W. Yan, V. L. Lowell, and L. Yang, “Developing English language learners' speaking skills through applying a situated learning approach in VR-enhanced learning experiences,” *Virtual Real.*, vol. 28, no. 4, p. 167, Nov. 2024. <https://doi.org/10.1007/s10055-024-01061-5>
- [22] C. L. Yuen and N. Schlote, “Learner Experiences of Mobile Apps and Artificial Intelligence to Support Additional Language Learning in Education,” *J. Educ. Technol. Syst.*, vol. 52, no. 4, pp. 507–525, Jun. 2024. <https://doi.org/10.1177/00472395241238693>
- [23] M. A. Mekheimer, “Beyond the textbook: A year-long exploration of VODcasts in EFL education,” *Educ. Inf. Technol.*, vol. 30, no. 4, pp. 4777–4793, Mar. 2025. <https://doi.org/10.1007/s10639-024-12900-y>
- [24] Q. Feng, H. Luo, W. Li, T. Chen, and N. Song, “Effects of gender diversity on college students' collaborative learning: From individual gender to gender pairing,” *Heliyon*, vol. 9, no. 6, p. e16237, Jun. 2023. <https://doi.org/10.1016/j.heliyon.2023.e16237>
- [25] J. Lin, “The structural relationships among L2 motivation, out-of-class informal learning, and oral proficiency: A multiple-group structural equation modeling study,” *Lang. Teach. Res.*, Aug. 2023. <https://doi.org/10.1177/13621688231189030>
- [26] S. A. Adeniran *et al.*, “Differentiated instruction for improved primary school learners' outcomes in basic science: a mixed-methods study,” *Int. J. Sci. Educ.*, pp. 1–26, Aug. 2025. <https://doi.org/10.1080/09500693.2025.2524868>
- [27] S. S. Bratlie, V. Grøver, R. Lekhal, S. Chen, and V. Rydland, “Home literacy environment, language use, and proficiency: Bilingual profiles in young learners,” *J. Appl. Dev. Psychol.*, vol. 96, p. 101728, Jan. 2025. <https://doi.org/10.1016/j.appdev.2024.101728>
- [28] L. Xu, T. Abdullah, and X. Tang, “Impacts of Family Environment on Chinese University Students' English Language Proficiency: The Mediating Effect of Parenting Styles,” *Sage Open*, vol. 14, no. 4, Oct. 2024. <https://doi.org/10.1177/21582440241289802>
- [29] J.-Y. Choi, S. Jeon, and F. Arabzadehjafari, “Learning Two Languages: Dual Language Learning Patterns, Predictors, and Outcomes,” *Early Educ. Dev.*, vol. 35, no. 6, pp. 1214–

- 1234, Aug. 2024. <https://doi.org/10.1080/10409289.2023.2233395>
- [30] Y. Liu, F. binti Ab Rahman, and F. binti Mohamad Zain, “A systematic literature review of research on automatic speech recognition in EFL pronunciation,” *Cogent Educ.*, vol. 12, no. 1, Dec. 2025. <https://doi.org/10.1080/2331186X.2025.2466288>
- [31] C. Tejedor-García, V. Cardeñoso-Payo, and D. Escudero-Mancebo, “Automatic Speech Recognition (ASR) Systems Applied to Pronunciation Assessment of L2 Spanish for Japanese Speakers,” *Appl. Sci.*, vol. 11, no. 15, p. 6695, Jul. 2021. <https://doi.org/10.3390/app11156695>
- [32] M. H. Al-khresheh, “Bridging technology and pedagogy from a global lens: Teachers' perspectives on integrating ChatGPT in English language teaching,” *Comput. Educ. Artif. Intell.*, vol. 6, p. 100218, Jun. 2024. <https://doi.org/10.1016/j.caeai.2024.100218>
- [33] N. Pellas, “The influence of sociodemographic factors on students' attitudes toward AI-generated video content creation,” *Smart Learn. Environ.*, vol. 10, no. 1, p. 57, Nov. 2023. <https://doi.org/10.1186/s40561-023-00276-4>
- [34] W. L. Q. Oga-Baldwin and L. K. Fryer, “Girls show better quality motivation to learn languages than boys: latent profiles and their gender differences,” *Heliyon*, vol. 6, no. 5, p. e04054, May 2020. <https://doi.org/10.1016/j.heliyon.2020.e04054>
- [35] S. Malakul and I. Park, “The effects of using an auto-subtitle system in educational videos to facilitate learning for secondary school students: learning comprehension, cognitive load, and satisfaction,” *Smart Learn. Environ.*, vol. 10, no. 1, p. 4, Jan. 2023. <https://doi.org/10.1186/s40561-023-00224-2>
- [36] S. Kim and H.-C. Kim, “The Benefits of YouTube in Learning English as a Second Language: A Qualitative Investigation of Korean Freshman Students' Experiences and Perspectives in the U.S.,” *Sustainability*, vol. 13, no. 13, p. 7365, Jun. 2021. <https://doi.org/10.3390/su13137365>