Issues in Applied Linguistics & Language Teaching
Volume 07, Issue 01, 2025, pp. 272-282
E-ISSN: 2597-9825
Open Access: https://doi.org/10.37253/iallteach.v7i1.10417

Video-assisted Instruction for Pronunciation Improvement: An Intervention Study with Indonesian EFL Students

Asma Mukarrama[1*], Kalsum[2], Syukron Fajriansyah[3]

[1,2,3] English Language Education, IAIN Parepare, Indonesia
E-mail addresses: syukron.fajriansyah98@gmail.com (Corresponding Author)

ARTICLE INFO

Article history:
Received May 20, 2025
Revised June 08, 2025
Accepted June 10, 2025
Available online June 17, 2025

Kata Kunci :
Pembelajaran Berbasis Video, Pengucapan, Vokal dan Konsonan, Pembelajar EFL

Keywords:
Video-based Instruction, Pronunciation, Vowel and Consonant, and EFL Learners

ABSTRAK

Penelitian ini mengkaji efektivitas penggunaan video dalam meningkatkan kemampuan pelafalan bahasa Inggris pada 32 siswa kelas XI di sebuah sekolah menengah atas negeri di Sulawesi Selatan. Dengan menggunakan desain pra-eksperimen satu kelompok pre-test–post-test selama satu bulan, siswa mengikuti delapan sesi pembelajaran yang memanfaatkan materi video mencakup vokal, konsonan, tekanan kata, intonasi, dan ritme. Data kuantitatif dari pre-test dan post-test dianalisis menggunakan SPSS. Hasil menunjukkan peningkatan skor pelafalan secara signifikan—dari rata-rata 48,50 ("rendah") menjadi 65,75. Uji t berpasangan (t = –10.213, p = 0.000) mengonfirmasi signifikansi statistik peningkatan tersebut. Siswa menunjukkan kemajuan nyata dalam mengucapkan fonem asing dan memahami tekanan kata, meskipun aspek intonasi dan ritme masih menjadi tantangan. Temuan ini menyarankan bahwa pembelajaran berbasis video efektif dalam mendukung peningkatan pelafalan siswa EFL. Implikasinya, guru disarankan untuk mengintegrasikan media video guna meningkatkan keterlibatan siswa, memberikan model otentik, dan menciptakan lingkungan belajar pelafalan yang lebih efektif dan inklusif.

ABSTRACT

This study examined the effectiveness of video-based instruction in improving the English pronunciation skills of 32 second-year students at a public senior high school in South Sulawesi. Using a one-group pre-test–post-test design over one month, students participated in eight sessions incorporating video materials that modeled vowels, consonants, stress, intonation, and rhythm. Quantitative data from pre- and post-tests were analyzed using SPSS. Results showed a significant improvement in pronunciation scores—from an average of 48.50 ("poor") to 65.75—with a paired-samples t-test (t = –10.213, p = 0.000) confirming the statistical significance. Students made notable gains in producing unfamiliar phonemes and recognizing word stress, though intonation and rhythm remained more challenging. These findings suggest that video-assisted learning can enhance pronunciation outcomes in EFL classrooms. Practically, this implies that teachers should integrate multimedia tools into instruction to increase student engagement, provide authentic models, and foster more effective and inclusive pronunciation learning environments.

1. INTRODUCTION

Pronunciation plays a crucial role in language learning, particularly for achieving fluency in speaking. The way words are pronounced in Indonesian differs markedly from English, yet pronunciation often receives insufficient attention in the classroom. Learners may understand vocabulary but struggle to articulate it correctly because English spelling does not always align with its pronunciation.

During a teaching practicum in Pontianak, the researcher observed that one of six classes exhibited significant pronunciation difficulties. In one public senior high school in South Sulawesi, traditional pronunciation instruction—where teachers model words or sentences and students simply repeat them—has proven largely ineffective. Students often appear confused when dealing with word stress, intonation, and phonemic distinctions. This conventional approach is typically monotonous and provides little room for interactive feedback or error correction, limiting students' ability to improve. Students were hesitant to read aloud or participate in speaking activities, and their lack of confidence prevented them from seeking clarification on how to pronounce words (Putri. et al., 2018). Similarly, many learners remain silent in class due to fear of mispronunciation, which hampers their participation and engagement. A mixed-method study revealed that when students receive more supportive instructional methods, they become more active, ask

and answer questions more readily, and interact more effectively in class. These observations motivated the researcher to explore ways to enhance pronunciation skills.

(AbdAlgane, 2020) notes that EFL students frequently err when producing sounds absent in Indonesian—such as final consonant clusters—and often substitute unfamiliar phonemes with ones they know. Such errors not only impede their spoken output but also their comprehension of spoken English. (Gilakjani, 2017) emphasizes that better pronunciation leads to improved listening comprehension, while (Grant, 2017) argues that accurate pronunciation promotes more successful communication by making speakers easier to understand and better able to understand others. Additionally, (Antaris & Omolu, 2019) revealed that Indonesian students face various internal and external obstacles, including low motivation, lack of confidence, limited exposure to English, unsupportive environments, and strong local accents. Similarly, (Lasabuda, 2017) found that students often confuse words with similar pronunciations, rarely practice English, and experience anxiety and nervousness when speaking. These psychological and contextual barriers inhibit consistent pronunciation improvement.

Phonological issues are another key factor affecting students' pronunciation. According to (Maiza, 2020), Indonesian learners frequently mispronounce certain English sounds such as /θ/, /ð/, /ʃ/, /ʒ/, /tʃ/, and /dʒ/ due to interference from their first language and a lack of understanding of the English phonological system. (Plailek & Essien, 2021) also noted that specific English phonemes—especially at initial and final syllable positions—are difficult for EFL learners due to differences between native and target language sound systems. Their study highlighted that inadequate foundational knowledge, insufficient teacher instruction, and low exposure to spoken English further compound these problems. (Toçi, 2020) added that lack of vocabulary, poor teaching methods, limited contact with native speakers, and low self-confidence significantly affect pronunciation accuracy. These findings emphasize the multifaceted nature of pronunciation issues among EFL learners and the need for innovative instructional approaches that can effectively address them.

Despite this growing awareness, many Indonesian classrooms still rely on traditional teaching methods in which the teacher models words and students repeat them with little variation. This approach, commonly practiced in one Indonesian public high school in Pinrang, has proven ineffective in addressing students' pronunciation problems. Learners continue to struggle with stress, intonation, and phoneme differentiation, and the method itself is often monotonous and lacks opportunities for immediate feedback or meaningful interaction.

To address these limitations, this study proposes the use of video-assisted instruction. This approach incorporates animated visuals and audio to create engaging, context-rich environments where students can observe, mimic, and interact with authentic pronunciation models. Video-based instruction offers advantages such as visual reinforcement, exposure to natural speech, and opportunities for repeated practice, making it a promising alternative in classrooms with multimedia access. Based on the situation described above, this study seeks to answer the following research questions:

a. What is the level of students' pronunciation skill before receiving video-based instruction?
b. What is the level of students' pronunciation skill after receiving video-based instruction?
c. Is there a significant improvement in students' pronunciation skill following the implementation of video-based instruction?

The Concept of Pronunciation

Pronunciation denotes the way in which sounds are articulated to form intelligible speech. It comprises two interrelated dimensions: segmental features and suprasegmental features (Murphy, 2017). Segmental features consist of the discrete phonemes—vowels and consonants—where vowels are produced with an open vocal tract and consonants with partial or total occlusion of airflow. Accurate segmental articulation is vital, since mispronouncing a single phoneme can change word meaning.

Suprasegmental features encompass prosodic elements that span multiple segments: stress (the relative prominence of syllables), intonation (pitch variation over phrases), and rhythm (temporal patterning of stressed and unstressed units). Mastery of these features is essential for conveying nuance, emotion, and pragmatic intent in spoken language (Darcy et al., 2021).

Physiologically, pronunciation relies on coordinated movements of the lips, tongue, jaw, and other articulators, coupled with precise airflow control (Ladefoged & Johnson, 2014). Learners must acquire correct tongue placement, lip shaping, and voicing to achieve clarity.

Pronunciation development is subject to individual and contextual factors: a learner's first language, age at exposure, extent of input, and personal motivation all influence outcomes (Derwing & Munro, 2015). Given its centrality to communicative competence, pronunciation warrants explicit instructional focus. As (Pennington, 2021) observes, "intelligible pronunciation enables speakers to convey their intended meanings effectively and to be understood by their interlocutors."

Pronunciation entails the articulation of speech sounds to achieve intelligible communication. Its components are typically divided into segmental features—individual sounds—and suprasegmental features—prosodic patterns (Jafari et al., 2021).

a. Segmental Features (Sounds)

Segmental features consist of vowels and consonants, which are distinguished by their modes of articulation. Vowels are produced with an unobstructed vocal tract, yielding resonant sounds that constitute the nucleus of syllables. They are classified by tongue height (high, mid, low), tongue advancement (front, central, back), and lip rounding; their precise formant frequencies underpin both perception and intelligibility (Zsiga, 2024). Consonants, in contrast, involve partial or complete occlusion of the airstream. They are described by place of articulation (e.g., bilabial, alveolar), manner of articulation (e.g., stop, fricative, nasal), and voicing. Accurate consonant production is essential for preserving lexical distinctions and ensuring clear speech (Odden, 2022).

b. Suprasegmental Features

Suprasegmentals—stress, intonation, and rhythm—extend over syllables, words, and phrases, shaping the melody and timing of speech.

1) Stress

Stress assigns prominence to syllables or words through increased loudness, pitch, and duration. In English, shifting primary stress can signal grammatical category changes (e.g., 'record vs. re'cord) and guides the listener's attention to salient information within an utterance (Darcy, 2018).

2) Intonation

Intonation refers to pitch movement across an utterance, serving to differentiate sentence types (declarative, interrogative), convey speaker attitude, and structure discourse. Common patterns include falling (statement, wh-question), rising (yes/no question, polite request), and fall–rise (contrastive or non-final) contours. Mastery of these patterns is crucial for expressing nuanced meanings and managing conversational turns (Cruttenden, 2014; Levis & Wichmann, 2015).

3) Rhythm

Rhythm emerges from the temporal arrangement of stressed and unstressed syllables. It contributes to speech fluency and affects comprehensibility, particularly for learners whose first language exhibits different timing conventions.

Physiologically, pronunciation depends on coordinated movements of lips, tongue, jaw, and controlled airflow (Ladefoged & Johnson, 2014). Learner variables—first-language phonological system, age of acquisition, exposure, and motivation—further influence the acquisition of both segmental and suprasegmental features (Derwing & Munro, 2015). Given its integral role in communicative competence, pronunciation merits dedicated pedagogical focus (Afshari & Ketabi, 2017).

Teaching Pronunciation

Teaching and learning are interdependent processes: as the Latin proverb "by learning you will teach, by teaching you will learn" suggests, educators and learners continuously inform one another. Effective teaching not only delivers content but also structures the environment to facilitate student learning (Sanjaya, 2017). Thus, instruction and acquisition occur simultaneously within a dynamic cycle.

In the domain of pronunciation instruction, teachers assume several key roles. First, they raise learners' phonological awareness by explaining the target language's sound system and articulatory mechanisms. They then model accurate pronunciation—through clear speech, audio recordings, and multimedia—to provide learners with reliable exemplars (Foote et al., 2011). Concurrently, instructors assess and diagnose individual pronunciation challenges, identifying common errors and tailoring corrective feedback to each student's needs. Importantly, teachers foster a supportive atmosphere, emphasizing intelligibility over accent perfection.

A range of classroom activities supports pronunciation development (Kang et al., 2018; McGregor & Reed, 2018; Qizi & Dilshodbekovna, 2024):

a. Dictation: learners transcribe spoken input, focusing on sound perception, stress, and connected speech.
b. Shadowing: students echo utterances immediately after hearing them, aiming to mirror pronunciation, rhythm, and intonation.
c. Multimedia instruction: videos, animations, and speech-visualization tools clarify articulatory processes and prosodic features.

d. Computer-Assisted Pronunciation Training (CAPT): software offers individualized visual feedback and self-paced practice.
e. Interactive Games: gamified, web-based activities engage learners through quizzes and challenges.
f. Pronunciation Journals: reflective logs help learners monitor their progress and set goals.
g. Spoken Production Tasks: discussions, debates, and presentations provide extended, spontaneous practice.
h. Focused Transcription: learners transcribe selected audio/video segments to target specific pronunciation phenomena.

Assessing Pronunciation

Pronunciation assessment evaluates a learner's ability to produce the sounds, stress patterns, and intonation of a target language relative to accepted norms. Common assessment techniques include:

a. Listening and Rating

Trained evaluators listen to read-aloud passages, word lists, or spontaneous speech and rate learners' pronunciation using rubrics or rating scales. Criteria typically cover accurate phoneme production, appropriate stress placement, correct intonation, rhythm, and overall intelligibility (Isaacs, 2008). While this method is straightforward and authentic, it relies on rater judgment and requires robust training and well-defined criteria to ensure reliability.

b. Minimal Pair Testing

By contrasting word pairs that differ by a single phoneme (e.g., "ship" vs. "sheep"), evaluators assess both perception and production of critical sound distinctions. Tests can target vowel contrasts (beat/bit), consonant contrasts (pat/bat), and stress shifts (INsight/inCITE). This focused approach diagnoses persistent segmental errors but does not address prosodic features (Derwing & Munro, 2015).

c. Phonetic Transcription

Learners' utterances are transcribed using the International Phonetic Alphabet (IPA) and compared to the target transcription. This reveals substitutions (e.g., [fɪŋk] for [θɪŋk]), deletions, insertions, and vowel inaccuracies at a granular level (Derwing & Munro, 2015). Although highly detailed, transcription demands substantial IPA training and is time-intensive.

d. Instrumental Analysis

Acoustic analysis tools measure parameters such as voice-onset time (VOT), formant frequencies, pitch contours, intensity, and spectrographic profiles. Comparing these metrics to native-speaker norms pinpoints both segmental and suprasegmental deviations (Derwing & Munro, 2015). This objective method provides precise data but requires specialized equipment and expertise.

e. Self-Assessment

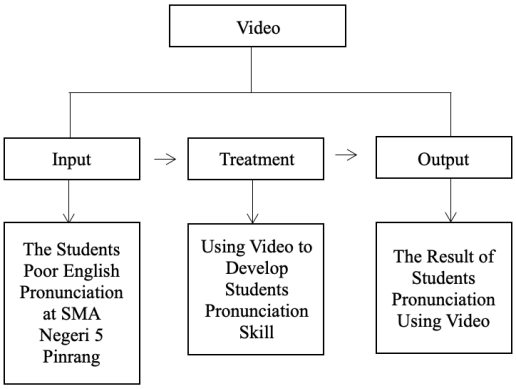
Learners use rubrics, model comparisons, reflective journals, and peer feedback to monitor and evaluate their own pronunciation (Derwing & Munro, 2015). This approach promotes autonomy and metacognitive skills but necessitates learner training to produce accurate self-evaluations.

f. Intelligibility Testing

In this functional assessment, proficient listeners transcribe learners' speech samples. The percentage of correctly recognized words yields an intelligibility score, reflecting real-world comprehensibility (Derwing & Munro, 2015). While context and listener familiarity can influence results, this method directly measures communicative effectiveness.

Conceptual Framework

The conceptual framework for this study—focused on enhancing second-year students' pronunciation skills through video-based instruction—is outlined as follows:



2. METHODS

This study employed a one-group pre-test–treatment–post-test pre-experimental design, as outlined by (Sugiyono, 2013). This design was selected because it allows for the measurement of change in participants' pronunciation abilities before and after the intervention, making it well-suited to classroom-based action research with practical constraints. Although the design lacks a control group, it is commonly used in exploratory educational research where the primary aim is to assess the potential impact of a specific instructional strategy—in this case, video-assisted pronunciation instruction (Syahroni, 2022).

The site of the research, a public senior high school in South Sulawesi, Pinrang, was purposefully selected due to its accessibility, relevance, and availability of multimedia facilities. The school had existing infrastructure (such as projectors, speakers, and a digital library) that supported the integration of video materials into language instruction. Furthermore, teachers at the site had expressed concerns about students' ongoing pronunciation challenges, which underscored the relevance and timeliness of the proposed intervention.

The participants were drawn from the population of second-year students at the selected school. Purposive sampling was used to select Class XI-6 as the study sample. This class was chosen because it reflected average proficiency levels and learning behaviors typical of the broader student population. Additionally, the class schedule aligned with the research timeline, and the students demonstrated sufficient availability and willingness to participate consistently throughout the study period.

Data were collected over the course of four weeks, consisting of eight instructional sessions and two testing sessions. At the outset, a pre-test (O1) was administered to assess students' baseline pronunciation abilities. The treatment (X) consisted of video-based instruction focusing on English phonemes, word stress, sentence intonation, and rhythm. The video materials provided both visual and auditory input, offering students contextualized models of authentic speech. Students were encouraged to observe, imitate, and engage with the video content through guided pronunciation practice and repetition exercises. After the instructional period, a post-test (O2) was conducted to evaluate the progress made by each student.

The primary instrument used was a pronunciation test, administered before and after the treatment. These tests required students to read aloud a set of words, phrases, and sentences that targeted specific pronunciation features. The test results were analyzed using quantitative methods, including descriptive statistics (mean and standard deviation) and inferential statistics via a paired-samples t-test to determine whether the changes in pronunciation scores were statistically significant (Syahroni, 2022). This methodological approach provided a structured yet flexible framework for assessing the effectiveness of video-based instruction and was appropriate given the classroom setting and research goals.

3. RESULT AND DISCUSSION

Results

a. Students Pronunciation before Using the Video

Results from the pre-test—analyzed via Excel and SPSS—provided baseline data on students' pronunciation skills prior to the video intervention.

Table 1. Students' Pronunciation Score and Classification in Pre-test

No.	Name	Students Point					Total Point	Total Score	Classification
		V	C	S	I	R			
1	AREM	4	3	4	3	3	17	68	GOOD
2	AB	3	2	2	2	1	10	40	POOR
3	AD	2	3	2	2	2	11	44	POOR
4	AFR	2	3	2	2	1	10	40	POOR
5	AIS	2	2	3	2	1	10	40	POOR
6	ALF	3	2	3	2	1	11	44	POOR
7	AMRD	2	2	2	2	2	10	40	POOR
8	FIT	2	3	2	2	2	11	44	POOR
9	FFM	2	2	2	1	1	8	32	VERY POOR
10	IB	2	2	2	2	2	10	40	POOR
11	JUN	2	3	2	3	2	12	48	POOR
12	KRA	2	3	2	2	1	10	36	VERY POOR
13	LSA	2	3	2	3	2	12	48	POOR
14	MS	2	3	2	1	1	9	40	POOR
15	MTH	2	2	2	2	2	10	40	POOR
16	MAM	2	3	2	3	2	12	48	POOR
17	MAS	3	3	3	3	3	15	60	FAIR
18	MAB	2	3	2	2	2	11	44	POOR
19	MFN	3	3	3	3	2	14	56	FAIR
20	MF	2	3	3	3	1	12	48	POOR
21	MUF	4	4	4	3	3	18	72	GOOD
22	MUN	2	2	3	3	2	12	48	POOR
23	NAD	3	2	2	2	1	10	40	POOR
24	NAS	4	3	3	3	2	15	60	FAIR
25	NRT	2	3	2	3	2	12	48	POOR
26	NSP	4	3	4	3	3	17	68	GOOD
27	NUR	3	2	3	2	2	12	48	POOR
28	NUE	2	2	2	2	2	10	40	POOR
29	PRA	3	4	4	2	2	15	60	FAIR
30	WIP	2	2	2	2	2	10	40	POOR
31	WIN	4	4	3	3	3	17	68	GOOD
32	ZUL	4	3	4	2	2	15	60	FAIR
	Total	83	87	83	75	60	388	1552	

From the pre-test conducted, vowels, consonants, stress, intonation, and rhythm were tested. Where it can be seen that students are still lacking in several aspects of pronunciation (vowel, consonant, stress, intonation, and rhythm).

Table 2. The Rate Percentage of the Frequency of the Pre-test

NO	Classification	Score	Score Frequency of Pre-Test	Percentage of Pre-Test
1	Very Good	80-100	0	0%
2	Good	66-79	4	12.5%
3	Fair	56-65	5	15.625%
4	Poor	40-55	21	65.625%
5	Very Poor	≤39	2	6.25%
	Total		32	100%

Based on the data analysis above, before students pronunciation learning strategies using the video only 4 students (12.5%) got the GOOD category than 5 students (15.625%) fell into the FAIR category, while 21 students (65.625%) fell into the POOR category and 2 students (6.25%) fell into FAIR POOR category. This shows that the students pronunciation in the pre-test activities are still in thr low category.

Based on the results of the data analysis above, the researcher presents descriptive statistics of students pronunciation before using the video on second years students.

Table 3. Descriptive Statistics of Pre-test

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Pre-test	32	32	72	48.50	10.64
Valid N (Listwise)	32				

Based on the data above, students pronunciation before using the video, in the pre-test is still in the low category with an average score of 48.50.

Prior to the video intervention, students' pronunciation performance averaged 48.50 (65.63%), classified as "poor" according to Arikunto's criteria for classical achievement. This low proficiency suggests that many learners struggled with accurate English articulation. Contributing factors likely included the use of monotonous, teacher-centered methods and limited opportunities for exposure to and practice with authentic pronunciation models.

b. Students' Pronunciation After Using Video

Following the video-based instruction, students completed a post-test assessing vowels, consonants, stress, intonation, and rhythm. Analysis of these results—using the same Excel and SPSS procedures—revealed their pronunciation performance after the treatment.

Table 4. Students' Pronunciation Score and Classification in Post-test

No.	Name	Students Point					Total Point	Total Score	Classification
		V	C	S	I	R			
1	AREM	5	4	5	4	3	21	84	VERY GOOD
2	AB	3	3	3	3	2	14	56	FAIR
3	AD	4	3	3	3	3	16	64	FAIR
4	AFR	5	4	5	4	3	21	84	VERY GOOD
5	AIS	4	4	4	3	3	18	72	GOOD
6	ALF	4	4	4	3	3	18	72	GOOD
7	AMRD	2	3	3	2	2	12	48	POOR
8	FIT	4	4	4	3	3	18	72	GOOD
9	FFM	3	2	3	3	2	13	52	POOR
10	IB	3	3	3	3	3	15	60	FAIR
11	JUN	4	4	4	3	3	18	72	GOOD
12	KRA	3	2	3	3	2	13	52	POOR
13	LSA	4	3	4	3	3	17	68	GOOD
14	MS	2	2	3	2	2	11	44	POOR
15	MTH	3	3	3	3	3	15	60	FAIR
16	MAM	3	3	3	3	2	14	56	FAIR
17	MAS	4	4	3	3	3	17	68	GOOD
18	MAB	4	3	3	3	3	16	64	FAIR
19	MFN	3	3	2	3	2	13	52	FAIR
20	MF	4	4	3	3	3	17	68	GOOD
21	MUF	5	4	5	5	4	23	92	VERY GOOD
22	MUN	4	3	4	3	3	17	68	GOOD
23	NAD	3	3	3	3	2	14	56	FAIR
24	NAS	4	4	5	4	3	20	80	VERY GOOD
25	NRT	4	4	4	3	3	18	72	GOOD
26	NSP	5	5	5	4	4	23	92	VERY GOOD
27	NUR	3	3	3	3	2	14	56	POOR
28	NUE	4	3	3	3	2	15	60	FAIR
29	PRA	4	4	3	3	3	17	68	GOOD
30	WIP	2	2	2	2	2	10	40	POOR
31	WIN	4	4	4	4	3	19	76	GOOD
32	ZUL	4	4	4	4	3	19	76	GOOD
	Total	117	108	113	101	87	526	2104	

From the post test that has been done tested vowel, consonant, stress, intonation, and rhythm. Where it can be seen that there is an increase in student scores in several aspects of pronunciation (vowel, consonant, stress, intonation, and rhythm).

Table 5. The Rate Percentage of the Frequency of the Post-test

NO	Classification	Score	Score Frequency of Pre-Test	Percentage of Pre-Test
1	Very Good	80-100	5	15.625%
2	Good	66-79	12	37.5%
3	Fair	56-65	9	28.125%
4	Poor	40-55	6	18.75%
5	Very Poor	≤39	0	0%
	Total		32	100%

After giving the treatment and conducting a post-test on the results after using the video to improve pronunciation skill, the results obtained showed that 5 students (15.625%) get a score in the "VERY GOOD" category, 12 students (37.5%) received the "GOOD" category than 9 students (28.125%) a got the "FAIR" category and there was only 6 student (18.75%) who still are in the "POOR" category.

Based on the above results, this researcher will present descriptive statistics of students' pronunciation skills after using the video technique on second-year students.

Table 6. The Rate Percentage of the Frequency of the Post-test

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Post-test	32	40	92	65.75	12.93
Valid N (Listwise)	32				

c. T-Test Independent

Paired T-Test is a parametric test that can be used on two data from pre- and post-test. The purpose of this test is to see if there is a mean difference between two paired datasets. The following are the results of the data description:

Table 7. T-test Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 Pre-test - Post test	-17.25	9.55	1.689	-20.6948	-13.8052	-10.213	31	0.000

The results of the paired-samples t-test yielded a t-value of –10.213. To determine statistical significance at the 95% confidence level ($\alpha = 0.05$), this value was compared against the critical t-value from the t-distribution table for 31 degrees of freedom ($n - 1 = 32 - 1$), which is 1.696. Because the observed t-value is less than –1.696, we reject the null hypothesis (H_0) and accept the alternative hypothesis (H_1), indicating a statistically significant effect of the video intervention on students' pronunciation.

Further confirmation comes from the two-tailed significance (p-value) reported in the paired-samples test: $p = 0.000$. Since $p \leq 0.05$, the difference between pre-test and post-test scores is significant. Consequently, the null hypothesis of no improvement is rejected in favor of the alternative hypothesis of improvement. These findings demonstrate that video-based instruction significantly enhances the pronunciation skills of second-year students.

Discussion

The findings of this study reveal a significant improvement in students' pronunciation skills following the implementation of video-assisted instruction. Prior to the intervention, students exhibited a low level of pronunciation proficiency, with an average score of 48.50, which falls into the "poor" category according to (Arikunto's, 2013) criteria. After the one-month treatment using video materials, the average score rose to 65.75, with a notable shift in the distribution of scores—more students achieving "good" and "very good" categories. These results are statistically significant, as confirmed by the paired-samples t-test ($t = -10.213$, $p = 0.000$), thereby affirming the positive impact of video-assisted instruction on students' pronunciation development.

These findings align with (Pourhosein Gilakjani et al., 2017), who argue that learners often face pronunciation difficulties due to insufficient access to accurate auditory models and a lack of structured production practice. This was evident in the pre-test phase, where students struggled particularly with suprasegmental features—stress, intonation, and rhythm—and problematic phonemes absent in their L1, such as /θ/, /ð/, /ʤ/, and /æ/ (Salam & Nurnisa, 2021). The limited effectiveness of traditional listen-and-repeat methods, as used in the research site prior to the intervention, also supports (Khan et al.'s, 2017) critique that such methods fail to actively engage learners or address their individual phonetic challenges.

In contrast, video-based instruction introduced a multisensory and contextualized approach to pronunciation learning. Drawing on (Shen's, 2023) cognitive theory of multimedia learning, the integration of visual and auditory input in video materials facilitates deeper cognitive processing, which likely contributed to improved retention and pronunciation accuracy. For example, through repeated viewing and mimicking of native speaker models, students demonstrated improved articulation of problematic vowels such as /æ/ and /oʊ/, and consonants like /ʤ/ and /θ/, consistent with (Juswandi et al., 2022; Ping & Tao, 2025) who emphasize the value of multimedia in mastering articulatory features. Moreover, students' improvements extended beyond individual sounds to include word and sentence stress, intonation, and rhythm. These aspects, which are typically challenging for learners due to L1 transfer, improved following exposure to video content illustrating natural prosody and speech patterns. This aligns with (Plailek & Essien's, 2021) findings, which identified suprasegmental features as persistent barriers in EFL pronunciation and highlighted the importance of exposure to authentic pronunciation models.

In addition to linguistic development, the study also suggests that video-assisted instruction positively influenced students' affective factors. Initially, learners expressed fear of making mistakes and lacked the confidence to practice aloud. However, the informal and engaging nature of the video content helped reduce anxiety and encouraged more active participation. This supports (Seger's, 2024) motivational framework and (Taş, 2024) who assert that video materials not only support language retention but also create enjoyable learning experiences that enhance motivation. The results of this study also reflect trends reported in other video-based pronunciation studies. For instance, (Martinsen et al., 2017) found that video-shadowing techniques led to statistically significant improvements in pronunciation tasks, with students valuing the authenticity and autonomy offered by the video format. Similarly, (Hidayati, 2021; Menggo et al., 2022) reported improved speaking skills, autonomy, and ICT literacy following video-assisted strategies in Indonesian EFL classrooms. These parallels reinforce the potential of video media to support pronunciation instruction across diverse learner populations and proficiency levels.

Despite the promising outcomes, some pronunciation features, particularly rhythm and intonation, showed only modest improvement. This indicates that while video tools can significantly enhance foundational pronunciation skills, mastering more complex prosodic features may require longer-term exposure and teacher-guided feedback, as also emphasized by (Juswandi et al., 2022). In summary, the current findings provide compelling evidence that video-assisted instruction is an effective strategy for improving EFL students' pronunciation. The approach not only addresses common segmental and suprasegmental errors but also fosters a positive, engaging, and confidence-boosting learning environment. These outcomes contribute to the growing body of literature affirming the value of multimedia resources in EFL instruction and offer practical insights for teachers seeking to enhance pronunciation teaching through innovative, learner-centered methods.

4. CONCLUSION

This study investigated the effectiveness of video-assisted instruction in improving the pronunciation skills of second-year students at a public senior high school in South Sulawesi, Pinrang. The findings clearly demonstrate that the use of video materials significantly enhanced learners' abilities across multiple pronunciation components—vowel and consonant articulation, stress, intonation, and rhythm. The statistically significant improvement in post-test scores, supported by both quantitative data and

qualitative observation, affirms that integrating audio-visual resources into pronunciation instruction can address common EFL learner challenges, particularly those rooted in L1 interference and limited exposure to authentic models. These results answer the central research problem: video-assisted instruction is not only effective but also pedagogically appropriate for developing students' English pronunciation skills in Indonesian EFL contexts.

Practically, this research highlights the importance of incorporating video-based materials into the English language curriculum, especially in schools equipped with basic multimedia facilities. Teachers are encouraged to move beyond passive "listen-and-repeat" methods and implement interactive techniques such as shadowing, voice replacement, and guided imitation of native speakers in communicative, real-life scenarios. Additionally, the engaging nature of videos helps reduce students' anxiety and increase their motivation—crucial affective factors in successful language learning. This suggests that video-based instruction can serve as both a linguistic and motivational tool, supporting a more inclusive and student-centered approach to pronunciation pedagogy. Future classroom applications should consider blending video content with structured feedback and regular practice to ensure lasting improvement, especially in more complex prosodic features like intonation and rhythm.

Recommendations

Students should consistently apply video-based pronunciation strategies in their independent study. By regularly watching and imitating authentic audiovisual materials—such as dialogues, news clips, or short films—learners can reinforce correct vowel and consonant articulations, stress patterns, intonation contours, and rhythmic phrasing. This sustained practice will not only solidify the gains achieved during classroom instruction but also enhance self-monitoring skills and boost confidence in spoken English.

English teachers are advised to integrate multimedia resources—particularly videos showcasing native-speaker pronunciation—into their lesson plans. Selecting materials that align with students' proficiency levels and interests encourages active engagement and provides clear, contextualized models of segmental and suprasegmental features. Incorporating guided imitation, subtitled playback, and peer feedback activities will further enable learners to self-correct and internalize accurate pronunciation patterns.

Future studies should build upon this video-based approach by extending the duration of interventions and diversifying the types of video content used. Focusing specifically on suprasegmental elements—intonation and rhythm—may uncover more nuanced strategies for addressing persistent prosodic challenges. Additionally, incorporating longitudinal designs or mixed-methods analyses could yield deeper insights into how video-assisted pronunciation training impacts learners over time and across different communicative contexts.

5. REFERENCES

AbdAlgane, M., & Idris, S. A. M. (2020). Challenges of pronunciation to EFL learners in spoken English. *Multicultural Education, 6*(5).

Afshari, S., & Ketabi, S. (2017). Current trends and future directions in teaching English pronunciation. *International Journal of Research Studies in Language Learning, 6*(2), 83-91.

Antaris, I., & Omolu, F. A. (2019). Factors affecting pronunciation difficulties of 8th grade students of MTSN Palu Barat. *Journal of Foreign Language and Educational Research, 2*(2), 10-19.

Cruttenden, A. (2014). *Gimson's pronunciation of English*. Routledge.

Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal, 30*(1).

Darcy, I., Rocca, B., & Hancock, Z. (2021). A window into the classroom: How teachers integrate pronunciation instruction. *RELC Journal, 52*(1), 110-127. https://doi.org/10.1177/0033688220964269

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals*. John Benjamins Publishing Company.

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada journal*, 1-22.

Gilakjani, A. P. (2017). English pronunciation instruction: Views and recommendations. *Journal of Language Teaching and Research, 8*(6), 1249-1255. http://dx.doi.org/10.17507/jltr.0806.30

Grant, L. (2017). *Well said: Pronunciation for clear communication* (4th ed.). Cengage Learning.

Hidayati, D. S. (2021). *The Implementation of Video-Assisted Guided Conversation to Improve Students' Speaking Skill at Eleventh Grade of MA Annuriyyah Jember* (Doctoral dissertation, Faculty of Tarbiyah and Teaching Training English Education Department).

Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *Canadian Modern Language Review, 64*(4), 555-580.

Jafari, S., Karimi, M. R., & Jafari, S. (2021). Beliefs and practices of EFL instructors in teaching pronunciation. *Vision: Journal for Language and Foreign Language Learning, 10*(2), 147-166.

Kang, O., Thomson, R. I., & Murphy, J. (Eds.). (2018). *The Routledge handbook of contemporary English pronunciation* (pp. 511-526). New York, NY: Routledge.

Lasabuda, N. (2017). An identification of students'difficulties in pronunciation. *Al-Lisan: Jurnal Bahasa (e-Journal), 2*(1), 17-24.

Ladefoged, P., & Johnson, K. (2014). *A Course in Phonetics*. 7th ed. Cengage Learning.

Levis, J. M., & Wichmann, A. (2015). *English intonation–Form and meaning*. The handbook of English pronunciation, 139-155.

Maiza, M. (2020). An Analysis of Students Pronunciation Errors. JOEEL Journal of English Education and Literature, 1(1), 18-23.

Martinsen, R., Montgomery, C., & Willardson, V. (2017). The effectiveness of video‐based shadowing and tracking pronunciation exercises for foreign language learners. *Foreign Language Annals, 50*(4), 661-680.

McGregor, A., & Reed, M. (2018). Integrating pronunciation into the English language curriculum: A framework for teachers. *The CATESOL Journal, 30*(1).

Menggo, S., Basir, A., & Halum, Y. S. (2022). Video-based tasks in strengthening speaking skills of EFL college students. *Indonesian Journal of EFL and Linguistics, 7*(2), 279.

Murphy, J. M. (2017). Teacher training in the teaching of pronunciation 1. In *The Routledge handbook of contemporary English pronunciation* (pp. 298-319). Routledge. https://doi.org/10.4324/9781315145006

Odden, D. (2022). Radical substance-free phonology and feature learning. *Canadian Journal of Linguistics/Revue canadienne de linguistique, 67*(4), 500-551.

Qizi, O. D. I., & Dilshodbekovna, A. Z. (2024). Teaching pronunciation in preschool and primary education. *Indonesian Journal of Primary Education, 8*(2), 247-254.

Plailek, T., & Essien, A. M. (2021). Pronunciation problems and factors affecting English pronunciation of EFL students. *Turkish Journal of Computer and Mathematics Education, 12*(12), 2026-2033.

Pennington, M. C. (2021). Teaching pronunciation: The state of the art 2021. *Relc Journal, 52*(1), 3-21.

Putri, Y. A., Sada, C., & Riyanti, D. (2018). Improving Students' Pronunciation By Using Tongue Twister Technique. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa (JPPK), 7*(11).

Sanjaya, W. (2017). *Paradigma baru mengajar. Kencana*.

Shen, W. W., Lin, J. M., Cheng, W. K., & Hong, Z. W. (2023). Developing and evaluating an online video-assisted collocation learning system for EFL students. *Interactive Learning Environments, 31*(8), 5292-5306.

Taş, Ş. (2024). Improving students' pronunciation skills through watching videos and audio recordings.

Toçi, A. (2020). Problems with pronunciation among students of English language and literature-SEEU. Seeu Review, 15(2), 113-125.

Zsiga, E. C. (2024). The sounds of language: An introduction to phonetics and phonology. John Wiley & Sons.