

Improving Students' English Pronunciation Skills Through Audio-Visual Media

Heddy

¹ English Lecturer Politeknik Negeri Medan, Indonesia

e-mail: heddytoruan@gmail.com

Abstract: Mastery of pronunciation skills is a crucial aspect in English oral communication. This study aims to describe the effectiveness of film and song-based audio-visual media in improving the pronunciation skills of English study program students. Using a descriptive qualitative approach with a case study design, data were obtained through observation, interviews, and documentation of pronunciation tests. The results showed that the integration of film and song media was able to significantly improve students' intelligibility, accuracy, and fluency. Films provide exposure to native speaker pronunciation in real contexts, while songs play a role in reducing anxiety and increasing self-confidence through sound repetition in a relaxed learning atmosphere. These findings are in line with recent studies that emphasize the importance of contextual and authentic media in teaching pronunciation. It is recommended that teachers utilize audio-visual media optimally in learning to create a more interesting and meaningful learning experience.

Keywords: Audio-Visual Media; English Language Learning; Pronunciation

1. Introduction

English pronunciation skills is a fundamental component that foreign learners must have in order to achieve fluent communication. This skill is often considered the main pillar that supports the effectiveness of oral interaction. More than just being able to pronounce words, good pronunciation includes understanding and proper application of phonemes (sound units), word stress, rhythm, and sentence intonation. Without mastery of these aspects, a learner, even with a wide vocabulary and a strong understanding of grammar, will have difficulty sounding natural and easily understood by native speakers and other learners.

The impact of good pronunciation is very significant in minimizing misunderstandings during the communication process. When a speaker has clear articulation and appropriate intonation, the listener can process information more easily and quickly, so that the focus can be diverted completely to the substance of the message. On the other hand, inaccuracies in pronunciation have the potential to cause disturbances in meaning or even total failure in conveying the message. For example, a small error in the sound of a vowel or consonant can change one word into another word with a very different meaning (e.g., ship becomes sheep). As emphasized by Yates (2021), this kind of disturbance not only hinders understanding but

Received: June 16, 2025

Revised: June 30, 2025

Accepted: July 01, 2025

Published: July 03, 2025

Curr. Ver.: July 03, 2025



Copyright: © 2025 by the authors.
Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>)

can also cause frustration and reduce the self-confidence of both parties involved in the conversation.

In Indonesia, pronunciation is still a major obstacle for many learners, including students in English study programs. This challenge stems from several factors, one of which is interference from the mother tongue (L1), namely Indonesian. The phonological system of Indonesian has significant differences from English, such as the absence of the *th* sound (/θ/ and /ð/) and differences in vowel sounds. In addition, teaching methods that have historically emphasized grammar and vocabulary for written tests often place pronunciation practice as a second priority. Lack of exposure to an authentic English-speaking environment also limits students' opportunities to practice listening and sound production naturally.

Addressing these pronunciation challenges, therefore, requires a shift in focus within the English language education system in Indonesia. It is important for educational institutions to integrate more intensive phonetics and phonology training into the curriculum. The use of technology, such as language learning apps and access to authentic media, can be effective tools to enhance exposure and practice. Ultimately, by equipping students with solid pronunciation skills, institutions will not only produce graduates who master language theory, but also competent, confident communicators who are ready to compete on a global stage where English is the *lingua franca*.

Several previous studies have shown that conventional methods in teaching pronunciation tend to be monotonous and provide less authentic exposure to native speaker models. Students often feel awkward, afraid of making mistakes, and lack confidence when they have to practice pronunciation in English (Putra et al., 2022). Therefore, innovative teaching strategies are needed that can increase students' learning motivation while providing authentic practice media.

Amidst the various challenges in teaching English pronunciation, especially in a foreign learner environment, the use of technology and learning media is a necessity. One alternative solution that has proven to be the most effective is the use of audio-visual media. This approach fundamentally changes conventional teaching methods that are often limited to textbooks and theoretical explanations. By presenting sound and image elements simultaneously, audio-visual media offers a more dynamic, multi-sensory learning experience that is relevant to real-world language use, making it a very powerful tool for overcoming pronunciation difficulties.

The primary strength of audio-visual media lies in its ability to provide direct, authentic exposure to native speakers' pronunciation models. Through media such as films, news videos, transcribed podcasts, and educational videos, students can hear over and over how words are pronounced correctly, including subtle aspects such as word stress, sentence rhythm, and elisions. Furthermore, the visual element allows students to observe native speakers' lip, tongue, and facial expressions, providing kinesthetic cues that are impossible to obtain from audio alone or text. This consistent exposure helps build an accurate "phonological map" in students' minds.

Furthermore, the advantage of this media lies in its ability to bring natural and contextual communication situations directly into the classroom. Unlike artificial dialogues in textbooks, scenes in a film or video clip present language used in specific social, cultural, and emotional contexts. Students learn not only 'how' to pronounce a sentence, but also 'when' and 'why'

certain intonations are used, for example to indicate sarcasm, friendliness, or urgency. As stated by Rahman et al. (2023), the ability of audio-visual media to simulate real interactions is crucial in bridging the gap between theoretical knowledge and practical communicative competence.

In addition to the technical aspects of pronunciation, students' psychological factors also play an important role in successful learning. This is where media such as English songs show their unique advantages. Pronunciation practice can be a stressful activity and cause anxiety because of the fear of making mistakes. Songs, with their relaxed, melodic, and fun nature, can significantly reduce this "affective filter". Students can sing along without feeling judged, practice pronunciation through natural repetition of lyrics, and get used to the rhythm and connected speech in English in an entertaining way.

Ultimately, the integration of audio-visual media in pronunciation teaching is not just a complement, but a comprehensive pedagogical strategy. Its implementation requires an active role from the teacher, not only as a material deliverer, but as a smart facilitator in selecting relevant materials and designing meaningful activities around them. By combining native speaker models from films, real communication contexts from videos, and a positive learning atmosphere from songs, pronunciation teaching can transform from a mechanical exercise into a lively, effective language learning experience that empowers students to communicate more confidently.

Several studies have shown that the integration of films and songs in the pronunciation learning process can increase students' motivation, active involvement, and courage in speaking (Kusumawati et al., 2023). Through the combination of these two media, students not only get exposure to sounds and intonations from native speakers but can also practice pronouncing song lyrics that have repetitions, thereby strengthening phonological memory.

With this background, this article aims to describe how the application of audio-visual media based on films and songs can improve students' pronunciation skills. In addition, this article will also describe supporting theories, research methods, analysis results, and conclusions related to the effectiveness of the media in pronunciation learning.

2. Theoretical basis

Pronunciation is defined as the way to pronounce language sounds correctly according to the phonetic rules that apply in a language (Yates & Zielinski, 2020). In learning English, pronunciation includes the pronunciation of vowel and consonant sounds, word stress placement, sentence intonation, and pronunciation rhythm. Mistakes in pronunciation can result in the meaning of the sentence being biased or even wrong.

According to Gilakjani (2021), pronunciation is not only about the clarity of sound articulation, but also about the intelligibility and accuracy of pronunciation in oral communication. Intelligibility refers to how easily the listener understands the speech, while accuracy relates to the conformity of pronunciation to the native speaker's sound standards. These two aspects are indicators of the success of pronunciation learning.

Audio-visual media is a learning tool that involves the senses of sight and hearing simultaneously, such as films, music videos, animations, or video podcasts (Rahman et al., 2023). This media is effective because it is able to present real and authentic communication

situations outside the classroom environment. Students can hear native speakers' pronunciation directly in various social contexts.

In addition to films, English songs have also been shown to be effective in practicing pronunciation. Rhyming and repetitive lyrical structures can help students strengthen their memory of English sounds and rhythms (Kusumawati et al., 2023). Songs can also create a relaxed learning atmosphere, thereby reducing anxiety about speaking a foreign language.

Based on these findings, it can be concluded that the integration of films and songs in pronunciation teaching has great potential in improving students' pronunciation skills, both in terms of clarity, accuracy, and fluency.

3. Research Methods

This study uses a descriptive qualitative approach with a case study design. This approach was chosen to gain an in-depth understanding of the process and results of implementing film and song-based audio-visual media in pronunciation learning (Creswell & Poth, 2021). The subjects of the study were second-semester students of the English Language Education Study Program at a university in Indonesia.

Data collection techniques were conducted through observation, interviews, and documentation. Observations were conducted during the learning process to record student activities, responses to media, and participation in pronunciation exercises. Interviews were conducted with students to explore their opinions and experiences during learning using films and songs.

Documentation in the form of pronunciation test results before and after learning is used to determine the increase in students' abilities. This test covers aspects of intelligibility, accuracy, and fluency, as recommended by Yates & Zielinski (2020). Data were analyzed descriptively qualitatively to identify patterns of change that occurred in students.

The research steps start from pre-observation, implementation of learning for two cycles, observation, reflection of results, and data analysis. Data validity is obtained through triangulation of sources and techniques, as well as member-checking of research subjects (Creswell & Poth, 2021).

4. Analysis and Discussion

Observation results show that at the beginning of learning, students still have difficulty pronouncing vowel and consonant sounds in English, especially the sounds [θ], [ð], and [æ]. They tend to use Indonesian pronunciation patterns when speaking (Rahman et al., 2023). In addition, most students seem less confident when asked to practice pronunciation in front of the class.

Of course, here is a development of that paragraph into five paragraphs that describe the process and its impacts in more detail.

The implementation of innovative learning media is a crucial step to overcome the classic challenges in teaching English pronunciation. Before the intervention using film media, many students tend to have a theoretical understanding of pronunciation, limited to what they read in books or hear from teachers who are sometimes not native speakers. Their exposure to authentic pronunciation is very minimal, so their sound production is often monotonous, stiff, and heavily influenced by the phonology of their mother tongue. The structured

introduction of film media into the curriculum is designed as a bridge to connect this passive knowledge with active and natural communicative skills.

In the early stages of implementing film media, the first impact felt is on students' auditory perception abilities. Gradually, they become accustomed to hearing authentic pronunciation from native speakers in various accents and natural speaking speeds. At first, the speed of dialogue in films may seem incredibly fast and difficult to understand, but through repeated exposure and guided tasks—such as watching short clips over and over again—their ears begin to train. This adaptation process is crucial because it helps students move from hearing "simplified" English for learners to being able to process language as it is in the real world.

Over time, this passive listening process evolves into a more active and in-depth analysis. Students no longer simply hear a stream of sounds, but they begin to consciously attend to the suprasegmental aspects that make up the music of language. They begin to identify how rising intonation signals a question or hesitation, and how falling intonation gives a statement a sense of finality. Furthermore, they begin to sense the rhythm of sentences and recognize which words are stressed to highlight important meanings in a dialogue. These “moments of insight” mark a significant cognitive leap, as students begin to deconstruct pronunciation not as a collection of sounds but as a complex system of meaning-carrying elements.

The positive developments observed in this class are not an anomaly, but rather a result that is in line with findings in pedagogical research. This is in accordance with the findings of Prambudi (2023) who stated that films are very effective in increasing authentic exposure in pronunciation learning. According to the study, authentic exposure through films provides rich and varied input, which cannot be imitated by teaching materials specifically designed for language learners. The dialogues in the film present language in its real context, complete with a variety of emotions, speeds, and sociolinguistic variations, thus providing a much more realistic and motivating target model for students.

Ultimately, the impact of the application of film media is manifested in the improvement of students' oral production competence. Armed with a new understanding of intonation, rhythm, and word stress, they began to experiment with applying it in their own speech. Although not immediately producing a perfect accent, their pronunciation became much more intelligible and sounded more natural. Their confidence in speaking also increased because they felt more connected to the way English is actually used. Thus, the application of film media has proven to be successful in transforming students from passive listeners to active observers and, most importantly, to more competent and confident communication participants.

After going through an initial learning cycle that may have focused more on conventional methods such as dialogue exercises and phonetic repetition, the application of new strategies becomes essential to overcome the remaining barriers, especially the psychological ones. In this second cycle, the integration of English songs is introduced as a pedagogical intervention specifically aimed at building students' self-confidence. This approach is based on the understanding that pronunciation proficiency does not only depend on cognitive and motor skills, but is also greatly influenced by the learner's affective or emotional state, such as anxiety and motivation levels.

The key to the effectiveness of this method lies in the careful selection of material. The English songs used in this cycle were deliberately chosen because they have relatively simple lyrics and repetitive structures, especially in the refrain. The simplicity of the lyrics reduces the cognitive load on students, allowing them to focus their full attention on the pronunciation aspect without having to grapple with complicated vocabulary or complex sentence structures. The repetitive nature of the songs naturally serves as a non-boring drill, helping students automate the movements of the mouth and tongue muscles to produce certain sounds until they become second nature.

The psychological impact of using this song has proven to be very significant. The classroom atmosphere that might have previously felt tense during pronunciation practice sessions, changed to be more fluid, relaxed, and fun. Music has the power to reduce anxiety levels and create an emotionally safe learning environment. In an atmosphere like this, students feel freer to experiment with their voices without being haunted by excessive fear of making mistakes. The courage to try to pronounce new words, even those considered difficult, arises naturally because the focus has shifted from "assessment" to "participation" in an entertaining activity.

This observation is not without foundation and has been validated by various studies in the field of language teaching. This is in line with a study conducted by Kusumawati et al. (2023), which found that students tend to be more relaxed and brave in pronouncing new words when singing songs compared to when engaging in regular dialogue exercises. According to the study, singing activities can lower the "affective filter", a mental filter that blocks the language acquisition process when learners feel anxious or insecure. When this filter is low, language input can be received and processed more effectively, so that the learning process becomes more optimal.

Ultimately, the increased confidence fostered through song has a lasting impact that goes beyond the singing activity itself. The success and positive feelings that students experience when they successfully imitate the pronunciation in a song gradually transfer to other, more formal speaking tasks. They become more vocal in class discussions, more willing to ask questions, and show greater initiative in conversation. Thus, the integration of song in this second cycle is not just a variation of methods, but a strategic bridge that successfully connects passive pronunciation knowledge with the courage to communicate actively and effectively.

The culmination of a series of learning interventions that have been implemented for two cycles is clearly reflected in the quantitative data obtained from the test results. The test data show convincing empirical evidence of the effectiveness of the approach used. The increase in intelligibility scores from an average of 60 in the initial test to 78 in the final test indicates a significant qualitative leap. An initial score of 60 can be interpreted as a level of intelligibility where communication is often hampered and the listener needs extra effort to understand the message. In contrast, a score of 78 indicates that students have reached a level where their speech is mostly easily understood, marking a transition from struggling communication to effective communication.

increase in intelligibility score is not a single number that stands alone, but rather the result of improvements in several fundamental components of pronunciation. Further analysis shows that the accuracy and fluency aspects have also increased significantly. The

increase in accuracy is seen from the students' ability to produce English segmental sounds (vowels and consonants) more precisely, reducing elementary errors that previously often occurred. On the other hand, the increase in fluency is reflected in the use of rhythm, word stress, and intonation that are more natural, making their speech no longer sound stiff and choppy. This combination of more accurate words and more flowing delivery is what synergistically builds a higher level of intelligibility.

These measurable improvements can be directly attributed to the two-cycle methodology that integrated audio-visual media. Exposure to authentic models through films in the first cycle sharpened students' ability to hear and analyze the elements of correct pronunciation. Furthermore, the use of songs in the second cycle successfully lowered psychological barriers, encouraging students to practice what they had learned in a relaxed and motivating atmosphere. This dual approach—combining critical analysis with confident practice—has proven to be a powerful formula for turning theoretical understanding into real, practical skills.

These findings not only stand as evidence of success in this classroom context, but also provide reinforcement to the broader corpus of research in the field of English language teaching. The results of this study directly reinforce the study conducted by Putra et al. (2022), who also concluded that the use of audio-visual media was able to improve aspects of students' pronunciation gradually. The keyword "gradually" is very relevant here, because the significant increase from a score of 60 to 78 did not happen overnight, but rather through a continuous and structured process. Thus, this study provides additional empirical evidence to support the argument that authentic media is a vital pedagogical tool for meaningful phonological development.

Overall, the data presented suggests a strong conclusion: systematic instructional interventions utilizing audio-visual media such as films and songs have been shown to significantly improve students' intelligibility, accuracy, and fluency in pronunciation. These numbers are more than statistics; they represent a transformation in students' communicative abilities. The implications for English language teachers and curriculum designers are clear: prioritize the integration of authentic media as a core component of pronunciation instruction, rather than as a secondary activity. The goal is to equip students with the skills that will enable them to communicate clearly and confidently on the global stage.

In addition to the quantitative evidence demonstrated by improved test scores, a deeper understanding of the efficacy of this learning method can be explored through qualitative data from the students themselves. Based on the results of interviews conducted after the completion of both learning cycles, a very positive and consistent perspective was revealed. Students collectively voiced that the approach that integrated film and song media had fundamentally changed their learning experience. This perspective is crucial because it highlights that the effectiveness of a method lies not only in the measurable end result, but also in the learning process and experience felt directly by the students.

One of the strongest themes that emerged from the interviews was in the affective domain, where students stated that pronunciation learning became much more enjoyable and less boring. This was a sharp contrast to traditional methods that they often found dry, monotonous, and stressful. The entertainment elements embedded in movies and songs successfully broke down the walls of boredom and created a more lively and dynamic learning

environment. When the learning process was wrapped up in an activity they enjoyed, their internal motivation to actively participate increased dramatically, transforming pronunciation sessions from an academic obligation to an activity they looked forward to.

From a cognitive perspective, students also feel that this method makes learning materials easier to follow and understand. The complexity of pronunciation, which often feels abstract when explained theoretically, becomes more concrete and real through audio-visual media. In films, they can see firsthand how native speakers' lip movements and facial expressions correlate with the sounds they produce. Meanwhile, the melody and rhythm in songs serve as powerful mnemonic aids, helping intonation patterns and word stress stick in their memory more easily and naturally than memorizing rules from a book.

More than just the technical aspects of pronunciation, students explicitly stated that the media had opened their eyes to the socio-cultural dimensions of language. Through the dialogues in the film, they were considered able to understand the social context of English use directly. They could observe how word choice, level of formality, and intonation changed depending on the situation, the relationship between speakers, and the purpose of communication. This was an authentic form of pragmatic learning, where they learned not only 'how to say' but also 'when and why to say it in a certain way'. This experience provided a more holistic understanding that language is a living communication tool, not just a collection of grammatical and phonetic rules.

In conclusion, student feedback confirms that the success of this method is driven by a combination of emotional engagement and cognitive ease of understanding. Their perspectives prove that a positive, enjoyable, and contextually relevant learning environment is a powerful catalyst for effective learning. Quantitatively measurable test score improvements were found to be rooted in qualitatively meaningful learning experiences. Therefore, listening to students and prioritizing methods that are not only academically effective but also personally engaging is key to designing truly successful and impactful language teaching programs.

Conclusion

This study concluded that the use of audio-visual media based on films and songs is effective in improving students' pronunciation skills. Through authentic exposure to films and pronunciation exercises through songs, students experienced significant improvements in intelligibility, accuracy, and fluency. Pronunciation learning becomes more interesting, contextual, and reduces speaking anxiety.

Films provide exposure to native speaker pronunciation in a variety of real-life communication contexts, while songs play a key role in building confidence through repetition and a relaxed learning environment. The combination of these two media creates a learning environment that is balanced between formal exposure and relaxed practice.

This study is in line with the latest findings from Yates (2021), Kusumawati et al. (2023), and Rahman et al. (2023) which emphasize the importance of using contextual and authentic media in pronunciation learning. It is recommended that English teachers utilize audio-visual media optimally in classroom learning activities.

The limitations of this study lie in the limited number of samples and the relatively short duration of the study. Therefore, further research is recommended to involve more diverse

subjects and a longer time so that the results are more representative. In the future, similar research can develop interactive digital media based on applications as a variation of audio-visual media.

References:

- 1) Creswell, J. W., & Poth, C. N. (2021). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.
- 2) Gilakjani, A. P. (2021). A review of factors affecting pronunciation learning. *International Journal of Research in English Education*, 6(1), 23–32.
- 3) Kusumawati, S., Nugroho, S. E., & Anggraeni, D. (2023). The use of English songs in improving students' pronunciation: A classroom action research. *ELT Forum: Journal of English Language Teaching*, 12(1), 12–22.
- 4) Putra, D. W., et al. (2022). Enhancing EFL students' pronunciation through audio-visual media: An experimental study. *Journal of Language Teaching and Research*, 13(5), 963–970.
- 5) Rahman, A., Syahrul, M., & Fitriani, R. (2023). Audio-visual media in English pronunciation teaching: A systematic review. *Journal of Language and Literature Education*, 8(2), 55–69.
- 6) Yates, L., & Zielinski, B. (2020). *Pronunciation for English as an international language*. Routledge.