The 4th Undergraduate Conference on Applied Linguistics, Linguistics, and Literature (ALLURE)
Vol xxx, March, 2026
Online ISSN: 2807-8926

Students' Perceptions of Using YouTube English Songs on Their Pronunciation Skills

Agung Haris Santoso [1], Siti Musarokah[2], Fitri Yulianti[3]
[1,2.3]Universitas PGRI Semarang
Email: harissan007@gmail.com

Abstract

The aims of this study are to identify the perception of ninth-grade students' regarding their use of YouTube English songs as a supplementary tool for pronunciation learning, and identified their learning strategies when employing YouTube Songs on their pronunciation. This research used a descriptive quantitative design with purposive sampling data were collected from Grade IX-B students at SMP Kartika III-2 Semarang in the 2025/2026 academic year through a 30-item Likert-scale questionnaire covering six indicators: Habit and Interest, Pronunciation Aspects, Confidence and Lyrics Comprehension, Motivation and Practice Habits, Comparison with Other Media, and Long-term Perception and Sustainability. Validity and reliability were confirmed through Pearson Product-Moment correlation and Cronbach's Alpha ($\alpha \geq 0.70$). Results showed perceptions ranged from moderate to high, with the highest scores in Habit and Interest (M = 3.58) and Pronunciation Aspects (M = 3.56), indicating strong engagement and perceived benefits in improving stress, intonation, and articulation. Motivation and Practice Habits (M = 3.44) and Comparison with Other Media (M = 3.45) were also high, while Confidence and Lyrics Comprehension (M = 3.40) and Long-term Perception and Sustainability (M = 3.35) were moderately high. Overall, YouTube English songs are perceived as engaging, motivating, and effective for pronunciation improvement, though long-term impact may require structured integration into formal instruction.

Keyword: YouTube, English songs, pronunciation, students' perception

Introduction

English, as a global lingua franca, plays a pivotal role in international communication, education, and career development. In Indonesia's EFL (English as a Foreign Language) context, speaking remains one of the most challenging skills to master, with pronunciation serving as a critical component for intelligibility (Celce-Murcia, Brinton, & Goodwin, 2010). Without accurate pronunciation, even grammatically correct utterances risk being misunderstood. However, in many Indonesian classrooms, pronunciation is often underemphasized, with greater focus placed on grammar and vocabulary. This limited exposure to natural spoken English leaves learners able to comprehend written texts but struggling to speak with clarity and confidence.

Pronunciation difficulties arise from both linguistic and pedagogical factors. The phonological system of English differs markedly from Bahasa Indonesia, introducing unfamiliar vowel sounds, stress patterns, and intonation (Jenkins, 2000). Moreover, many teachers lack specialized training in pronunciation instruction, resulting in insufficient practice and feedback for students. Authentic, repeated listening input is essential for developing accurate speech (Ur, 1984), yet classroom materials often rely on scripted dialogues that fail to reflect natural English usage.

Technology offers opportunities to address this gap, with YouTube emerging as a prominent tool for language learning (Al-Jarf, 2007). Its audiovisual resources, ranging from interviews to lyric videos, provide learners with authentic input, visual support, and flexible, self-paced learning (Almurashi, 2016). Among its diverse content, English songs are particularly effective for pronunciation practice. Their rhythm, melody, and

repetition help reinforce stress, intonation, and connected speech (Murphey, 1992; Medina, 1993). Songs also integrate emotional engagement, which lowers learners' affective filter (Krashen, 1982) and promotes active participation.

YouTube's lyric video features allow learners to connect spoken and written forms, enhancing both listening and pronunciation skills. Learners can replay content, adjust playback speed, and choose songs aligned with their interests, fostering autonomy and motivation (Richards & Renandya, 2002). This is particularly relevant at Junior High School Kartika III-2 Semarang, where students face limited English exposure and pronunciation challenges. Integrating English songs from YouTube into instruction can bridge the gap between formal lessons and real-world language use, leveraging students' existing familiarity with the platform.

Previous studies have examined English songs in EFL learning (Rahmah, 2019; Puspitasari, 2020) and YouTube as a supplementary tool (Lestari, 2018), but few have focused on students' perceptions of using YouTube English songs specifically for pronunciation improvement. Understanding these perceptions is critical, as they influence learners' engagement, motivation, and sustained use of the medium (Bandura, 1997; Zhang, 2016).

Therefore, this study investigates ninth-grade students' perceptions of using YouTube English songs to enhance pronunciation. It explores perceived benefits for specific pronunciation features such as word stress, intonation, and articulation, students' learning strategies, and the frequency of use beyond the classroom. By capturing learners' voices, the research aims to inform English teachers, curriculum designers, and future researchers on the potential of integrating music-based digital media into pronunciation instruction for Indonesian junior high school contexts.

Research Methodology

This study employed a descriptive quantitative research design to objectively present students' perceptions in numerical form without manipulating variables (Sugiyono, 2014; Gay, Mills, & Airasian, 2012). The research was conducted at Junior High School Kartika III-2 Semarang in the 2025/2026 academic year, with the participants consisting of Grade IX-B students selected through purposive sampling due to their similar characteristics and familiarity with listening to English songs on YouTube.

The primary research instrument was a closed-ended questionnaire comprising 30 statements distributed across six indicators: (1) Habit and Interest (students' habits and interest in listening to English songs on YouTube), (2) Pronunciation Aspects (including word pronunciation, word stress, intonation, letter sounds, and recalling the pronunciation of new words), (3) Confidence and Lyrics Comprehension (students' confidence in pronunciation and understanding of song lyrics), (4) Motivation and Practice Habit (motivation and habits in practicing pronunciation), (5) Comparison with Other Media (perceptions of YouTube compared to other learning media), and (6) Long-term Perception and Sustainability (long-term views and sustainability of using YouTube songs for learning).

Each item was measured using a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). The instrument's validity was tested using the Pearson Product-Moment correlation and was confirmed valid (Sig. < 0.05), while its reliability was examined using Cronbach's Alpha, indicating that all indicators were reliable ($\alpha \geq 0.70$).

Data collection was carried out through coordination with the school administration and English teacher, followed by the distribution of the questionnaire via Google Forms. The collected data were analyzed using Microsoft Excel for preliminary processing (mean scores, frequency, and percentage) and SPSS for further analysis, including validity and reliability testing, as well as categorization of scores based on the guidelines by Best & Kahn (2006) and Gay, Mills, & Airasian (2012) into five interpretation levels: very low, low, moderate, high, and very high.

Findings

This section presents the findings of the study regarding students' perceptions of using YouTube English songs to improve their pronunciation skills. The analysis focuses on six indicators: Habit and Interest, Pronunciation Aspects, Confidence and Lyrics Comprehension, Motivation and Practice Habits, Comparison with Other Media, and Long-term Perception and Sustainability. For each indicator, the mean score and its interpretation category were calculated based on the Likert scale classification (Best & Kahn, 2006; Gay, Mills, & Airasian, 2012).

To provide a concise overview, the results are summarized in Table 1, which shows the mean scores and corresponding categories for all six indicators. This summary enables a clear comparison of students' perceptions across different aspects and highlights the areas where YouTube English songs were perceived as most and least beneficial

Table 1.
The Category of Students Perception

Indicator	Mean	Category
Habit and Interest	3.58	High
Pronunciation Aspects	3.56	High
Confidence & Lyrics Comprehension	3.40	Moderate
Motivation & Practice Habits	3.44	High
Comparison with Other Media	3.45	High
Long-term Perception & Sustainability	3.35	Moderate

The results in Table X indicate that students' perceptions of using YouTube English songs for pronunciation improvement generally fall within the moderate to high categories. The highest mean score was recorded for Habit and Interest (3.58, high), suggesting that listening to English songs on YouTube is already a common and well-accepted activity among students. Pronunciation Aspects (3.56, high) also received a strong score, reflecting students' recognition of the songs' usefulness in enhancing elements such as word stress, intonation, and articulation.

Similarly, Comparison with Other Media (3.45, high) and Motivation and Practice Habits (3.44, high) indicate that YouTube English songs are perceived as more engaging and motivating than traditional learning resources, encouraging students to maintain consistent practice. In contrast, Confidence and Lyrics Comprehension (3.40, moderate) and Long-term Perception and Sustainability (3.35, moderate) scored slightly lower, suggesting that while students value the method, its impact on confidence, lyric comprehension, and sustained long-term use may require further reinforcement through guided practice and integration into formal instruction.

In Summary, these findings highlight the positive role of YouTube English songs in supporting pronunciation learning, with the strongest contributions seen in habit formation, direct pronunciation improvement, and motivation, while confidence and long-term sustainability remain areas for potential development.

a. **Students' Perceptions of the Use of YouTube English Songs in Learning English Pronunciation**

The findings indicate that students' perceptions of using YouTube English songs for pronunciation learning are predominantly positive. The Pronunciation Aspects indicator (M = 3.56, high) demonstrates that learners perceive tangible benefits in improving segmental features (e.g., articulation of consonants and vowels) and suprasegmental features (e.g., stress and intonation) through repeated exposure to songs. Similarly, the Comparison with Other Media indicator (M = 3.45, high) reflects students' preference for YouTube over traditional learning tools, suggesting that its multimodal affordances combining audio, visual, and textual input provide richer pronunciation models and more engaging learning experiences.

The Long-term Perception and Sustainability indicator (M = 3.35, moderate) scored comparatively lower, which may indicate that while students recognize the immediate benefits of using YouTube English songs, their sustained commitment could be influenced by factors such as the novelty effect, accessibility constraints, or the absence of structured follow-up activities in formal instruction.

These findings are consistent with previous research underscoring the pedagogical value of music-based and digital platforms for pronunciation learning. Setiawan et al. (2019) reported that integrating English songs into classroom activities significantly enhanced learners' mastery of stress and intonation patterns, while Susanti and Mubarok (2021) found that YouTube lyric videos improved both articulation and learner engagement due to the simultaneous presentation of auditory and visual cues. Furthermore, Alimemaj (2022) emphasized the role of digital platforms like YouTube in enabling

autonomous, repetitive practice, which fosters both pronunciation accuracy and learner confidence over time.

In line with these studies, the current research reinforces the argument that YouTube English songs serve as an effective supplementary medium for pronunciation instruction, offering authentic input, repeated exposure, and motivational value. Nevertheless, the relatively lower score in long-term perception highlights the need for intentional integration of this medium into sustained learning strategies to ensure its enduring impact on students' pronunciation development.

b. **Students' Perceptions of the Usefulness of YouTube English Songs in Learning Specific Pronunciation Aspects**

The results for the Pronunciation Aspects indicator (M = 3.56, high) suggest that students perceive YouTube English songs as a valuable medium for mastering both segmental and suprasegmental pronunciation features. The rhythmic and melodic structure of songs appears to facilitate the internalization of word stress, intonation, and articulation patterns, aligning with Medina's (1993) assertion that music provides a natural context for practicing authentic pronunciation, rhythm, and intonation. This finding is reinforced by Jenkins (2000), who emphasizes that suprasegmental mastery often has a greater impact on intelligibility than segmental accuracy alone.

The Confidence and Lyrics Comprehension indicator (M = 3.40, moderate) reveals that while learners benefit from lyric comprehension as a complementary skill, its influence on pronunciation development may be less direct than that of focused phonological practice. Nonetheless, students' ability to process song lyrics in real time may support their alignment with native-like speech patterns, a point supported by Mora and Valls-Ferrer (2018), who found that enhanced phonological awareness contributes to both accuracy and fluency.

These findings parallel the outcomes of Saito and Plonsky's (2019) meta-analysis, which demonstrated that repeated exposure to authentic listening materials, including songs, significantly improves both segmental accuracy and prosodic features such as stress and intonation. Similarly, Lestari and Hartono (2020) observed that learners using English songs as pronunciation models exhibited clearer articulation and more accurate speech rhythm compared to those taught through conventional drills.

In the context of the present study, the combination of auditory (listening to songs), visual (reading lyrics), and kinesthetic (singing along) engagement aligns with Mayer's (2001) Cognitive Theory of Multimedia Learning, which posits that dual-channel input enhances comprehension and retention. Consequently, while lyric comprehension alone may not be the strongest driver of pronunciation improvement, its integration with active listening and imitation tasks appears to strengthen learners' overall phonological competence.

c. **Frequency of Students' Use of YouTube English Songs as a Tool for Learning Pronunciation Outside the Classroom**

The findings for the Habit and Interest indicator (M = 3.58, high) and the Motivation and Practice Habits indicator (M = 3.44, high) demonstrate that students frequently engage with YouTube English songs beyond the classroom setting. This pattern reflects not only established listening habits but also the integration of these activities into students' daily routines, indicating that YouTube functions as both an entertainment medium and an educational resource. The high mean scores suggest that learners are intrinsically motivated to maintain regular exposure to authentic English pronunciation models, supporting Dörnyei's (2003) claim that motivation is a significant predictor of sustained engagement in language learning.

These results align with the findings of Yuliana and Susanto (2020), who reported that students incorporating English songs into their daily routines achieved greater improvement in pronunciation due to extended exposure and opportunities for informal learning. Similarly, Fitria (2019) found that frequent engagement with music-based input outside the classroom correlates positively with pronunciation accuracy, as repeated listening facilitates imitation, self-monitoring, and correction of speech production.

The role of learner autonomy is also evident in these findings. Shvidko (2020) emphasized that when learners exercise control over the selection and frequency of digital content such as YouTube English songs they are more likely to develop consistent practice habits and achieve long-term retention of phonological features. In this study, students' reported tendencies to actively search for new English songs and to prefer them over songs in other languages indicate a degree of self-directed learning that contributes to the improvement of both segmental and suprasegmental pronunciation skills.

In summary, the frequent use of YouTube English songs outside formal instructional contexts highlights its potential as a supplementary tool for pronunciation learning that fosters learner autonomy, enhances language exposure, and sustains motivation. To optimize these benefits, it is recommended that educators design structured follow-up activities that channel this informal engagement into focused pronunciation practice, thereby ensuring that habitual listening leads to measurable and enduring improvements in learners' pronunciation competence.

Discussion

The results of this study demonstrate that students hold a generally positive perception toward the use of YouTube English songs for enhancing their pronunciation skills. While the descriptive data revealed high mean scores across most indicators, a deeper analysis suggests several nuanced findings that warrant further discussion in relation to existing theories and previous studies. One of the most notable observations is the predominance of "Neutral" responses across many questionnaire items. Although the overall mean scores fall within the "High" category, the prevalence of neutrality may indicate a lack of strong self-awareness among learners regarding the actual benefits of engaging with English songs on YouTube. Students may enjoy the content and recognize it as useful, yet they may not possess sufficient metacognitive reflection to fully articulate how their listening habits contribute to specific improvements in pronunciation. Zhang (2016) argues that students often experience passive enjoyment of digital content without recognizing its potential for language acquisition unless guided explicitly to make linguistic connections. This notion is supported by Gilakjani (2016), who asserts that pronunciation instruction must incorporate conscious noticing strategies to convert input into meaningful phonological development.

Furthermore, the slightly lower mean score in the Confidence and Lyrics Comprehension indicator reveals a potential affective barrier in the learning process. Although students appear motivated and engaged, their limited self-confidence may hinder them from practicing pronunciation actively. According to Bandura's (1997) theory of self-efficacy, learners must believe in their own capacity to succeed before they are willing to undertake challenging tasks, such as imitating pronunciation or speaking aloud. In this context, students who do not feel confident in understanding lyrics or identifying pronunciation features may choose to remain passive listeners. Millington (2011) emphasizes the need for structured guidance and feedback in music-based learning, as simply listening to songs does not automatically lead to productive speech unless supported by meaningful interaction and reinforcement.

The study also indicates that students' long-term perception regarding the use of YouTube English songs is relatively moderate. Although the platform is perceived as beneficial in the short term, there is uncertainty about its sustained use over time. This suggests that, while students appreciate the medium's entertainment value and accessibility, their commitment to consistently using it for pronunciation development may diminish without formal integration into classroom instruction. Mora and Valls-Ferrer (2018) argue that pronunciation gains require gradual, sustained exposure accompanied by structured progression in task complexity. Without curriculum-based support or ongoing teacher feedback, students may lose motivation or shift their focus to other media. This finding reinforces the idea proposed by Almurashi (2016), who recommends incorporating YouTube as a formal supplementary resource within lesson planning to ensure both engagement and sustainability.

Another important consideration lies in the students' comparison of YouTube with other media. Although the results show that students generally favor YouTube, the margin of preference over other platforms is not as pronounced as might be expected. This could be attributed to the limited range of media tools students have been exposed to. Shvidko (2020) notes that digital literacy plays a crucial role in shaping students' perceptions of learning tools. If learners are only familiar with YouTube and textbooks, their ability to critically evaluate the comparative effectiveness of each medium may be constrained. Thus, it is essential for educators to introduce a variety of digital learning platforms, including podcasts, pronunciation apps, and interactive AI tools, to foster critical reflection and broader engagement.

The absence of qualitative data in this study also presents a limitation in understanding students' individual learning behaviors. While the closed-ended questionnaire provided valuable quantitative insights, it did not capture more personal aspects, such as students' preferred song genres, tempo preferences, or strategies used when listening. These details are significant because musical features such as rhythm, melody, and tempo have been shown to influence pronunciation acquisition. Fitria (2019) and Yuliana & Susanto (2020) found that students' progress in pronunciation can vary depending on their familiarity with certain music genres and

their ability to follow the lyrical rhythm. Including open-ended responses or conducting interviews in future studies would provide richer context and allow for a more comprehensive understanding of learner behavior.

In other hand, the data suggest that while YouTube is a promising and engaging tool for pronunciation learning, its pedagogical value depends greatly on how it is implemented. The findings emphasize the need for a balanced approach one that combines student autonomy with structured support, guided reflection, and diverse resource exposure. As Mayer (2001) points out in his multimedia learning theory, meaningful learning occurs when students actively process content through multiple channels visual, auditory, and textual. YouTube English songs inherently provide this multimodal input, but without guided instructional strategies, the cognitive benefits may not be fully realized.

Moreover, it is important to consider these findings within the broader context of English language education in Indonesia, where students often have limited exposure to authentic English input outside the classroom. In such environments, platforms like YouTube provide an accessible and cost-effective solution to supplement formal instruction. However, without proper pedagogical framing, these tools risk being underutilized or reduced to passive entertainment. As Robbins and Judge (2013) argue, behavior is influenced not only by access to resources but also by reinforcement, guidance, and organizational structure. Therefore, language teachers must move beyond recommending YouTube songs as homework and instead curate purposeful learning tasks such as lyric gap-fills, pronunciation shadowing, and peer review sessions using selected video content.

In this regard, the teacher's role evolves from being a transmitter of knowledge to a facilitator of learning autonomy and digital navigation. By scaffolding learners' engagement with YouTube content, educators can help students transition from passive listeners to active participants in their pronunciation development. This is particularly important in promoting noticing, a key concept in second language acquisition theory proposed by Schmidt (1990), which emphasizes that learners must consciously notice linguistic features in order to acquire them. Through guided viewing, reflective discussions, and follow-up speaking activities, teachers can create an environment where learners notice, analyze, and apply pronunciation patterns embedded in songs.

Additionally, the implementation of YouTube as a learning tool offers opportunities for learner differentiation. Given the diversity of student learning styles and language proficiency levels, YouTube allows students to select content based on their interests, preferred accents, or difficulty level. This aligns with Krashen's (1982, 1985) input hypothesis, which asserts that language acquisition occurs most effectively when learners are exposed to input that is both comprehensible and slightly above their current level (i+1). When teachers help students identify videos that meet these criteria, the platform can serve as a powerful means of personalized learning, addressing both cognitive and affective domains.

In conclusion, the discussion of these findings demonstrates that the effectiveness of YouTube English songs in pronunciation learning is multifaceted. It is not solely the availability of the content that ensures learning, but the combination of learner awareness, confidence, structured use, and long-term engagement. Educators, therefore, must play an active role in designing meaningful tasks around YouTube content, fostering learner reflection, and sustaining motivation to ensure this digital tool becomes a lasting asset in pronunciation development.

Conclusion

This study concludes that students generally hold positive perceptions toward the use of YouTube English songs as a supplementary medium for pronunciation learning. Indicators such as Pronunciation Aspects and Comparison with Other Media were consistently rated in the high category, suggesting that learners find YouTube an engaging and effective platform for improving key pronunciation features, including word stress, intonation, and articulation. The Habit and Interest and Motivation and Practice Habits indicators also scored highly, indicating that students frequently integrate this medium into their daily routines and view it as a motivating tool for sustained pronunciation practice.

However, the Confidence and Lyrics Comprehension and Long-term Perception and Sustainability indicators received moderately high scores, implying that while students recognize the immediate benefits of this learning method, its long-term effectiveness may depend on continued integration into formal instruction and the provision of structured follow-up activities. Overall, the findings support the potential of YouTube English songs to enhance pronunciation skills, foster learner autonomy, and increase motivation, provided that their use is strategically incorporated into ongoing language learning programs.

References

Alimemaj, Z. (2022). *YouTube as a learning tool for improving EFL pronunciation*. International Journal of Applied Linguistics, 32(2), 54–63.
Al-Jarf, R. (2007). *YouTube as a tool for EFL classrooms*. Proceedings of the 2007 International Conference on the Future of Education, 1–5.
Almurashi, W. A. (2016). The effective use of YouTube videos for teaching English language in classrooms as supplementary material at Taibah University in Alula. *International Journal of English Language and Linguistics Research, 4*(3), 32–47.
Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge: Cambridge University Press.
Dörnyei, Z. (2003). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.
Fitria, T. N. (2019). The use of English songs in improving students' pronunciation skills. *Journal of English Language Studies, 4*(1), 1–10.
Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
Lestari, D., & Hartono, R. (2020). The effect of English song media on students' pronunciation mastery. *Journal of English Education, 8*(2), 145–153.
Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
Medina, S. L. (1993). The effect of music on second language vocabulary acquisition. *National Network for Early Language Learning, 6*(3), 1–8.
Mora, J. C., & Valls-Ferrer, M. (2018). Oral fluency, accuracy, and complexity in L2 speech production: The role of phonological awareness. *Language Learning, 68*(1), 31–66.
Murphey, T. (1992). *Music and song*. Oxford: Oxford University Press.
Richards, J. C., & Renandya, W. A. (Eds.). (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge University Press.
Saito, K., & Plonsky, L. (2019). Effects of instruction on L2 pronunciation: A meta-analysis. *Applied Linguistics, 40*(3), 1–27.
Setiawan, H., Sulistyo, G. H., & Nurhayati, D. A. W. (2019). The use of songs to improve students' pronunciation. *Journal of English Teaching and Learning, 9*(1), 50–60.
Shvidko, E. (2020). Language learner autonomy in the digital era. *TESOL Journal, 11*(3), 1–15.
Susanti, D., & Mubarok, H. (2021). The effectiveness of YouTube lyric videos in improving students' pronunciation. *Jurnal Pendidikan Bahasa Inggris, 8*(2), 120–129.
Ur, P. (1984). *Teaching listening comprehension*. Cambridge: Cambridge University Press.
Yuliana, E., & Susanto, D. (2020). The influence of listening to English songs on students' pronunciation ability. *Journal of English Language Teaching, 8*(1), 35–44.
Zhang, Y. (2016). Student perceptions of multimedia classroom teaching. *English Language Teaching, 9*(4), 21–29.