

FROM SCROLL TO SKILL 2.0: INTEGRATING YOUTUBE AND AI TOOLS TO ENHANCE STUDENTS' EXTENSIVE LISTENING SKILLS

Dodie Ardian Sakuta¹, Jusriati², Puspa Sari³

¹ Pendidikan Bahasa Inggris, Universitas Muhammadiyah Palopo,

² Pendidikan Bahasa Inggris, Universitas Muhammadiyah Palopo,

³ Pendidikan Bahasa Inggris, Universitas Muhammadiyah Palopo,

¹ dodieardiansakuta@gmail.com, ²jusriati@umpalopo.ac.id,

³puspasari@umpalopo.ac.id

ABSTRACT

This study investigates how the integration of YouTube and artificial intelligence (AI) tools enhances students' extensive listening skills in English as a Foreign Language (EFL) contexts. The research addresses a key issue in digital language learning: while students frequently use YouTube for exposure to authentic language, the absence of structured scaffolding and strategy instruction often limits comprehension gains. Focusing on this problem, the study explores how AI-assisted tools—such as real-time transcription, automatic subtitles, summarization, and adaptive feedback—can transform YouTube-based listening from passive consumption into active, autonomous learning. Using a qualitative descriptive approach, data were collected through semi-structured interviews with four purposively selected students from We Learn English Students Course in Palopo, an English language training center. Data collection and analysis followed a three-phase procedure: transcription, member checking, and coding to uncover learners' strategies, motivations, and perceived skill development. The findings reveal that AI-supported YouTube activities significantly improved learners' comprehension, pronunciation, and self-regulated learning behaviors. Students demonstrated increased metacognitive awareness, utilizing AI tools to plan, monitor, and evaluate their listening performance. Furthermore, intrinsic motivation and learner autonomy were strengthened, as participants selected content aligned with their interests and received personalized feedback. The study concludes that AI-mediated YouTube engagement fosters a shift “from scroll to skill,” transforming entertainment-oriented digital habits into meaningful language learning experiences. The results underscore the pedagogical potential of integrating AI-driven scaffolding into extensive listening instruction to promote motivation, competence, and learner independence in EFL settings.

Keywords: extensive listening, youtube, artificial intelligence, learner autonomy

A. Introduction

In recent years, the integration of digital media into language education has significantly reshaped the way learners develop their communicative competencies. Platforms such as YouTube have emerged as potent tools for enhancing extensive listening, a crucial skill in second language acquisition (Widiyanto et al., 2021; Lestari et al., 2023). Extensive listening enables learners to absorb authentic input at their own pace, fostering comprehension and fluency through repeated exposure. Research has emphasized that digital video content provides not only linguistic models but also rich cultural contexts, which are instrumental in reinforcing vocabulary, intonation, and pragmatic awareness (Shafwati et al., 2021; Gönülal, 2022). In parallel, the rise of artificial intelligence (AI) technologies has introduced new dimensions to learner support, offering personalized guidance and immediate feedback through intelligent tutoring systems and real-time transcription tools (Suryana et al., 2020; Khotimah et al., 2024). Together, these technologies are transforming passive viewing into

interactive and strategic learning experiences. The convergence of YouTube and AI tools offers unprecedented opportunities to improve not only linguistic input exposure but also learner autonomy, motivation, and engagement—key principles in modern language pedagogy.

Within the Indonesian context, these developments hold particular relevance. The national education policy emphasizes digital literacy, learner autonomy, and communicative competence as core goals to prepare students for global challenges. The widespread accessibility of mobile devices and internet connectivity in Indonesia has catalyzed the popularity of platforms like YouTube among youth learners, transforming them into informal yet powerful learning environments (Ate et al., 2025; Fradana et al., 2025). The implementation of multimodal pedagogy through these platforms aligns with Indonesia's broader vision of fostering independent learners equipped for international discourse. However, while learners engage enthusiastically with digital content, studies suggest that the lack of instructional scaffolding often limits

the pedagogical potential of such media (Ramadhani & Samsi, 2023; Rahila et al., 2023). The challenge thus lies not in the availability of resources but in designing learning frameworks that harness these tools effectively to foster cognitive, metacognitive, and motivational dimensions of listening comprehension. The present study is situated within this context of evolving educational practices and technological affordances.

Despite the ubiquity of YouTube and the increasing prevalence of AI applications in language learning, a significant research gap remains regarding their structured integration to enhance students' extensive listening skills. Current literature underscores the motivational and experiential benefits of using digital videos; however, passive engagement with such content often results in superficial learning gains (Afrizal & Herlina, 2023; Gönülal, 2022b). Students may enjoy watching English-language videos but frequently lack the metacognitive strategies necessary to convert this exposure into effective comprehension practice (Mahdavy & Namavar, 2023; Wang & Treffers-

Daller, 2017). Moreover, the overload of unfiltered content can hinder cognitive processing, especially among learners unfamiliar with structured language input. Without pedagogical design, digital tools may fall short of facilitating meaningful, long-term improvement in listening proficiency. These issues are exacerbated in contexts where learners are expected to self-regulate their engagement without clear instructional direction.

In response to these challenges, educational researchers and practitioners have begun advocating for a structured, strategic approach to digital listening practices. This involves aligning video content with pedagogical goals, incorporating pre-, while-, and post-listening tasks, and embedding metacognitive strategy instruction (Goh, 2008; Vandergrift, 2005). The integration of AI tools in this framework can amplify its impact by providing personalized, adaptive support that enhances learners' autonomy and competence (Walker et al., 2025; Mokhtari, 2025). For instance, AI applications can generate summaries, speech-to-text conversions, and real-time comprehension feedback, thereby

promoting active listening and reflection. This general solution—merging the motivational appeal of YouTube with the cognitive support of AI—has the potential to transform extensive listening into a strategic, learner-centered practice. Nonetheless, empirical research is needed to validate these pedagogical claims, particularly in the Indonesian educational landscape where such tools are becoming increasingly prevalent but remain underutilized in structured language instruction.

One promising pedagogical response to the aforementioned challenges lies in the structured integration of YouTube videos within listening instruction supported by AI-enhanced tools. Scholars have advocated for guided listening frameworks that sequence learning into pre-listening, while-listening, and post-listening stages, thereby optimizing cognitive processing and retention (Yan, 2023). In this approach, students are prepared for comprehension through contextual clues or vocabulary support before watching a video; they are then engaged in active listening through targeted tasks; and finally, they reflect and consolidate understanding via

summaries, discussions, or comprehension questions. This method ensures that learners interact meaningfully with content rather than merely consuming it passively. When YouTube content is curated around specific linguistic targets—such as speech speed, accent diversity, or thematic relevance—its educational value is significantly enhanced (Fadillah et al., 2023; Saputra & Fatimah, 2018). Moreover, AI-powered tools offer a critical layer of cognitive and metacognitive support during the listening process. For example, real-time subtitles and speech recognition software facilitate comprehension by bridging auditory and textual inputs (Hosogoshi, 2016). AI also enables adaptive learning through tools that analyze student performance and provide immediate feedback, thus fostering metacognitive awareness and personalized learning paths (Suvorova & Demirbilek, 2025). Recent research emphasizes the value of integrating AI at each phase of the listening task. In the pre-listening phase, AI tools can generate summaries or anticipate key themes; during the task, they provide real-time assistance; and after listening, they

offer analytics and reflective prompts to deepen comprehension (Yifan et al., 2025). These tools not only mitigate cognitive overload but also encourage learners to develop autonomy and self-regulated learning strategies—capabilities aligned with the goals of 21st-century education.

Despite growing support for the integration of YouTube and AI in language learning, empirical studies reveal persistent limitations in current instructional practices. One of the most significant gaps concerns the absence of scaffolding and strategic listening instruction in digital media use. For instance, Ramadhani and Samsi (2023) emphasize that students often consume YouTube content passively without engaging in structured listening strategies. Similarly, Afrizal and Herlina (2023) report that while students exhibit increased confidence in speaking after watching YouTube videos, the absence of guided tasks undermines listening comprehension gains. These studies reveal that without metacognitive instruction or purposeful task design, digital media becomes a source of entertainment rather than an effective learning tool. Moreover, learners may experience

cognitive overload due to the sheer volume and complexity of online content, which further impairs the development of listening proficiency (Vellanki et al., 2022). Additionally, there is limited research that systematically explores how AI can be used to support metacognitive processes in listening instruction. While existing literature acknowledges AI's potential in providing feedback and learner analytics, few studies have investigated its application in guiding pre-listening prediction, real-time comprehension scaffolding, and post-listening reflection (Walker et al., 2025). This represents a critical gap, as language learners often struggle with managing their listening tasks effectively. Structured digital programs that combine the motivational appeal of YouTube with the adaptive capacity of AI remain under-researched, particularly in the context of Indonesian secondary and tertiary education. As Indonesia seeks to enhance learners' communicative competencies through digital innovation, it is imperative to investigate how such integration can foster not only listening proficiency but also learner autonomy and engagement. This study, therefore,

aims to fill this void by offering an empirical examination of AI-assisted YouTube listening practices.

The objective of the present study is to investigate how the integration of YouTube and artificial intelligence (AI) tools can enhance students' extensive listening skills, with particular attention to the development of learner motivation, comprehension, and autonomy. Specifically, the study explores the pedagogical effectiveness of structured listening frameworks supported by AI tools such as real-time transcription, adaptive feedback, and post-listening analytics. By examining student interaction with YouTube content facilitated by AI, the study aims to determine the extent to which these digital resources contribute to improved listening proficiency in English as a Foreign Language (EFL) contexts. The novelty of this study lies in its systematic approach to combining the immersive nature of YouTube with the adaptive, metacognitive support of AI technologies—a pedagogical intersection that remains underexplored in current literature. While prior studies have independently examined the benefits

of YouTube or AI tools in language learning, few have addressed their integrated application within a structured listening framework. Furthermore, this study is situated within the Indonesian educational context, which has unique challenges related to digital literacy, language instruction infrastructure, and learner autonomy. Thus, the research not only contributes to the growing body of literature on digital learning in EFL contexts but also responds directly to national priorities concerning digital competence and communicative skill development.

B. Research Methodology

This study employed a qualitative descriptive research design to explore how students engage with YouTube videos enhanced by AI tools for extensive listening practice. The qualitative approach was chosen for its suitability in capturing participants' lived experiences and contextual interpretations of their learning behaviors (Creswell, 2012). By focusing on real-world interactions with digital learning tools, the study aimed to construct a comprehensive understanding of how AI-supported

YouTube practices foster language development. The research was conducted over a two-month period at the We Learn English Students Course in Palopo City, an English language training center where learners actively incorporate digital technologies into their independent study routines. The study specifically focused on examining the cognitive and behavioral strategies used by learners when engaging in extended listening activities using YouTube in combination with AI applications such as auto-captioning, transcription tools, and summarization platforms.

The participants were selected using purposive sampling, a technique suitable for qualitative research in which individuals are chosen based on specific criteria relevant to the research objectives (Ulfert-Blank & Schmidt, 2022). Four students from the 2017 English Language Education (ELED) cohort were chosen to participate in the study. These participants had prior experience with listening skill courses and reported frequent engagement with YouTube to enhance their English listening proficiency outside of classroom settings. Pseudonyms were assigned to each participant to ensure

confidentiality and to protect personal data. The study instruments included a semi-structured interview guide, a pen and notebook for note-taking, and a smartphone used to digitally record the interviews. The interviews were conducted in Bahasa Indonesia via WhatsApp and Zoom, depending on participant availability, and lasted between 6 to 17 minutes. Open-ended questions were employed to elicit detailed responses about participants' motivations, strategies, and reflections related to AI-assisted extensive listening.

Data collection and analysis followed a three-phase procedure: transcription, member checking, and coding. Interviews were first transcribed to transform spoken dialogue into written text with pseudonyms used throughout to preserve anonymity. Subsequently, member checking was conducted by sending transcripts to participants for validation and accuracy assurance. The coding process, as defined by Creswell (2012), involved organizing the data into meaningful categories through a three-step procedure: constructing a narrative from the data, generating and assigning thematic codes using Microsoft Word, and

writing memos to justify code selections and support analytical interpretation. This method allowed the researcher to systematically identify patterns, themes, and insights related to how AI-supported YouTube activities contribute to learners' listening skill development.

C. Results and Discussion

YouTube as an AI-Enhanced Authentic Learning Source

The first finding from the qualitative data reveals that all participants perceived YouTube as more than just an entertainment platform. They described it as an expansive, global learning environment that facilitates exposure to authentic English usage across diverse genres, including vlogs, tutorials, interviews, and academic lectures. Participants unanimously valued YouTube's real-world content, emphasizing its effectiveness in providing contextual and meaningful language input. Significantly, they highlighted the integration of AI-driven features—such as automatic captioning, speech recognition, and content summarization—as key elements that enhance listening comprehension and accessibility.

Participants specifically mentioned how these AI tools aid in recognizing diverse accents, decoding rapid speech, and clarifying unfamiliar vocabulary. One participant (P2) noted: *"The subtitles generated by YouTube really help me catch words that I miss. Sometimes I also use ChatGPT to summarize the video and check my understanding."* This suggests that AI technologies embedded in or used alongside YouTube serve not only as support mechanisms but also as scaffolding tools that make listening practice more manageable and effective. These findings underscore the adaptive potential of AI-supported YouTube engagement, aligning well with the principles of comprehensible input and ecological learning. The platform, when mediated through AI, becomes an immersive site where learners interact with authentic language in real-time, facilitating deep linguistic exposure.

The findings affirm the central research question: How do students' activities using YouTube with AI support facilitate the development of their listening skills? Participants' reflections indicate that AI-enhanced YouTube content provides a rich and

adaptive environment that significantly contributes to their listening proficiency. These findings align closely with Krashen's Comprehensible Input Hypothesis, which emphasizes the value of slightly challenging input ("i+1") in language acquisition (Nguyen, 2025; Lichtman & VanPatten, 2021). By offering real-world linguistic contexts, YouTube—augmented with AI tools—delivers language that is both comprehensible and cognitively stimulating.

Compared to traditional textbook audio or scripted language materials, the real-life variability and complexity of YouTube videos offer broader exposure to spoken English in authentic contexts (Huang et al., 2018). This aligns with Đurić and Gandhi (2024), who affirm that AI-supported platforms create customized listening experiences tailored to individual learner levels. The current study extends these assertions by demonstrating how AI tools—like summarization applications and auto-captioning—enhance the learning process by scaffolding understanding and promoting active engagement. This research also supports the ecological perspective of language acquisition,

wherein learners' interaction with their environment—here, YouTube augmented by AI—plays a crucial role in constructing meaningful learning (Li, 2022). Unlike studies emphasizing passive exposure, the present findings reveal how learners actively utilize tools such as ChatGPT to engage with and reflect on video content, a behavior that supports metacognitive development and language internalization (Nowbakht & Shahnazari, 2015).

This research builds upon the findings of Ma'rufah et al. (2024), who highlighted that extensive listening through podcasts enhances vocabulary and pronunciation. However, this study goes beyond by illustrating how YouTube, supported by AI, not only offers repeated exposure to authentic speech but also provides personalized, responsive feedback mechanisms that deepen understanding and learner autonomy—features often lacking in static podcast formats. Contrastingly, some literature notes the limitations of YouTube when used without pedagogical oversight (Yang, 2020; Khotibi & Hadi, 2022). Learners may experience cognitive overload or become distracted by irrelevant

content. However, the current study suggests that when learners are guided by self-directed goals and utilize AI tools strategically, they can mitigate these challenges and transform YouTube into a productive learning space. This marks a significant contribution to existing scholarship, proposing that structured autonomy—rather than teacher-centered instruction—is sufficient when coupled with intelligent tools.

Students' Listening Strategies in AI-Supported YouTube Engagement

Participants in this study demonstrated a range of self-developed strategies to support their listening comprehension while engaging with YouTube and AI tools. What stands out is the natural emergence of both cognitive and metacognitive behaviors, not as a result of formal instruction, but through individual experimentation. Learners described adjusting playback speeds, enabling AI-generated captions, and re-watching segments to manage comprehension difficulties. Beyond that, they took notes on new vocabulary, used AI dictionaries and translators to explore meanings, and engaged in active imitation practices

such as shadowing and repeating after native speakers in the videos. More advanced metacognitive strategies were also evident. Participants spoke about summarizing video content using AI summarization tools like ChatGPT or Otter.ai, and even generating comprehension questions to self-assess their understanding. One student (P3) reflected, “*After watching a video, I usually ask ChatGPT to make questions about it. It helps me test whether I really understood the content.*” This statement encapsulates a shift from passive video consumption to active, reflective, and autonomous learning. These strategies reflect a high level of learner agency and awareness, indicating that the integration of AI tools did not just enhance access to content, but also reshaped how students approached their own learning processes.

This finding shows a natural alignment with Andringa et al. (2012) framework, which identifies effective listening as a multi-step process involving planning, monitoring, and evaluating comprehension. What makes this case unique, however, is that participants independently enacted these steps—without explicit

strategy instruction—using AI tools as scaffolds. For instance, using ChatGPT to generate comprehension questions or to rephrase content mimics the evaluation and reflection phases of Vandergrift’s model. This organic strategy development offers evidence that digital tools, when appropriately leveraged, can foster deep metacognitive engagement even in informal or out-of-class learning contexts. Other studies, such as those by Liu (2025), emphasizes the importance of structured metacognitive instruction in the classroom to improve listening comprehension. Yet, what this research adds is a compelling case for learner-driven metacognition—facilitated not by the teacher, but by the learner’s interaction with AI tools and authentic content. This expands the narrative around listening strategy development and shows that structured instruction, while valuable, is not the only path toward successful strategy use.

In contrast to studies that underscore the teacher’s role in modeling or teaching listening strategies (Rong & Xue, 2018; Cross, 2009), this study reveals how learners, motivated by interest and need,

creatively develop techniques when given control and the right tools. This reflects the principles of Self-Determination Theory (SDT), particularly the dimensions of autonomy and competence. As students explored tools like captioning, summarization, and speech imitation, they began exercising control over their learning, reinforcing their sense of competence through self-evaluation and performance feedback (Ntoumanis, 2005; Veas et al., 2018).

This interdependence between autonomy and competence is echoed in more recent findings (Lin et al., 2024), where students’ perception of control and their ability to self-manage learning were linked to higher motivation and sustained engagement. Our findings mirror this: students who felt empowered to choose their videos, manipulate playback, generate AI prompts, and evaluate their understanding were not just more engaged, but also more reflective. This level of metacognitive sophistication in an informal learning context is a strong argument for integrating AI tools more consciously into listening instruction. Nonetheless, it’s worth noting that not all learners

may naturally adopt such strategies. The participants in this study were advanced learners with prior experience in extensive listening, and this may have influenced their ability to experiment and reflect. Less proficient students may still require explicit guidance to benefit from similar tools, as highlighted by Vandergrift and Tafaghodtari (2010).

Motivation and Learner Autonomy

An analysis of participant responses revealed that motivation—particularly intrinsic motivation—played a pivotal role in sustaining students' engagement with YouTube-based listening activities. Learners consistently expressed a sense of enjoyment and personal investment in selecting content aligned with their interests, such as music, podcasts, or specific YouTube channels. This preference-based engagement, facilitated by platform algorithms and AI-assisted recommendations, created a dynamic and self-directed learning experience. Participants felt empowered by the freedom to curate their own listening paths, which, in turn, kept them emotionally invested and consistently involved.

A recurring theme was the way AI tools contributed to this sustained motivation. Learners reported that personalized suggestions and interactive feedback systems transformed passive viewing into meaningful exploration. One participant (P1) remarked: *“Because AI recommends new videos based on what I like, I don’t get bored. I feel like I’m still learning but also enjoying the process.”* This interplay between enjoyment and goal-oriented behavior reflects a shift from incidental exposure to intentional, skill-focused learning. The integration of AI tools not only made listening more accessible but also more rewarding, as learners perceived tangible progress and increased their confidence. In essence, the learning journey evolved from “scrolling through” content to a more deliberate process of skill cultivation—encapsulated in the phrase, “From Scroll to Skill.”

The findings provide a compelling answer to the research inquiry concerning the factors that motivate learners to sustain their extended listening practices in AI-mediated YouTube environments. Intrinsic motivation, supported by personalized content delivery and

learner autonomy, emerged as the dominant force. This aligns seamlessly with the core tenets of Self-Determination Theory (SDT), particularly the psychological needs for autonomy, competence, and relatedness (Jang et al., 2010). Students' ability to select content based on personal interests fulfilled the need for autonomy, while AI features offering real-time feedback or comprehension aids reinforced feelings of competence. The learners' engagement, thus, was not merely habitual but intentional, structured around internal satisfaction and the perception of growth. What sets the current study apart from earlier literature is the active role of AI not just in content delivery, but in sustaining learners' motivation. While previous studies (Karataş et al., 2015; Ishak et al., 2020) emphasized autonomy and competence as classroom constructs facilitated by teacher support, this study demonstrates how AI algorithms independently simulate that support structure. AI-generated video suggestions and adaptive feedback mechanisms function as motivational scaffolds, simulating the autonomy-supportive environment essential for intrinsic motivation. These findings

resonate with recent arguments that adaptive technologies—when used appropriately—can enhance learner agency and foster sustained engagement (Xie et al., 2023).

Competence, another critical pillar of SDT, was supported through AI's ability to provide digestible, tailored feedback. Platforms offering subtitles, playback control, and comprehension tools helped learners feel more in control of the input, thus reinforcing their perceived efficacy. This sense of mastery is closely linked to deeper motivational engagement, as evidenced in prior research (Pelikan et al., 2021; Frikha et al., 2022). The learners in this study reported a confidence boost when they noticed progress, which in turn encouraged them to persist with challenging content—a feedback loop that fosters deeper, more meaningful learning. A particularly novel aspect of this study lies in its illumination of how AI bridges the gap between autonomous content consumption and intentional language learning. While many digital learning studies warn about the risks of passive scrolling (Lee & Wong, 2022; Zheng et al., 2020), the current findings suggest that when integrated thoughtfully, AI

can reorient learners from disengaged browsing to focused skill development. The phrase “*From Scroll to Skill*” encapsulates this transformation, describing how learners’ interaction with AI-supported content evolves into a goal-driven, reflective, and self-motivated practice.

It is important to recognize, however, that these motivational effects may not be uniform across all learner profiles. The participants in this study had already cultivated an interest in improving their listening skills and were relatively advanced in their English studies. Therefore, while AI may facilitate intrinsic motivation for such learners, others may require more structured interventions to experience similar motivational gains. Nonetheless, this does not undermine the broader implication: that AI, when aligned with learner interests and embedded in authentic content, can meaningfully operationalize the psychological components of SDT in digital self-directed environments.

The Impact of AI-Assisted YouTube Listening on Skill Development

The final theme emerging from the participants’ narratives emphasizes the tangible

improvements they experienced in their listening-related competencies. Students consistently reported that after sustained engagement with AI-supported YouTube content, their listening skills improved across four key domains: comprehension, pronunciation, confidence, and academic learning behavior. They no longer regarded YouTube as merely a source of entertainment, but as a structured environment for language development—one made more accessible and adaptable through AI integration.

In terms of comprehension, students described quicker recognition of speech patterns and a better grasp of spontaneous conversation, especially when dealing with unfamiliar topics or accents. Pronunciation gains were attributed to repeated imitation of native speakers—practices like shadowing and pausing videos to mimic intonation patterns. Notably, participants also reported higher levels of confidence and adaptability in authentic listening situations, suggesting that repeated exposure helped them internalize both linguistic forms and communicative strategies. Additionally, the shift toward more

deliberate learning was reflected in self-regulated academic behaviors—students began summarizing content, reflecting on their learning, and applying strategies across contexts. One participant (P4) summarized this shift: *“Before, I just watched for fun. Now I use subtitles, AI summaries, and note-taking. It feels more like learning.”* This perception highlights a transition from passive consumption to active skill development, supported by AI tools that allowed learners to replay content, adjust speed, and receive immediate assistance.

The finding affirms that regular engagement with AI-enhanced YouTube content can support substantial improvements in listening-related skills, particularly when learners actively integrate strategies such as repetition, summarization, and reflection. In comparison to prior studies that explored the effectiveness of podcasts or classroom listening tasks, this study situates itself uniquely within a learner-driven, technology-enhanced framework that aligns authentic content with structured cognitive support.

From a comprehension perspective, the results echo findings by Shaojie et al. (2022), who

demonstrated that multimodal input significantly boosts comprehension and retention. YouTube videos, rich with visual and auditory cues, offer precisely this multimodal scaffolding. The added layer of AI tools, such as real-time subtitles and summarizers, functions as an extension of the scaffolding process, enabling learners to digest and analyze language input more effectively. This dual structure—authentic content and AI-driven support—stands in contrast to the often linear and rigid delivery of textbook-based audio tasks (Aziz et al., 2025). Pronunciation and intonation improvements reported by participants further validate the efficacy of shadowing as a technique, corroborating prior research by Hamada (2015) and Mishima & Cheng (2017). However, what this study contributes beyond existing literature is the observation that learners adopted shadowing voluntarily and strategically—a behavior fostered not through formal instruction but through the interplay of interest, motivation, and technological facilitation. This self-initiated prosodic practice, coupled with AI feedback, bridges the divide between informal learning and skill acquisition. Moreover, the growth

in learners' confidence and adaptability must be understood through the lens of emotional and psychological engagement. Echoing Muñoz (2022), who emphasized that repeated exposure to authentic audiovisual content fosters psychological readiness, this study shows that learners not only feel more prepared but also act on this readiness by experimenting with varied content and learning strategies. The freedom to explore materials tailored to individual preferences—amplified by AI recommendation systems—creates a feedback loop of competence and confidence.

In terms of self-regulation, this study resonates with the conclusions of Razaghi et al. (2019) and Read et al. (2021), who highlighted the role of cognitive scaffolding in managing complex language tasks. The learners in this study employed tools such as ChatGPT and note-taking apps not merely for passive summarization, but as part of a metacognitive routine: to monitor comprehension, evaluate progress, and adjust learning tactics. This layer of strategic engagement distinguishes the current study from earlier works that focused solely on language exposure. In addition, these

outcomes shed light on the broader pedagogical implications. The integration of AI into listening practice illustrates a paradigm where language acquisition is not teacher-dependent but systemically supported—where learners draw on digital scaffolds to construct their own pathways of growth. This is a significant shift from traditional classroom practice and underscores the urgency of integrating autonomous learning tools into contemporary language instruction.

D. Conclusion

This study concludes that integrating YouTube with artificial intelligence (AI) tools represents a transformative approach to enhancing extensive listening skills among EFL learners. The findings reveal that students not only improved their listening comprehension but also developed stronger pronunciation, prosodic awareness, and metacognitive regulation. AI-supported features—such as automatic subtitles, summarization, playback control, and interactive feedback—functioned as effective scaffolds, enabling learners to process authentic linguistic input more efficiently and independently.

Furthermore, learners' intrinsic motivation and autonomy were strengthened, as they actively curated content aligned with their interests, reflecting the principles of Self-Determination Theory. Through AI-mediated personalization, learners shifted from passive viewing to intentional, reflective engagement—transforming entertainment-based digital practices into structured learning opportunities.

The implications of these findings are twofold. Pedagogically, this study underscores the potential of AI to extend the boundaries of traditional listening instruction by promoting learner autonomy, sustained motivation, and contextualized comprehension. Practically, it offers a framework for integrating digital media into language learning environments, demonstrating how YouTube—when combined with AI tools—can serve as a low-cost yet powerful platform for autonomous skill development. The results advocate for educators to incorporate AI-driven scaffolding and reflective tasks into extensive listening programs to maximize learners' engagement and comprehension outcomes.

Despite its contributions, the study acknowledges certain limitations. The small sample size and specific context limit generalizability, and the absence of longitudinal data constrains the ability to assess long-term effects of AI-assisted listening. Future research should therefore expand to diverse learner populations, incorporate quantitative measures of proficiency gains, and explore the integration of teacher-mediated support alongside AI tools. Additionally, examining the balance between learner autonomy and guided instruction could yield deeper insights into optimizing digital learning frameworks.

DAFTAR PUSTAKA

- Afrizal, A., & Herlina, R. (2023). Students' Experiences on Learning Listening Through Youtube Media. *Journal of English Education Program (Jeep)*, 10(1), 25. [https://doi.org/10.25157/\(jeep\).v10i1.9682](https://doi.org/10.25157/(jeep).v10i1.9682)
- Andringa, S., Olsthoorn, N., Beuningen, C. v., Schoonen, R., & Hulstijn, J. H. (2012). Determinants of Success in Native and Non-Native Listening Comprehension: An Individual Differences Approach. *Language Learning*, 62(s2), 49–78. <https://doi.org/10.1111/j.1467-9922.2012.00706.x>

- Ate, C. P., Anuno, A. D. C., & Muliani, M. (2025). Self-Directed Learning With Youtube: Enhancing Listening Skills on Students of English Education Study Program. *J. Innov. Technol. Learn. Educ.*, 2(2), 58–68.
<https://doi.org/10.37792/jitle.v2i2.1525>
- Aziz, M., Tauchid, A., Zainuddin, Moh., & Fatoni, M. (2025). Exploring University EFL Learners' Experiences in Using Podcasts for Listening and Vocabulary Development. *Intens*, 2(02), 1–14.
<https://doi.org/10.32665/intens.v2i02.5145>
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed). Pearson.
- Cross, J. (2009). Effects of Listening Strategy Instruction on News Videotext Comprehension. *Language Teaching Research*, 13(2), 151–176.
<https://doi.org/10.1177/1362168809103446>
- Đurić, I., & Gandhi, S. M. G. (2024). Enhancing English Listening Skills in Indian Engineering Students Through Comprehensible Input: An Experimental Study. *Kuey*.
<https://doi.org/10.53555/kuey.v30i4.8551>
- Fadillah, A. S. N., Nasrullah, N., & Rosalina, E. (2023). The Use of Watching Youtube Videos for Acquiring Students' Listening Comprehension. *International Journal of Educational Best Practices*, 7(1), 92.
<https://doi.org/10.31258/ijebp.v7n1.p92-104>
- Fradana, A. N., Suwarta, N., & Sotlikova, R. (2025). Strengthening Multimodal Literacy Through Digital Literary Text Learning: Enhancing Students' Achievement in the Age of Disruption. *Journal of Language and Literature Studies*, 5(1), 216–225.
<https://doi.org/10.36312/jolls.v5i1.2635>
- Frikha, M., Chaâri, N., Mezghanni, N., Alhumaid, M. M., & Alibrahim, M. S. (2022). Students' Perceptions in Online Physical Education Learning: Comparison Study of Autonomy, Competence, and Relatedness in Saudi Students During the COVID-19 Lockdown. *International Journal of Environmental Research and Public Health*, 19(22), 15288.
<https://doi.org/10.3390/ijerph192215288>
- Goh, C. C. M. (2008). Metacognitive Instruction for Second Language Listening Development. *Relc Journal*, 39(2), 188–213.
<https://doi.org/10.1177/0033688208092184>
- Gönülal, T. (2022a). Improving Listening Skills With Extensive Listening Using Podcasts and Vodcasts. *International Journal of Contemporary Educational Research*, 7(1), 311–320.
<https://doi.org/10.33200/ijcer.685196>
- Gönülal, T. (2022b). Improving Listening Skills With Extensive Listening Using Podcasts and Vodcasts. *International Journal of Contemporary Educational Research*, 7(1), 311–320.

- <https://doi.org/10.33200/ijcer.685196>
- Hamada, Y. (2015). Shadowing: Who Benefits and How? Uncovering a Booming EFL Teaching Technique for Listening Comprehension. *Language Teaching Research*, 20(1), 35–52.
<https://doi.org/10.1177/1362168815597504>
- Hosogoshi, K. (2016). Effects of Captions and Subtitles on the Listening Process: Insights From EFL Learners' Listening Strategies. *The Jalt Call Journal*, 12(3), 153–178.
<https://doi.org/10.29140/jaltcall.v12n3.j206>
- Huang, B. H., Chang, Y. S., Zhi, M., & Niu, L. (2018). The Effect of Input on Bilingual Adolescents' Long-Term Language Outcomes in a Foreign Language Instruction Context. *International Journal of Bilingualism*, 24(1), 8–25.
<https://doi.org/10.1177/1367006918768311>
- Ishak, T., Kurniawan, R., & Zainuddin, Z. (2020). The Role of Pre-Class Asynchronous Online Video Lectures in Flipped-Class Instruction: Identifying Students Perceived Need Satisfaction. *Journal of Pedagogical Research*, 4(1), 1–11.
<https://doi.org/10.33902/jpr.v4i1.145>
- Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging Students in Learning Activities: It Is Not Autonomy Support or Structure but Autonomy Support and Structure. *Journal of Educational Psychology*, 102(3), 588–600.
<https://doi.org/10.1037/a0019682>
- Karataş, H., Alcı, B., Yurtseven, N., & Yüksel, H. G. (2015). Prediction of ELT Students' Academic (Language) Achievement: Language Learning Orientation and Autonomous Learning. *International Online Journal of Educational Sciences*.
<https://doi.org/10.15345/iojes.2015.01.014>
- Khotibi, A. J., & Hadi, M. S. (2022). Students' Problem of Using Youtube Video as English Learning Media During Covid-19 Pandemic. *Jurnal Studi Guru Dan Pembelajaran*, 5(3), 221–225.
<https://doi.org/10.30605/jsgp.5.3.2022.1501>
- Khotimah, K., Rusijono, & Mariono, A. (2024). Enhancing Metacognitive and Creativity Skills Through AI-Driven Meta-Learning Strategies. *International Journal of Interactive Mobile Technologies (Ijim)*, 18(05), 18–31.
<https://doi.org/10.3991/ijim.v18i05.47705>
- Lee, I. C. J., & Wong, P. (2022). Are There Differences Between Online and in-Person Team-Based Learning Classes? Through the Lens of Self-Determination Theory and Learners' Characteristics. <https://doi.org/10.21203/rs.3.rs-1298958/v1>
- Lestari, P., Rohliah, L., Ningsih, W., & Hutagalung, G. U. (2023). Youtube: A Teaching Media to Optimize Students' English Listening Comprehension Skill. *Global Expert Jurnal Bahasa Dan Sastra*, 11(1), 17–24.
<https://doi.org/10.36982/jge.v11i1.3210>

- Li, P. (2022). A Study on Using Noticing Hypothesis to Improve Learners' Language Input in Language Learning. *Journal of Contemporary Educational Research*, 6(1), 68–71. <https://doi.org/10.26689/jcer.v6i1.2896>
- Lin, I., Yang, C., Lin, E. C. L., Chung, F., & Yu, C. (2024). The Determinants of Mental Illness Self-Management for Patients With Schizophrenia and Type 2 Diabetes: A Self-Determination Theory Approach. *Journal of Clinical Nursing*, 34(8), 3206–3213. <https://doi.org/10.1111/jocn.17502>
- Liu, X. (2025). Reconceptualizing Foreign Language Learning Through Artificial Intelligence Within the Framework of the Zone of Proximal Development. *GBP Proc. Ser.*, 8, 9–15. <https://doi.org/10.70088/790gnb90>
- Mahdavy, B., & Namavar, M. M. (2023). Listening Strategies and L2 Listening Comprehension: Does the Test Method Matter? *Theory and Practice of Second Language Acquisition*, 9(1), 1–22. <https://doi.org/10.31261/tapsla.12097>
- Ma'rufah, A., Samanhudi, U., & Ardiasih, L. S. (2024). Project-Based Learning Using Podcast-Assisted Media in the Teaching of English Speaking Skill. *Lectura Jurnal Pendidikan*, 15(2), 521–532. <https://doi.org/10.31849/lectura.v15i2.20706>
- Mishima, M., & Cheng, L. (2017). The Impact of a Computer-Mediated Shadowing Activity on ESL Speaking Skill Development: A Pilot Study. *L2 Journal*, 9(1). <https://doi.org/10.5070/l29132493>
- Mokhtari, K. (2025). *Thinking With Machines: Leveraging Artificial Intelligence (AI) to Foster Metacognitive Reading Comprehension*. <https://doi.org/10.5772/intechopen.1012695>
- Muñoz, C. (2022). Audiovisual Input in L2 Learning. *Language Interaction and Acquisition*, 13(1), 125–143. <https://doi.org/10.1075/lia.22001.mun>
- Nguyen, Q. N. (2025). Beyond Comprehensible Input: A Neuro-Ecological Critique of Krashen's Hypothesis in Language Education. *Frontiers in Psychology*, 16. <https://doi.org/10.3389/fpsyg.2025.1636777>
- Nowbakht, M., & Shahnazari, M. (2015). The Comparative Effects of Comprehensible Input, Output and Corrective Feedback on the Receptive Acquisition of L2 Vocabulary Items. *Advances in Language and Literary Studies*, 6(4). <https://doi.org/10.7575/aiac.all.v.6n.4p.103>
- Ntoumanis, N. (2005). A Prospective Study of Participation in Optional School Physical Education Using a Self-Determination Theory Framework. *Journal of Educational Psychology*, 97(3), 444–453. <https://doi.org/10.1037/0022-0663.97.3.444>
- Pelikan, E., Korlat, S., Reiter, J., Holzer, J., Mayerhofer, M.,

- Schober, B., Spiel, C., Hamzallari, O., Uka, A., Chen, J., Välimäki, M., Puharić, Z., Anusionwu, K. E., Okocha, A. N., Zabrodskaia, A., Salmela-Aro, K., Käser, U., Schultze-Krumbholz, A., Wachs, S., ... Lüftenegger, M. (2021). Distance Learning in Higher Education During COVID-19: The Role of Basic Psychological Needs and Intrinsic Motivation for Persistence and Procrastination—a Multi-Country Study. *Plos One*, *16*(10), e0257346.
<https://doi.org/10.1371/journal.pone.0257346>
- Rahila, C. D. I., Batubara, M. H., & Marlina, M. (2023). The Influence of YouTube on Listening Skills. *Journal of Linguistics Literature and Language Teaching (Jlllt)*, *3*(1), 1–7.
<https://doi.org/10.37249/jlllt.v3i1.507>
- Ramadhani, G. P., & Samsi, Y. S. (2023). Efl Junior High School Students' Perception Towards Using Youtube in Learning Process of English Listening Skills. *Edusaintek Jurnal Pendidikan Sains Dan Teknologi*, *11*(1), 351–366.
<https://doi.org/10.47668/edusaintek.v11i1.970>
- Razaghi, M., Bagheri, M. S., & Yamini, M. (2019). The Impact of Cognitive Scaffolding on Iranian EFL Learners' Speaking Skill. *International Journal of Instruction*, *12*(4), 95–112.
<https://doi.org/10.29333/iji.2019.1247a>
- Read, T., Kukulska-Hulme, A., & Bárcena, E. (2021). Supporting Listening Comprehension by Social Network-Based Interaction in Mobile Assisted Language Learning. *Porta Linguarum Revista Interuniversitaria De Didáctica De Las Lenguas Extranjeras*, *35*, 295–309.
<https://doi.org/10.30827/portalin.v0i35.15341>
- Rong, Z., & Xue, W. (2018). The Application of Metacognitive Strategies to EFL Listening Instruction—A Study on Metacognitive Strategy-Centered Instructional Model in Listening Classroom. *Destech Transactions on Social Science Education and Human Science*, *icessh*.
<https://doi.org/10.12783/dtssehs/icessh2018/23795>
- Saputra, Y., & Fatimah, A. S. (2018). The Use of TED and YOUTUBE in Extensive Listening Course: Exploring Possibilities of Autonomy Learning. *Indonesian Jelt Indonesian Journal of English Language Teaching*, *13*(1), 73–84.
<https://doi.org/10.25170/ijelt.v13i1.1451>
- Shafwati, D., Sholihah, L., Prakoso, G. H., & Riyantika, F. (2021). *The Use of YouTube Video Toward Students' Listening Ability*.
<https://doi.org/10.4108/eai.16-10-2020.2305198>
- Shaojie, T., Samad, A. A., & Ismail, L. (2022). Systematic Literature Review on Audio-Visual Multimodal Input in Listening Comprehension. *Frontiers in Psychology*, *13*.
<https://doi.org/10.3389/fpsyg.2022.980133>
- Suryana, I., Asrianto, A., & Murwantono, D. (2020). Artificial

- Intelligence to Master English Listening Skills for Non-English Major Students. *Journal of Languages and Language Teaching*, 8(1), 48. <https://doi.org/10.33394/jollt.v8i1.2221>
- Suvorova, T., & Demirbilek, M. (2025). *Beyond Adaptive Learning*. 1–24. <https://doi.org/10.4018/979-8-3693-8332-2.ch001>
- Ulfert-Blank, A.-S., & Schmidt, I. (2022). Assessing digital self-efficacy: Review and scale development. *Computers & Education*, 191, 104626. <https://doi.org/10.1016/j.compedu.2022.104626>
- Vandergrift, L. (2005). Relationships Among Motivation Orientations, Metacognitive Awareness and Proficiency in L2 Listening. *Applied Linguistics*, 26(1), 70–89. <https://doi.org/10.1093/applin/ah039>
- Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 Learners How to Listen Does Make a Difference: An Empirical Study. *Language Learning*, 60(2), 470–497. <https://doi.org/10.1111/j.1467-9922.2009.00559.x>
- Veas, A., Costa, J. L. C., Miñano, P., & Corbí, R. G. (2018). Relationship Between Parent Involvement and Academic Achievement Through Metacognitive Strategies: A Multiple Multilevel Mediation Analysis. *British Journal of Educational Psychology*, 89(2), 393–411. <https://doi.org/10.1111/bjep.12245>
- Vellanki, S. S., Mond, S., Khan, Z. K., & Nair, L. G. (2022). Teachers' Viewpoint of Metacognitive Strategy Instruction in Listening During Remote Teaching in Oman: Challenges and Strategies. *International Journal of Learning Teaching and Educational Research*, 21(7), 82–106. <https://doi.org/10.26803/ijlter.21.7.5>
- Walker, P. B., Haase, J., Mehalick, M. L., Steele, C. T., Russell, D. W., & Davidson, I. N. (2025). Harnessing Metacognition for Safe and Responsible AI. *Technologies*, 13(3), 107. <https://doi.org/10.3390/technologies13030107>
- Wang, Y., & Treffers-Daller, J. (2017). Explaining Listening Comprehension Among L2 Learners of English: The Contribution of General Language Proficiency, Vocabulary Knowledge and Metacognitive Awareness. *System*, 65, 139–150. <https://doi.org/10.1016/j.system.2016.12.013>
- Widiyanto, S., Mulyadi, M., Wulansari, L., & Ati, A. P. (2021). Youtube as a Media for Students' Comprehensive Listening Ability at Junior High Schools in Bekasi. *Jurnal Pajar (Pendidikan Dan Pengajaran)*, 5(6), 1699. <https://doi.org/10.33578/pjr.v5i6.8504>
- Xie, Y., Zhou, R., Chan, A. H. S., Jin, M., & Qu, M. (2023). Motivation to Interaction Media: The Impact of Automation Trust and Self-Determination Theory on Intention to Use the New Interaction Technology in Autonomous Vehicles. *Frontiers*
-

in *Psychology*, 14.
<https://doi.org/10.3389/fpsyg.2023.1078438>

Yan, B. (2023). Exploration of College English Listening Teaching Model in the Context of New Curriculum Reform. *International Journal of New Developments in Education*, 5(14).
<https://doi.org/10.25236/ijnde.2023.051414>

Yang, F.-Y. (2020). EFL Learners' Autonomous Listening Practice Outside of the Class. *Studies in Self-Access Learning Journal*, 12(4), 328–346.
<https://doi.org/10.37237/110403>

Yifan, P., Hashim, H., & Said, N. E. M. (2025). Cognitive Amplification: Harnessing Artificial Intelligence to Augment Metacognitive Learning Strategies. *Dirasat Human and Social Sciences*, 53(1), 7967.
<https://doi.org/10.35516/hum.2025.7967>

Zheng, F., Ping-ying, H., Lian, Z., Wang, Y., Wu, S., & Li, H. (2020). Contributing Factors to the Improvement of International Students' Health Literacy in China: A Self-Determination Theory Perspective. *Frontiers in Public Health*, 8.
<https://doi.org/10.3389/fpubh.2020.00390>