



ElevenLabs-Based Critical Listening Instruction: Exploring Students' Social-Emotional Competencies in Indonesian Language Education

Eka Sartika¹, Meilina Widya Dali²

^{1,2}Pendidikan Bahasa dan Sastra Indonesia, Universitas Negeri Gorontalo

Article Info	Abstract
Received: 2026-09-18 Revised: 2026-09-23 Accepted: 2026-09-24 Keywords: <i>critical listening;</i> <i>ElevenLabs;</i> <i>artificial intelligence;</i> <i>social-emotional competence;</i> <i>language learning</i> DOI: 10.24256/ideas.v14i1.12562 Corresponding Author: Eka Sartika eka@ung.ac.id Pendidikan Bahasa dan Sastra Indonesia, Universitas Negeri Gorontalo	<i>The integration of artificial intelligence in language learning provides new opportunities for developing more contextual and meaningful listening experiences. This study explores the utilization of ElevenLabs in critical listening instruction and students' social-emotional experiences during ElevenLabs-based listening activities in Indonesian Language and Literature Education. This study employed a descriptive qualitative approach, supported by descriptive numerical data to enrich the interpretation of learning responses and critical listening performance. Data were collected through classroom observations, critical listening ability analysis, in-depth interviews, and documentation. The participants consisted of 25 students involved in classroom observation and critical listening assessment, while eight students participated in interviews related to their social-emotional experiences. Data were analyzed thematically through data reduction, data display, and conclusion drawing. The findings indicate that students responded positively to ElevenLabs-based listening activities, particularly regarding voice clarity, listening focus, and emotional engagement. Descriptive assessment data showed that most students were categorized as highly developed in critical listening performance following the learning activity, with the highest achievement observed in the understanding and reflecting dimensions, while interpreting remained an area requiring further support. The interview findings revealed that social awareness was identified in the largest number of informants, reflected in students' ability to understand situations and consider others' perspectives. These findings suggest that ElevenLabs can function as an audio stimulus within critical listening activities and provide opportunities for students' social-emotional reflection. However, its implementation requires integration with pedagogical strategies, such as discussion, reflective questioning, and argumentative activities, to facilitate deeper meaning-making processes.</i>

1. Introduction

Listening is one of the fundamental skills in language learning because it involves more than receiving auditory information. Listening requires learners to actively construct meaning through processes of understanding, interpretation, evaluation, and reflection. In higher education contexts, particularly among students of Indonesian Language and Literature Education (PBSI), listening competence needs to be developed beyond basic comprehension toward critical listening. Critical listening refers to an active cognitive process in which listeners analyze information, evaluate messages, recognize implicit meanings, and construct interpretations based on the information received (Rost, 2016). Therefore, critical listening plays an important role in developing students' higher-order thinking abilities.

The development of critical listening requires learners to engage in metacognitive processes during listening activities. Vandergrift and Goh (2012) explain that effective listening involves learners' awareness of their own comprehension processes through planning, monitoring, and evaluating their understanding. In this perspective, students are not positioned as passive recipients of information but as active participants who regulate and construct meaning from spoken messages. Thus, listening instruction should provide learning experiences that encourage students to analyze, interpret, evaluate, and reflect on the information they receive.

Beyond cognitive processing, critical listening activities also involve interpersonal and emotional dimensions. When students listen to narratives, experiences, or perspectives of others, they encounter different emotions, values, and viewpoints. These processes create opportunities for students to connect cognitive understanding with social-emotional responses. According to the Collaborative for Academic, Social, and Emotional Learning (CASEL, 2020), social-emotional learning consists of five core competencies: *self-awareness*, *self-management*, *social awareness*, *relationship skills*, and *responsible decision-making*. These competencies can be connected to critical listening processes. Interpretation of messages may encourage perspective-taking and contribute to social awareness, while reflection on messages may support recognition and understanding of emotions related to self-awareness. Furthermore, evaluation activities require learners to consider consequences and viewpoints, which are related to responsible decision-making. Collaborative discussions following listening activities may also provide opportunities for communication and relationship skills development.

However, listening instruction often remains focused on comprehension outcomes rather than deeper engagement with meaning and reflection. Many learning activities emphasize identifying explicit information, while opportunities for interpretation, evaluation, and personal reflection are still limited. This condition indicates the need for innovative instructional approaches that transform listening from a receptive activity into a meaningful learning process involving cognitive, reflective, and social-emotional engagement.

The development of digital technology provides new opportunities for designing more contextual and engaging listening experiences. Digital audio and audiovisual media enable learning materials to incorporate various auditory features, including voice quality, intonation, stress, pauses, and emotional expressions. These elements can influence how learners perceive and construct meaning from spoken messages. Mayer (2021) explains that multimedia learning can support meaningful learning processes when instructional messages are designed according to how learners process verbal and auditory information.

The emergence of artificial intelligence (AI) has further expanded possibilities for developing digital learning resources. AI technologies can support educational innovation by assisting teachers in designing learning materials, providing personalized experiences, and creating new forms of interaction between technology and learners (Holmes et al., 2019). UNESCO (2023) emphasizes that the use of generative AI in education should be considered from pedagogical perspectives, ensuring that technology contributes to meaningful, ethical, and learner-centered educational practices.

One AI-based technology that has potential for listening instruction is ElevenLabs. This technology enables the generation of synthetic voices with human-like characteristics, including variations in tone, expression, and speech patterns. In critical listening instruction, these characteristics may function as auditory stimuli within learning activities because students are required not only to understand spoken information but also to interpret speakers' intentions, evaluate messages, and reflect on meanings. Therefore, ElevenLabs is positioned not as a replacement for pedagogical interaction but as an instructional resource that may support richer listening experiences.

Previous studies have demonstrated the potential of digital media and technology in supporting listening instruction. Sartika et al. (2026) examined the integration of audiovisual folktales with Blooket-based gamification in listening instruction and found that students showed positive engagement and perceived the learning experience as more interactive. The study indicates that digital integration can create more participatory listening environments. Similarly, Izzati et al. (2025) found that storytelling approaches provide opportunities for learners to engage more deeply with narratives and messages, while Ubaidillah et al. (2025) emphasized that listening ability serves as an important foundation for developing broader language competencies.

Although previous studies have shown the potential of digital media in listening instruction, several research gaps remain. First, existing studies have generally focused on audiovisual media, storytelling, or digital platforms, while the use of AI-generated audio technology for critical listening instruction remains limited. Second, previous research has primarily emphasized cognitive outcomes such as listening comprehension and engagement, whereas the emergence of social-emotional responses during technology-supported listening activities has received less attention. Third, empirical studies examining the use of ElevenLabs as an AI-based audio stimulus in university-level Indonesian language learning are still limited.

Based on these gaps, this study positions ElevenLabs not merely as a tool for producing audio materials but as an instructional resource within critical listening pedagogy. The novelty of this study lies in examining how AI-generated audio is implemented in critical listening activities and how students respond cognitively and socially-emotionally during the learning process. The study integrates three important dimensions: language learning, critical thinking, and social-emotional learning.

Therefore, this study aims to describe the practice of utilizing ElevenLabs in critical listening instruction among Indonesian Language and Literature Education students, identify students' engagement during the learning process, and explore students' social-emotional experiences emerging during ElevenLabs-based critical listening activities.

2. Method

2.1 Research Design and Context

This study employed a descriptive qualitative approach to explore the implementation of ElevenLabs-based critical listening instruction and students' social-emotional experiences during the learning process in Indonesian Language and Literature Education (PBSI). A qualitative approach was selected because it enables researchers to

examine learning phenomena in natural settings by exploring participants' experiences, interactions, responses, and meanings constructed during instructional activities (Creswell & Creswell, 2023).

The study was conducted in the Indonesian Language and Literature Education (PBSI) program at the university level. The participants were 25 third-semester students from the 2025 cohort who participated in a critical listening learning activity. The learning activity was conducted in one classroom session with a duration of 3×50 minutes (150 minutes). During the session, students participated in critical listening activities using AI-generated audio materials produced through ElevenLabs. The audio materials functioned as instructional stimuli to support students in understanding, interpreting, analyzing, evaluating, and reflecting on the messages conveyed.

Although descriptive numerical data were obtained from observation and critical listening assessment, these data were used only as supporting information to describe students' responses and performance during the learning activity. The main analysis remained qualitative, focusing on students' experiences, engagement, and social-emotional responses throughout the learning process.

2.2 Data Collection

Data were collected through classroom observations, critical listening ability analysis, in-depth interviews, and documentation. Before data collection, the researchers prepared observation guidelines, interview protocols, critical listening assessment instruments, and learning materials in the form of AI-generated audio produced using ElevenLabs.

The data collection process consisted of four stages: (1) implementation of ElevenLabs-based critical listening activities, (2) classroom observation, (3) critical listening assessment, and (4) interviews regarding students' learning experiences and social-emotional responses.

2.3 ElevenLabs Audio Materials and Learning Procedure

The learning materials consisted of audio narratives generated using ElevenLabs artificial intelligence voice technology. The audio was designed as a listening stimulus to provide students with contextual spoken messages containing vocal elements such as intonation, emphasis, pauses, and emotional expressions.

The learning activity was conducted through three stages. First, in the **pre-listening stage**, students were introduced to the topic of the audio and guided to predict the possible content and purpose of the message. Second, in the **while-listening stage**, students listened to the ElevenLabs-generated audio and identified important information, interpreted meanings, analyzed messages, and responded to questions related to the content. Third, in the **post-listening stage**, students engaged in discussion, reflection, and argumentation activities to express interpretations, evaluate messages, and connect the listening material with broader perspectives. The use of ElevenLabs in this study was positioned as an instructional resource that provided auditory stimuli, while the development of critical listening and social-emotional responses was explored through the overall learning process, including interaction, reflection, and discussion activities.

2.4 Classroom Observation and Critical Listening Assessment

Classroom observations were conducted during the implementation of critical listening instruction. The observation focused on eight aspects: ElevenLabs voice clarity, students' responses to intonation, responses to word stress, the use of pauses, emotional

responses, listening focus, critical thinking activities, and classroom interaction. The observation instrument used a four-point rating scale ranging from 1 to 4. The scale was used descriptively to support interpretation of students' responses during learning activities. The score categories were:

Tabel 2.1 The score categories

Score Range	Category
3.26–4.00	Highly developed
2.51–3.25	Developed
1.76–2.50	Developing
1.00–1.75	Beginning

The observation results were not interpreted as evidence of causal effects of ElevenLabs but were used to describe students' responses during the implementation of AI-based listening activities. In addition to observation, students' critical listening abilities were analyzed based on five dimensions: *understanding*, *interpreting*, *analyzing*, *evaluating*, and *reflecting*. These dimensions were used to examine how students comprehended information, interpreted meanings, analyzed messages, evaluated information, and reflected on the messages received.

The critical listening assessment consisted of tasks related to the audio content, including identifying explicit information (*understanding*), explaining implied meanings (*interpreting*), analyzing relationships among ideas (*analyzing*), evaluating messages or viewpoints (*evaluating*), and providing personal reflections based on the listening experience (*reflecting*). The assessment results were presented descriptively to support qualitative interpretation of students' critical listening performance.

2.5 Interview and Social-Emotional Analysis

To explore students' social-emotional experiences, eight students were selected as interview informants. The selection was conducted purposively to obtain variations in students' responses, participation, and experiences during the learning activity. The informants represented different levels of engagement observed during the learning process. The interviews explored students' emotional experiences, self-reflection, understanding of others' perspectives, social interaction, and decision-making processes after participating in ElevenLabs-based listening activities. The analysis of students' social-emotional experiences was based on five CASEL dimensions: *self-awareness*, *self-management*, *social awareness*, *relationship skills*, and *responsible decision-making* (CASEL, 2020).

2.6 Data Analysis

The research data were analyzed thematically through the processes of data reduction, data display, and conclusion drawing and verification (Miles et al., 2020). Observation notes, assessment results, interview transcripts, and documentation were organized and categorized according to the research focus.

Interview data were transcribed, coded, and grouped into emerging themes related to students' learning experiences and social-emotional responses. Meanwhile, descriptive numerical data from observation and critical listening assessments were used to complement qualitative findings by illustrating patterns of student engagement and performance. The final interpretation focused on how ElevenLabs-based audio functioned as an instructional stimulus within critical listening activities and how students demonstrated cognitive engagement and social-emotional responses during the learning process.

3. Result

This section presents the findings regarding the implementation of ElevenLabs-based critical listening instruction, students' critical listening performance, and students' social-emotional experiences during the learning activities. The findings were derived from classroom observations involving 25 students, descriptive analysis of critical listening performance, and in-depth interviews with eight students. The numerical data obtained from observations and critical listening assessments were used as supporting descriptive information to complement the qualitative interpretation of students' responses and learning experiences.

3.1 Students' Responses during ElevenLabs-Based Listening Activities

Classroom observations were conducted to explore students' responses during critical listening activities using ElevenLabs-generated audio materials. The observation focused on eight aspects: voice clarity, response to intonation, response to word stress, use of pauses, emotional response, listening focus, critical thinking activity, and classroom interaction. The observation results were summarized using descriptive scores on a four-point scale. These scores were used to describe patterns of students' responses during the learning activity rather than to measure the effect of ElevenLabs on students' learning outcomes.

Table 3.1 Students' Responses during ElevenLabs-Based Listening Activities

Observed Aspect	Mean Score	Category
Voice clarity of ElevenLabs	3.68	Highly Developed
Response to intonation	3.20	Developed
Response to word stress	3.36	Highly Developed
Use of pauses	3.24	Developed
Emotional response	3.60	Highly Developed
Listening focus	3.64	Highly Developed
Critical thinking activity	3.32	Highly Developed
Classroom interaction	3.36	Highly Developed

As shown in Table 1, students demonstrated positive responses across the observed aspects during the ElevenLabs-based listening activity. Voice clarity obtained the highest descriptive score ($M = 3.68$), followed by listening focus ($M = 3.64$) and emotional response ($M = 3.60$). These findings indicate that students perceived the AI-generated audio as sufficiently clear and engaging for supporting listening activities.

The relatively lower scores appeared in response to intonation ($M = 3.20$) and use of pauses ($M = 3.24$). These aspects suggest that although students were able to follow the audio content, interpreting subtle vocal features such as intonation patterns and pauses required further instructional support. Overall, the observation data indicate that ElevenLabs functioned as an auditory stimulus that supported students' engagement during critical listening activities.

3.2 Students' Critical Listening Performance during the Learning Activity

Students' critical listening performance was examined based on five dimensions: *understanding*, *interpreting*, *analyzing*, *evaluating*, and *reflecting*. The assessment results were used as descriptive supporting data to illustrate students' performance after participating in the learning activity. The distribution of students' performance categories is presented in Table 2.

Table 3.2 Distribution of Students' Critical Listening Performance

Category	Number of Students	Percentage
Highly Developed	21	84%
Developed	4	16%
Total	25	100%

Based on Table 2, 21 students (84%) were categorized as highly developed, while four students (16%) were categorized as developed in the critical listening assessment. These categories describe students' performance following the learning activity and are not interpreted as evidence of improvement caused by ElevenLabs. The performance across the five critical listening dimensions is presented in Table 3.

Table 3.3 Students' Critical Listening Performance by Dimension

Dimension	Total Score	Mean	Category
<i>Understanding</i>	94	3.76	Highly Developed
<i>Interpreting</i>	78	3.12	Developed
<i>Analyzing</i>	85	3.40	Highly Developed
<i>Evaluating</i>	85	3.40	Highly Developed
<i>Reflecting</i>	94	3.76	Highly Developed

Table 3 shows that understanding and reflecting obtained the highest descriptive scores ($M = 3.76$). This finding indicates that students were able to identify information from the listening material and connect the messages with personal thoughts or experiences. Meanwhile, analyzing and evaluating reached the same score ($M = 3.40$), showing students' ability to examine information and consider viewpoints presented in the audio.

The *interpreting* dimension obtained the lowest score ($M = 3.12$), indicating that interpreting implicit meanings and deeper messages remained a more challenging aspect. This finding suggests that students may require additional scaffolding activities, such as guided questioning and discussion, to strengthen interpretation skills.

3.3 Students' Social-Emotional Experiences during ElevenLabs-Based Listening Activities

To explore students' social-emotional experiences during the learning process, in-depth interviews were conducted with eight students. Interview data were analyzed based on the five CASEL dimensions: *self-awareness*, *self-management*, *social awareness*, *relationship skills*, and *responsible decision-making* (CASEL, 2020).

Table 3.4 Social-Emotional Themes Identified from Student Interviews

CASEL Dimension	Number of Informants	Main Theme
<i>Self-awareness</i>	6 of 8	Awareness of emotions and personal change
<i>Self-management</i>	6 of 8	Managing academic pressure
<i>Social awareness</i>	7 of 8	Understanding others' perspectives
<i>Relationship skills</i>	5 of 8	Positive communication
<i>Responsible decision-making</i>	6 of 8	Responsible decision-making

The interview findings show that all five CASEL dimensions emerged during students' reflections on the listening activities. Among these dimensions, *social awareness* appeared most frequently, identified by seven of the eight informants. Students described how listening to narratives enabled them to reconsider situations from different perspectives and understand the experiences of others.

One student stated:

"After listening to the narrative, I felt that I understood my own feelings better and realized that everyone has different problems."

This statement reflects the emergence of *self-awareness*, as the student connected the listening experience with personal emotions and reflection.

Another student explained:

"At first, I saw the lecturer's assignments as a burden, but after listening to the story, I understood the lecturer's purpose in giving the assignments."

This response illustrates *social awareness* through the student's ability to reconsider another person's perspective.

Overall, the findings indicate three patterns. First, students demonstrated positive engagement during ElevenLabs-based listening activities, particularly regarding audio clarity, listening focus, and emotional responses. Second, descriptive assessment data showed that students demonstrated stronger performance in understanding and reflecting than in interpreting implicit meanings. Third, interview data revealed the emergence of social-emotional experiences across all CASEL dimensions, with social awareness being the most frequently identified theme.

This section presents the findings concerning the use of ElevenLabs in critical listening instruction, students' critical listening performance, and their social-emotional competencies. The findings were obtained from classroom observations involving 25 students, an analysis of their critical listening performance, and interviews with eight students concerning their social-emotional experiences.

4. Discussion

4.1 ElevenLabs as an Instructional Stimulus for Critical Listening Activities

One of the main findings of this study concerns students' critical listening performance during ElevenLabs-based listening activities. Descriptive assessment data showed that most students were categorized as highly developed, with 84% reaching this category and 16% categorized as developed. Among the five dimensions assessed, *understanding* and *reflecting* obtained the highest mean scores, both reaching 3.76. These findings indicate that students were able to identify important information from the listening material and relate the messages to broader meanings or personal experiences after participating in the learning activity.

These findings support the view that listening is not merely a process of receiving verbal information but involves active meaning construction. Critical listening requires learners to understand information, interpret meanings, analyze messages, evaluate ideas, and reflect on the messages received. In this study, ElevenLabs-generated audio functioned as an instructional stimulus by providing listening materials containing vocal features such

as voice clarity, word stress, intonation, pauses, and emotional expressions. These characteristics became part of the auditory information that students processed when constructing meaning during listening activities.

The findings are consistent with previous studies highlighting the potential of digital audio and audiovisual media in listening instruction. Rahmayani et al. (2025) reported the relevance of audio media in supporting listening activities, while Khoirunnisaa (2025) emphasized the role of audiovisual media in developing listening skills. Similarly, Sartika et al. (2026) demonstrated that the integration of audiovisual folktales and digital gamification created more interactive listening environments. The present study extends previous research by examining AI-generated audio, particularly ElevenLabs, within the context of critical listening rather than focusing only on general listening comprehension.

However, differences were observed among the critical listening dimensions. The *interpreting* dimension obtained the lowest mean score of 3.12 and remained in the developed category. This finding suggests that students were relatively able to understand explicit information and reflect on the messages presented, but interpreting implicit meanings remained more challenging.

Interpretation requires learners to move beyond literal comprehension by connecting information with context, identifying implied meanings, recognizing perspectives, and making inferences. Therefore, the use of AI-generated audio alone may not be sufficient to support all aspects of critical listening. Additional pedagogical activities, including guided questioning, discussion, comparison of perspectives, and argumentative activities, are needed to help students move from understanding information toward deeper interpretation.

4.2 Voice Characteristics, Listening Focus, and Emotional Engagement

The observation findings provide further understanding of students' responses during ElevenLabs-based listening activities. Voice clarity obtained the highest descriptive score (3.68), followed by listening focus (3.64) and emotional response (3.60). These findings suggest that the auditory characteristics of the learning material supported students' engagement during the listening process.

From a multimedia learning perspective, the way information is presented influences how learners process instructional messages. Mayer (2021) explains that meaningful learning occurs when instructional materials are designed according to learners' cognitive processing systems. In this study, the clear and expressive characteristics of ElevenLabs-generated audio provided an auditory representation that students could follow during listening activities. The relatively high listening focus score indicates that students were able to maintain attention while engaging with the listening material.

The emotional response identified in the observation also provides an important perspective. Listening involves not only processing factual information but also interpreting how messages are expressed. Voice quality, intonation, pauses, and word stress may influence how learners perceive emotional meanings within spoken messages. This finding corresponds with Pujawati (2024), who highlights the importance of listening experiences beyond literal comprehension, and with Izzati et al. (2025), who emphasize that narrative-based learning can encourage deeper engagement with messages.

Nevertheless, students showed relatively lower responses to intonation and pauses. These findings indicate that students may require explicit instructional guidance to recognize the communicative function of prosodic features. Intonation and pauses may represent emphasis, emotional states, uncertainty, or changes in meaning; therefore, lecturers need to direct students' attention toward these features through analytical listening tasks.

The findings suggest that the educational value of ElevenLabs does not rely solely on the quality of AI-generated voices. Instead, its contribution depends on how the audio material is integrated into instructional activities. AI-generated audio can provide flexible and expressive listening resources, while lecturers remain responsible for selecting materials, designing learning tasks, facilitating discussion, and guiding interpretation.

4.3 Social-Emotional Experiences Emerging during Critical Listening Activities

Another important finding concerns students' social-emotional experiences during ElevenLabs-based listening activities. All five CASEL dimensions emerged from the interview data: *self-awareness, self-management, social awareness, relationship skills, and responsible decision-making*. Among these dimensions, *social awareness* appeared most frequently, identified in seven of the eight interviews.

The emergence of social awareness indicates that listening activities may provide opportunities for students to consider perspectives beyond their own. When students engage with narratives containing experiences, emotions, and conflicts, they may reflect on different viewpoints and reconsider situations from other perspectives.

This process can be understood through the connection between critical listening and social-emotional competencies. Interpretation activities may involve perspective-taking, which is closely related to social awareness. Reflection activities may encourage students to recognize and evaluate their own emotions, supporting self-awareness. Meanwhile, evaluation activities require students to consider consequences and alternative viewpoints, which relate to responsible decision-making.

One student stated:

"After listening to the narrative, I felt that I understood my own feelings better and realized that everyone has different problems."

This response reflects self-awareness because the student connected the listening experience with personal emotional understanding.

Another student explained:

"At first, I saw the lecturer's assignments as a burden, but after listening to the story, I understood the lecturer's purpose in giving the assignments."

This response illustrates social awareness because the student reconsidered the situation from another person's perspective.

The findings are consistent with Aulia (2024), who emphasizes the connection between listening activities and social-emotional development. However, these findings should not be interpreted as evidence that ElevenLabs directly developed students' social-emotional competencies. Rather, the findings indicate that social-emotional responses emerged during learning activities where ElevenLabs-generated audio was used as a stimulus, together with narrative content, reflective questions, classroom interaction, and discussion.

4.4 Position and Novelty of the Study

Previous studies have demonstrated the usefulness of audio, audiovisual materials, storytelling, and digital platforms in listening instruction (Rahmayani et al., 2025; Khoirunnisaa, 2025; Izzati et al., 2025; Sartika et al., 2026). However, previous research has generally focused on listening comprehension, engagement, or the use of digital media as instructional support.

This study extends previous research in three ways. First, it examines AI-generated voice technology within critical listening activities rather than general listening comprehension. Second, it combines descriptive analysis of listening performance with exploration of students' social-emotional experiences. Third, it positions listening instruction as a learning space where language literacy, critical thinking, digital literacy, and social-emotional reflection interact.

The findings suggest that the educational potential of AI-generated audio lies not merely in producing human-like speech but in how the technology is embedded within meaningful learning activities. When accompanied by analytical questions, reflection, and classroom interaction, AI-generated audio may provide opportunities for students to understand, analyze, evaluate, and reflect on messages while considering emotional and interpersonal perspectives.

4.5 Pedagogical Implications and Limitations

The findings suggest several implications for listening instruction. Lecturers may utilize ElevenLabs to develop varied listening materials with expressive voices, particularly when authentic recordings are difficult to obtain. However, AI-generated audio should be accompanied by carefully designed learning activities that encourage higher levels of listening processing. Activities focusing on implied meaning, speaker intention, evidence, perspective, and emotional cues are particularly important because the *interpreting* dimension showed lower performance compared with other critical listening dimensions.

Furthermore, reflective activities are important when listening instruction aims to involve social-emotional learning. Students can be encouraged to identify emotions, consider alternative perspectives, relate messages to personal experiences, and explain changes in their viewpoints after listening.

Several limitations should be considered. This study involved 25 students in one educational context, while social-emotional experiences were explored through interviews with eight students. In addition, the descriptive qualitative design does not allow conclusions regarding causal effects of ElevenLabs on critical listening or social-emotional competencies. Instead, the findings describe patterns of engagement and responses that emerged during ElevenLabs-based listening activities. Future studies may involve larger and more diverse participants, compare AI-generated audio with human-recorded audio, or employ mixed-method designs to examine changes in critical listening performance more systematically.

5. Conclusion

This study explored the implementation of ElevenLabs-based audio as an instructional stimulus in critical listening activities among Indonesian Language and Literature Education students and examined students' cognitive engagement and social-emotional experiences during the learning process. The findings indicate that students responded positively to the use of ElevenLabs-generated audio, particularly regarding voice clarity, listening focus, and emotional engagement during listening activities.

Descriptive assessment data showed that most students demonstrated highly developed performance in critical listening following the learning activity, with 84% categorized as highly developed and 16% categorized as developed. Among the five critical listening dimensions, *understanding* and *reflecting* obtained the highest mean scores, indicating that students were able to identify information from the listening material and connect the messages with broader meanings or personal experiences. Meanwhile, the *interpreting* dimension showed a relatively lower performance, suggesting that students still required further instructional support to identify implicit meanings and deeper messages.

The interview findings revealed that students' social-emotional experiences emerged across all five CASEL dimensions: *self-awareness*, *self-management*, *social awareness*, *relationship skills*, and *responsible decision-making*. Among these dimensions, *social awareness* appeared most frequently, reflecting students' ability to consider different perspectives and understand experiences beyond their own. These findings suggest that critical listening activities involving AI-generated audio may provide opportunities for cognitive processing and social-emotional reflection.

However, this study does not conclude that ElevenLabs directly causes improvements in students' critical listening abilities or social-emotional competencies. Instead, the findings indicate that ElevenLabs functioned as an instructional resource that supported listening activities when integrated with pedagogical strategies such as reflective questioning, discussion, and analytical tasks. The role of the lecturer remains essential in designing learning activities, guiding interpretation, and facilitating meaningful reflection.

This study contributes to the discussion of AI integration in language learning by illustrating how AI-generated audio can be incorporated into critical listening pedagogy. Future research may involve larger and more diverse participants, compare AI-generated audio with human-recorded materials, or employ experimental and mixed-method approaches to further investigate how different characteristics of AI-generated voices support listening development and social-emotional learning.

6. References

- Collaborative for Academic, Social, and Emotional Learning (CASEL). (2020). *CASEL's SEL Framework: What are the core competence areas and where are they promoted?* Chicago, IL: CASEL. <https://casel.org/fundamentals-of-sel/what-is-the-casel-framework/>
- Creswell, J. W., & Creswell, J. D. (2023). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (6th ed.). Thousand Oaks, CA: SAGE Publications.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston, MA: The Center for Curriculum Redesign. <https://www.curriculumredesign.org/our-work/artificial-intelligence-in-education/>
- Izzati, S. S., Putri, N. E., & Bhakti, W. P. (2025). Penerapan metode storytelling untuk meningkatkan keterampilan menyimak berkelanjutan bagi siswa kelas rendah sekolah dasar. *Jurnal Pendidikan Bahasa*, 15(4). <https://doi.org/10.37630/jpb.v15i4.3544>
- Mayer, R. E. (2021). *Multimedia Learning* (3rd ed.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108894333>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative Data Analysis: A Methods Sourcebook* (4th ed.). Thousand Oaks, CA: SAGE Publications.
- Rost, M. (2016). *Teaching and Researching Listening* (3rd ed.). New York: Routledge. <https://doi.org/10.4324/9781315834988>

- Sartika, E., Salam, & Irawati, W. O. (2026). Implementing audio-visual folktales with Blooket-based gamification in listening instruction: A descriptive study of high school students. *International Journal of Educational Practice and Policy*, 4(3), 1–15. <https://doi.org/10.66314/ijepp.v4i3.974>
- Ubaidillah, M. I., Masripah, & Holis, A. (2025). Kemampuan menyimak sebagai pondasi pengembangan keterampilan berbahasa siswa kelas rendah pada mata pelajaran Bahasa Indonesia. *Jurnal Pendidikan dan Pembelajaran Indonesia (JPPI)*, 5(1), 439–448. <https://doi.org/10.53299/jppi.v5i1.1404>
- UNESCO. (2023). *Guidance for Generative AI in Education and Research*. Paris: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and Learning Second Language Listening: Metacognition in Action*. New York: Routledge. <https://doi.org/10.4324/9780203843370>