

INVESTIGATING THE USE OF MULTIMODAL RESOURCES IN LISTENING SKILLS AT JUNIOR HIGH SCHOOL

Ade Lia Atikah^{1*}, Muhammad Handi Gunawan²

^{1,2}University of Education Indonesia, Indonesia

Email: adeliaatkh21@upi.edu¹, handi_gunawan@upi.edu²

First Received: August 2025

Final Proof Received: September 2025

ABSTRACT

This study examines junior high school students' perceptions of the use of multimodal resources in their listening learning experiences, specifically focusing on the integration of video, images, and text. Although, research on multimodal listening is growing, studies focusing on junior high school students' perceptions of listening activities are limited. This study employs a qualitative case study design involving open-ended questionnaires in the form of students' reflections, semi-structured interviews, and classroom observations to gather insight into junior high school students' perceptions and challenges related to multimodal materials in listening instruction. The findings revealed that students had positive views on the use of multimodal resources and the integration of several elements, such as audio, images, and text, can improve their comprehension and engagement in class. On the other hand, challenges such as fast speech speed, unfamiliar vocabulary, and external distractions can hinder their listening comprehension process. To improve teaching, educators should effectively integrate multimodal resources, considering speech speed and the learning environment. In addition, considering students' perceptions and challenges can help to incorporate multimodal resources more effectively in junior high school environments.

Keywords: *Listening skills, multimodal resources, students' perception*

INTRODUCTION

Listening plays a crucial role in English language acquisition, as it forms the basis for understanding spoken input and supports the development of other language skills (Gilakjani & Sabouri, 2016). Listening requires active involvement, where students understand messages that originate from incoming data directly starting from recognizing sounds or voices, then identifying words, phrases, clauses, and sentences, until finally deciphering the meaning (bottom-up processing) and predicting the meaning based on context and prior knowledge (top-down processing) (Richards, 2008). In second language learning, students have crucial problems in listening comprehension because universities and schools tend to focus more on writing, reading, and vocabulary (Gilakjani & Sabouri, 2016).

Various teaching strategies can be used to improve students' listening skills. One of them is intensive listening. Intensive listening is considered as paying attention to certain sounds, words, phrases, and grammatical units and information correctly (Ahmadpour & Asadollahfard, 2018). Latifa et al. (2025) pointed out that intensive listening helps students improve their English language skills. By using certain audio materials repeatedly, students can strengthen their comprehension skills and improve their pronunciation and vocabulary.

Therefore, teaching listening in EFL can be a challenging task. Listening comprehension is often challenging due to several factors, such as speed of speech, unfamiliar vocabulary,

noises, long-spoken text and variety of accents, which require students to develop listening skills. In addition, the commonly used listening resources in the classroom only come from CDs, cassettes, and textbooks. These listening resources can make the learning experience monotonous and less interesting (Winiharti & Herlina, 2017). To overcome monotonous teaching in listening class, various modes or modalities of teaching can be offered. The involvement of two or more modes of communication in language learning is referred to as multimodality.

Multimodality, according to Kress and van Leeuwen (2006) is defined as the combination of several modes utilized to convey meaning. In the context of teaching and learning, multimodality refers to the various forms and resources that can be used to create meaning in any language, including sounds, music, speech, images, gestures and postures (Jewitt, 2013). Shaojie et al. (2022) also stated that audio-visual input can be used as a type of multimodal input because it includes three modes, namely images, text, and actions, and is mainly realized in sounds, subtitles, and images. Winiharti and Herlina (2017) recognized that students can see how intonation aligns facial expressions, gestures, and speech by watching videos. Thus, students can see how the speaker conveys the language.

In the field of language education, the integration of multimodality has become a focal point. Various studies have shown that the use of multimodality in teaching materials can effectively improve students' listening comprehension. Masinde et al. (2023) stated that students who learned with multimodal integration showed better engagement and learning outcomes compared to traditional methods. In addition, multimodal integration in teaching is preferred by students and improves their overall understanding (Onyebuchi, 2017; Ruan, 2015). The results of research conducted by Asilestari (2016) also show that learning using interactive multimedia is more effective than audio only. Nata and Suharyanto (2024) also revealed that multimodal integration can adjust various learning styles of students, thus making the learning process easier and more interesting. Furthermore, the use of multimodal in listening learning, such as videos, images, and other visual elements, can increase students' understanding, motivation, and engagement (Fan & Qi, 2023; Gjinali & Piri, 2023; Sudewi & Fadilah, 2023). Audio-visual integration in learning is considered more interesting and helps students understand pronunciation and vocabulary. Similarly, Ma'fiyah et al. (2021) suggested that the use of virtual learning could be a variable method for teaching listening skills during the pandemic. Fitri et al. (2025) found that visual elements, subtitles, and gestures aided comprehension and increased engagement in extensive listening learning. However, they also stated that cognitive overload, technical interference, and distraction were their challenges when multimodal was integrated in extensive listening.

Although research on multimodality has been popular and in high demand, especially in technology and English listening classes. There is still a gap in scholarly research focusing on how junior high school students perceive the use of combined modes in listening learning. This emphasizes the importance of examining students' perceptions and the challenges they may face when using multimodality to develop listening skills. This research focuses on two main questions: 1) what are junior high school students' perceptions of using multimodal resources in listening class? 2) what challenges do junior high school students face when learning listening skills through multimodal resources?

Multimodality in ELT

Multimodality refers to the integrated use of multiple communication modalities to convey meaning (Kress & van Leeuwen, 2006). Additionally, understanding texts, especially their linguistic components, has evolved into a complex endeavor, requiring an appreciation of the multiple features involved in the overall meaning of a text (Kress, 2003). In the context of

teaching and learning, multimodality refers to the variety of forms and resources that can be used to create meaning in any language, including images, gestures, sounds, music, speech, and body postures (Jewitt, 2013). In line with this, research conducted by Shaojie et al. (2022) asserts that audio-visual input can be viewed as multimodal, encompassing three meta-functions (images, texts, and actions) expressed through sound, subtitles, and dynamic images. In addition, effective approaches to teaching listening skills involve interactive exercises and a variety of multimedia resources (Ma'fiah, Sitoresmi, & Yulianto, 2021).

Multimodal Resources for Listening Skills

Multimodality in listening refers to different elements (sound, visual, movement) working together in listening activity to help the listener understand the message thoroughly (Campoy-Cubillo, 2019). In addition, Campoy-Cubillo (2019) also emphasizes that in order to use multimodal communication, teachers need to consider the elements of other modes that may occur in a particular order, such as multimodal in listening activities, some materials can provide access to audio and visual communication, such as music, video clips, DVDs, or watching online while listening (Ivan & Indah, 2021). Additionally, the combination of various modes of communication such as spoken dialog with visual cues in the use of videos, such as facial expressions, body movements, and background settings, can help to understand (Winiharti & Herlina, 2017). Campoy-Cubillo (2019) also pointed out that watching videos that combine facial and visual expressions with speech creates a more immersive learning experience that support comprehension.

Furthermore, the advantages in the use of video in listening activities is that many cues are provided through several elements of communication modes, such as intonation, facial expressions, gesture, and speech. Thus, it provides a more accessible opportunity for students to observe and interpret language use in an authentic and natural context, which can allow them to gain knowledge through a combination of auditory and visual cues (Campoy-Cubillo, 2019; Winiharti & Herlina, 2017).

Audio-visual Modes in Listening Classrooms

Regarding the different modes used in the EFL classrooms, Li (2015) claims that the function of visual images in videos is to focus attention, helping language students to reduce the cognitive load that arises when performing listening tasks, by providing real context as well as cultural interaction with native speakers that can fill in the background knowledge that they do not get. Winiharti and Herlina (2017) showed that the use of visual modes such as images, text, and subtitles can improve students' comprehension in listening compared to using only audio modes. Furthermore, research conducted Rachmaniputri et al. (2020), suggests that the use of podcasts in English listening class provides authentic listening material that foster students' enthusiasm for learning English, facilitates the development of understanding vocabulary development, refines pronunciation, and increases recognition of various English accents.

Listening Skills

Rost (2011) defines listening as the process of constructing meaning from spoken input by applying cognitive and linguistic knowledge. Furthermore, Brown (2004) argues that listening skills is an active process involving attention, comprehension, and the ability to derive meaning from auditory input. In the context of second language acquisition, listening plays an essential role in facilitating comprehension of spoken discourse (Richards, 2008). According to Nemtchinova (2020), listeners understand the incoming information in two different ways. Bottom-up refers to obtaining meaning from individual lexical, grammatical, and pronunciation items. It applies the decoding process, from sounds to words and the grammatical

relationships between the two, to sentences that lead to a comprehensive understanding. On the other hand, in a top-down process, listeners rely on their knowledge to guide their understanding of what they hear to understand the meaning of communication. Listening skills can be developed through various learning strategies and approaches that are centered on learners, as stated by Brown and Abeywickrama (2019), there are four types of listening performance in general, namely: 1) intensive listening, listening to understand language components (phonemes, words, intonation, discourse markers) from a wider range of language. 2) responsive listening, listening to relatively short language stretches (such as greetings and questions). 3) selective listening, processing discourse stretches such as short monologues to identify specific information. 4) extensive listening, listening to understand the overall meaning or main idea, and making conclusions from what has been heard.

METHOD

Research Design

This study used a qualitative case study to investigate the use of multimodal resources in listening skills among junior high school students. According to Creswell and Creswell (2017), qualitative data is an approach to exploring information from participants, and the result is in the form of words. Case study was chosen because the design focused on exploring processes, activities, and events (Creswell, 2014).

Participants

According to Guest et al. (2013), the implication of the qualitative case study data collection approach is that participants are selected based on specific characteristics so that the sample selected is small. The research was conducted in one public junior high school in Majalengka, involving 24 eighth-grade students. Five students were selected as samples to be interviewed semi-structurally. Each participant was given a pseudonym (S1 to S24) for anonymity.

Instruments

The data collection instruments in this study were observation, open-ended questionnaires, and semi-structured interviews. Although the primary sources of data were open-ended questionnaires and interviews, the observation is presented first in this section since it was conducted earlier, during the teaching process. The study was conducted around four class observations, adjusted to the school schedule, in which the researcher also acted as the teacher. Observation is a process of collecting information directly and openly by observing people and places at the research location (Creswell & Guetterman, 2019). Data collected from these observations include video recordings and field notes. Next, open-ended questionnaires were distributed by the researcher to collect qualitative data on their experiences with multimodal materials, with 26 students responding. These open-ended questionnaires took the form of students' reflections, providing space for participants to express how they interpreted their experiences during learning with multimodal materials in the classroom, thus providing rich and detailed data. Similarly, five participants were selected to answer questions in semi-structured interviews, which were conducted in Indonesian and audio-recorded for detailed insights. The primary data came from the open-ended questionnaires and interviews, while the observations ensured validity. The data collected through open-ended questionnaires, semi-structured interviews, and classroom observations were then transcribed, coded, and analysed using thematic analysis (Braun & Clarke, 2006).

Procedures

The first step in this study was to contact the school that would be the research site, then explain the purpose of the study and ask for approval from the school. The researcher then began preparing a lesson plan that focused on listening skills for four meetings. The researcher then asked the participants to complete students' reflection sheet about their experiences during four listening class meetings, which utilized multimodal materials, including the challenges they encountered. After the students' reflection were collected, the researcher selected five students to be asked for further explanations to be interviewed them for 15 minutes each using Indonesian, so that students felt more comfortable. The next stage was to transcribe the interview recordings for further analysis. After the transcription was complete, the collected data were analysed using thematic analysis to identify the main themes related to their experiences and challenges when presented with multimodal materials.

Data Analysis

The data collected through observation, students' reflection, and interviews were then analyzed using thematic analysis. After attending four learning meetings, students were asked to complete open-ended questionnaires in the form of reflections. Furthermore, five of them were selected for semi-structured interviews to obtain further explanations. The interview recordings were carefully transcribed and analyzed using thematic analysis.

FINDINGS AND DISCUSSION

The findings of this study show that most participants expressed an overall positive view of the use of multimodal resources in listening learning in the classroom. Specifically related to multimodal resources used in learning activities, they felt that the use of various modes provides an overall classroom experience and makes listening learning activities varied and interesting. Participants stated that the use of multiple modes made it easier for them to understand the context of the listening material. In addition, the majority of students also stated that the application of multimodal elements in listening materials can enhance learning engagement and motivation.

On the other hand, although participants gave positive responses to the use of multimodal resources in listening learning in class, some participants also revealed that they experienced some challenges when multimodal resources were applied in class. Fast-speed speech and unfamiliar vocabulary in the learning process as a result of using multimodal resources in listening materials were the most frequently mentioned problems. Furthermore, distractions such as lack of focus and noises were also found. The following section describes the recurring themes identified in students' responses, followed by examples of open-ended questionnaires and semi-structures interview excerpts from them. Therefore, the following discussion elaborates the findings in two main categories: first, it will explore students' perceptions and experiences regarding the application of multimodal resources in the listening classroom, and second, it will discuss the challenges they face when multimodal resources are applied in listening classroom.

Students' Perceptions of the Use of Multimodal Resources in Listening Learning ***Multimodal Resources in Listening Learning Improve Students' Understanding***

Data obtained from open-ended questionnaires, semi-structured interviews, and classroom observations revealed that the perceptions obtained from the majority of students regarding the use of multimodal resources in listening learning are positive. Students argue that listening to materials presented through various modes, including audio, visual, and text, helps improve

their listening skills. They argue that visual and text, which complement audio, can make it easier for them to process information and build their understanding more accurately. Students argue that listening to learning materials becomes easier to understand and comprehend due to the use of multiple modes. Multimodal resources allow students to combine information acquired auditory with visual and textual representations, creating a more complete understanding. The application of multimodal resources in listening learning in the classroom helps students to understand listening materials, especially for those who may have difficulty in capturing information that comes from audio only. Considering that, in this context, images, intonation, and text can be very useful in conveying meaning. Some example of these statement can be seen through the open-ended questionnaire and interview excerpts below:

Stated on the open-ended questionnaires:

“When I was learning to listen, I saw two people talking. I heard their voices talking, and I read the text slowly so that I could understand it. I felt happy because the pictures and text in the video made me understand better.” (S1)

“In my opinion, learning using multiple modes such as sound, images, text, and video can be more understandable and fun because we learn while listening to the sound, looking at the images in the video, and reading the text.” (S8)

Stated on the semi-structured interviews:

“I think learning with multiple modes using sound, images, text, and video is very exciting and easier to understand. Especially for me who is less able to learn English. Learning by combining several modes makes me understand more quickly.” (S4)

“... Learning using multimodal is easier to understand because there is a lot of help in the form of sound, text, and images.” (S22)

Based on the excerpts above, it was found that students generally perceive multimodal resources in listening class positively as they engage with multimodality elements such as sounds, images, and texts in listening materials. Overall, it was said that students' learning experience and comprehension were enhanced by the integration of these varied multimodal elements. Students are encouraged to actively engaged with and interpret information through multiple modes when they are exposed to diverse content and encourage a deeper understanding of the context.

In addition, the researcher's observations confirmed the previously mentioned claims. During class a discussion in the first meeting, the teacher displayed several images taken from the video and then asked students to predict the content of the story based on the image. Most students predicted that the content of the video was about a trip to the beach. As evidenced by these results, students' perception of the context of the video material was enhanced by visual cues.

These results are in line with previous relevant studies, which define multimodal resources as a variety of forms and resources, such as images, sounds, music, and speech that can be used to convey meaning in any language and in any learning and teaching context. Therefore, the combination of multiple additional modes such as text can improve students' listening comprehension more than using only audio modes (Campoy-Cubillo, 2019; Fitri et al., 2025; Gjinali & Piri, 2023; Shaojie et al., 2022; Sudewi & Fadilah, 2023; Winiharti & Herlina, 2017).

Multimodal Resources Create Engaging and Enjoyable Learning Experiences

In addition to the advantages that students gain when interacting with multimodal elements in listening learning materials, some students also mention that multimodal resources creates a more interesting, enjoyable learning experience and eliminate boredom. The presence of stimulus from several modes applied in listening learning in this class not only prevents boredom but also makes the learning process feel dynamic. This statement is proven by the following excerpts:

Stated on the open-ended questionnaires:

“... learning is not boring. I feel that is more interesting because I can hear, see, and read together.” (S3)

“I think learning with video, pictures, sound, and text is more interesting and fun because I learn while watching video, seeing pictures, listening to sounds, and reading text.” (S23)

Stated on the semi-structured interviews:

“When I study with you, I watch the video, listen to the dialog, read the materials. This experience is very exciting, making me quickly understand the material. It is not boring. It is fun.” (S9)

From the excerpts above, it can be concluded that when teaching and learning is delivered in an innovative and varied way, students feel more interested and listen more easily after using several modes to teach. Apart from just listening to audio, students can also see images and text, so they do not feel bored when learning listening, this is line with Hawema (2023), which highlights that the use of audio visuals can attract and reduce students' boredom while learning. The use of multimodal resources not only helps students to better understand the content but also creates an enjoyable learning experience.

In addition, this is also supported by classroom observation at the first meeting, in learning activities, the teacher integrates several mode elements simultaneously, such as visual in the form of images from video footage, audio, verbal in the form of oral instructions, and physical movement activities in listen and do activities. Students focus on listening and seem enthusiastic. After watching the video, students are introduced to the new vocabulary, record it, and then imitate the correct pronunciation.

This finding is consistent with previous research, which underlines how the use of audio and visual together can help students in learning activities, as well as emphasizing how audio visual can enhance their learning experience (Arrahma et al., 2024; Fadhillah et al., 2021; Fan & Qi, 2023; Onyebuchi, 2017).

Multimodal Resources Increase Learning Engagement and Motivation

In addition to the advantages experienced by students when interacting with multimodal elements in listening materials, such as improving comprehension and providing positive learning experiences for students, some students also mentioned that multimodal resources can motivate their learning and have a positive impact on engagement. The data obtained from open-ended questionnaire showed that some students thought that the use of several modes, such as images, sound, and text, simultaneously was important in listening learning activities. In addition, data obtained from semi-structured interviews also concluded that varied learning

modes can make the listening learning process more enjoyable. This claim is shown by the following excerpts:

Stated on the open-ended questionnaires:

“The utilization of video, images, and text make it easier for me to understand and engage with the lessons.” (S2)

“... video and images make learning easier to understand and increase enthusiasm for learning.” (S11)

Stated on the semi-structured interviews:

“Yes. It is motivating, for example, because it is varied, there are pictures and text, so it is fun. I like it because it is not boring, and it is entertaining.” (S16)

“Yes, it is exciting. From the images and sounds, because we get bored easily, we can also look at the images. Sometimes we get sleepy, but we can hear the sounds in the video while looking at the images.” (S21)

Furthermore, observation data confirms the statement mentioned earlier. Classroom observation at the fourth meeting showed that students worked on assignments independently and then discussed with their peers to check their answers guided by the teacher. It shows that students are actively involved in the listening learning process. Overall, data obtained from open-ended questionnaires, interviews, and classroom observation concluded that in addition to improving understanding and providing an enjoyable learning experience for students, multimodal resources also provide students' involvement and motivate them in learning activities. This is in line with research conducted by Gjinali and Piri (2023), the results of which were based on a 5-point Likert scale questionnaire that investigated students' attitudes toward their difficulties in understanding listening material. The results showed that the highest percentage for all items was in the “strongly disagree” column, indicating that most students did not experience difficulties in understanding listening material. Therefore, the researchers indicated that third-year Italian language students demonstrated confidence when learning listening material. Another study conducted by Sudewi and Fadilah (2023) also showed that audio-visual make students enthusiastic, excited, and help them understand listening lessons more easily since the explanations provided through audio-visual are easy to understand.

However, in addition to providing positive and beneficial impacts, some students also have negative responses to the use of multimodal resources in the context of listening materials. Their negative responses are described in the section below.

Students Challenges in Learning Listening Materials through Multimodal Resources

The use of multimodal resources in listening learning activities can positively increase engagement and provide an enjoyable learning experience. However, students still have some challenges when they learn listening comprehension. Data analyzed from students' reflections, interviews, and classroom observation, revealed several challenges experiences by students when learning listening materials through multimodal resources. The challenges they face when learning listening material are mentioned as fast speech speed, unfamiliar vocabulary, and distractions that arise.

Fast Speed Speech

Generally, most of the challenges faced by the students were the speed of speech when listening to the material delivered through the video. In this case, the main obstacle was that the speed of speech was too fast due to difficulties in understanding the overall information. They felt confused when listening to the sound from the video, which was too fast, so they did not have enough time to process the information provided. The following excerpts show this statement: Stated on the open-ended questionnaires:

“... sometimes, when the video and speech are too fast, I have trouble focusing on listening to them.” (S1)

“... the speech and the video are fast, I do not understand English if the speech is fast, so I get dizzy looking at the pictures or want to listen to the voice in the video fast.” (S7)

Stated on the semi-structure interviews:

“My difficulty is that the voice in the video is sometimes too fast.” (S16)

Moreover, the observation data obtained from the fourth meeting class observation showed that most of the students when doing the listening task, they asked to replay the video several times due to not getting some information from the material in the video.

This result is in line with Astika & Kurniawan's (2020) study, which conducted similar research related to the difficulties experienced by students in listening activities. The findings of this study indicate that students are distracted and unable to continue interpreting information due to the speed of speech. Consequently, students have difficulty recognizing the meaning of words when the speaker is talking.

Unfamiliar Vocabulary

Second, unfamiliar vocabulary is another challenge for students when encountering listening materials using multimodal resources. In some cases, students find it difficult to understand unfamiliar vocabulary when multimodal resources are used in listening activities. Most students assume that unfamiliar new words hinder their ability to understand the content of the video. Despite the integration of multimodal resources, such as visual images and text during listening activities, this challenge persists for the majority of students. The following excerpts show students' statements regarding issues related to unfamiliar vocabulary.

Stated on the open-ended questionnaires:

“...sometimes, I do not understand because there are many new words that I do not know. Even though there is text, I still sometimes get confused about the meaning of the words in the video.” (S14)

“I only had a little trouble with the listening part when the audio was not clear and sometimes, I got confused by parts of the text that I did not understand. Sometimes, I was not sure where to focus, especially when the words were new to me.” (S8)

Stated on the semi-structure interviews:

“If the video is too fast, sometimes I cannot understand new words because they are unclear. The fast pronunciation makes it difficult to understand the video’s content.”
(S9)

“I feel challenged because there are new words that I do not know the meaning of, which makes learning difficult.” (S16)

Nevertheless, the data obtained from the third-class observations meeting, when working on listening tasks through multimodal resources, some students asked the teacher when they found new vocabulary that they did not know yet. They asked the teacher how to write the new vocabulary and find out its meaning.

Overall, these excerpts and classroom observations describe problems experienced by most students. Unfamiliar vocabulary hinders their listening comprehension, even when multimodal resources are integrated. These finding is in line with the research conducted by Fadilah et al. (2024), who found that some students face challenges in listening to English speakers, one of which is encountering unfamiliar vocabulary and lack of knowledge regarding its meaning. The limitations in understanding such vocabulary cause students to be unable to capture information comprehensively.

Distraction

Finally, the open-ended questionnaire data revealed that the utilization of multiple modes in listening materials or distractions from the surrounding environment caused some students to feel disturbed. Some students mentioned that they found it difficult to stay focused when multimodal resources were used. In addition, students found it difficult to obtain information from various modes simultaneously. They said that they felt confused about where to get the information first, whether from pictures, text, or sound. Furthermore, data from semi-structured interviews indicated that a noisy classroom environment was also distracting students’ concentration when learning to listen in class. This suggests that other factors, such as noise from classmates, can affect students’ ability to focus on listening material delivered through multimodal resources.

Stated on the open-ended questionnaires:

“When I practice listening with videos, images, audio, and text, I tend to focus only on the images, so I forget to read the text and listen to the audio.” (S3)

“I felt confused when participating in the listening lesson because the video was too fast and I had difficulty understanding the conversation. I was also confused about whether I should look at the images or read the text first.” (S20)

Stated on the semi-structure interviews:

“... it is hard to concentrate because my friends are being noisy, so I get distracted”
(S4)

Distractions, such as concentration problems, are one of the challenges that most students experience when learning to listen through multimodal resources. This aligns with the findings of Fitri et al. (2025), which highlights that multimodal resources not only offer numerous benefits in students’ learning experiences but also that the simultaneous integration of multiple

modalities can disrupt students' concentration. Some students reported that they struggled to divide their focus between listening, reading subtitles, and observing images or the speaker's facial expressions, leading them to miss important information.

In addition, noise from the classroom environment, such as noisy classmates, can also distract students' focus and attention to the listening material delivered through these multimodal resources. Students lose their concentration due to the noise, making it difficult for them to capture important information from the listening material. This is in line with the findings of Hardiyanto et al. (2021), who stated that one of the difficulties students face in listening activities in the classroom is noise. The researchers argued that the noise comes from the classroom, especially from classmates when they are learning to listen. This noise can disrupt students' concentration when listening and cause them to not understand the content of the material being presented.

CONCLUSION

The findings highlight the importance of using multimodal resources in listening activities in the classroom. The application of multimodal resources in listening classes has a positive impact on improving students' learning experiences. The combination of multimodal elements, such as visual, audio, and text, can significantly improve students' understanding of the material. Additionally, the combination of several multimodal elements also creates a more engaging and enjoyable learning experience, improving students' engagement and motivation in learning listening materials. However, although multimodal resources provide many benefits, students also face several challenges when learning listening materials using multimodal resources. The challenges include fast speed speech, unfamiliar vocabulary, and distractions from the physical environment. Students often mention the challenge of fast-speed speech, which makes it difficult for them to process information from listening materials. Moreover, unfamiliar vocabulary also acts as a challenge for students in listening activities, thereby hindering their ability to understand the context of the video. Furthermore, distractions such as noise in the classroom and difficulty in dividing their focus between multiple modes, are also become challenges for them in obtaining information from listening materials.

These insights provide practical implications for teaching practices in listening materials at the junior high school level. Educators should strive to include multimodal resources to make students' learning experiences more engaging and enjoyable. In addition, educators also need to consider selecting the right learning materials related to audio and fast speed speech in the learning content. Similarly, creating a conducive learning environment can help students' with their listening skills in the classroom.

To build on these findings, future research should further examine students' perceptions of multimodal resources presented in the classroom and the strategies that can be used to overcome the challenges encountered. Future research could explore the most effective combinations of multimodal resource elements for students and how educators can implement these strategies in their classrooms. By focusing on this area, we can design more inclusive and effective learning environments for listening instruction.

REFERENCES

- Ahmadpour, S., & Asadollahfam, H. (2018). A Study of the Effect of Extensive and Intensive Listening on the Tense Use among EFL Learners in Iran. *The Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, 6(2), 141–161. <https://doi.org/10.22049/jalda.2019.26383.1098>

- Arrahma, N., Harahap, S. R., Salwa, M., & Lubis, Y. (2024). The Effects Of Audio-Visual Aids On Listening Comprehension. *Journal of Education, Arabic and Islamic Studies*, 2(1), 27–37. <https://doi.org/https://doi.org/10.58355/qwt.v2i1.38>
- Asilestari, P. (2016). An Improving Students' Listening Skill Through Interactive Multimedia at XI Grade of SMK Negeri 1 Bangkinang. *Journal of English Language and Education*, 1(2). <https://doi.org/https://doi.org/10.31004/jele.v1i2.16>
- Astika, G., & Kurniawan, A. (2020). The challenges of using TED Talks as authentic resources of academic listening for EFL university students. *Indonesian Journal of Applied Linguistics*, 9(3), 589–598. <https://doi.org/10.17509/ijal.v9i3.23209>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson.
- Brown, H. D., & Abeywickrama, P. (2019). *LANGUAGE ASSESSMENT PRINCIPLES AND CLASSROOM PRACTICES* (3rd ed.). Pearson.
- Campoy-Cubillo, M. C. (2019). Functional diversity and the multimodal listening construct. *European Journal of Special Needs Education*, 34(2), 204–219. <https://doi.org/10.1080/08856257.2019.1581402>
- Creswell, J. W. (2014). *Research Design: Qualitative, quantitative, and mixed methods approach*. SAGE Publications.
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage publications.
- Creswell, J. W., & Guetterman, T. C. (2019). *EDUCATIONAL RESEARCH PLANNING, CONDUCTING, AND EVALUATING QUANTITATIVE AND QUALITATIVE RESEARCH* (6th ed.). Pearson.
- Fadhillah, J. N., Muniroh, S., & Rahmaningtyas, H. (2021). English Department Students' Perceptions of Using English Talk Show Videos on YouTube to Improve Listening Skill. *JoLLA: Journal of Language, Literature, and Arts*, 1(10), 1451–1465. <https://doi.org/10.17977/um064v1i102021p1451-1465>
- Fadilah, L., Islam, Muhammad, H., & Kholili, A. (2024). Investigating Students' Difficulties in Learning Listening Compreheson Skills at Uswatun Hasanah Senior High School. *EDUSAINTEK: Jurnal Pendidikan, Sains Dan Teknologi*, 11(4), 1664–1676.
- Fan, K., & Qi, Y. (2023). Effects of Multimodal Listening Instruction on Listening Performance of Rural Middle School Students from East China. *Global Research in Higher Education*, 6(1), 21–35. <https://doi.org/10.22158/grhe.v6n1p21>
- Fitri, S. A., Gunawan, M. H., & Nurlaelawati, I. (2025). MULTIMODALITY IN EXTENSIVE LISTENING : A CASE STUDY OF INDONESIAN EFL STUDENTS. *English Review: Journal of English Education*, 13(1), 269–278. <https://doi.org/https://doi.org/10.25134/erjee.v13i1.9706>
- Gilakjani, A. P., & Sabouri, N. B. (2016). Learners' Listening Comprehension Difficulties in English Language Learning: A Literature Review. *English Language Teaching*, 9(6), 123–133. <https://doi.org/10.5539/elt.v9n6p123>
- Gjinali, A., & Piri, S. (2023). Usage of Video and Multimedia to Improve Listening Skills: The

- Perception of Foreign Language Students. *Journal of Educational and Social Research*, 13(6), 160–173. <https://doi.org/10.36941/jesr-2023-0155>
- Guest, G., Namey, E. E., & Mitchell, M. L. (2013). *Collecting qualitative data: A field manual for applied research*. Sage publications.
- Hardiyanto, A., Tanjung, M., & Suharjono, S. (2021). Listening Comprehension Difficulties; a Case Study of Efl Students in Listening Class. *ETERNAL (English, Teaching, Learning, and Research Journal)*, 7(1), 168–179. <https://doi.org/10.24252/eternal.v7i1.2021.a12>
- Hawema, S. (2023). The Utilization Of Audio Visual Media In Increasing Students Listening Skills. *ATHLA : Journal of Arabic Teaching, Linguistic and Literature*, 4(1), 62–71. <https://doi.org/10.22515/athla.v4i1.6840>
- Ivan, M. N., & Indah, R. N. (2021). Music and Seventh Graders ' Listening Proficiency : Does Islamic song work ? *Journal of Arabic Teaching, Linguistic And Literature*, 6(2), 51–58. <https://doi.org/https://doi.org/10.31004/jele.v6i2.153>
- Jewitt, C. (2013). Multimodal teaching and learning. In *The Encyclopedia of Applied Linguistics* (pp. 1–5). Blackwell Publishing.
- Kress, G., & van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design* (2nd ed.). Routledge.
- Latifa, A. H., Suryana, D., Danuwijaya, A. A., & Gunawan, M. H. (2025). EXPLORING STUDENTS ' PERCEPTIONS OF INTENSIVE LISTENING IMPLEMENTATION IN INDONESIAN HIGHER EDUCATION. *Indonesian EFL Journal*, 11(1), 21–30. <https://doi.org/https://doi.org/10.25134/ieflj.v11i1.11336>
- Li, C. H. (2015). A comparative study of video presentation modes in relation to L2 listening success. *Technology, Pedagogy and Education*, 25(3), 301–315. <https://doi.org/10.1080/1475939X.2015.1035318>
- Ma'fiyah, I., Sitoresmi, U., & Yulianto, I. (2021). Efl Students' Perception of Virtual Learning in Listening Class: a Classroom Implementation. *English Review: Journal of English Education*, 10(1), 149–158. <https://doi.org/10.25134/erjee.v10i1.5365>
- Masinde, R., Barasa, D., & Mandillah, L. (2023). Effectiveness of using multimodal approaches in teaching and learning listening and speaking skills. *Nairobi Journal of Humanities and Social Sciences*, 7(2), 91–112. <https://doi.org/10.58256/njhs.v7i2.1333>
- Nata, D. F. M., & Suharyanto, S. P. (2024). *Multimodal Approach in Teaching Listening in EFL Classroom of Indonesia: Teacher's Perspective and Experience*.
- Nemtchinova, E. (2020). *Teaching listening*. TESOL Press.
- Onyebuchi, C. (2017). Effect of multimodal strategies on the learning of listening skills among junior secondary school students in Sokoto Metropolis. *Sokoto Educational Review*, 17(1), 39–47. <https://doi.org/https://doi.org/10.35386/ser.v17i1.15>
- Rachmaniputri, F. I., Purnawarman, P., & Gunawan, M. H. (2020). Students' Perceptions Towards the Use of Podcasts in Listening Class. *Advances in Social Science, Education and Humanities Research*, 546, 179–187. <https://doi.org/10.2991/assehr.k.210427.028>
- Richards, J. C. (2008). *Teaching Listening and Speaking: From theory to Practice (RELC Portfolio Series)*. Cambridge University Press.
- Rost, M. (2011). *Teaching and Researching Listening Skills*. Routledge.

<https://doi.org/10.4324/9781351043281-2>

- Ruan, X. (2015). The Role of Multimodal in Chinese EFL Students' Autonomous Listening Comprehension & Multiliteracies. *Theory and Practice in Language Studies*, 5(3), 549–554. <https://doi.org/10.17507/tpls.0503.14>
- Shaojie, T., Samad, A. A., & Ismail, L. (2022). Systematic literature review on audio-visual multimodal input in listening comprehension. *Frontiers in Psychology*, 13, 1–11. <https://doi.org/10.3389/fpsyg.2022.980133>
- Sudewi, P. W., & Fadilah, N. (2023). Second Graders' Perception of Audio Visual Use in Teaching Listening Skills at SMP Negeri 3 Majene. *Journal of English Teaching*, 9(1), 98–107. <https://doi.org/https://doi.org/10.33541/jet.v9i1.4578>
- Winiharti, M., & Herlina, C. (2017). Audio only or Video?: Multimodality for Listening Comprehension. *Advances in Social Science, Education and Humanities Research*, 82, 189–192. <https://doi.org/10.2991/conaplin-16.2017.41>