MULTIMODALITY IN EXTENSIVE LISTENING: A CASE STUDY OF INDONESIAN EFL STUDENTS

Sherly Annisa Fitri
English Education Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia
Email: sherlyannisa@upi.edu

Muhammad Handi Gunawan*
English Education Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia
Email: handi_gunawan@upi.edu

Iyen Nurlaelawati
English Education Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia
Email: iyenn@upi.edu

Abstract: This study explores 26 EFL students' perception towards the use of multimodality in their extensive listening (EL) experiences, specifically the use of videos, images and text (subtitles). Despite the growing research on multimodality in listening, research focusing on multimodality in extensive listening (EL) remains limited. This research employed a case study design and utilized course reflections which served as open-ended questionnaires, semi-structured interviews, and classroom observations to gather insights into students' perceptions and challenges related to multimodal materials in EL instruction. The findings reveal that multimodality can enhance students' understanding in EL materials but can become distracting if overused. To improve teaching, teachers should carefully balance the use of multiple modes, ensuring they assist rather than overwhelmed learners. In addition, considering students' perceptions and challenges can lead to more effective multimodal integration in EL environments.
Keywords: *extensive listening; listening comprehension; multimodality; students' perception.*

INTRODUCTION

Extensive Listening (EL) in the context of English as a Foreign Language (EFL) is a teaching and learning approach that encourages language learners to listen to a large quantity of enjoyable and easy-to-understand audio materials in the target language over a long period (Ivone & Renandya, 2019; Pamuji et al., 2019). This method allows students to enjoy the listening experience without focusing on the specific details of the content or the language used, which makes it different from intensive listening. Through the practice of extensive listening, scholars noted that students are encouraged to engage with texts in the target language which they can choose based on their potential relevance, interest, and enjoyment instead of their linguistic content, emphasizing that students are encouraged to complete comprehension-focused, non-intrusive tasks to talk about their reaction on the chosen listening topic in their native language at their own pace without assistance from the teacher, and prioritize meaning over form (Chang & Renandya, 2023; Ivone & Renandya, 2019; Setyawan et al., 2022).

However, students might find it challenging when the majority of the listening resources come from audio-only sources like CDs, cassettes, and textbooks. These materials often repeat standard listening exercises, which can make the learning experience feel monotonous and less-engaging (Winiharti & Herlina, 2017). In contrast, incorporating multimodality into the listening classroom—by integrating visual, textual, and interactive elements—offers a dynamic approach that might enhance their comprehension and engagement. For example, using videos, animations, pictures, or subtitles alongside audio materials can provide students with a richer, more varied learning experience, helping to maintain their interest and improve understanding (Arrahma et al., 2024; Chaichompoo, 2019; Dheghu et al., 2021; Efruan et al., 2023; Hussain & Khan, 2022; Napikul et al., 2018; Pribadi et al., 2024).

Multimodality, according to Kress and Leeuwen (2006), is defined as a combined utilization of several communication modalities to convey meaning. Multimodality in the context of teaching and learning refers to the diversity of

forms and resources that can be used to create meanings in any language, including picture, sound, music, gesture, gaze, speech, action with books and boards, and body posture (Grapin, 2019; Jewitt, 2013). In line with this, some scholars also pointed out that audio-visual input can be thought of as a type of multi-modal input as it embodies the three meta-functions of image, text, and action and primarily manifests as sound, subtitles, text-to-speech, and images (Oktalia & Drajati, 2018; Quyen, 2020; Shaojie et al., 2022).

Multimodality in listening, as noted by Campoy-Cubillo (2019), refers to how different elements (like sounds, visuals, or gestures) work together during a listening activity to help us fully understand the message. It involves considering all these elements as part of the overall communication process. In reference to the application of multimodality in listening, there are several materials that can give access to both audio and visual communication such as film clips on video, DVD or online watch while listening (Ivan & Indah, 2021; Widyaningrum et al., 2022). When compared to audio-only listening activities, the main advantage of using video in listening activities is the numerous cues provided by the various communicative modes, such as intonation, facial expressions, gestures, and speech. This tends to allow learners to more easily observe and interpret language usage in natural contexts, gaining knowledge from both auditory and visual cues (Campoy-Cubillo, 2019; Winiharti & Herlina, 2017).

Extensive listening (EL) has been widely studied as an approach to improving listening skills in language learning. Scholars have explored various methods and materials to enhance EL, such as podcasts, TED talks, and audiobooks, which provide learners with ample exposure to authentic language input. For instance, the use of podcasts and extensive listening journals has been investigated, highlighting their role in helping students plan, observe, and reflect on their listening progress (Alm, 2013; Rachmaniputri et al., 2020; Gunawan et al., 2023c; Gonzalez-Torres & Solano, 2024). These studies demonstrate that EL materials, when effectively integrated into the classroom, can significantly improve students' listening proficiency and overall learning experience.

In the realm of language education, the integration of multimodality has become a focal point, particularly in listening classrooms. Previous research has delved into various approaches, employing a diverse range of communication modes to enhance the learning experience. Notably, scholars have explored the use of audio-visual elements, including video presentations, podcasts, TED talks, talk shows, still images, audiobooks, and even music, recognizing the potential of visual cues alongside auditory stimuli (Fadhilah et al., 2021; Fatimatuzzahro et al., 2024; Febriani et al., 2022; Gunawan et al., 2023a; Gunawan et al., 2023b; Hui & Jamaludin, 2024; Lestary, 2019; Nadiyya & Suryadi, 2024; Rachmaniputri et al., 2021; Solanon et al., 2020; Setiani et al., 2024; Shaojie et al., 2022; Winiharti & Herlina, 2017; Yıldırım & Taşpınar, 2023). Additionally, some studies have even ventured into the synergistic combination of these modalities, acknowledging the multifaceted nature of language learning and the importance of catering the diverse learner preferences (Alm, 2013; Gunawan et al., 2023a; Gunawan et al., 2023b).

With regard to different modes used in EFL classes, Lee (2014) suggested that using different methods in English classes, like using visuals and other ways of teaching, can help EFL students learn better, and it can also help EFL teachers plan their lessons. In addition, scholars claimed that visual images in videos serve as a focus of attention, and that content and context visuals in videos help language learners relieve the cognitive strain that comes with listening tasks, observe real-world language and cultural interactions with native speakers, and—above all—visually fill in the background knowledge that they are missing, which can help students' development in learning (Li, 2015; Polat & Erişti, 2019).

To provide an example, numerous studies on the application of audio-visual modes in listening classrooms have been conducted. Winiharti and Herlina (2017) emphasized that utilizing multiple modalities is crucial for enhancing the efficiency and effectiveness of English language learning, enabling students to acquire the language more quickly and proficiently, as they highlighted in the results of their study that the use of visual mode such as images, text, and occasional subtitles in addition to audio, enhances students' listening comprehension compared to relying solely on audio mode. In line with this, interactive exercises and multimedia resources can be modern and efficient ways to teach listening skills, as suggested by Maqableh and Alia (2021). With these approaches, students are involved in activities that focus on the process of learning as well as improving and gaining effective listening skills.

Furthermore, the use of technology like podcasts (audio and video), TED talks, and Google

Sites, which employs several modes, is also one of the examples of how multimodality is applied in English listening classrooms. The use of podcasts and extensive listening journals were investigated in related studies, highlighting its use in helping students plan, observe, and reflect on their English listening progress (Alm, 2013; Rachmaniputri et al., 2020; Gonzalez-Torres & Solano, 2024; Gunawan et al., 2023c). Prior research has shown that the application of podcasts in extensive listening classrooms greatly improved the overall quality of learning for students enrolled in these listening courses. Likewise, scholars discovered that the use of Google Sites in students' EL materials empowers them to efficiently organize their materials, assignments, and completed work in one central location, providing them with full control over their materials and learning process (Gunawan et al., 2023b; Hanifa et al., 2024).

Despite growing interest in multimodality, there remains a gap in understanding EFL students' perceptions, particularly in extensive listening (EL) classrooms. Although studies on the use of multimodality have become more popular, especially in the field of technology and the use of particular combined modes in English listening classes, there is a lack of scholarly research focused on how students perceive the combination of different modes in the context of EL instruction. This research gap emphasizes the necessity of a more thorough examination of students' perceptions as well as the possible challenges that they may encounter while engaging with multimodal materials in the context of EL materials.

Hence, the main objective of this study is to offer perceptive answers to the following questions that have been determined to be worth looking into: (1) What are the EFL students' perceptions towards the use of multimodality in the extensive listening materials they engage with? (2) What benefits and challenges do the EFL students encounter when engaging with multimodal elements in the extensive listening materials?

METHOD

This study employs a case study approach to investigate the use of multimodality in extensive listening (EL) classrooms. The case study methodology is chosen for its ability to provide an in-depth exploration of a specific phenomenon within its real-life context (Yin, 2018). It focuses on students' perceptions of multimodal materials in the 'Listening for Academic Purposes' course, employing qualitative methods like open-ended questionnaires, interviews, and classroom observations.

The participants were 26 fifth-semester English majors from an Indonesian public university in West Java, all enrolled in the 'Listening for Academic Purposes' course. The selection of participants and course was intentional, as the course integrates extensive listening (EL) practices, allowing students to select listening materials on topics related to the lecturer's content, such as Global Englishes. Each participant was assigned a pseudonym (S1 to S26) for anonymity.

Participants first received consent forms to confirm their voluntary participation, and all 26 students agreed and signed the forms. Next, open-ended questionnaires were distributed to collect qualitative data about their experiences with multimodal materials, with 40 students responding. Five participants were then selected for semi-structured interviews, conducted in Indonesian and audio-recorded for detailed insights. To validate the data, nine classroom observations were conducted, using field notes, observation sheets, and video recordings to capture real-time behaviors. The primary data came from the questionnaires and interviews, while the observations ensured validity. All data were transcribed, coded, and analyzed using thematic analysis (Kiger & Varpio, 2020; Preiser et al., 2021) to identify recurring themes and patterns, linking the findings to previous research on multimodality in EL materialss.

RESULTS AND DISCUSSION

The results and discussion of this research shed light on the participants' perspectives regarding the integration of multimodality in EL materials. Overall, a significant majority of the participants expressed a positive outlook on the use of multimodality in the EL materials. When it comes to participating in activities and assignments, in particular, they felt that having a range of modes enhanced their overall classroom experience and made the activities more varied and engaging. The utilization of different modes, according to the participants, made it easier to comprehend the context included in the listening materials. Additionally, some students indicated that the application of multimodal elements in extensive listening materials can help them explore their preferred learning styles (visual, kinesthetic, or auditory).

On the other hand, despite these positive responses, some of them also expressed that they encountered some difficulties when interacting

with multimodal materials. Distraction and a sense of inefficiency in the learning process as a result of the overuse of modes in the EL materials were the most frequently mentioned problems. Furthermore, technical problems like poor connection and accessibility were also found. The following section will elaborate on recurring themes identified in the students' responses, succeeded by sample open-ended questionnaires and interview excerpts from them. As such, the discussion that follows the break down the findings into two main categories: first, it will explore the participants' perceptions and experiences regarding the application of multimodality in an extensive listening classroom, and second, it deals with the challenges they face when utilizing these resources and how they deal with them.

Students' perceptions of the use of multimodality in EL materials
Multimodality in EL materials enhances students' understanding
The data from open-ended questionnaires, semi-structured interviews, and classroom observations revealed that the majority of the students perceive the use of multimodality in EL to be positive. They believe that the use of multimodal elements such as audio, visual, and text help them understand the context of the video materials that they watch on YouTube better. They also argued that the use of intonations helps them greatly in interpreting the emotions and feelings of the speakers in the weekly video materials they watch. Students are able to infer from the speaker's tone whether they are trying to be humorous, sarcastic, or mocking. Considering that, in this context, the focus of the EL materials is Global Englishes, and watching videos of non-native English speakers was part of the syllabus, body language, intonation, facial expressions, and lip movements can all be very useful in conveying meaning. A few examples of this statement can be seen through these open-ended questionnaire and interview excerpts below:
Stated on the open-ended questionnaires:

"... I find multimodal elements very helpful. For example, when watching YouTube videos that are accompanied by animations, it is easier for me to understand. Then, subtitles that are available are very helpful for me, who sometimes still find it quite difficult when faced with speakers who have a fast pace of speech. The use of videos to explain the material also helped me in capturing the key points. Thus, all these elements made it easier for me to do the assignments in this course. One of them is doing

the Extensive Listening Journal." (S7)

Stated on the semi-structured interviews:

"I personally believe that having a variety of media types increases my engagement when it comes to learning materials. … Even though the content is not particularly interesting to me at first, I find that I can understand it better if it includes visuals like graphics or illustrations in addition to audio or video materials, either in the e-book or through supplemental media." (S1)

"I use subtitles most of the time. more so in the case of English. especially when the native speaker is speaking quickly, like in a TED talk. … I believe that subtitles are essential, especially for TED talks and interviews. Nonetheless, I frequently avoid using subtitles for videos with voiceovers that are clear and understandable." (S4)

Based on the excerpts above, it was found that the students generally perceive multimodality in EL positively as they are used to accessing multimodal elements like gestures, facial expressions, sounds, images, and text in the EL materials. It was said that the students' overall learning experience is enhanced and their comprehension is improved by this varied combination of multimodal inputs. Students are encouraged to actively engage with and interpret information through multiple channels when they are exposed to such diverse content, which creates a dynamic learning environment and promotes deeper understanding about the context.

Apart from that, the researcher's observation confirmed the previously mentioned claims. Following the viewing of a video featuring non-native English speakers, a student (S23) mentioned during the class discussions in the fourth meeting that they assumed the first speaker might be a celebrity because of her confident manner of speaking. As evidenced by this result, students' perception of the video materials' context is improved by non-verbal cues.

These results show that they are consistent with the previous relevant studies, which defined multimodality as the variety of forms and resources—such as images, sounds, music, gestures, speech, and body posture—that can be employed to convey meaning in any language in the context of teaching and learning. Moreover, it is also said that additional modes of communication, such as subtitles, gestures, physical distances, and facial expressions can

improve students' listening comprehension more than relying only on audio mode (Winiharti & Herlina, 2017; Campoy-Cubillo, 2019; Hao et al., 2021; Shaojie et al., 2022; Efruan et al., 2023; Gjinali & Piri, 2023).

Multimodality in EL materials support students' learning preferences
In addition to the advantages that students experienced when interacting with multimodal elements in EL materials, a few students also mentioned that multimodality could be used to support students' personal learning preferences. Interesting enough, students find multimodal elements useful in exploring their preferred learning style—a finding that comes as a surprise. This claim is demonstrated by the following excerpts:
Stated on the open-ended questionnaires:

"Apart from enhancing my comprehension of specific concepts in the extensive listening classroom, multimodality may also help in accommodating the diverse learning styles of every person, given that there exist multiple types of learners, including auditory, visual, and kinesthetic." (S6)

"The use of diverse modes is important because it can increase student participation and help students understand the learning material. In addition, the use of diverse modes can also facilitate students' diverse learning styles, ranging from visual, audio, and kinesthetic learning styles." (S11)

From the above excerpts, it can be inferred that a way to accommodate students' preferred learning styles in an EL materials could be to use multiple modes, which aligns with Masinde et al.'s study (2023), highlighting the fact that multimodal approaches can facilitate students' multiple or different learning styles. Students can interact with the content in a way that suits their unique learning preferences when both visual and auditory modes are offered in a single, comprehensive listening piece. This adaptability not only accommodates a range of learning preferences but also improves comprehension and retention in general. In addition, this finding is consistent with the existing studies, which highlighted how using both auditory and visual modes helped students learn more thoroughly, highlighting the results of the audio-visual aids helped the students to improve their learning activities and achievements (Dheghu et al., 2021; Efruan et al., 2023; Arrahma et al., 2024).

Similarly, Arrahma et al. (2024) noted that multimedia materials can stimulate the multiple senses in the human's body, resulting in making the whole learning process seem more engaging and effective.

However, few students also had negative reactions to the use of multimodality in the context of an EL materials, despite its beneficial application. Their negative responses, however, will be broken down in the section below.

Students' preference in engaging with EL materials
Even though most of the students acknowledged that multimodality can be useful for them in accessing online information, one of the students (S9) mentioned the less urgency of using modes other than audio in a listening class. Another student (S12) expressed that they would prefer not to use multimodality in EL materials if given the choice. The following excerpts provide data for these claims:
Stated on the open-ended questionnaires:

"... In extensive listening, the work requirement should be audio only. It doesn't always have to be multimodal." (S9)

"My experience of using multimodal in extensive listening activities is actually not something I would choose if given the choice. But the use of multimodal in extensive listening activities has definitely increased my understanding and knowledge so I don't think it's a bad thing and I'm just a bit lazy." (S12)

Moreover, while some students believe that the EL materials' use of gestures, intonation, facial expressions, and lip movements assist in their understanding of the subject matter, some of them disagree with this claim for a variety of reasons. For instance, a student (S21) stated that they would rather focus on the information conveyed rather than the speaker's gestures, facial expressions, intonation, and lip movements in order to comprehend the content of EL materials.
Stated on the open-ended questionnaire:

"I did not think that the content was well understood by the use of gestures, facial expressions, intonation, and lip movements, and this was still the case when I paid attention to them. I find that when I pay attention to what is being discussed, I understand things better." (S21)

Parallel to this, a different student (S5)

mentioned that, out of personal preference, they prefer to pay attention to the message or content being presented in the video material rather than the speaker's non-verbal cues. The following excerpt supports this assertion:

Stated on the semi-structured interview:

> "Personally, I focus more on the content rather than paying attention to gestures, facial expressions, or intonation. In terms of gestures, it's good that the speaker presents the material more (visually) engaging or so on, but, in terms of the material itself, because I'm more content-oriented than (focusing on) how the speaker delivers, I prefer to pay attention to the content." (S5)

Students' challenges in accessing extensive listening materials using multimodality
Generally, the students mentioned that they encountered no significant difficulties interacting with multimodal elements in the EL materials. However, when it comes to the difficulties in interacting with EL materials, technical difficulties, accessibility issues, and distractions were mentioned.

Technical issues while accessing EL materials
The majority of the students' challenges were related to technical issues, such as a slow internet connection and poor audio or video quality. In this case, technical difficulties such as these are probably among the issues because the majority of the class meetings were held virtually using Zoom meetings. Since watching videos and participating in class or group discussions make up the majority of the activities conducted in this classroom, a bad internet connection may cause lag or low-quality videos. The following excerpts demonstrate this statement:

Stated on the open-ended questionnaires:

> "During the lectures, I found it difficult to understand the content of the video because often my signal was unstable when the lecturer played the video. … To solve this problem, I tried to find the video on the internet and watched it again by myself." (S5)

> "So far, no difficulties have occurred with the use of multimodal elements, the only obstacle is the poor internet connection. This is indeed difficult to avoid. What I usually do is replace the WiFi network by using a mobile hotspot from a cellphone or vice versa." (S17)

These results are in line with Mandiri et al.

(2022)'s study who conducted a similar study related to the difficulties encountered by the students in EL activities. According to the study's findings, the most common technical issues that participants faced were playback errors and a bad internet connection, which caused video and audio to buffer.

Accessibility issues while accessing EL materials
Secondly, accessibility issues were brought up as the next challenge. It can occasionally be difficult for some students to find suitable videos for the weekly Extensive Listening Journal (ELJ) assignment. Besides, the platform Google Sites, where students are supposed to gather all of their assignments, can only be edited using a laptop or computer. Because of this, every student needs to bring a laptop to class if lectures are held offline on campus, which can be a bother for some students. The following excerpts will highlight the students' statements regarding accessibility-related issues.

Stated on the open-ended questionnaires:

> "Another difficulty I faced was deciding what kind of videos to use. The videos chosen were often on related topics but not in English. This made me have to look for other videos that were more suitable." (S7)

> "The problem I usually experience is the accessibility of Google Sites. Unfortunately, until now Google Sites has not been editable other than through a laptop or computer. So, when I want to work on assignments from this class, I have to bring a laptop which seems quite fragile." (S11)

This finding is in line with Khairat et al. (2024)'s study, which discovered that some of the students' challenges in watching EL materials online such as *YouTube* are related to the content quality as well as the accessibility issues.

Distractions while accessing EL materials
Finally, some students mentioned being distracted by the material's use of multiple modes or by their surroundings. It is expected to lose focus when watching videos that integrate multiple modes at the same time. Moreover, it can be difficult to concentrate on EL materials when there are other distractions brought on by the surrounding noises. Stated on the open-ended questionnaires:

> "The numerous distractions around me sometimes made it difficult for me to focus when listening to material." (S3)

"Regarding the obstacles while doing extensive listening using multimodal, there was nothing that made it difficult. However, sometimes I got distracted or cannot focus on various modes at once. For example, when watching a video in the form of audio-visual, sometimes I only focused on listening, but I did not pay attention to the video. It also happened the other way around." (S10)

In conclusion, there are a lot of advantages to using multimodal materials in EL classes, but there is still a risk of distraction. These interruptions may result from background noises in the surrounding area or from the complexity of handling several modes at once. It is critical to acknowledge these difficulties because they draw attention to the need for strategies that support students in staying focused and attentive so they can get the most out of the multimodal learning environment.

CONCLUSION
The findings clearly highlight the importance of multimodality in improving students' experiences in EL materials. By combining text, visual, and audio modes, students' comprehension of materials is significantly enhanced. Non-verbal elements, such as gestures, lip movements, intonation, and facial expressions, also play a key role in helping learners grasp the context and emotions in videos. Allowing students to choose their own videos adds another layer of engagement, giving them the chance to connect with the material in a more personal and meaningful way.

An interesting and unexpected discovery is how multimodality supports students in exploring their preferred learning styles. It appears that multimodal approaches do more than improve understanding—they give students the flexibility to try out different methods, like visual, auditory, or kinesthetic learning, to find what works best for them. This kind of adaptability can lead to more effective and personalized learning experiences.

These insights have practical implications for teaching practices in EL materials. Educators should aim to include a wide range of multimodal resources to cater to diverse learning styles, ensuring all students—whether they learn best through seeing, hearing, or doing—feel supported. Multimodal strategies can also promote greater learner independence, allowing students to discover the approaches that suit them best.

To build on these findings, future research should take a closer look at how multimodality can accommodate different learning preferences.

Studies could explore which multimodal combinations are most effective for various learners and how teachers can incorporate these strategies into their classrooms. Additionally, examining both student and teacher perspectives could provide valuable insights into how multimodality impacts engagement, understanding, and long-term learning outcomes. By focusing on these areas, we can better design inclusive and effective learning environments in EL education.

REFERENCES
Alm, A. (2013). Extensive listening 2.0 with foreign language podcasts. *Innovation in Language Learning and Teaching, 7*(3), 266-280. http://dx.doi.org/10.1080/17501229.2013.836207.

Arrahman, 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/10.58355/qwt.v2i1.38.

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.

Chaichompoo, P. (2019). Impact of subtitles in audiovisual media on english learning. *Journal of Liberal Arts Maejo University, 7*(1), 200-213. https://so03.tcithaijo.org/index.php/liberalartsjournal/article/view/199358.

Chang, C-S., & Renandya, W. A. (2023). The multidimensional effects of extensive listening on EFL learners. *Journal of Research on English and Language Learning, 5*(1), 1–12. https://doi.org/10.33474/j-reall.v5i1.21190

Creswell, J.W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.

Dheghu, Y. P., Mata, R., & Rita, P. (2021). Improving students' listening skill using audio visual aids. *Advances in Social Science, Education and Humanities Research, 544.* https://doi.org/10.2991/assehr.k.210424.079

Efruan, E. N., Sukmawati, Yahrif, M., & Yusuf, D. (2023). The impact of Youtube English videos on students' listening comprehension skills at senior high school in Gowa. *Education, Language, and Culture (EDULEC), 3*(3). 377-391. https://doi.org/10.56314/edulec.v3i3

Fadhillah, J. N., Muniroh, S., & Rahmaningtyas, H. (2021). English department students' perceptions of using English talk show videos on YouTube to improve listening skill. *Journal of Language, Literature, and Arts, 1*(10), 1451–1465. http://doi.org/

10.17977/um064v1i102021p1451-1465

Fatimatuzzahro, S., Rodiyah, S., Rakhmawati, D. M., Wahdini, L., & Susandi, D. G. (2024). The use of English songs to improve students listening skill at SMP Negeri 242 Jakarta. *ULIL ALBAB: Jurnal Ilmiah Multidisiplin, 3*(6), 129-136. https://doi.org/10.56799/jim.v3i6.3465

Febriani, H., Asvio, N., Rahmadoni, J., & Vivekanantharasa, R. (2022). Improving EFL learners listening skills by using audio visual aids. *Linguists: Journal Of Linguistics and Language Teaching, 8*(2), 216-228. http://dx.doi.org/10.29300/ling.v8i2.7662

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.

Gonzalez-Torres, P. F., & Solano, L. M. (2024). Video listening journals and enhanced EFL listening skills. *Journal of University Teaching and Learning Practice, 21*(1). https://doi.org/10.53761/stbpjr54

Grapin, S. E. (2019). Multimodal teaching and learning. In Chapelle, C. A. (Ed.) *The concise encyclopedia of applied linguistics* (pp. 811-815). Wiley Blackwell.

Gunawan, M. H., Mandiri, I. C. P., Suherdi, D., Hanifa, E. S. D., & Gunawan, A. P. (2023a). Exploring EFL students' perceptions of the use of podcasts/TED Talks in extensive listening class. *Indonesian EFL Journal, 9*(1), 43-52.

https://doi.org/10.25134/ieflj.v9i1.7486

Gunawan, M. H., Hanifa, E. S. D., Gunawan, A. P., & Suherdi, D. (2023b). The use of Google sites in extensive listening classroom: Students' voices. *English Review: Journal of English Education, 11*(2), 365-378.

https://doi.org/10.25134/erjee.v11i2.8225

Gunawan, M. H., Damayanti, H., Hanifa, E. S. D., Gunawan, A. P., & Suherdi, D. (2023c). Enhancing EFL learning through podcast audio-assisted listening. *Indonesian EFL Journal, 9*(2), 109-122. https://doi.org/10.25134/ieflj.v9i2.8782.

Hanifa, E. S. D., Yusuf, F. N., & Gunawan, M. H. (2024). Students' voices on e-portfolio-based English listening assessment using Google Sites. *Journal of Applied Linguistics and Literacy, 8*(1), 102-131.

Hao, T., Sheng, H., Ardasheva, Y., & Wang, Z. (2021). Effects of dual subtitles on Chinese students' English listening comprehension and vocabulary learning. *Asia-Pacific Edu Res, 31*(5), 529–540. https://doi.org/10.1007/s40299-021-00601-w

Hui, C. W. J., & Jamaludin, K. A. (2024). The use of songs in English language teaching: A systematic literature review to identify advantages and appropriateness of age

categories. *International Journal of Academic Research in Progressive Education and Development, 13*(1), 1680-1693. http://dx.doi.org/10.6007/IJARPED/v13-i1/20729

Hussain, S., & Khan, H. K. (2022). The role of images in the teaching and learning of English: Practices, issues, and possibilities.Pakistan Languages and Humanities Review, 6(4), 338-348. https://doi.org/10.47205/plhr.2022(6-IV)31

Ivan, M. N., & Indah, R. N. (2021). Music and seventh graders' listening proficiency: Does Islamic song work?. *Journal of English Language and Education, 6*(2), 51-58. https://doi.org/10.31004/jele.v6i2.153.

Ivone, F. M., & Renandya, W. A. (2019). Extensive listening and viewing in ELT. *TEFLIN Journal, 30*(2), 237-256. https://doi.org/10.15639/teflinjournal.v30i2/237-256.

Jewitt, C. (2012). Multimodal teaching and learning. In Chapelle, C.A. (Ed.). *The encyclopaedia of applied linguistics*, 1-5. Blackwell Publishing.

Khairat, N., Muth'im, A., & Nasrullah. (2024). The use of YouTube to enhance learners' English listening skills. *Surakarta English and Literature Journal, 7*(2), 134-147. https://doi.org/10.52429/selju.v7i2.265

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No.131. *Medical Teacher, 42*(8), 846–854. https://doi.org/10.1080/0142159X.2020.1755030.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

Lee, H. C. (2014) Using an arts-integrated multimodal approach to promote English learning. *English Teaching: Practice and Critique, 13*(2), 55-75.

Lee, Y. J., & Cha, K. W. (2020). Writing listening logs and its effect on improving L2 students' metacognitive awareness and listening proficiency. *International Journal of Contents, 16*(4), 50-67. https://doi.org/10.5392/IJoC.2020.16.4.050

Lestary, N. L. G. W. (2019). The use of songs to improve students' listening comprehension ability. *Yavana Bhasha: Journal of English Language Education, 2*(2), 34-45. https://doi.org/10.25078/yb.v2i2.1024.

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. http://dx.doi.org/10.1080/1475939X.2015.1035318.

Mandiri, I. C. P., Gunawan, M. H., & Suharno. (2022). EFL student's difficulties in extensive listening activities: a thematic content analysis. *English Review: Journal of English Education, 10*(2),

393-400.
https://doi.org/10.25134/erjee.v10i2.6240.

Maqableh, M., & Alia, M. (2021). Evaluation online learning of undergraduate students under lockdown amidst COVID 19 pandemic: the online learning experience and students' satisfaction. *Children and Youth Services Review, 128.* https://doi.org/10.1016/j.childyouth.2021.106160.

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

Nadiyya, R., & Suryadi. (2024). The power of English music to improve students' listening skills. *Ideguru: Jurnal Karya Ilmiah Guru, 9*(2), 683-688. https://doi.org/10.51169/ideguru.v9i2.772.

Napikul, S., Cedar, P., & Roongrattanakool, D. (2018). The effects of film subtitles on English listening comprehension and vocabulary. *International Journal of Applied Linguistics & English Literature, 7*(6), 104-111. http://dx.doi.org/10.7575/aiac.ijalel.v.7n.6p.104

Oktalia, D., & Drajati, N. A. (2018). English teachers' perceptions of text to speech software and Google Site in an EFL Classroom: What English teachers really think and know. *International Journal of Education and Development using Information and Communication Technology (IJEDICT), 14*(3), 183-192.

Pamuji, K. D., Waring, R. & Kurniawan, E. (2019). EFL teachers' experiences in developing L2 proficiency through extensive listening. *TEFLIN Journal, 30*(2), 257-273. http://dx.doi.org/10.15639/teflinjournal.v30i2/257-273.

Polat, M. & Erişti, B. (2019). The effects of authentic video materials on foreign language listening skill development and listening anxiety at different levels of English proficiency. International *Journal of Contemporary Educational Research, 6*(1), 135-154. https://doi.org/10.33200/ijcer.567863.

Preiser, R., García, M. M., Hill, L., & Klein, L. (2021). Qualitative content analysis. In Biggs, R., de Vos, A., Preiser, R., Clements, H., Maciejewski, K., & Schlüter (Eds.). *The Routledge handbook of research methods for social-ecological systems* (pp. 270–281). Routledge.

Pribadi, R. J., Chozin, N. F., Aini, S. Q., Yulianti, I., & Tarihoran, N. (2024). The use of pictures and its effectiveness in English language teaching: A systematic review. Jurnal Pendidikan: SEROJA, 3(2), 1-13. Retrieved from https://www.jurnal.anfa.co.id/index.php/seroja/article/view/1856

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.

Quyen, B. T. T. (2020). English majors' perceptions of their teachers' use of ICT in teaching of Listening-Speaking courses at Ho Chi Minh City Open University. *Ho Chi Minh City Open University Journal of Science, 10*(2), 68-77. https://doi.org/10.46223/HCMCOUJS.soci.en.10.2.550.2020.

Setiani, A. A., Apandi, & Prasasti, R. (2024). Using English talk show videos on Youtube to enhance listening skill: A learners' voices. *Cirebon Annual Multidiciplinary International Conference (CAMIC).* Retrieved from https://ejournalugj.com/index.php/camic/article/view/9078

Setyawan, D., Ivone, F. M., & Suryati, N. (2022). Learners' preferences of multimedia resources in an extensive listening program. *LLT Journal: A Journal on Language and Language Learning, 25*(1), 23-37.

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.

Solanon, S. O., Idong, C. A. D., San Jose, A. E., & Concepcion, M. G. R. (2020). Heeding to English music and songs enhances listening skills. *International Journal of New Economics and Social Sciences, 2*(12). https://doi.org/0.5604/01.3001.0014.6887.

Widyaningrum, L., Rizal, D., & Prayogo, A. Multimodal project-based learning in listening and speaking activities: Building environmental care?. *Journal of English Education, Literature, and Culture, 7*(2). 209-223. http://dx.doi.org/10.30659/e.7.2.209-223

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.

Yıldırım, Ş. & Taşpınar, E. (2023). A case study investigating the impact of music listening on vocabulary development during foreign language learning. *Journal of Education and Recreation Patterns, 4*(2), 479-494. https://doi.org/10.53016/jerp.v4i2.176

Yin, R. K. (2018). *Case study research and its applications: Design and methods* (6th ed.). Sage Publications.