EDUCATIONAL : Jurnal Inovasi Pendidikan & Pengajaran
Vol. 5 No. 4, November 2025
E-ISSN : 2775-2593
P-ISSN : 2775-2585
Online Journal System : https://jurnalp4i.com/index.php/educational

EXPLORING THE USE OF YOUTUBE VIDEOS WITH TEXTS AND WITHOUT TEXTS IN TEACHING LISTENING TO BEGINNER EFL STUDENTS

Najwa Rizki Amalia[1], Siti Fatimah Purmasari[2], Diar Mukti Asri[3], Nisrina Farida Ramadhanti[4], Cahaya Ningrum[5], Abdul Syahid[6]
Universitas Islam Negeri Palangka Raya[1,2,3,4,5,6]
e-mail: najwara@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk mengeksplorasi penggunaan video YouTube dengan teks (subtitle) dan tanpa teks dalam pengajaran keterampilan menyimak (listening) bagi siswa EFL tingkat pemula. Latar belakang penelitian ini berangkat dari perkembangan teknologi digital yang membuka peluang baru dalam pembelajaran bahasa, salah satunya melalui media video yang memadukan unsur audio, visual, dan teks. Metode yang digunakan adalah quasi-eksperimen dengan desain pretest-posttest nonequivalent control group. Subjek penelitian terdiri atas 73 siswa kelas VIII di Berkah Integrated Junior High School yang dibagi menjadi dua kelompok: kelas eksperimen menggunakan video dengan teks dan kelas kontrol menggunakan video tanpa teks. Data diperoleh dari hasil pretest dan posttest yang dianalisis menggunakan statistik deskriptif dan uji-t independen dengan taraf signifikansi 0,05. Hasil penelitian menunjukkan bahwa tidak terdapat perbedaan signifikan antara kedua kelompok ($t = –1.720$; $p = 0.090$). Namun, secara deskriptif, kelompok tanpa teks menunjukkan peningkatan skor, sedangkan kelompok dengan teks mengalami sedikit penurunan. Temuan ini mengindikasikan bahwa penggunaan subtitle dapat menambah beban kognitif bagi siswa pemula karena mereka harus membagi perhatian antara membaca teks dan mendengarkan audio. Sebaliknya, video tanpa teks memungkinkan siswa fokus pada input pendengaran secara penuh. Dengan demikian, penggunaan subtitle sebaiknya disesuaikan dengan tingkat kemahiran siswa agar tidak menghambat proses pemahaman.
Kata Kunci: YouTube, subtitle, listening

ABSTRACT

This study aims to explore the use of YouTube videos, both with and without subtitles, in teaching listening skills to beginner-level EFL learners. The research is motivated by the evolution of digital technology, which provides new opportunities in language learning, particularly through video media that integrates audio, visual, and textual elements. A quasi-experimental method with a pretest-posttest nonequivalent control group design was employed. The subjects consisted of 73 8th-grade students at Berkah Integrated Junior High School, divided into two groups: an experimental class using videos with subtitles and a control class using videos without subtitles. Data were collected from the results of the pretest and posttest, analyzed using descriptive statistics and an independent t-test at a 0.05 significance level. The results indicated no statistically significant difference between the two groups ($t = –1.720$; $p = 0.090$). However, descriptively, the group without subtitles showed an increase in scores, while the group with subtitles experienced a slight decrease. This finding suggests that the use of subtitles might increase the cognitive load for beginner students as they have to divide their attention between reading the text and listening to the audio. Conversely, videos without subtitles allow students to fully focus on aural input. Therefore, the application of subtitles should be adjusted according to students' proficiency levels to avoid hindering the comprehension process.
Keywords: *YouTube, subtitle, listening comprehension, EFL learners, cognitive load*

EDUCATIONAL : Jurnal Inovasi Pendidikan & Pengajaran
Vol. 5 No. 4, November 2025
E-ISSN : 2775-2593
P-ISSN : 2775-2585
Online Journal System : https://jurnalp4i.com/index.php/educational

INTRODUCTION

The rapid and massive advancement of digital technology has significantly influenced the landscape of English as a *Foreign Language* (EFL) learning, offering educators and students new, dynamic avenues for enhancing listening skills. In the traditional classroom, listening was often confined to audio-only recordings that lacked visual context, but the digital era has introduced video-based platforms such as YouTube which have gained immense popularity. These platforms are particularly effective because they seamlessly combine audio, visual, and textual elements, enabling learners to engage with the target language in a manner that is far more authentic and contextualized than textbook dialogues. Through these multimodal features, students are not only exposed to the sounds of the language but can also hear natural pronunciation, observe non-verbal cues like gestures and facial expressions, and associate spoken language with its written forms. This technological shift necessitates a re-evaluation of teaching materials, moving towards resources that better mimic real-world communication and provide richer sensory input for language acquisition.

The pedagogical benefits of using video platforms have been substantiated by recent academic inquiries, which demonstrate that YouTube videos can substantially improve learners' comprehension and overall engagement in the learning process. Visualizing the speaker and the setting provides semantic clues that aid in decoding meaning, which is often lost in audio-only formats. For instance, a study by Audina et al. (2022) found that Indonesian EFL students exhibited significantly higher levels of motivation and attentiveness when authentic YouTube videos were integrated into their learning during the challenging period of the COVID-19 pandemic. Similarly, research conducted by Yuyun and Simamora (2021) highlighted that repeated exposure to natural speech patterns through YouTube videos provided critical support for the development of listening skills, particularly for beginner learners. These findings suggest that the engaging nature of video content helps to lower the affective filter, allowing students to absorb language input more effectively and with greater enthusiasm than traditional methods allow.

However, despite the acknowledged potential benefits of multimedia tools, the practical application of YouTube in EFL classrooms varies considerably among practitioners, creating a disparity in learning experiences. While some teachers actively incorporate videos that include textual support, such as captions, subtitles, or highlighted keywords, others prefer to use videos without any on-screen text to simulate natural listening environments. This discrepancy in pedagogical approach can significantly influence learners' ability to process language effectively, especially at lower proficiency levels. Beginners often struggle immensely to recognize new vocabulary, differentiate between similar phonemes, or understand speech delivered at a natural, native-speaker pace. These challenges may lead to reduced comprehension accuracy and a subsequent drop in learner confidence (Manurung et al., 2024; Telaumbanua et al., 2022). Therefore, understanding the impact of these different presentation modes is essential for optimizing instructional strategies.

The specific challenges faced by beginner learners underscore the need for scaffolding mechanisms within listening instruction. Without visual text, the stream of speech can appear as an unbreakable wall of sound, making it difficult for novices to parse individual words or phrases. This difficulty is often compounded by the anxiety associated with missing critical information during the listening process. Comparing the effectiveness of YouTube videos with and without textual cues is, therefore, crucial to determine how different modalities influence beginner learners' listening outcomes. If the gap between the learner's current proficiency and the difficulty of the input is too wide, frustration may occur; conversely, appropriate support can bridge this gap. The presence or absence of text acts as a variable that alters the cognitive

load imposed on the student, potentially freeing up mental resources for comprehension or, alternatively, splitting attention between reading and listening.

The role of textual support in multimedia learning is further emphasized and validated by various empirical studies and pedagogical theories regarding cognitive processing. Videos equipped with captions or subtitles are generally found to enhance comprehension, facilitate vocabulary recognition, and boost overall engagement to a greater degree than non-captioned videos. According to Marthafany (2023) and Qomariyah et al. (2021), the presence of text allows learners to verify what they believe they have heard, acting as a confirmatory feedback loop that reinforces learning. Additionally, textual support creates a form of cognitive scaffolding that helps students map the auditory input to the orthographic representation of words, which is particularly beneficial for languages with deep orthographies like English. This multimodal input ensures that if a student misses the auditory cue, the visual text serves as a backup, ensuring that the flow of comprehension is not entirely broken.

Nevertheless, the effectiveness of textual support is not uniform and can be influenced by a complex interplay of variables. Recent scholarship suggests that factors such as the speech rate of the video, the specific type of captioning used (keyword vs. full text), and the learners' existing proficiency level can all affect the magnitude of improvement. Research by Abbas et al. (2025), Almusharraf et al. (2024), and Li (2025) indicates that while captions generally help, their impact may vary; for example, advanced learners might find them distracting, whereas beginners find them indispensable. These findings suggest that integrating textual support is not merely about turning on subtitles but involves a strategic decision to provide cognitive scaffolding that facilitates listening comprehension. For beginner EFL students, who are still building their bottom-up processing skills, this support is likely critical in helping them segment the continuous stream of speech into meaningful units.

For beginner learners specifically, textual support in videos serves a dual purpose: it not only aids immediate comprehension but also significantly increases motivation and engagement with the learning material. Students tend to participate more actively and are more willing to repeat listening activities when captions and visual cues are available, as these features reduce the fear of failure. This supportive environment helps them internalize vocabulary and sentence structures more effectively (Marthafany, 2023). Moreover, the synergy of combining audio, visual, and textual modalities has been shown to improve overall comprehension and enhance both verbal and nonverbal communication skills (Al-Jumaily & Alazzawi, 2025). Building on this cumulative evidence, the present study aims to explore the use of YouTube videos with and without textual support in teaching listening to beginner EFL students, providing vital insights into how different modes of video presentation affect listening comprehension outcomes.

RESEARCH METHOD

This study applied a quasi-experimental design with a pretest-posttest nonequivalent control group design. The purpose of this study was to evaluate the impact of the use of YouTube videos along with text on the listening ability of EFL students at the beginner level. The subjects in this study are grade VIII students at Berkah Integrated Junior High School which totals 73 students. The experimental group consisted of 37 students in class A, while the control group consisted of 36 students in class B. The research was conducted in two different sessions but with identical procedures and materials, carried out on different days of the same week. Each session begins with a pre-test to measure students' initial abilities, followed by the provision of materials and treatments. Class A gets learning by using YouTube videos accompanied by subtitles, while class B watches the same video without text. After completing the learning and explanation activities, the two classes immediately carried out a post-test that

had the same level of difficulty to assess the improvement of listening skills after the treatment was given. The data generated from this study were obtained from pre-test and post-test scores, then analyzed quantitatively using descriptive statistics such as mean, median and standard deviation, as well as through independent t-tests with a significance level of 0.05 (5%) to determine whether there was a significant differences in the listening achievement between two classes.

RESULTS AND DISCUSSION

To give a full picture of the research results, this section shares the findings in a clear and organized way. It starts with basic summaries of the data, then looks at how the scores are spread out, and finally checks if the differences between the groups are meaningful using a statistical test. The summaries show the average scores and how much they vary for both groups, class VIII A, which used YouTube videos with text, and class VIII B, which used videos without text—helping to show how well the students performed before and after the treatment. Next, graphs are used to show how the scores are distributed, making it easier to see how listening skills changed after the intervention. Lastly, a statistical test called the independent samples t-test is used to see if the difference in scores between the two groups is significant. All these steps together give a clear and complete understanding of how using YouTube videos with and without text affected the listening abilities of English as a foreign language students.

Results

1. Descriptive Statistics

The results of the pre-test and post-test for classes VIII A and VIII B are shown in Table 1.

Table 1. Descriptive Statistics of Pre-Test and Post-Test Scores for Class VIII A and VIII B

	Pre-Test		Post-Test		Selisih	
	VIII A	VIII B	VIII A	VIII B	VIII A	VIII B
Valid	37	31	37	31	37	36
Missing	0	5	0	5	0	0
Median	5.000	5.000	4.000	5.000	0.000	0.000
Mean	5.216	5.226	4.514	5.355	0.838	-0.111
Std. Deviation	1.566	1.359	2.077	1.924	2.386	1.864

According to Table 1, the average pre-test scores for class VIII A (M = 5.216) and VIII B (M = 5.226) show nearly similar initial abilities. After the treatment, the average post-test score for VIII A slightly decreased, while VIII B showed an increase. This suggests there is a difference in learning outcomes between the two groups.

2. Independent Samples T-Test

To examine whether the difference in post-test scores between the two classes was statistically significant, an independent samples *t*-test was conducted.

Independent Samples T-Test

Table 2. Independent Samples t-Test for Post-Test Scores of Class VIII A and VIII B

					95% CI for Cohen's d	
T		**df**	**p**	**Cohen's d**	**SE Cohen's d**	**Lower**	**Upper**
Post-Test	**-1.720**	**66**	**0.090**	**-0.419**	**0.248**	**-0.900**	**0.065**

Based on the results of the inferential statistical analysis listed in Table 2, hypothesis testing using the Independent Samples t-Test shows that there is no statistically significant

EDUCATIONAL : Jurnal Inovasi Pendidikan & Pengajaran
Vol. 5 No. 4, November 2025
E-ISSN : 2775-2593
P-ISSN : 2775-2585
Online Journal System : https://jurnalp4i.com/index.php/educational

difference between the post-test scores of class VIII A and class VIII B. This is evidenced by the probability or significance value (p) of 0.090, where this figure is greater than the standard significance level of 0.05 which is the reference for rejecting the null hypothesis. With a calculated t value of -1.720 and degrees of freedom (df) of 66, this result indicates that although there is a difference in the average scores between the two groups, the difference is not strong enough to be considered significant in generalizing the population. In addition, the calculation of the effect size using Cohen's d produces a value of -0.419 with a 95% confidence interval between -0.900 and 0.065. The negative number on Cohen's d and the t value confirms the direction of the difference in scores where the second group has a higher average, but statistically this variation is most likely due to chance factors alone

Descriptives

Table 3. Descriptive Statistics for Post-test Scores for Class VIII A and Class VIII B

Group Descriptives

	Group	N	Mean	SD	SE	Coefficient of variation
Post-Test	VIII A	37	4.514	2.077	0.341	0.460
	Group	N	Mean	SD	SE	Coefficient of variation
	VIII B	31	5.355	1.924	0.346	0.359

The descriptive data presented in detail in Table 3 provides a comprehensive overview of the distribution of post-test scores from both sample groups, namely class VIII A and VIII B. Specifically, class VIII B consisting of 31 students showed a superior average performance with a mean value of 5.355, compared to class VIII A consisting of 37 students with an average score of only 4.514. In addition to the difference in average values, the distribution of the data also shows different variability characteristics, where class VIII A has a standard deviation of 2.077 and a coefficient of variation of 0.460, which is higher than class VIII B with a standard deviation of 1.924 and a coefficient of variation of 0.359. This indicates that the values of students in class VIII A tend to be more varied or heterogeneous with a wider distribution range than the average, while the abilities of students in class VIII B appear more homogeneous or uniform in their achievement of learning outcomes on the final test.

Discussion

Relation to Multimedia Learning Theory and Cognitive Load

The findings of this study indicate that the mean post-test score of the experimental group (Class A, using YouTube videos with text) decreased from 5.216 to 4.514, while that of the control group (Class B, using videos without text) increased from 5.226 to 5.355. Although the difference between groups was not statistically significant ($t = -1.720$; $p = 0.090$), the trend observed is noteworthy and warrants analysis through the lens of multimedia learning theory, particularly the Cognitive Theory of Multimedia Learning (CTML) by Richard E. Mayer and the Cognitive Load Theory (CLT).

According to CTML, when creating multimedia learning that combines words and pictures, it should be designed to help learners understand better by using three key thinking processes: choosing, arranging, and combining information in their short-term memory before it moves to long-term memory. In the case of English as a Foreign Language (EFL) listening, videos with subtitles can be seen as a type of media that uses two types of input: listening through sound and understanding through text. CLT says that our short-term memory has a limit, so it's important to consider cognitive load, which includes unnecessary work, the natural difficulty of the material, and the effort needed to truly understand something.

In the present study, the lower scores in Class A might indicate that subtitles added

extra or even essential mental work for beginner English as a Foreign Language students. When subtitles are used, learners have to listen, read, and link written words with spoken words, which can create a split-attention effect. This happens when learners' attention is pulled between text, sound, and images. According to multimedia theory, using too many different types of information at once can be bad for learning if not handled well, as it takes up working memory with unnecessary tasks. While captioned videos can help with understanding in some situations, this has mostly been seen with more advanced learners, like university students learning Mandarin (Guo et al., 2020).

On the other hand, the group without subtitles (Class B) showed some improvement, even if it wasn't significant. This suggests that for beginner EFL learners, focusing only on the audio might reduce extra mental effort, allowing them to concentrate more on listening and understanding without being distracted by text. While text and audio together could theoretically help learning through dual coding, for beginners, it might actually make their thinking process more difficult and slow down how they organize and use the information (Wu et al., 2022).

Analysis of Attention, Motivation, and Proficiency Factors

When it comes to attention, students in Class A had to split their focus between listening to the audio and reading the subtitles. For beginners with low language skills, this can lead to quicker mental tiredness or cause them to focus more on one task while neglecting the other—like reading but not paying attention to what is being said. Subtitles don't always help low-proficiency learners automatically because they might depend more on reading than on actively listening (Li, 2025).

From a motivational perspective, subtitles might make students feel more confident because they can follow along with the text. But when the combination of text and sound becomes too much to handle mentally, it can reduce their motivation since the learning process might feel too difficult. On the other hand, videos without subtitles (Class B) let students focus entirely on the audio, making the task a suitable challenge and helping them feel proud of understanding spoken English without any text to guide them. Research shows that media that is both enjoyable and not too hard can boost motivation in English as a foreign language classrooms (Karim et al., 2023).

Language proficiency also plays a big role in how these results are interpreted. This study included eighth-grade students who are generally at a beginner level in English as a foreign language. Their ability to manage both audio and text at the same time might not be very strong yet. Subtitles are usually more helpful for learners with intermediate or advanced skills, while for beginners, they can be less helpful or even not useful at all (Pujadas & Muñoz, 2020 in Wu et al., 2022). Therefore, the lower performance in Class A could be linked to the lower language skills of the students, making the combination of audio and text less effective for them.

Relation to Previous Studies

The results of this study are in line with past research indicating that captions or subtitles don't always greatly improve listening comprehension, particularly for learners with lower proficiency. Studies have found that captioned videos offer only small benefits for beginners (Guo et al., 2020). However, other research suggests that English-subtitled videos can help EFL learners, although these studies mostly looked at mixed-language abilities instead of focusing specifically on beginner listening skills (Oktapiani, 2024). These differences can be explained by factors such as the learners' proficiency level, the type of subtitles used (full or partial), and how the lessons are structured.

This study also adds to the existing research on using videos in EFL listening classes.

Some studies have shown that captions on YouTube videos during online lessons are connected to better listening performance, but their effectiveness depends on how they are used and the learners' proficiency (Boltiziar & Munková, 2024). These varying results show that when using subtitles in teaching, it's important to consider the learners' readiness and how they process information. The results suggest that for beginning English as a Foreign Language (EFL) students in junior high school, using videos with subtitles does not always lead to better listening skills. When students have limited language ability, subtitles can add extra mental effort and make it harder to understand the content. As a result, listening lessons should manage cognitive effort carefully by focusing first on hearing the language, and later adding text as students' skills improve. Support structures should be provided to help students process information more effectively, in line with Communicative Language Teaching (CLT) and Cognitive Theory of Multimedia Learning (CTML). These findings highlight the importance of matching instructional materials to students' language levels and mental capacity.

CONCLUSION

This study concludes that the use of YouTube videos with subtitles in listening instruction for eighth-grade beginner EFL students actually has a counterproductive effect, indicated by a decrease in the average post-test score in the experimental group, in contrast to an increase in the control group without subtitles. This phenomenon can be comprehensively explained through the Cognitive Theory of Multimedia Learning and Cognitive Load Theory, which highlight the limitations of beginner students' working memory capacity. The presence of subtitles creates a cognitive overload and triggers a split-attention effect, where students' attention is inefficiently divided between decoding visual text and simultaneously processing auditory input. For low-proficiency students, the text element does not function as a helpful scaffolding but rather becomes a distraction that hinders the process of selecting and organizing information in short-term memory. Conversely, the removal of subtitles has been shown to reduce extraneous mental load, allowing students to focus their cognitive resources entirely on the audio aspect, which theoretically aligns more with their mental readiness to process foreign language information.

Further analysis implies that the effectiveness of learning media is highly dependent on students' language proficiency and motivation, where overly complex materials can lower the confidence of beginning learners. These findings emphasize the need for educators to design materials that prioritize audio processing before integrating text, in line with the principles of Communicative Language Teaching. For future research, it is recommended that researchers expand the scope of the study by comparing the effectiveness of different types of subtitles, such as full captions versus keyword captions, to identify which format is most optimal in balancing visual aid and cognitive load. Furthermore, longer-term longitudinal studies are highly recommended to test whether the negative effects of subtitles on beginners are permanent or will change as students adapt. Investigations into the influence of individual learning styles, such as visual or auditory preferences, on the effectiveness of using captioned videos are also needed to provide more specific insights for developing adaptive and effective EFL curricula.

REFERENCES

Abbas, N. F., et al. (2025). YouTube as a learning tool among EFL learners: A systematic review. *Arab World English Journal (AWEJ) Special Issue on CALL*, *11*, 96–110. https://doi.org/10.24093/awej/call11.6

Al-Jumaily, I. H. A., & Alazzawi, I. T. J. (2025). The influence of multimodal visual

methodologies on EFL university students' audio-visual comprehension, verbal and nonverbal communication. *Journal of Language and Education*, *11*(2), 35–56. https://doi.org/10.17323/jle.2025.22731

Almusharraf, A., et al. (2024). Video captioning and subtitles in second language listening comprehension: Fast-paced versus slow-paced speakers. *Journal of Psycholinguistic Research*, *53*(2), 29. https://doi.org/10.1007/s10936-024-10070-z

Audina, I. P., et al. (2022). YouTube videos for listening skills during COVID-19 pandemic: EFL students' views. *Jurnal Penelitian dan Pengembangan Sains dan Humaniora*, *6*(3), 327–334. https://doi.org/10.23887/jppsh.v6i3.55809

Boltiziar, J., & Munkova, D. (2024). Emergency remote teaching of listening comprehension using YouTube videos with captions. *Education and Information Technologies*, *29*, 11367–11383. https://doi.org/10.1007/s10639-023-12282-7

Guo, S., et al. (2020). Effects of captioned videos on learners' comprehension. *Journal of Global Literacies, Technologies, and Emerging Pedagogies*, *6*(1), 1062–1082.

Hochstrasser, K., & Stoddard, H. A. (2022). Use of cognitive load theory to deploy instructional technology for undergraduate medical education: A scoping review. *Medical Science Educator*, *32*, 553–559. https://doi.org/10.1007/s40670-021-01499-1

Karim, S. A., et al. (2023). Exploring EFL students' views on using YouTube videos for improving listening skills in English. *English Review: Journal of English Education*, *11*(3), 757–766. https://doi.org/10.25134/erjee.v11i3.7763

Li, Y. (2025). Listen or read? The impact of proficiency and visual complexity on learners' reliance on captions. *Behavioral Sciences*, *15*(4), 542. https://doi.org/10.3390/bs15040542

Manurung, I. D., et al. (2024). The effect of using video subtitles on EFL students' listening comprehension. *Holistic Science*, *4*(2), 317–327. https://doi.org/10.56495/hs.v4i2.664

Marthafany, I. D. M. (2023). Using video-captioning-based activities to develop EFL secondary school learners' vocabulary. *Journal of English Teaching and Learning*, *3*(1), 12–19. https://doi.org/10.62734/jetling.v3i1.213

Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. *Educational Psychology Review*, *36*(8). https://doi.org/10.1007/s10648-023-09842-1

Oktapiani, I. U., et al. (2024). The effectiveness of English-subtitled videos in enhancing speaking, listening, and vocabulary skills: A meta-analysis. *Journal of Languages and Language Teaching*, *12*(1), 1–15. https://doi.org/10.33394/jollt.v12i1.9501

Qomariyah, S. S. A., et al. (2021). The effect of YouTube video on students' listening comprehension performance. *Jo-ELT (Journal of English Language Teaching)*, *8*(1), 67–73. https://doi.org/10.33394/jo-elt.v8i1.3837

Telaumbanua, Y., et al. (2022). "YouTubeGo": Developing EFL learners' listening skills and social learning process. *International Journal of English and Applied Linguistics (IJEAL)*, *2*(1), 1–12. https://doi.org/10.47709/ijeal.v2i1.1329

Wu, H., et al. (2022). Video captioning effects on EFL listening comprehension and vocabulary learning. *International Journal of Computer-Assisted Language Learning and Teaching*, *12*(2), 1–16. https://doi.org/10.4018/IJCALLT.291534

Yuyun, I., & Simamora, F. Y. (2021). The use of YouTube to support EFL students' listening skills. *English Linguistics and Language Teaching Research Journal*, *2*(2), 1–12. https://doi.org/10.22236/ellter.v2i2.7512