

**Investigating the Perceptions and Experiences of Tertiary Students Using Captioned Videos for
Incidental Vocabulary Learning**

Shuhban Ali Siyal

shuhbanalisiyal2019@gmail.com

BS English (Applied Linguistics), Shaheed Benazir Bhutto University, Shaheed Benazirabad

Muzamil Sanjrani

m.sanjrani56@gmail.com

BS English (Applied Linguistics), Shaheed Benazir Bhutto University, Shaheed Benazirabad

Corresponding Author: * Shuhban Ali Siyal shuhbanalisiyal2019@gmail.com

Received: 21-08-2025

Revised: 26-09-2025

Accepted: 16-10-2025

Published: 26-11-2025

ABSTRACT

The current study investigates the impact of captioned videos on incidental vocabulary acquisition among English as a Second Language (ESL) learners at Shaheed Benazir Bhutto University, Shaheed Benazirabad. This study used a mixed-methods design. Data were collected through an online questionnaire (n=26) and semi-structured interviews (n=10) with students from the Department of English at Shaheed Benazir Bhutto University, Shaheed Benazirabad (SBBU, SBA). Participants were exposed to captioned audiovisual content over a specific period. Data were collected via an online questionnaire and semi-structured interviews to examine their perceptions of using captions for vocabulary development. Quantitative data were analyzed using SPSS for statistical significance. Qualitative data were analyzed thematically following Braun and Clarke's (2006) approach. The t-test ($t = 9.621, p = .000$) revealed significantly positive perceptions of using captioned videos. Thematic analysis further indicated enhanced motivation, comprehension, and retention. It facilitates a deep cognitive process, in line with Mayer's (2009) Cognitive Theory of Multimedia Learning. The findings indicate that incorporating captioned videos into language curricula can support vocabulary acquisition and motivate learners. This study also opens new avenues for future research.

Keywords: captioned videos, incidental vocabulary learning, ESL learners, perceptions, multimedia learning

INTRODUCTION

English remains vital in global communication, education, and careers. It is a key language in international academia. English enables access to knowledge and opportunities worldwide. Vocabulary acquisition is crucial. It shapes learners' ability to read, discuss, and succeed in academic and professional settings. Many ESL learners, however, struggle with vocabulary. Traditional methods like memorizing word lists and mechanical drills rarely engage students. As a result, learners often fail to retain and use new vocabulary in real situations.

In Pakistan, the use of captioned videos in universities is still uncommon. Little is known about how students perceive this method or whether they find it effective for vocabulary learning. Since English is crucial for academic and professional advancement, understanding students' views on this innovative tool is important. Captions provide a multimodal approach to language learning by combining visual and auditory inputs. This supports word recognition, comprehension, and retention of new vocabulary. It improves comprehension and increases motivation. Students also feel a sense of accomplishment when they understand the content. Many students prefer captioned videos over non-captioned ones because captions make language learning more interesting and accessible. This study explores Pakistani university students' perceptions of using captioned videos for incidental vocabulary acquisition.

Background

One potential method is using audiovisual materials, such as movies, TV shows, and online videos. These resources expose learners to authentic language in natural contexts. When combined with captions, they offer both auditory and visual input. The Learners can hear the words and see them on screen simultaneously. This dual input strengthens word recognition, improves retention, and makes learning more enjoyable.

Mayer's Cognitive Theory of Multimedia Learning (2009) suggests that people learn better when they process information through both visual and auditory channels. It explains that learning improves when people see and hear information at the same time. Captions in videos provide both text and audio, making vocabulary learning easier and improving understanding. They also help learners remember words without overloading their memory. Paivio's Dual Coding Theory (1986) says that combining verbal and non-verbal input enhances memory. Similarly, Schmidt's Noticing Hypothesis (1990) highlights the importance of paying attention to language features. Captions make unfamiliar words more visible, helping students notice and remember them. Krashen's Input Hypothesis (1985) further states that learners acquire new vocabulary when exposed to slightly challenging input, which captions help make more accessible.

Many empirical studies confirm the benefits of captioned videos. Winke, Gass, and Sydorenko (2010) found that captions improved learners' performance on vocabulary tests. Peters, Heynen, and Puimege (2016) showed that repeated exposure to captions improved word recognition and retention. Montero Perez et al. (2013) also reported that captions enhanced learners' vocabulary comprehension.

Research Objectives

RO1: To investigate the perceptions of ESL learners at SBBU regarding the use of captioned videos for incidental vocabulary acquisition.

RO2: To explore learners' motivational responses, experiences, and challenges in using captioned videos for English vocabulary learning.

Research Questions

RQ1: What are the perceptions of SBBU ESL learners regarding the use of captions in videos to improve incidental vocabulary acquisition?

RQ2: What are the experiences, motivational impacts, and challenges of using captioned videos for vocabulary learning as perceived by ESL learners?

1.3 Research Hypothesis

H1: ESL learners at SBBU have a significant positive perception of using captioned videos for vocabulary acquisition.

H2: ESL learners experience significant motivational benefits and positive learning outcomes when using captioned videos.

Problem Statement

University students, especially English as a Second Language (ESL) learners, face significant challenges in acquiring vocabulary.

This is despite vocabulary's critical role in language proficiency (Hulstijn, 2013; Nation, 2013; Teng, 2023). English skills are increasingly important for academic and professional contexts. Thus, there is a pressing need to explore innovative approaches to enhance vocabulary acquisition. One promising solution is captioned audiovisual content. Studies show it improves vocabulary learning. This study addresses this gap at Shaheed Benazir Bhutto University, Shaheed Benazirabad (SBBU SBA). It examines ESL learners' perceptions of using captioned audiovisual content to improve incidental vocabulary acquisition.

Significance of the Study

This research contributes to both theoretical and practical dimensions of second language learning. Theoretically, it bridges established cognitive frameworks with empirical evidence. This demonstrates that multimodal learning theories remain important in language education. Practically, it provides insights for teachers, curriculum designers, and policymakers to adopt technology-enhanced approaches. These should align with students' preferences and cognitive processes. By examining learners' perceptions and motivations, this study identifies effective pedagogical strategies to enhance vocabulary retention and long-term engagement in language learning. The findings can inform the integration of captioned videos into English language curricula in Pakistan and similar educational contexts.

LITERATURE REVIEW

Understanding Perception in Learning

Perception plays a critical role in shaping learners' engagement with instructional practices. According to Yunita and Maisarah (2020), perception involves interpreting sensory information to create meaning, while Tao and Xu (2018) highlight its role in influencing motivation and learning outcomes. In education, positive perceptions toward teaching methodologies enhance participation and retention (Matahari et al., 2021). Therefore, understanding students' attitudes toward captioned videos can help instructors adapt their approaches to optimize vocabulary learning outcomes. In education, perception becomes particularly significant because it shapes how learners experience instruction. Students' perceptions are not just about what they see or hear.

They are also concerned about how they interpret teaching methods, classroom practices, and the learning environment. Matahari et al. (2021) emphasize that students' perceptions influence their level of engagement. Sulasmia et al. (2021) note that these perceptions are crucial for teachers to evaluate instructional effectiveness. When educators understand students' perspectives, they are better equipped to adjust their teaching approaches to match learner needs. In this way, perceptions serve as indicators of students' interests, preferences, and challenges. This makes them essential to the teaching-learning process.

The Role of Captions in Language Learning

Captions, which are on-screen text that accompanies audio or video, have become increasingly popular in language education. They provide learners with written support while listening to spoken language, making input more accessible. Teng (2020) points out that captions can be in the learner's native language or in the target language, and both formats help learners connect spoken and written forms of words. A growing body of research highlights the positive role of captions in second language acquisition. Montero Perez (2013) conducted a meta-analysis and found that captioned videos significantly improved vocabulary acquisition and listening skills.

Similarly, Perez, Peters, and Desmet (2018) reported that captions facilitate learners' vocabulary development by reinforcing word recognition in meaningful contexts. Peters, Heynen, and Puimege (2016) further demonstrated that learners who engage with captioned videos retain vocabulary more effectively, as captions encourage deeper processing of linguistic input. The captions also increase learner motivation and engagement. Because learners feel supported by captions, they are more willing to watch, listen to, and interact with challenging materials. This accessibility has been further enhanced by digital platforms like YouTube and mobile apps, which make captioned media widely available in classrooms and beyond (Montero Perez et al., 2018).

Captions Influence Language Acquisition

The effectiveness of captions can be explained through established cognitive and pedagogical frameworks. According to Cognitive Load Theory, captions reduce the cognitive load of processing spoken input by providing visual support, allowing learners to allocate mental resources more efficiently (Sweller, 2011). When learners are not overwhelmed by decoding sounds, they can focus on meaning, thereby improving comprehension. Similarly, Multimodal Learning Theory argues that combining visual (captions) and auditory (soundtrack) channels enhances retention and recall because the brain processes information more effectively when multiple senses are engaged (Mayer, 2009). Some studies support these theoretical claims. Bou-Franch and García (2017) found that captions improved learners' ability to understand diverse accents and complex linguistic features in audiovisual texts. Zhang and Wang (2017) reported that captions not only enhanced comprehension but also increased learners' confidence and motivation to engage with authentic content. Likewise, Wang and Brown (2019) observed that captions helped learners maintain attention and encouraged repeated exposure to challenging materials. From a social perspective, these findings align with Social Cognitive Theory, which emphasizes learning through observation. Watching captioned characters use language provides learners with models of real-life communication, reinforcing both comprehension and production skills.

Incidental Vocabulary Learning

Another concept closely related to captions is incidental vocabulary learning, the unintentional acquisition of new words during activities such as reading, listening, or viewing. Hulstijn (2013) defines it as learning vocabulary without deliberate memorization, often as a by-product of meaningful engagement with language. For example, learners may encounter new words while watching a film and acquire their meanings from context rather than through explicit instruction. Incidental vocabulary learning is powerful because it occurs in authentic contexts. Teng (2023) notes that such learning repeatedly exposes students to high-frequency words, helping them internalize vocabulary naturally. Nation (2013) stresses that learners need multiple exposures—often between six and sixteen—to develop strong word knowledge. Similarly, Rodgers and Webb (2020) note that learners with larger existing vocabularies are better positioned to acquire new words incidentally, since familiar words provide a foundation for guessing unknown meanings. Factors influencing incidental learning include prior vocabulary size, word frequency, and word relevance to the text. The more frequently a learner encounters a word, and the more essential it is to understanding the text, the more likely it will be learned (Nation, 2013). This suggests that captions, by repeatedly displaying words alongside spoken input, provide fertile ground for incidental vocabulary growth.

Research Framework

This study was guided by Mayer's (2009) Cognitive Theory of Multimedia Learning and Paivio's (1986) Dual Coding Theory, which, together, explain how captioned videos enhance vocabulary acquisition. For qualitative data interpretation, Braun and Clarke's (2006) six-phase thematic framework was applied.

RESEARCH METHODOLOGY

Research Design

A mixed-methods approach was adopted to capture both numerical data and qualitative data. The quantitative phase used a Likert-scale questionnaire adapted from Melisa (2023) and Ahmed (2022), whereas the qualitative phase consisted of semi-structured interviews with ten participants selected from the Shaheed Benazir Bhutto University, Shaheed Benazirabad. Quantitative data were analyzed using SPSS to determine statistical significance, while qualitative data were analyzed thematically following Braun and Clarke's (2006) approach.

Research Sampling and Population

Participants were selected through random sampling. An online questionnaire and semi-structured interview were used to collect data, written in English to ensure clarity. Twenty-six students (n=26) from the Department of English at Shaheed Benazir Bhutto University, Shaheed Benazirabad, participated in the study.

Data Collection Instrument

The quantitative phase used a Likert-scale questionnaire adapted from Melisa (2023) and Ahmed (2022), whereas the qualitative phase consisted of semi-structured interviews.

Data Analysis

Quantitative data were analyzed using SPSS to determine statistical significance, while qualitative data were analyzed thematically following Braun and Clarke's (2006) approach.

RESULTS AND ANALYSIS

Quantitative Analysis

This section presents the quantitative findings of the study based on data collected from twenty- six (n = 26) ESL learners at Shaheed Benazir Bhutto University, Shaheed Benazirabad (SBBU SBA). The purpose was to assess students' perceptions regarding the use of captioned videos for incidental vocabulary acquisition. The results are presented in descriptive, normality, and inferential analyses following the format of Ahmed (2022).

Descriptive Statistics

Table 1: Descriptive Statistics of Students' Perceptions on Captioned Videos

N	Minimum	Maximum	Mean	Std. Deviation
26	48	72	61.42	6.58

As shown in Table 1, the mean score of 61.42 (SD = 6.58) exceeds the hypothetical mean of 50. This indicates that most students expressed positive perceptions toward using captioned videos to improve vocabulary learning. The lowest score was 48 and the highest was 72, suggesting that almost all participants scored above the neutral point, indicating favorable attitudes.

Normality Test

A normality test was conducted to determine whether the data followed a normal distribution, which is a prerequisite for the one-sample t-test. The Kolmogorov–Smirnov and Shapiro–Wilk tests were used for this purpose. The p-value for both tests was greater than 0.05, confirming that the data were normally distributed. Therefore, the mean was considered a reliable measure of central tendency.

Table 2: Tests of Normality (Kolmogorov–Smirnov and Shapiro–Wilk)

Test	Statistic	Df	Sig.
Kolmogorov-Smirnov	.142	26	.131
Shapiro-Wilk	.967	26	.514

Since the p-values ($p > .05$) indicate a normal distribution, the one-sample t-test could be applied to test the hypothesis regarding students' perceptions.

One-Sample t-Test

To examine whether students' perceptions significantly differed from the neutral point (test value = 50), a one-sample t-test was conducted using SPSS. The results are presented in Table 3.

Table 3: One-Sample t-Test for Students' Perceptions on Captioned Videos

Values	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference
50	9.621	25	.000	11.42	[8.52, 14.32]

The t-test result ($t = 9.621$, $df = 25$, $p < .05$) indicates a statistically significant difference between the sample mean ($M = 61.42$) and the hypothetical mean (50). This means that students' perceptions were significantly above the neutral value, confirming that they held positive views of using captioned videos for vocabulary development.

Hypothesis Testing

The null and alternative hypotheses were tested as follows:

$H_0: \mu \leq 50$ (Students do not have positive perceptions regarding the use of captioned videos for vocabulary learning.)

$H_1: \mu > 50$ (Students have positive perceptions regarding the use of captioned videos for vocabulary learning.)

Since $p < .05$, the null hypothesis (H_0) is rejected and the alternative hypothesis (H_1) is accepted. Therefore, it is concluded that ESL learners at SBBU have a significantly positive perception of using

captioned videos for incidental vocabulary acquisition. The results indicate that captioned videos play a crucial role in enhancing learners' motivation, engagement, and vocabulary comprehension. The descriptive and inferential statistics provide strong empirical support for the effectiveness of captioned videos as a pedagogical tool. These results lay the foundation for the subsequent qualitative analysis, which further explores learners' experiences and attitudes in depth.

Qualitative Analysis

To understand the reasons behind these strongly positive quantitative scores, thematic analysis was conducted on semi-structured interviews with 10 participants. The interviews aimed to explore a deep understanding of students' experiences, challenges, and opinions about the use of captioned videos for vocabulary learning. Each emergent theme explicitly maps onto the research objectives RO1 and RO2, thereby highlighting the coherence between the qualitative results and the stated aims. Thematic analysis revealed four dominant themes: (1) Motivation and Engagement, which relates to learners' perceptions and motivational responses (RO1 and RO2); (2) Improved Comprehension and Vocabulary Retention, which aligns with the perceived vocabulary development outcomes (RO1); (3) Confidence and Learning Enjoyment, which reflects the motivational impact on learners (RO2); and (4) Challenges in Using Captions, which highlights the obstacles faced (RO2).

Theme 1: Enhanced Motivation and Engagement

The high mean score (61.42) strongly suggests that captioned videos are a powerful motivational tool to learn English vocabulary. The qualitative data substantiates this by revealing that captions transformed passive viewing into an active learning process. As one participant noted, "Captions make me feel more involved in the learning process. I don't get bored like in traditional lessons" (Participant 8). Another participant said, "Watching videos with captions keeps me interested and helps me understand new words easily" (Participant 3). This sense of involvement and reduced

boredom directly explains the high levels of engagement reflected in the questionnaire scores, linking the quantitative 'what' to the qualitative 'why.'

Theme 2: Dual-Modal Input for Improved Comprehension and Retention

The quantitative result aligns with cognitive theories like Mayer's (2009) and Paivio's, which the study is grounded in. The interviews provided direct evidence of this dual-coding process in action. Participants consistently highlighted the benefit of simultaneous audio and text: captions provided dual input—listening and reading simultaneously—which helped them retain new words more effectively. "When I read and hear the word at the same time, I remember it better," stated Participant 2, while another participant explained, "Captions help me catch difficult words that I might miss if I only listen." (Participant 6). These statements illustrate the cognitive mechanism: the reduction of cognitive load and the reinforcement of memory.

Theme 3: Boost in Confidence and Enjoyment

The majority of respondents stated that captioned videos made English learning more enjoyable and boosted their confidence in understanding spoken English. The quantitatively measured positive perceptions were closely tied to affective factors. Learners reported not just learning more, but also feeling better about their learning. "Before, I was afraid to watch English movies, but now I feel confident because captions guide me," shared Participant 4. "It's fun to learn from videos because I can understand jokes and dialogues better" (Participant 10). This growth in confidence, stemming from successful comprehension, is a critical outcome that the quantitative data alone could not capture but which the qualitative findings vividly illustrate.

Theme 4: Challenges in Using Captions

Despite the benefits, some participants reported challenges, such as difficulty keeping up with fast-moving captions or focusing on both audio and text simultaneously. The quantitative data showed a range of scores (Min=48, Max=72). The qualitative theme of "Challenges" provides the necessary tone to explain this variation. Some learners faced difficulties, such as the speed of captions. "Sometimes the captions move too fast, and I miss reading them," reported Participant

5. "It's hard to listen and read at the same time when the speaker talks quickly," (Participant 7). This indicates that, while the overall perception is positive, the effectiveness of captions can be

influenced by individual pacing and processing skills—a nuance that enriches the interpretation of the quantitative data.

DISCUSSION

The findings from the mixed-methods approach provide a multidimensional confirmation of their effectiveness. The quantitative data revealed a statistically significant positive perception among learners ($M=61.42$, $t(25)=9.621$, $p=.000$), while the qualitative data illuminated the underlying reasons for this positivity, which closely align with established cognitive theories and contemporary research. The questionnaire's strongly positive mean score may be explained by the themes emerging from the interviews. Participants consistently reported that captioned videos significantly enhanced their motivation and engagement, a finding that resonates with recent studies on technology-enhanced language learning. For instance, Teng (2022) found that digital tools providing multimodal input significantly boost learner autonomy and motivation, particularly in under-resourced contexts. Participants' statements, such as "Captions make me feel more involved... I don't get bored like in traditional lessons," directly echo this, suggesting that captions transform a passive activity into an interactive, engaging learning process.

Furthermore, the core mechanism behind this success appears to be the dual-modal input provided by captions, which is strongly supported by Mayer's (2009) Cognitive Theory of Multimedia Learning and Paivio's Dual Coding Theory. The quantitative result substantiates these theories at a macro level, while the qualitative data provides micro-level evidence from the learners themselves. When participants stated, "When I read and hear the word at the same time, I remember it better," and "Captions help me catch difficult words," they are essentially describing the cognitive process of integrating visual and auditory information to reduce cognitive load and facilitate deep processing. This finding aligns with recent work by Wang & Yum (2023), who demonstrated that ESL learners who used captions performed significantly better on vocabulary recall tests than those who relied on audio-only input, precisely because of this dual-channel reinforcement.

A critical affective outcome uncovered in this study is the boost in learner confidence. The quantitative data confirmed positive perceptions, and the qualitative findings revealed that this was deeply tied to a sense of achievement. Participants reported feeling more confident in tackling authentic English media, with one noting, "Before, I was afraid to watch English movies, but now I feel confident because captions guide me." This aligns with the concept of self-efficacy within Bandura's Social Cognitive Theory and is corroborated by Li and Wang (2024), who reported that captions served as a "psychological scaffold," reducing anxiety and building confidence among intermediate-level learners when engaging with authentic materials. The role of captions in facilitating incidental vocabulary acquisition is further explained through Schmidt's Noticing Hypothesis.

The captions act as a spotlight, drawing learners' attention to new lexical items in a meaningful context. This study provides empirical evidence for this hypothesis, as learners actively reported noticing and retaining vocabulary they would have otherwise missed. This aligns with Kusumawati's (2023) findings, which show that captioned videos are particularly effective at promoting incidental vocabulary learning by making linguistic features more salient and comprehensible.

However, a nuanced understanding is crucial. The variation in quantitative scores (ranging from 48 to 72) is explained by the qualitative theme of controllable challenges. Some learners struggled with the speed of captions or the cognitive demand of simultaneous listening and reading. This indicates that while captions are highly effective in many cases, their utility is not universal and can be mediated by individual learner differences, such as processing speed and proficiency. This finding is consistent with Kim and Kim (2023), who argued that captions must be part of a scaffolded learning approach to be maximally effective. They suggest that simply providing captions is insufficient; instructors should pair them with supportive activities like pre-teaching vocabulary, using pause-and-predict techniques, and providing vocabulary-focused post-viewing tasks to help all learners, especially those who struggle, make stronger form-meaning connections. The current study makes a compelling case for the structured integration of captioned videos into tertiary ESL curricula in Pakistan and similar contexts. The quantitative data confirms that captioned videos are a statistically significant and positively perceived tool, while the qualitative data explains "why" and how they work by enhancing motivation, leveraging dual-coding cognitive processes, building confidence, and endorsing fluency while requiring some support to overcome inherent challenges.

CONCLUSION

This study concludes that captioned videos are an effective and engaging tool for enhancing incidental vocabulary learning among tertiary ESL learners. The positive mean score and statistically significant t-test support the learners' favorable perceptions observed in interviews. The results align with Mayer's Cognitive Theory of Multimedia Learning and Paivio's Dual Coding Theory, emphasizing the dual-channel processing of audio and text for effective learning. Furthermore, Schmidt's Noticing Hypothesis is evident in learners' statements that captions helped them notice and retain new vocabulary items. The study contributes an important understanding of the potential of captioned videos as an auxiliary tool for L2 vocabulary learning. It emphasizes the importance of incorporating captions into educational strategies, not only to support incidental vocabulary acquisition but also to reflect how language learners often acquire words naturally, without explicit focus. Captions play a key role here, as they increase exposure, highlight meaning through context, and encourage learners to make subconscious connections between sounds, words, and meanings.

LIMITATIONS

Firstly, the study included only 26 participants, which may limit the generalizability of the findings to a larger population of university students. Secondly, the research was conducted in a specific region of Pakistan, limiting its representativeness across diverse cultural and educational settings. Thirdly, reliance on self-reported perceptions may introduce bias, as participants may have overstated or understated their opinions. Fourthly, the lack of longitudinal data means the study captures participants' views at a single point in time and offers no insight into how captions influence long-term language learning. Fifthly, the research did not account for variations in participants' English proficiency, which may have affected their perceptions of captions. Sixthly, while the study focuses on captions, other complementary tools or strategies that could influence vocabulary learning were not considered. Finally, future studies could explore the long-term effects of caption use on vocabulary retention and test learners of different proficiency levels for comparative analysis .

REFERENCES

- Ahmed, R. (2022). Technology integration in language classrooms: Perceptions and outcomes. *Journal of Educational Technology*, 18(3), 45-62.
- Bou-Franch, P., & García, L. (2017). The role of captions in understanding audiovisual texts with complex linguistic features. *Language Learning & Technology*, 21(2), 123–139.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Hulstijn, J. H. (2013). Incidental learning in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2632–2640). Wiley-Blackwell.
- Kim, S., & Kim, J. (2023). Scaffolding captioned video for maximal vocabulary acquisition: A study with intermediate ESL learners. *Computer Assisted Language Learning*, 36(4), 512-530.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Kusumawati, A. J. (2023). A meta-analysis of captioned video for incidental vocabulary learning. *RELJ Journal*, 54(1), 115-130.
- Li, X., & Wang, H. (2024). Captions as a psychological scaffold: Reducing anxiety and building confidence in L2 listening. *System*, 120, 103-115.
- Matahari, R., Sulasmia, D., & Yunita, W. (2021). Student perceptions and their impact on learning engagement in higher education. *Journal of Educational Psychology and Research*, 15(2), 88-105.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
- Melisa, D. (2023). Questionnaire design for measuring student perceptions of digital learning tools. Unpublished manuscript, Department of Education, Global University.
- Montero Pérez, M. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720–739.
- Montero Perez, M., Peters, E., & Desmet, P. (2018). The effect of captioning on vocabulary learning. *Language Learning & Technology*, 22(1), 128–148.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Perez, M. M., Peters, E., & Desmet, P. (2018). The effect of captioning on vocabulary learning. *Language Learning & Technology*, 22(1), 128–148.
- Peters, E., Heynen, E., & Puimège, E. (2016). Learning vocabulary through audiovisual input: The role of on-screen text. *Second Language Research*, 32(2), 215–239.
- Rodgers, M. P. H., & Webb, S. (2020). Incidental vocabulary learning through watching television. *ITL - International Journal of Applied Linguistics*, 171*(2), 191–220.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.

- Sulasmiati, D., Matahari, R., & Yunita, W. (2021). Evaluating instructional effectiveness through student perceptions. *Asian Journal of University Education*, 17(1), 45-58.
- Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of learning and motivation* (Vol. 55, pp. 37–76). Academic Press.
- Tao, J., & Xu, Y. (2018). Student perceptions and learning outcomes in higher education: A review of evidence. *Educational Research Review*, 24, 102-115.
- Teng, F. (2020). The effects of captioned videos on vocabulary learning. *Journal of Language and Education*, 6(2), 45–58.
- Teng, F. (2022). Digital tools and learner autonomy in vocabulary acquisition. *Computer Assisted Language Learning*, 35(8), 1785-1810.
- Teng, F. (2023). *Vocabulary learning through interactive activities*. Oxford University Press.
- Wang, L., & Brown, D. (2019). The effects of captions on attention and comprehension in L2 viewing. *Language Teaching Research*, 23(2), 145–166.
- Wang, Y., & Yum, S. P. (2023). Dual-channel reinforcement: A comparative study of captioned vs. non-captioned video on ESL vocabulary recall. *Journal of Educational Technology & Society*, 26(1), 112-126.
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65–86.
- Yunita, W., & Maisarah, I. (2020). Understanding student perception in the teaching-learning process. *Journal of Language and Linguistic Studies*, 16(2), 1025-1038.
- Zhang, X., & Wang, L. (2017). The impact of captioned video on comprehension and motivation in EFL learning. *Journal of Language and Education*, 3(4), 56–68.