

Structured Scaffolding with Digital Graded Audiobooks: Impacts on L2 Listening Comprehension and Learner Perceptions

Davika Sakkampang¹, Pilanut Phusawisot¹ & Massaya Rachawong²

¹ Faculty of Humanities and Social Sciences, Mahasarakham University, Maha Sarakham 44150, Thailand

² Faculty of Humanities and Social Sciences, Loei Rajabhat University, Loei, 42000, Thailand

Correspondence: Pilanut Phusawisot, Faculty of Humanities and Social Sciences, Mahasarakham University, Thailand.
E-mail: pilanut.p@msu.ac.th

Received: August 17, 2025

Accepted: October 15, 2025

Online Published: November 14, 2025

doi:10.5430/wjel.v16n1p419

URL: <https://doi.org/10.5430/wjel.v16n1p419>

Abstract

This study employed a mixed-methods research design to investigate the effects of structured scaffolding with digital graded audiobooks on the listening comprehension of low-proficiency Thai EFL undergraduates and to explore their perceptions of audiobook use within a structured extensive listening framework. Seventy-four students participated, including thirty-eight in the experimental group and thirty-six in the control group. Research instruments comprised a listening comprehension pretest and posttest, together with semi-structured interviews. The experimental group engaged with twelve Oxford Bookworms digital graded audiobooks integrated into pre-, while-, and post-listening scaffolding activities over twelve weeks, supplemented by autonomous practice through the Oxford Reading Club platform. The control group received conventional textbook-based listening instruction. Quantitative results showed that the experimental group's mean score improved from 13.03 (SD = 2.17) to 17.68 (SD = 2.76), while the control group increased from 12.69 (SD = 2.59) to 14.86 (SD = 3.39). A significant between-group difference was found at posttest, $t(72) = 3.94$, $p < .001$, $d = 0.92$, indicating a large effect size. Qualitative findings revealed that students perceived digital audiobooks as enhancing comprehension, sound discrimination, accent awareness, vocabulary development, and authentic language use. The study highlights the pedagogical value of digital audiobooks in under-resourced, low-input contexts where access to authentic English and technology is limited. Overall, the results demonstrate that a guided-to-autonomous structured scaffolding model can promote accessible, sustainable, and independent listening development in EFL learning environments.

Keywords: digital graded audiobooks, listening comprehension, listening fluency, Thai EFL learners, guided-to-autonomous listening

1. Introduction

Listening is a core skill in English as a Foreign Language (EFL) learning, underpinning academic achievement, communicative competence, and professional readiness. In interpersonal contexts, effective listening fosters rapport, trust, and empathy (Brownell, 2012), supports accurate message interpretation, and reduces the likelihood of misunderstanding (Bodie et al., 2015). In educational settings, it facilitates the acquisition, retention, and application of new knowledge (Wolvin & Coakley, 2019) and contributes to conflict resolution and perspective-taking (Barker, 2018). Professionally, it strengthens leadership, collaboration, and decision-making by enabling a clearer understanding of team dynamics and client needs (Hargie, 2011). Listening competence also enhances emotional intelligence through accurate interpretation of emotions and intentions (Goleman, 1998) and supports cognitive control via selective attention (Sussman & Kane, 2015).

For EFL university students, listening forms the foundation for vocabulary expansion, pronunciation accuracy, and grammatical awareness (Vandergrift & Goh, 2012). Exposure to fluent speakers provides the comprehensible input necessary for language acquisition (Krashen, 1985) and enables learners to engage with lectures, follow instructions, and participate in academic discussions (Rost, 2016). Furthermore, listening contributes to the development of pragmatic competence, critical thinking, and intercultural communication skills (Field, 2008; Jenkins, 2000).

Listening comprehension can be defined as the ability to construct meaning from spoken input by identifying main ideas, recognising supporting details, and making inferences in real time (Rost, 2016). It relies on the integration of bottom-up processes, such as decoding phonological forms and recognising lexical items, with top-down processes, such as activating prior knowledge and applying contextual cues (Field, 2008; Vandergrift & Goh, 2012). Successful comprehension requires flexibility in handling variations in accent, speech rate, and discourse organisation. However, in low-input or under-resourced contexts, EFL learners often experience difficulties in coordinating these processes due to limited vocabulary, reduced exposure to authentic spoken English, and insufficient training in listening strategies (Khamkhien, 2010; Noom-ura, 2013; Sun & Li, 2023).

These challenges are particularly pronounced in provincial Thai universities, where opportunities for authentic listening are scarce (Khamkhien, 2010; Noom-ura, 2013). Instruction frequently prioritises reading and grammar over listening development (Wiriyaichitra,

2002), leaving students with restricted access to level-appropriate listening resources. Consequently, many learners report feeling overwhelmed by rapid speech, struggling to adapt to unfamiliar accents, and lacking both the motivation and the means to systematically improve their listening proficiency. Recent research confirms that such low-input conditions remain common in Southeast Asian EFL classrooms, where limited technological infrastructure and teacher-centred pedagogy restrict exposure to extended English input (Nguyen & Boers, 2021; Liu & Wang, 2024).

One potential solution to these challenges lies in the use of audiobooks, which provide sustained, level-appropriate exposure to spoken language in a controlled and engaging format. Traditionally, audiobooks are recorded readings of written texts, narrated by professional voice actors or authors, and offer consistent models of fluent reading, accurate pronunciation, and natural prosody. Their origins date back to 1932, when the American Foundation for the Blind launched the Books for the Adult Blind project to provide visually impaired adults with access to literature (Rizal et al., 2020). Early recordings on vinyl were constrained by limited storage capacity, but advances in cassette technology during the mid-20th century increased portability and affordability (Have & Pedersen, 2015). By the 1980s and 1990s, audiobooks had gained popularity in literacy instruction, particularly for children and struggling readers, as they combined aural and textual input to enhance comprehension, fluency, and engagement (Serafini et al., 2008; Whittingham et al., 2013).

The advent of digital media in the late 1990s and early 2000s transformed distribution formats from cassettes to MP3 files, CD-ROMs, and eventually online streaming, making large audio libraries widely accessible (Hodges et al., 2014). In language education, this shift coincided with the development of graded readers, which are simplified texts carefully adapted to match specific proficiency levels. Publishers began pairing these texts with professionally narrated audio recordings, giving rise to graded audiobooks that provide level-appropriate and comprehensible input aligned with the Common European Framework of Reference for Languages (CEFR) (Nation & Newton, 2009). At lower proficiency levels, graded audiobooks typically feature slower narration, simplified vocabulary, and clear articulation. At higher levels, they introduce more complex grammar, faster pacing, and greater lexical variety. Many are accompanied by transcripts, glossaries, playback controls, and comprehension tasks (Kartal & Simsek, 2017; Vandergrift & Goh, 2012). These features support vocabulary retention, pronunciation development, prosodic awareness, and fluency through repeated and meaningful exposure. Recent mobile-learning studies have further demonstrated that interactive audiobook applications foster learner motivation and comprehension through multimodal engagement and self-paced practice (Hsu & Chen, 2022; Bewersdorff et al., 2023).

In under-resourced or low-input educational contexts such as rural Thai universities, digital graded audiobooks offer an equitable and sustainable means of providing extended listening exposure. Their mobile accessibility and offline functionality allow learners to practise beyond classroom hours without dependence on high-speed internet or costly materials (Sun & Li, 2023; Liu & Wang, 2024). By combining comprehensible input with autonomy-supportive technology, digital audiobooks can mitigate inequities in resource distribution and help bridge the exposure gap between provincial and urban institutions.

While prior research has demonstrated that graded audiobooks can enhance listening comprehension (Chang & Millett, 2015; Yeh, 2014) and fluency (Chang, 2011; Nation & Newton, 2009), most investigations have been conducted with intermediate or advanced learners in resource-rich environments. Limited attention has been given to adapting these resources for low-proficiency learners who require structured scaffolding through explicit decoding support, segmented listening, and frequent replay options (Field, 2008; Vandergrift & Goh, 2012; Nguyen & Boers, 2021). Furthermore, existing studies have largely relied on post-intervention test scores or self-assessments, offering insufficient insight into the real-time cognitive and metacognitive processes activated during listening (Vandergrift & Goh, 2012). Implementation in low-resource educational settings, particularly where digital literacy and equipment are limited, remains underexplored (Liu & Wang, 2024). This gap underscores the need to investigate how a structured pre-, while-, and post-listening sequence, operationalised within a guided-to-autonomous framework, can foster both comprehension development and autonomy in low-input EFL environments.

This study investigates the effects of digital graded audiobooks on listening comprehension and examines learner perceptions in a low-input Thai EFL context. Specifically, it aims:

- 1) To examine the effects of digital graded audiobooks on the listening comprehension of Thai EFL undergraduate students.
- 2) To investigate the perceptions of Thai EFL undergraduate students toward the use of digital graded audiobooks in enhancing their listening comprehension.

The study makes a theoretical contribution by integrating established second-language-acquisition (SLA) frameworks with audiobook-supported extensive listening, highlighting how a guided-to-autonomous model can be operationalised to optimise comprehension development. It also clarifies the pedagogical significance of applying digital audiobook scaffolding in under-resourced settings, where equitable access and learner independence are critical. Pedagogically, it offers a structured, replicable approach tailored to the needs of lower-proficiency learners in under-resourced contexts, ensuring that both cognitive and affective dimensions of listening are addressed. Practically, it provides an adaptable instructional model that can be implemented across diverse educational settings to enhance listening skills while sustaining learner engagement, motivation, and autonomy.

2. Method

2.1 Participants

Seventy-four non-English-major undergraduates from a provincial Thai university participated in this study. All were enrolled in a

compulsory general English course and had no prior experience with digital graded audiobooks or formal structured extensive-listening programs. Participants were assigned to either the experimental group ($n = 38$) or the control group ($n = 36$) according to intact class sections to avoid disruption of the existing timetable. Their ages ranged from 18 to 21 years, with a balanced representation of male and female students. CEFR proficiency levels, assessed using the Oxford Quick Placement Test Version 2, showed that most participants were at the A1 level (73.7 % in the experimental group and 80.6 % in the control group), while the remaining students were at the A2 level (26.3 % and 19.4 %, respectively). The university is located in a rural province of Thailand, where opportunities for authentic English exposure are minimal. Outside the classroom, students reported little to no interaction with English speakers and rare engagement with English media. Their prior English-learning experiences were largely shaped by grammar-focused instruction and reading-based assessments, with limited emphasis on listening or speaking. As a result, many entered the course with low listening-comprehension proficiency, restricted vocabulary, and limited confidence in processing spoken English. These characteristics reflect the low-input environment typical of many Thai provincial universities and highlight the relevance of investigating interventions that can provide sustained, comprehensible listening input supported by structured scaffolding. In addition to the main sample, twelve participants from the experimental group were purposively selected for post-intervention semi-structured interviews. They represented three proficiency levels: four high, four mid, and four low scoring students, based on post-test results. This stratified selection ensured diverse perspectives and enabled comparison of experiences across proficiency levels. All participants were informed about the research objectives and procedures, provided written consent, and were assured that participation was voluntary and would not affect their course grades.

2.2 Intervention: Digital Graded Audiobooks

The intervention utilised twelve digital graded audiobooks from the Oxford Bookworms Library Series, a widely adopted graded reader collection designed for English as a Foreign Language (EFL) learners. Selection criteria considered both the CEFR proficiency levels of participants in the low-input Thai EFL context and the availability of supportive digital features to facilitate listening comprehension development. The set included six Starter level titles (CEFR A1) with simplified vocabulary of approximately 250–300 headwords and slower narration speeds of about 100–110 words per minute, as well as six Level 1 titles (CEFR A1–A2) with an expanded vocabulary range of approximately 400–450 headwords and slightly faster narration at 110–120 words per minute.

All digital graded audiobooks incorporated synchronized audio-text display to support word recognition and orthographic-phonological mapping, adjustable playback speed to enable adaptation to varied speech rates, replay functions for targeted listening to challenging sections, and clear, expressive narration by professional voice actors to model natural intonation and prosody. Thematically, the collection encompassed adventure, mystery, and human-interest stories to cater to diverse learner preferences and sustain motivation. All titles were delivered through the Oxford Reading Club platform.

However, in this study their primary function was to serve as guided instructional materials during in-class scaffolded activities. Students were introduced to the audiobooks in class under teacher guidance, but they were not required to engage with all twelve titles outside the classroom. Out-of-class use of the platform was optional and supplementary, depending on individual learner preference. Content selection was guided by the principle of providing comprehensible input (Krashen, 1985) while being matched to the learners' CEFR proficiency levels and maintaining linguistic authenticity through natural-sounding English, thereby creating optimal conditions for simultaneous gains in listening comprehension through structured classroom scaffolding.

2.3 Instrument

2.3.1 Listening Comprehension Pretest and Posttest

Listening comprehension was measured using parallel forms of a standardized test adapted from the Oxford Listening Test Series, selected for its alignment with the A1–A2 proficiency range according to the CEFR. Both the pretest and posttest consisted of 25 multiple-choice items assessing the ability to identify main ideas, recognize supporting details, and make inferences from short dialogues and narrative passages. Each test included a balanced distribution of items targeting factual recall, contextual inference, and gist comprehension. Audio recordings were delivered at a controlled speech rate of 120–140 words per minute, consistent with Nation and Newton's (2009) recommendations for lower-proficiency EFL learners. The pretest was administered one week before the intervention to establish baseline proficiency, while the posttest was administered in the final week of the twelve-week instructional period to measure learning gains. Both assessments were conducted under standardized classroom conditions, with instructions provided in Thai to ensure clarity.

2.3.2 Interview Protocol

Semi-structured interviews were designed to explore learners' perceptions of how digital graded audiobooks supported their listening-comprehension development within the structured-scaffolding framework. The protocol consisted of open-ended questions focusing on perceived benefits, challenges, motivational factors, and strategy use. Probing prompts were included to elicit detailed examples, such as references to specific audiobooks, listening strategies, or scenes that enhanced comprehension. Each interview lasted approximately 20–30 minutes, was conducted in Thai to ensure natural expression, and was audio-recorded with participant consent. Recordings were transcribed verbatim and subsequently translated into English, with translations cross-verified by two bilingual researchers to preserve accuracy and intended meaning. Interview sessions were conducted during the final week of the intervention and followed consistent procedures for scheduling, questioning, and data handling to maintain reliability.

2.4 Procedure

The 12-week intervention commenced in Week 1 with the administration of a listening comprehension pretest to all participants, ensuring baseline comparability between the control and experimental groups. Immediately after the pretest, the experimental group received an orientation session introducing the Oxford Reading Club (ORC) platform. This training familiarized learners with key digital graded audiobook features such as synchronized audio-text display, adjustable playback speed, section replay, and embedded glossaries, enabling them to navigate the platform independently.

The integrated digital graded audiobook listening activities followed a guided-to-autonomous instructional framework, integrating in-class and out-of-class components.

In-class activities for the experimental group were conducted weekly and began with vocabulary previews designed to pre-teach high-frequency and story-specific words. Prediction tasks were used to activate background knowledge and set a purpose for listening. Students then engaged in focused listening segments of 3–5 minutes, segmented to match narrative events, with comprehension checks and strategy scaffolding provided between segments. Post-listening discussions encouraged learners to summarise content, clarify ambiguities, and relate the storyline to personal experiences, fostering deeper comprehension and engagement within Vygotsky's ZPD framework.

Out-of-class activities required each learner to complete a minimum of 30 minutes per week of autonomous listening via the ORC platform. Students could select any digital graded audiobook aligned with their CEFR level, with the option to explore beyond their comfort zone to meet Krashen's $i+1$ principle. They were instructed to replay challenging sections, adjust playback speed for listening comprehension and listening fluency development, and use the glossary feature to support vocabulary acquisition. Weekly listening logs were submitted to track activity completion and monitor learner engagement.

The control group followed the same curriculum topics but used conventional textbook-based instruction. Their listening practice involved commercially provided audio tracks accompanying the coursebook, supplemented by comprehension exercises, without digital graded audiobook exposure.

In Week 12, a listening comprehension posttest was administered to both groups under identical conditions. Additionally, 12 students purposive selected from the experimental group participated in semi-structured interviews. These interviews aimed to capture learner perceptions of the usefulness, challenges, and motivational impact of the digital graded audiobook activities.

2.5 Data Analysis

Quantitative data were analyzed using paired- and independent-samples t tests ($\alpha = .05$) to assess within- and between-group differences, and effect sizes (Cohen's d) were calculated to indicate magnitude of change. Qualitative data from the twelve interviews were analyzed using Braun and Clarke's (2006, 2021) six-phase thematic-analysis framework. The process began with repeated reading of transcripts for familiarization, followed by initial open coding to identify meaningful units related to cognitive, affective, and behavioral dimensions of learning. Two trained coders independently generated and refined codes, which were then collated into broader categories and synthesized into overarching themes through iterative discussion. Inter-coder reliability was calculated (Cohen's $\kappa = 0.87$), indicating high agreement. Discrepancies were resolved through consensus meetings, and member-checking with four participants was conducted to validate interpretations. This systematic procedure enhanced the credibility, dependability, and confirmability of the qualitative findings.

3. Results

3.1 The Effect of Digital Graded Audiobooks on Students' Listening Comprehension

Descriptive statistics (Table 1) indicated that the control group's mean listening-comprehension score increased from 12.69 (SD = 2.59) at pretest to 14.86 (SD = 3.39) at posttest, while the experimental group improved from 13.03 (SD = 2.17) to 17.68 (SD = 2.76). Although both groups demonstrated gains, the improvement was substantially greater for the experimental group, reflecting the pedagogical value of the structured scaffolding approach.

Paired-samples t test results (Table 1) confirmed that the increase in scores was statistically significant for both the control group, $t(35) = 7.05$, $p < .001$, and the experimental group, $t(37) = 15.53$, $p < .001$. Independent-samples t test analysis showed no significant difference in baseline performance between groups at pretest, $t(72) = 0.60$, $p = .552$, with a negligible effect size ($d = 0.14$). However, at posttest, the experimental group significantly outperformed the control group, $t(72) = 3.94$, $p < .001$, with a large effect size ($d = 0.92$). This demonstrates that the structured scaffolding framework integrated with digital graded audiobooks led to a measurable and educationally meaningful impact on listening comprehension performance in the low-input EFL context.

Table 1. Independent Samples T-Test of the Pretest and Posttest Scores on the Listening Comprehension

Groups	N	Tests	Mean	SD	t-value	p-value	Effect size (d)
Pre-test	36	CG	12.694	2.594	.598	.552	.014
	38	EG	13.026	2.174			
Post-test	36	CG	14.861	3.390	3.937	.000	0.92
	38	EG	17.684	2.761			

Discription: Control Group: CG, Experimental Group: EG

3.2 Perceptions toward the Use of Digital Graded Audiobooks in Developing Students' Listening Comprehension

Drawing on in-depth interviews with twelve purposively selected participants (four high-, four mid-, and four low-score students from the experimental group), the findings indicate that learners across proficiency levels perceived the structured scaffolding framework with digital graded audiobooks as both beneficial and motivating. Reported gains encompassed cognitive, affective, and behavioral domains. Thematic analysis following Braun and Clarke's (2006, 2021) six-phase procedure identified five interrelated themes: (1) Enhancing Listening and Deep Story Understanding, (2) Improving Sound Discrimination, (3) Improving Understanding of Accents, (4) Acquiring New Vocabulary Knowledge, and (5) Learning Language in Real-Life Contexts.

Learners consistently described how digital graded audiobooks enabled them to engage more fully with stories, not only by following the storyline but also by deepening their understanding of meaning. Their experiences revealed a clear developmental progression across proficiency levels.

For low-proficiency learners, the main impact was a reduction of stress and a stronger focus on meaning rather than on individual sounds. One participant explained how the clarity and pacing of audiobooks helped shift attention away from worry and towards story comprehension:

"With audiobooks, the sound is always clear, and the pronunciation is easy to catch. I feel more comfortable because the language matches my level, so I can listen without feeling stressed and can focus on understanding the story instead of worrying about what I missed. (L1)"

For mid-proficiency learners, the focus moved from overcoming stress to achieving smoother, more fluent listening. These students reported that repeated exposure improved their ability to recognize words quickly and link them to meaning, which in turn allowed them to process longer stretches of language:

"Listening to audiobooks helps me concentrate on the words and the way they are spoken. Over time, I notice that I can recognize words more quickly and connect them to the meaning without stopping to think, which makes listening feel more natural and less tiring. (M3)"

For higher-proficiency learners, the emphasis was on deeper narrative engagement. These students were able to connect ideas across story sections, recall details, and even anticipate upcoming events. One participant described this process as follows:

"The stories are divided into sections, which makes it easier to remember and reflect on what I've heard. I can connect the events from one part to another more clearly, which helps me follow the whole story. (H1)"

These findings illustrate how learners at varying proficiency levels benefited from progressively deeper engagement when structured scaffolding guided their audiobook use, ranging from stress reduction for beginners to inferential understanding for advanced learners.

3.2.2 Improving Sound Discrimination

Another key theme from the interviews was that students felt they improved in hearing and distinguishing English sounds. Repeated listening to digital graded audiobooks helped them notice subtle pronunciation features such as final consonants, word stress, and sound connections in fast speech. This process was not only about following the story but also about training their ears to recognize patterns more accurately, consistent with phonological-noticing processes highlighted in multimodal listening research (Hsu & Chen, 2022).

For low-proficiency learners, the main benefit was avoiding confusion between words that sounded similar. One participant explained that the shorter listening sections allowed focus on particular sounds without being overwhelmed:

"The story is divided into small parts, so I can follow each part and understand what is happening. If it was all in one long piece, I would get confused and probably miss important parts. This way, I can focus on one scene at a time and hear the words more clearly. (L1)"

At the mid-level, students began to reflect more on their listening, noticing when certain sounds or stress patterns were unclear and replaying them until they became familiar. One student described how this process helped build awareness of connected speech:

"When I listen again to the same part, I start to hear how the words join together. At first it sounds like one long sound, but after I repeat it, I can separate the words. This makes me more confident that I really hear the right thing. (M2)"

For higher-level learners, the focus extended beyond recognition of sounds to interpreting differences in rhythm and intonation. One student noted how this practice improved both comprehension and pronunciation awareness:

"When I listen many times, I notice how the voice goes up and down, and how some words are stronger than others. Before, I did not pay attention, but now I try to copy it. It helps me understand the feeling in the story and also how to say it myself. (H3)"

These perspectives suggest that digital graded audiobooks within a structured scaffolding framework enhanced sound discrimination across proficiency levels. Beginners used segmentation to avoid confusion, mid-level learners replayed sections to identify word boundaries, and advanced learners focused on intonation and rhythm to strengthen comprehension and pronunciation.

3.2.3 Improving Understanding of Accents

A third theme was learners' increased comfort in understanding different English accents, especially British and American. They noted

that digital graded audiobooks, with clear articulation and adjustable speed, offered a supportive way to handle accent variation. Exposure to diverse narrations in the Oxford Bookworms series helped them adapt within a controlled pace, while multimodal pairing enhanced word recall, consistent with findings by Nguyen and Boers (2021).

For lower-level learners, the main concern was simply keeping up when the accent was new. One student shared that clear and deliberate pronunciation was essential for following the story:

"When the speaker pronounces clearly, I can still understand even if the accent is different. Without clear speech, I get lost very quickly. Clear pronunciation gives me more time to think about the meaning and follow the story. (L1)"

At the mid-level, learners described becoming more flexible. They began to notice that slowing down or repeating parts of the audio helped them grow accustomed to the accent, and once familiar, they could increase the speed without losing comprehension. One student explained:

"Slower speed helps me understand the accent and the meaning at the same time. I can focus on the way the words sound and still follow the story. Later, when I am used to it, I can listen faster without losing the meaning. (M2)"

Higher-level learners spoke of using accent exposure as a kind of training. They described listening slowly at first to grasp the accent's rhythm and then challenging themselves with faster narration. One student reflected:

"When the speech is slow at first, I can catch the pattern of the accent. Then, when I hear it faster, I already know what to expect. It's like I am training my ears to adjust from slow to normal speed. (H1)"

Another high-score student described developing sharper sensitivity to subtle accent differences:

"The more I hear different British or American speakers, the easier it is to catch the meaning right away. I start to notice the small sound changes that are part of the accent, and this makes it easier for me to understand even if the speaker talks quickly. (H2)"

These accounts indicate that structured scaffolding with digital graded audiobooks enhanced accent comprehension progressively, as beginners relied on clarity, mid-level learners improved recognition across variations, and advanced learners developed adaptability to different accent patterns.

3.2.4 Acquiring New Vocabulary Knowledge

Participants widely noted that digital graded audiobooks promoted vocabulary growth through repeated, contextualized exposure. Leveled content and recurring key words enhanced recognition, recall, and overall listening comprehension.

For lower-level learners, vocabulary learning was closely tied to keeping comprehension manageable. One student explained that the right level of difficulty helped them enjoy the story without being discouraged:

"The words are not too difficult, so I can understand the story and learn some new words at the same time. If there are too many hard words, I cannot follow the story, and I lose interest. With this level, I can enjoy the story and still learn. (L1)"

At the mid-level, learners described noticing how a small number of new words were repeated in different parts of the story, which made it easier to guess and later recall their meaning. One participant explained:

"Each story has some new words, but most of the words I already know. This makes it easier to guess the meaning and remember the new ones because I can focus on just a few at a time, not too many. (M2)"

For higher-level learners, the process extended beyond recognition toward using new words productively. One student noted that hearing the same expressions many times encouraged them to experiment with applying them in their own communication:

"The focus words remind me of vocabulary I can try to use in speaking or writing. It's not just knowing the meaning, but also how to use them naturally in conversation or essays. When I hear the same word many times in the story, it becomes much easier to remember and apply. (H1)"

These reflections show that structured scaffolding with digital graded audiobooks supported vocabulary growth across levels by offering manageable words for beginners, reinforcing contextual guessing for mid-level learners, and deepening usage understanding for advanced learners.

3.2.5 Learning Language in Real-Life Contexts

Finally, many participants perceived digital graded audiobooks as a bridge to real-life communication. They highlighted that exposure to natural-sounding expressions, idioms, and conversational language in meaningful contexts helped them transfer new language into their own speaking and writing.

Learners at the lower level tended to notice short, ready-made expressions that they could repeat directly. One student explained how satisfying it was to hear sentences that could be immediately reused:

"I hear sentences that I can use, like greetings or short questions. When the character says it in the story, I know I can say it to people too. It feels real, not just practice. (L1)"

At the mid-level, learners described recognising phrases that felt authentic and sounded different from what they usually found in

textbooks. Because these expressions were tied to realistic storylines, they were easier to recall and more enjoyable to practice. One student shared:

“Some of the sentences sound like what real people say, not like in the textbook. I remember them better because I can imagine using them in the same way when I talk with friends or classmates. (M3)”

For more advanced learners, audiobooks offered opportunities to notice variation in phrasing and to adapt language creatively. One high-score participant explained how exposure to multiple ways of expressing the same idea supported flexibility in communication:

“I learn different ways to say the same thing. For example, instead of only saying ‘I am sorry,’ I hear other ways people apologise. It helps me choose the right words in different situations, like formal or casual. (H2)”

These accounts show a clear progression: beginners repeated simple expressions, mid-level learners reused authentic phrases, and advanced learners adapted language creatively. This indicates that digital graded audiobooks within a structured scaffolding framework supported both receptive and productive language use.

4. Discussion

4.1 The Effects of Structured Scaffolding with Digital Graded Audiobooks on Listening Comprehension of Thai EFL Undergraduates

The quantitative findings of this study provide clear evidence that digital graded audiobooks, when integrated into a structured scaffolding framework of three instructional stages, substantially enhanced the listening comprehension of Thai EFL undergraduates in a low-input context. The experimental group’s mean listening comprehension score increased from 13.03 to 17.68, a significantly greater gain than that of the control group, which improved from 12.69 to 14.86. This statistical difference, with a large effect size ($d = 0.92$), confirms the intervention’s measurable impact.

These findings both confirm and extend results from previous studies. Similar to Chang (2011) and Waring and Takaki (2003), this study found that repeated listening improved comprehension of storylines and character actions. The results also align with Yeh (2014) and Gonulal (2020), who demonstrated that graded audiobooks provide accessible input and support comprehension through simplified vocabulary and structured pacing. However, the present study moves beyond prior research by focusing specifically on low-proficiency learners in a low-input Thai context. Unlike earlier investigations conducted in urban or advanced-level environments (Chang & Millett, 2015; Gonulal, 2020), the present findings demonstrate that beginner-level learners can achieve significant comprehension gains when digital graded audiobooks are combined with scaffolded pre-, while-, and post-listening activities. In this respect, the results challenge the assumption in earlier research that extensive listening alone is sufficient to develop comprehension, instead showing that structured scaffolding is necessary for low-level learners to benefit fully. However, the present findings further resonate with more recent research (Hsu & Chen, 2022; Sun & Li, 2023; Liu & Wang, 2024), showing that digital audiobook technologies, when embedded in pedagogically guided frameworks, can compensate for limited input exposure in resource-constrained settings.

A key finding is that comprehension gains resulted not only from level-appropriate input but from embedding digital graded audiobooks within structured pre-, while-, and post-listening activities. In the pre-listening stage, vocabulary brainstorming, prediction, and activation of prior knowledge helped learners process language more effectively. This preparation was crucial, as students might not revisit the same audiobook during autonomous listening, making in-class readiness essential. Learners reported that this stage improved their ability to anticipate storylines and reduced anxiety, allowing focus on meaning. The process aligns with Schema Theory (Anderson, 1984), which primes mental frameworks, and Cognitive Load Theory (Sweller, 1988), which lowers extraneous cognitive load before listening.

In the while-listening stage, the structured use of audiobook playback, first for gist and then for detail, was central to comprehension development. Here, learners engaged with adjustable playback speed, replay functions, and segmented story structures to manage pacing, confirm predictions, and notice features missed during initial listening. Quantitative improvement in comprehension scores was consistent with learner accounts that repeated listening helped them capture previously missed sentences and better understand plot developments. This stage operationalized Krashen’s Input Hypothesis by delivering comprehensible input slightly beyond learners’ current level ($i+1$), and reinforced Noticing Theory (Schmidt, 1990) by directing attention to salient lexical, grammatical, and discourse markers during repetition.

The post-listening stage further consolidated comprehension through comprehension questions, vocabulary review, and group discussion. These classroom-based reflection activities were again essential, since students might not have engaged in follow-up work when listening independently, making it critical to process and consolidate understanding while the material was still fresh. Students reported that discussing characters’ decisions with peers made the story clearer and easier to remember. This stage encouraged germane cognitive load (Sweller, 2010) by fostering schema construction and linking new information to prior knowledge. The combination of immediate feedback, explicit clarification, and reflective writing also enhanced long-term retention, as supported by Nation’s (2001) emphasis on repeated, meaningful use of vocabulary and content.

The outcomes thus validate that structured scaffolding transforms digital audiobooks from passive exposure tools into dynamic, interactive learning systems. This model supports Cognitive Load Theory (Sweller, 1988, 2010) by managing task complexity, Noticing Theory (Schmidt, 1990) through repeated listening and focus on form, and Sociocultural Theory (Vygotsky, 1978) by situating listening development within guided interaction before autonomy. The evidence also aligns with recent multimodal-learning studies indicating that audio-text synchronisation enhances comprehension by reducing split-attention effects (Mayer & Moreno, 2003; Bewersdorff et al.,

2023).

The integration of the three stages formed a continuous scaffolding cycle: pre-listening prepared learners, while-listening guided focused processing, and post-listening reinforced comprehension. This sequence maximized the audiobooks' multimodal benefits, including clear narration, synchronized text, and visual cues, by promoting active engagement. It enhanced both detailed understanding and higher-order skills such as inference and thematic interpretation, with advanced learners using playback features to explore nuance in tone and meaning.

4.2 Students' Perceptions of the Use of Digital Graded Audiobooks on Listening Comprehension

The qualitative findings showed that students perceived digital graded audiobooks as highly effective for developing listening comprehension, with reflections centering on five areas: story comprehension, sound discrimination, accent awareness, vocabulary acquisition, and real-life language use. These perceptions provided explanatory depth to the quantitative gains by illustrating not only what improved but also how learners experienced the process.

Learners across proficiency levels reported clearer understanding of storylines, improved recall of details, and deeper engagement with narratives. Features such as synchronized text, controlled pacing, and expressive narration were viewed as particularly supportive. This confirms Chang's (2011) argument that multimodal pairing enhances comprehension and extends Bui and Teng's (2021) claim that multimodality sustains attention. Unlike Bui and Teng's participants, learners valued playback control as a means of adjusting input to individual pace, thereby reducing processing overload. While Serafini and Youngs (2013) suggested narration may restrict interpretive autonomy, students here saw expressive narration as a model for interpretive strategies.

Students also emphasized improvements in distinguishing phonemes, stress, and intonation, attributing this to repeated, high-quality input paired with synchronized text. This aligns with Vanderplank (2016) on phonological awareness and Brown, Waring, and Donkaewbua (2008) on reduced-form recognition. In contrast to Brown et al.'s participants, who struggled without visual reinforcement, learners in this study found the multimodal format sufficient for auditory discrimination. These findings partly challenge Rasinski and Yildirim's (2017) concern that audiobooks without explicit phonemic instruction may yield limited gains, suggesting instead that scaffolded and leveled input can independently strengthen sound awareness.

Participants further reported growing confidence with accent variation after exposure to narrators with British, American, and other native-like accents. They noted greater ease in following rhythm, intonation, and pronunciation differences. This supports Ockey and French's (2016) argument that accent exposure develops adaptable listening skills and Goh's (2000) view that such exposure reduces anxiety. Unlike Goh's focus on unscripted speech, this study shows that graded narration with controlled pacing provides a manageable entry point. Learners also acknowledged Derwing and Munro's (2009) point that familiarity with one accent may not automatically transfer but valued audiobooks as a stepping stone toward global English adaptability.

Across proficiency levels, students highlighted vocabulary growth through repeated contextual encounters. They reported learning pronunciation, spelling, and usage, reinforced by audio-text pairing. This confirms Nation's (2013) principle of repetition in context and Vidal's (2011) finding that contextualized listening enhances retention. Unlike Webb and Chang's (2012) warning that gains may be limited without post-listening tasks, learners here found repeated encounters sufficient for incidental acquisition.

Finally, students appreciated exposure to natural expressions, idioms, and pragmatic language transferable to everyday interactions. They recalled phrases such as greetings and apologies successfully applied with peers. This supports Nguyen and Boers' (2019) emphasis on contextualized input and Mason's (2004) advocacy of story-based listening. In contrast with Gilmore's (2007) preference for fully authentic materials, students valued the graded format for balancing accessibility with authenticity, which built confidence to progress toward unsimplified input.

5. Implications

The findings of this study carry significant implications for second language listening development across theoretical, methodological, and pedagogical domains. By examining how structured scaffolding with digital graded audiobooks enhances listening comprehension among low-proficiency EFL learners in a low-input Thai context, the study not only confirms the value of scaffolded input but also advances theoretical perspectives by proposing a refined framework that integrates guided and autonomous listening.

The key theoretical contribution of this study lies in refining views on extensive listening and learner autonomy. While frameworks such as the Input Hypothesis (Krashen, 1982, 1985), Cognitive Load Theory (Sweller, 1988), Automaticity Theory (LaBerge & Samuels, 1974), Noticing Theory (Schmidt, 1990), and Sociocultural Theory (Vygotsky, 1978) have long shaped pedagogy, the findings show that in low-input contexts, autonomy does not develop when learners begin independently. Instead, a guided-to-autonomous model is supported, where teacher-led scaffolding through vocabulary activation, prediction, discussion, and comprehension checks within the Zone of Proximal Development (Vygotsky, 1978; Lantolf & Thorne, 2006) equips learners before responsibility is gradually released. This staged approach challenges assumptions of immediate autonomy (Chang & Millett, 2015; Renandya & Farrell, 2011) and extends Cognitive Load, Noticing, and Automaticity theories.

Methodologically, this study demonstrates the value of a mixed-methods design for investigating L2 listening. Combining pretest and posttest data with qualitative interviews provided both measurable outcomes and explanatory depth. Multiple test formats, including multiple choice, matching, true or false, and timed listening tasks, captured diverse dimensions of comprehension and minimized ceiling

and floor effects, aligning with best practices (Buck, 2001; Lynch, 2017). Semi-structured interviews further illuminated learner perceptions, while the Oxford Reading Club platform functioned as both an intervention tool and a naturalistic research environment, supporting calls for ecologically valid designs in applied linguistics (Godwin-Jones, 2018). Future research could build on this framework through longitudinal designs to examine retention, strategy transfer, and the sustainability of autonomous listening.

Pedagogically, the findings emphasize embedding digital graded audiobooks within a structured scaffolding sequence that integrates guided classroom practice with autonomous listening. Pre-listening activities such as vocabulary activation and prediction prepare comprehension, while strategies like segmented listening, gist-to-detail sequencing, and playback control deepen processing. Post-listening tasks, including comprehension checks, discussion, and reflection, consolidate learning and promote retention. With explicit strategy instruction and CEFR-aligned, culturally relevant materials, audiobooks effectively support the transition from guided to autonomous listening, enhancing both comprehension and learner confidence.

6. Conclusion

This study examined the effects of digital graded audiobooks on the listening comprehension of low-proficiency Thai EFL undergraduates within a structured scaffolding framework for extensive listening. Using a mixed-methods design that combined pretest–posttest measures with qualitative interviews, the study provided insights into how comprehension can be developed in low-input contexts. Quantitative results showed significant gains in listening comprehension for the experimental group, with strong effect sizes. Qualitative findings further revealed that improvements were closely linked to learners' engagement with multimodal features such as clear narration, synchronized text, visual cues, and playback control, when integrated into pre-, while-, and post-listening activities. This structured sequence allowed learners to activate vocabulary, process input through repetition and segmentation, and consolidate understanding through reflection and application.

Several limitations must be noted. The participant pool was confined to a single rural Thai university, limiting generalizability. The twelve-week intervention captured short-term progress but may not reflect long-term retention or transfer. Self-reported interviews carried risks of recall bias, and variability in out-of-class listening input was introduced by students' free choice of audiobooks. Future research should include more diverse participants, longer intervention periods, and triangulated data sources such as digital logs, classroom observations, and analytic tools like speech recognition or eye tracking. Expanding audiobook genres and accent models, including American, Australian, and non-native English, would enhance adaptability to global English. Applying the guided-to-autonomous model with higher proficiency learners could further extend its theoretical and pedagogical relevance in EFL contexts.

Acknowledgments

The authors would like to express their sincere gratitude to their colleagues for their constructive comments and thoughtful suggestions, which greatly enhanced the quality of this manuscript. Their guidance and encouragement throughout the research process were invaluable. Special thanks are also extended to the student participants for their enthusiastic cooperation and meaningful contributions. This study would not have been possible without their active engagement and support.

Authors' contributions

Davika Sakkampang, Dr. Pilanut Phusawisot, and Massaya Rachawong were responsible for the study design and data analysis. Davika Sakkampang primarily contributed to manuscript structuring, critical revision, and overall coordination of the research. Massaya Rachawong mainly managed the literature review and data collection. Dr. Pilanut Phusawisot provided ongoing guidance and consultation throughout the entire research process. All three authors jointly analyzed the data, interpreted, and discussed the findings. Davika Sakkampang and Massaya Rachawong drafted the manuscript, and Dr. Pilanut Phusawisot revised it. All authors read and approved the final manuscript. The three authors contributed equally to this study. AI-assisted tools were used only for language polishing, structural alignment, and consistency checking. All academic ideas, data interpretation, and conceptual contributions were solely produced by the authors.

Funding

This research project was financially supported by Mahasarakham University.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Sciedu Press.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

References

- Anderson, R. C. (1984). Role of the reader's schema in comprehension, learning, and memory. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), *Learning to read in American schools: Basal readers and content texts* (pp. 243-257). Lawrence Erlbaum Associates.
- Barker, L. (2018). *Communication in organizations* (5th ed.). Routledge. <https://doi.org/10.4324/9780429494335>
- Bewersdorff, B., Paniagua, A., & Moebius, K. (2023). Digital storytelling for language learning: A review of theoretical perspectives and empirical evidence. *Language Learning & Technology*, 27(1), 1-19. <https://doi.org/10.1016/j.compedu.2023.104771>
- Bodie, G. D., Vickery, A. J., Cannava, K., & Jones, S. M. (2015). The role of "active listening" in informal helping conversations: Impact on perceptions of listener helpfulness, sensitivity, and supportiveness and discloser emotional improvement. *Western Journal of Communication*, 79(2), 151-173. <https://doi.org/10.1080/10570314.2014.943429>
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136-163. <https://doi.org/10.1111/j.1944-9720.2008.tb00625.x>
- Brownell, J. (2012). *Listening: Attitudes, principles, and skills* (5th ed.). Pearson.
- Carrell, P. L., & Eisterhold, J. C. (1983). Schema theory and ESL reading pedagogy. *TESOL Quarterly*, 17(4), 553-573. <https://doi.org/10.2307/3586613>
- Chang, A. C. S. (2011). The effect of reading while listening to audiobooks: Listening fluency and vocabulary gain. *Asian Journal of English Language Teaching*, 21, 43-64.
- Chang, A. C. S., & Millett, S. (2014). The effect of extensive listening on developing L2 listening fluency: Some hard evidence. *ELT Journal*, 68(1), 31-40. <https://doi.org/10.1093/elt/cct052>
- Chang, A. C. S., & Millett, S. (2015). Improving reading rates and comprehension through audio-assisted extensive reading for ESL learners. *Reading Matrix: An International Online Journal*, 15(2), 43-55.
- Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 63-103). Blackwell. <https://doi.org/10.1002/9780470756492.ch4>
- Field, J. (2008). *Listening in the language classroom*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511575945>
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97-118. <https://doi.org/10.1017/S0261444807004144>
- Goleman, D. (1998). *Working with emotional intelligence*. Bantam Books. <https://doi.org/10.1002/ltl.40619981008>
- Griffiths, R. (1992). Speech rate and listening comprehension: Further evidence of the relationship. *TESOL Quarterly*, 26(2), 385-390. <https://doi.org/10.2307/3587015>
- Hargie, O. (2011). *Skilled interpersonal communication: Research, theory and practice* (5th ed.). Routledge. <https://doi.org/10.4324/9780203844830>
- Have, I., & Pedersen, B. S. (2015). Digital audiobooks: New media, users, and experiences. *Journal of the International Association for Media and Communication Research*, 2(1), 1-19. <https://doi.org/10.4324/9781315743080>
- Hodges, G. C., Simpson, L., & Stahl, S. A. (2014). Listening to learn: Audiobooks supporting literacy. *Reading Teacher*, 67(8), 610-614. <https://doi.org/10.1002/trtr.1258>
- Hsu, T. C., & Chen, C. M. (2022). Effects of interactive audiobook system on English listening comprehension and learning interest of elementary school students. *Educational Technology Research and Development*, 70, 1885-1908. <https://doi.org/10.1007/s11423-022-10142-x>
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Kartal, G., & Simsek, E. (2017). The effects of audiobooks on EFL learners' listening comprehension and attitudes towards them. *TESOL*

- Journal*, 8(3), 1-13. <https://doi.org/10.1002/tesj.288>
- Khamkhien, A. (2010). Teaching English speaking and English speaking tests in the Thai context: A reflection from Thai perspective. *English Language Teaching*, 3(1), 184-190. <https://doi.org/10.5539/elt.v3n1p184>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293-323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
- Liu, Y., & Wang, L. (2024). Digital audiobooks and EFL listening development: A technology-enhanced approach for low-resource contexts. *RELC Journal*, 55(1), 112-128. <https://doi.org/10.1177/00336882231189456>
- Lynch, T. (2017). *Listening: Second edition*. Oxford University Press.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43-52. https://doi.org/10.1207/S15326985EP3801_6
- Nasir, M., Ahmad, N., & Fatima, A. (2021). Digital resources for EFL learners: Opportunities and challenges. *Journal of Language and Education*, 7(3), 24-38. <https://doi.org/10.17323/jle.2021.11684>
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge. <https://doi.org/10.4324/9780203891704>
- Nguyen, C. D., & Boers, F. (2021). The effect of content retelling on vocabulary uptake from a TED talk. *TESOL Quarterly*, 55(1), 293-304. <https://doi.org/10.1002/tesq.57>
- Noom-ura, S. (2013). English-teaching problems in Thailand and Thai teachers' professional development needs. *English Language Teaching*, 6(11), 139-147. <https://doi.org/10.5539/elt.v6n11p139>
- Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart & Winston.
- Renandya, W. A., & Farrell, T. S. C. (2011). 'Teacher, the tape is too fast!' Extensive listening in ELT. *ELT Journal*, 65(1), 52-59. <https://doi.org/10.1093/elt/ccq015>
- Rizal, A., Prasetyo, T., & Sari, D. (2020). The evolution of audiobooks: Historical perspectives and contemporary trends. *International Journal of Emerging Technologies in Learning*, 15(7), 123-135. <https://doi.org/10.3991/ijet.v15i07.13215>
- Rodgers, M. P. H., & Webb, S. (2011). Narrow viewing: The vocabulary in related television programs. *TESOL Quarterly*, 45(4), 689-717. <https://doi.org/10.5054/tq.2011.268062>
- Rost, M. (2011). *Teaching and researching listening* (2nd ed.). Pearson.
- Rost, M. (2016). *Teaching and researching listening* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315732862>
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158. <https://doi.org/10.1093/applin/11.2.129>
- Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 382-408). Blackwell. <https://doi.org/10.1002/9780470756492.ch13>
- Serafini, F., Youngs, S., & Gregory, A. (2008). Supporting boys as readers. *Reading Teacher*, 62(5), 394-403. <https://doi.org/10.1598/RT.62.5.3>
- Stockwell, G. (2007). A review of technology choice for teaching language skills and areas in the CALL literature. *ReCALL*, 19(2), 105-120. <https://doi.org/10.1017/S0958344007000225>
- Sun, Y., & Li, J. (2023). Enhancing English listening comprehension through mobile audiobooks: A study with Chinese EFL learners. *Computer Assisted Language Learning*, 36(3), 456-478. <https://doi.org/10.1080/09588221.2021.1946563>
- Sussman, T. J., & Kane, M. J. (2015). The role of working memory capacity in listening comprehension. *Journal of Educational Psychology*, 107(2), 349-363. <https://doi.org/10.1037/a0037763>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123-138. <https://doi.org/10.1007/s10648-010-9128-5>
- Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge. <https://doi.org/10.4324/9780203843376>

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.
<https://doi.org/10.1093/applin/aml048>
- Whittingham, J., Huffman, S., Christensen, R., & McAllister, T. (2013). Use of audiobooks in a school library and positive effects of struggling readers' participation in a library-sponsored audiobook club. *School Library Research*, 16, 1-19.
- Wiriyachitra, A. (2002). English-language teaching and learning in Thailand in this decade. *Thai TESOL Focus*, 15(1), 4-9.
- Yeh, C. C. (2014). The impact of multimedia reading materials on L2 reading: Learners' perception and performance. *English Language Teaching*, 7(7), 38-45. <https://doi.org/10.5539/elt.v7n7p38>