



**JOURNAL OF INFORMATION  
SYSTEM AND TECHNOLOGY  
MANAGEMENT  
(JISTM)**


[www.gaexcellence.com/jistm](http://www.gaexcellence.com/jistm)




## **THE IMPACT OF COMPUTERIZED ADVANCE ORGANIZERS ON LISTENING COMPREHENSION: A SCOPING REVIEW**


Sangeetha Thiarajah<sup>1</sup>, Hema Vanita Kesevan<sup>2\*</sup>


<sup>1</sup>Fakulti Bahasa dan Komunikasi, Universiti Pendidikan Sultan Idris Malaysia

 [p20241000162@siswa.upsi.edu.my](mailto:p20241000162@siswa.upsi.edu.my)

 <https://orcid.org/0009-0001-8277-555X>

<sup>2</sup>Professor, Faculty of Arts and Science, International Universiti Malaya-Wales, Malaysia,

 [hemakesevan@fbk.upsi.edu.my](mailto:hemakesevan@fbk.upsi.edu.my)

 <https://orcid.org/0000-0002-6731-451X>

\*Corresponding Author

### **Article Info:**

#### **Article history:**

Received date: 28.04.2026

Revised date: 17.05.2026

Accepted date: 11.06.2026

Published date: 30.06.2026

#### **To cite this document:**

Thiarajah, S., & Kesevan, H. V. (2026). The Impact Of Computerized Advance Organizers On Listening Comprehension: A Scoping Review. *Journal of Information System and Technology Management*, 11 (43), 182-204.

**DOI:** 10.35631/JISTM.1143011

### **Abstract:**

The integration of advance organizers in a computerized setting has renewed research interest in language education as institutions shift toward technology-enhanced learning environments. However, despite their growing relevance, the pedagogical impact of Computerized Advance Organizers (CAOs) on English as a Second Language (ESL) or English as a Foreign Language (EFL) learners' listening comprehension remains insufficiently understood. Therefore, a scoping review was conducted to examine the extent, range, and nature of existing research on the use of CAOs and CAOs aligned instructional supports in ESL and EFL listening contexts. The databases, Scopus and Education Resources Information Center (ERIC) were used to gather a comprehensive body of empirical evidence. The six-stage methodological framework of Arksey and O'Malley guided the review process. Findings indicate that CAOs positively contribute to learners' listening comprehension by activating prior knowledge, reducing cognitive load, and providing structured pre-listening support. Nevertheless, results also reveal significant variations in technological integration across studies. The current state of research suggests that a more systematic understanding of CAO-based interventions is still emerging, particularly in virtual and primary ESL or EFL settings. This review may assist educators, researchers, and curriculum developers in strengthening CAO-based listening pedagogy and identifying directions for future investigation.

### **Keyword:**

Computerized Advance Organizers, ESL/EFL Listening Comprehension, Pre-listening Support, Scoping Review



© The authors (2026). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [jistm@gaexcellence.com](mailto:jistm@gaexcellence.com).

## Introduction

The increasing integration of digital technologies into classrooms has reformed how listening skills are imparted and learned. In technology-enhanced learning environments, listening comprehension often involves learners to process spoken input alongside visual and textual information (Shamsi & Bozorgian, 2024). While such environments offer rich learning opportunities, Teng (2022) added that learners need to go through substantial cognitive processing demands during listening. Empirical findings show that learners experience difficulties related to speed of speech, prosody, and processing spoken input during listening tasks (Shamsi & Bozorgian, 2024). This complexity also applies when it comes to computer-based delivery of listening tests in assessment contexts. In such situations, the students have to process the auditory information in addition to the visual information provided by the speaker and the slide presentation, which can affect their test results (Kwon & Yu, 2024). Shamsi and Bozorgian (2024) referred that the collaborative metacognitive teaching and learning with the use of multimedia input can alleviate the learners' difficulties.

Apart from that, research on listening with the support of multimedia shows that the use of audiovisual and captioned materials can increase comprehension and vocabulary learning (Pujadas & Muñoz, 2024; Teng, 2022). The literature also warns about the negative aspects of multimedia input. The impact of visual cues on listening, for example, is mixed, with some cues helping to understand. While some may distract learners, others may distract learners based on the way in which they are approached during the task (Kwon & Yu, 2024). Likewise, if the elements in subtitle or video-based interventions are not well organized or well designed, it will result in less effectiveness of the interventions (Mulyani et al., 2022). These results indicate that technology is not a sufficient means for better listening comprehension. Instead, instructional design is a must. To address this observation, 15 reviewed articles have been analyzed with the conceptual lens of Computerized Advanced Organizers (CAOs) to look at how technology-based scaffolds can be systematically organized to support the learner before listening. This term, "computerized," is used in this review to highlight the fact that many of today's educational scaffolding are delivered in a digital format, are multimodal and are embedded in technology-mediated learning environments while grounded in Ausubel's (1960) advanced organizer theory.

For the purposes of this review, CAOs are considered to be numerically represented as digital pre-instructional supports which help learners activate previous knowledge in preparation for future learning activities and ensure cognitive readiness. These supports can include visual previews, graphic representations, textual introductions, multimedia prompts, and other digitally mediated instructional supports that are presented prior to learning. Some of the interventions in the studies reviewed might not be considered traditional interventions in the theoretical sense of advanced organizers, but have similar instructional goals to prepare learners for future learning activities. Therefore, for this review, CAOs are defined in terms of

their instruction function, rather than by their technological form. In addition, studies that addressed the more general language learning uses of CAOs were included when these studies offered theoretical or pedagogical insights that were applicable to the preparation for listening, the process of listening comprehension, the readiness of the language learner and the activation of the language knowledge. This wider view allows for a deeper understanding of the role of CAOs in today's technology-enhanced language learning contexts.

Based on this definition, CAOs are defined as digital pre-listening tools, including vocabulary preview, guiding questions, visual representation, transcript or concept map, that are intentionally designed to elicit prior knowledge and set up a conceptual framework for the incoming input for the purpose of this review. As seen in Table 1, Wang and Hsiao (2022) and Kajiura et al. (2025) clearly implemented the advance organizer framework, which sees instructional materials like text, graphics, and pre-listening script as scaffolds to be used before listening. Advance organizers (AOs) were developed in Wang and Hsiao (2022) to help comprehension and to alleviate cognitive load by providing vocabulary, sentences and visual images before watching the video. At the same time, Kajiura et al. (2025) found that reading the transcript before listening acts as an "advance organiser" which "wakes up" the prior knowledge and helps to predict what to listen to. In addition to these specific uses, the literature cited in these studies mentions other multimodal supports such as captions, transcripts and visual cues. But, their effectiveness is dependent on the design and integration. Notably, these supports have been identified in both studies as supports that occur prior to listening, and not concurrent or post-listening, highlighting the importance of these supports in preparing learners for the input. Overall, the results indicated that the use of multimedia elements alone does not influence the improvement of students' listening comprehension skills. Instead, they rely on how much such supports are routinely designed as advanced organizers that will activate prior knowledge, focus attention and diminish the cognitive load during listening.

While the research above shows some interesting results, research on CAOs in ESL listening remains disjointed. Empirical research shows that there are a lot of variations in instructional design, such as the type of support materials used (textual, visual, or multimodal materials), learners' proficiency level, and the outcome criteria used to measure learners' comprehension (Wang & Hsiao, 2022; Pujadas & Muñoz, 2024). Some studies found positive impacts of CAO-related supports like transcripts and text-based scaffolds on listening comprehension and meaning construction, while Aldukhayel (2023) and Kajiura et al. (2025) pointed out that the impact is conditional and depends on the structure and delivery of the support. For example, Aldukhayel (2023) found that using comments as an extra language support tool on social media led to a significant increase in the number and accuracy of students' summarization of their comprehension. In the same vein, Kajiura et al. (2025) found that a deferred multimodal transcript support enhanced listening performance and adaptation to the rapid speech, decreasing the cognitive load and enhancing the sound-meaning integration.

The use of advanced organizers and multimodal supports is effective based on learners' proficiency and cognitive processing capacity (Wang & Hsiao, 2022; Pujadas & Muñoz, 2024). For instance, Wang and Hsiao (2022) found that the cognitive load of advanced organizers was beneficial to the comprehension of higher-proficient learners but might create an excessive cognitive load to the comprehension for lower-proficient learners when more advanced organizers were added. Furthermore, Pujadas and Muñoz (2024) evaluated the effect of captioned input on comprehension, and it was found that the impact of captioned input is more related to the level of proficiency and the extent of vocabulary knowledge. Moreover, most of

the research has been carried out in the EFL context among secondary or tertiary students, and there has been little synthesis of studies conducted in different ESL contexts and levels of education (Wang & Hsiao, 2022; Pujadas & Muñoz, 2024). Thus, it is challenging to draw conclusions that can be generalized in other classroom contexts. This lack of consolidated evidence poses challenges for teachers when they try to use a CAO-based intervention in the digital or hybrid ESL classroom. With the growing use of digital learning in educational practice, the need for synthesising existing research is apparent to clarify good design principles, identify regularities and patterns, and identify gaps for future research. In this regard, the mapping of the state of the art of empirical studies on CAOs in ESL/EFL listening education is carried out using the scoping review approach. The main purpose is to explore the impact of CAOs on ESL/EFL learners' listening comprehension results, as well as to find thematic trends in the design and implementation of CAOs. In addition, this review attempts to identify the trends, inconsistencies, and emerging directions which can guide future research and facilitate more effective incorporation of CAOs in current ESL/EFL classrooms. This article has the following structure. The literature reviewed in part 1 explores the research on CAOs and ESL/EFL listening comprehension. Part 2, in the meantime, provides the methodological steps for the scoping review. The results are presented and discussed in part 3. Part 4 concludes with limitations, implications for practice, recommendations for future research, and final reflections.

## Materials and Method

This study adopted a scoping review method to map the scope, characteristics, and thematic trends of empirical research on different types of CAOs in ESL/EFL listening contexts. A scoping review is appropriate for this area of inquiry. This is because research on technology-enhanced listening scaffolds spans diverse digital tools, instructional designs, and outcome measures. Recent empirical studies have examined a range of listening support mechanisms. For instance, advance organizers in video-based listening have been shown to function as scaffolding, helping learners connect prior knowledge with new input and reduce cognitive load (Wang & Hsiao, 2022). Similarly, transcript-based pre-listening support, particularly when presented prior to listening, has been determined to reduce cognitive load and facilitate comprehension by allowing learners to use prior information to process incoming auditory input more efficiently (Kajiura et al., 2025). Moreover, captioned audiovisual input has been consistently linked to improvements in comprehension outcomes, as the on-screen text helps to decode and comprehend spoken language, especially for learners with difficulties in processing the continuous speech (Pujadas & Muñoz, 2024). It has also been found that captioned input can improve the results of learning vocabulary, and the students who were exposed to captioned input had significantly higher vocabulary learning scores than the students who were not exposed to captioned input (Teng, 2022).

In addition to text-based scaffolds, there has been recent research on interactive digital scaffolds. For instance, the comments left on social media posts on videos have been proven to improve listening comprehension, where students make more accurate and detailed summaries after listening to videos with the presence of social media comments (Aldukhayel, 2023). Similarly, social annotations can enhance students' learning engagement and understanding by allowing them to make comparisons, adjust their learning strategies, and connect with their previous knowledge (Lin & Wei, 2024). Overall, these studies suggest that listening can be supported through the use of digitally mediated supports to help students process the input and gain access to content. For example, the use of multimedia along with metacognitive instruction

has been found to enhance listening comprehension through planning, monitoring and evaluating (Shamsi & Bozorgian, 2024). Furthermore, a multi-modal setting, which combines visual, auditory, and textual inputs, offers multiple representations that help in understanding and meaning-making (Shamsi & Bozorgian, 2024). Since these studies vary in terms of the research design, target group of learners and digital tools used, a scoping review is conducted to comprehensively map these studies and serve as a basis for identifying patterns and gaps in research related to CAO. A six-stage scoping review process was employed, which included: (1) identifying research questions, (2) identifying relevant studies, (3) selecting studies for inclusion, (4) charting the data, (5) collating and summarising the results, and (6) discussing the findings. This process helped to maintain transparency and methodological rigor in the review.

### ***Step 1: Identifying the Research Questions***

This scoping review's primary purpose was to explore the design and implementation of CAOs to facilitate ESL/EFL learners' listening comprehension. Previous studies have shown that the use of multimedia input and structured multimodal support can improve listening comprehension and engage learners through both auditory and visual inputs (Shamsi & Bozorgian, 2024; Kajiura et al., 2025). Furthermore, pre-listening or prior knowledge activation using multimodal inputs, such as reading the transcript before listening, can help to reduce cognitive load and facilitate comprehension by supporting the mapping of sounds and meaning, and predictive processing (Kajiura et al., 2025). Moreover, visual-verbal tools like concept mapping can be used to organize and integrate information, which can help to improve multimodal learning and communication processes (Tsai, 2025).

The following research questions were therefore addressed in the review:

1. How do CAOs influence ESL/EFL learners' listening comprehension outcomes?
2. What thematic patterns emerge in the design and implementation of CAOs across digital listening contexts?
3. What future research directions are identified in the literature for CAOs in ESL/EFL listening contexts?

These questions guided the systematic mapping of empirical evidence and framed the synthesis of findings across studies.

### ***Step 2: Identifying Relevant Studies***

A systematic search strategy was used to ensure a comprehensive review of literature, which followed the guidelines of Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) and the five-stage framework proposed by Arksey and O'Malley (2005). Systematic searches were conducted on two known academic databases, Scopus and ERIC, that include high-quality journals in the field of applied linguistics, language education, and educational technology. The search was done to identify studies published in the last five years (2022-2026) that focused on the use of computerized advanced organizers (CAOs) and other preparatory instructional supports in the context of listening comprehension and language learning. The search was directed by the terminologies used in the previous studies about the advanced organizers and listening comprehension. The search string used was: (("computerized organizer" OR "advance organizer" OR "graphic

organizer" OR "visual organizer") AND ("listening comprehension" OR "auditory understanding" OR "listening skills" OR "listening ability"))

Search results were downloaded from both databases, and before screening, duplicate records were eliminated. Then, a screening was done on the retrieved studies according to the pre-defined inclusion-exclusion criteria. A PRISMA flow diagram was used to document the screening and selection process to give a clear description of the identification, screening, eligibility, and inclusion process. Through this process, empirical studies examining the use of different types of CAOs and related pre-listening supports for ESL/EFL learners' listening comprehension were identified and included in the review.

### ***Step 3: Selecting Studies for Inclusion***

The study selection process involved three stages: title screening, abstract screening, and full-text review. Retrieved records were assessed against predetermined inclusion and exclusion criteria to ensure consistency and alignment with the objectives of the scoping review. Studies were included if they (a) were peer-reviewed journal articles published between 2022 and 2026, (b) reported empirical findings, (c) were conducted within language-learning contexts, and (d) examined computerized or digitally mediated instructional supports designed to prepare learners before engaging in learning tasks. Eligible studies also reported outcomes related to language learning, including comprehension, vocabulary acquisition, learner engagement, cognitive processes, or metacognitive development. Studies that did not involve a preparatory instructional intervention, were not situated within language-learning contexts, lacked empirical data, or were published as conference proceedings, theses, book chapters, editorials, or review articles were excluded. Studies that did not involve either the use of a computer instructional support or learning preparation were also not included.

The main criteria for this review were that the intervention in the studies be a CAO, but not every study specifically labeled their intervention as an advanced organizer. Hence, studies were included if the intervention met the basic functional features suggested by Ausubel (1960), which are the activation of prior knowledge, the introduction of a conceptual framework, organization of the incoming information, or preparation for further learning. This decision was made to reflect the recent advances in technology-based language learning, in which language learning supports are often contained within a multimedia context and might be referred to by other names, like captions, transcripts, vocabulary preview, guiding questions, visual support, social annotation tools, and metacognitive prompts. Therefore, the review focuses on the instructional role of the intervention as opposed to just the language that was used by the original authors. Moreover, research on general language learning outcomes such as vocabulary learning, reading comprehension, pronunciation development, learner engagement, social annotation and metacognition development was kept if it included theoretical or pedagogical insights that could be applied to the preparation and comprehension of language learning. This holistic approach allowed the review to identify new varieties of digital preparatory supports that could help language learning in ways that are compatible with advanced organizer theory. The screening and selection process was done by the author as per the pre-defined eligibility criteria. Figure 1 summarizes the final study selection process.

## ***Quality Appraisal***

The methodological quality or effectiveness of individual studies was not assessed, as the purpose of this scoping review was to map the extent, nature and key trends of the evidence to date on CAOs. This is in line with the methodological framework proposed by Arksey and O'Malley (2005) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines, which highlight the exploratory nature of scoping reviews. Therefore, methodological quality was not a criterion for inclusion or exclusion, and all studies that met the pre-defined eligibility criteria were included. The results should therefore not be seen as a reflection of the comparative methodological quality of the individual studies, but rather as an amalgamation of the evidence available.

### ***Step 4: Charting the Data***

After the selection of studies, the relevant data were systematically extracted and organized in MS. Excel. Information extracted from the publications was the context of the teaching, pedagogical design elements, support type and reported learning outcomes. To allow for cross-study comparisons, recurring patterns related to listening support and multimodal mediation were documented as well. The charting process involved determining where instructional supports were located with respect to the listening task, with a specific focus on how the supports were used to activate prior knowledge, to decrease processing demands, to provide metacognitive guidance, and to include multimodal supports, including text, visuals, and structured representations. These analytical dimensions are not pre-determined but are of an inductive nature and are based on the included studies. For example, Wang and Hsiao (2022) showed that advance organizers (AO) in the form of text-based and text-plus-visual supports before the video viewing can be used as scaffolding to help the learners comprehend the video and to reduce their cognitive load. Likewise, Kajiura et al. (2025) evaluated the effect of deferred multimodal input, where the learner reads the transcript and then listens to it, as it reduces split attention and allows the learner to leverage the previously created linguistic representation. Shamsi and Bozorgian (2024) demonstrated the benefits of using multimedia input in listening comprehension in another instructional setting by providing metacognitive instruction and engaging students in collaborative and strategic regulation. Since there was some variation in the focus of instruction across studies, charting focused on the pedagogical functions of each of the supports and did not assume the concepts to be the same across contexts. This facilitated an accurate representation of the way in which the various kinds of instructional support are related to listening and learning outcomes, which fed into the subsequent synthesis.

### ***Step 5: Collating, Summarizing, and Reporting the Results***

The data collected were then analyzed systematically, using an inductive approach to look for patterns and relationships in the studies reviewed. Instead of categorizing the instructional supports, the synthesis identified three major and interwoven dimensions of the role of instructional supports in listening and language-related learning: cognitive support, metacognitive regulation, and multimodal mediation. Throughout the literature, the instructional supports were consistently found to be related to the facilitation of comprehension, with the reduction of processing demands, the direction of learners' attention to input and the facilitation of the interaction between verbal and visual information all identified as being relevant to the facilitation of comprehension. Cognitively, supports before

or beside listening tasks supported the activation of relevant knowledge structure and cognitive load management. For instance, Wang and Hsiao (2022) found that advance organizers facilitated listening comprehension in video-based contexts and reduced cognitive load. Likewise, Kajiura et al. (2025) found that transcript-supported listening (sequential format) significantly contributed to enhancing listeners' listening proficiency and their ability to adjust to the fast speech rate.

Metacognitively, the use of explicit instruction of the listening strategies was very important in influencing the students' strategy in listening. Shamsi and Bozorgian (2024) showed that the metacognitive instruction integrated into the listening process using multimedia facilitated the listeners to plan, monitor and evaluate their comprehension process, especially in a collaborative learning environment. Further, the findings indicated the significance of multimodal mediation, in which the combination of textual, auditory and visual elements was important for making the input more accessible and meaningful. Although not every study focused on listening comprehension alone, some studies showed that visual structuring tools like concept maps can be used to strengthen the visual-verbal coherence in integrated language tasks and help the learners to organize and communicate the information, as evidenced by Tsai (2025). The results of these studies, when combined, indicate that technology alone does not indicate the effectiveness of instructional supports in listening contexts. Instead, it relies on the pedagogical organisation of these supports to match the cognitive processing and regulation of learners. Table 1 shows the synthesis of these patterns.

## Findings

The initial search of the Scopus and ERIC databases yielded 634 records (572 from Scopus and 62 from ERIC). The search strategy employed the following keywords: (computerized organizer OR "advance organizer" OR "graphic organizer" OR "visual organizer") AND ("listening comprehension" OR "auditory understanding" OR "listening skills" OR "listening ability"). Following a preliminary screening of titles and abstracts, 577 records were excluded for failing to meet the predetermined inclusion criteria. Exclusions were based on publication type (e.g., systematic reviews, meta-analyses, books, and book chapters), language (non-English articles), publication date (articles published before 2022), and subject area (articles not within the social sciences). This screening process resulted in 57 records eligible for full-text review. After removing one duplicate, the full texts of 56 articles were assessed for eligibility. A further 41 articles were excluded for the following reasons: lack of empirical data, conference proceedings, focus on non-educational hard sciences, or lack of alignment with the core review focus on CAOs and listening-related language outcomes. Consequently, 15 articles were included in the final synthesis. A final quality assessment of these articles was conducted to ensure their methodological rigor and relevance to the research questions. This selection process is summarized in the PRISMA flow diagram (Figure 1).

### *Article Selection Rationale and Screening Criteria*

This scoping review used a well-defined set of criteria to guide the selection of articles that met the requirements of relevance, quality and contemporaneity in the field of education.

### ***Source and Publication Type***

Only peer-reviewed journal articles were included in the review, and they were obtained from the major academic databases (Scopus and ERIC). This was done to provide access to research that was of good quality, vetted and to capitalize on the full and multi-disciplinary indexing of such research.

### ***Research Methodology***

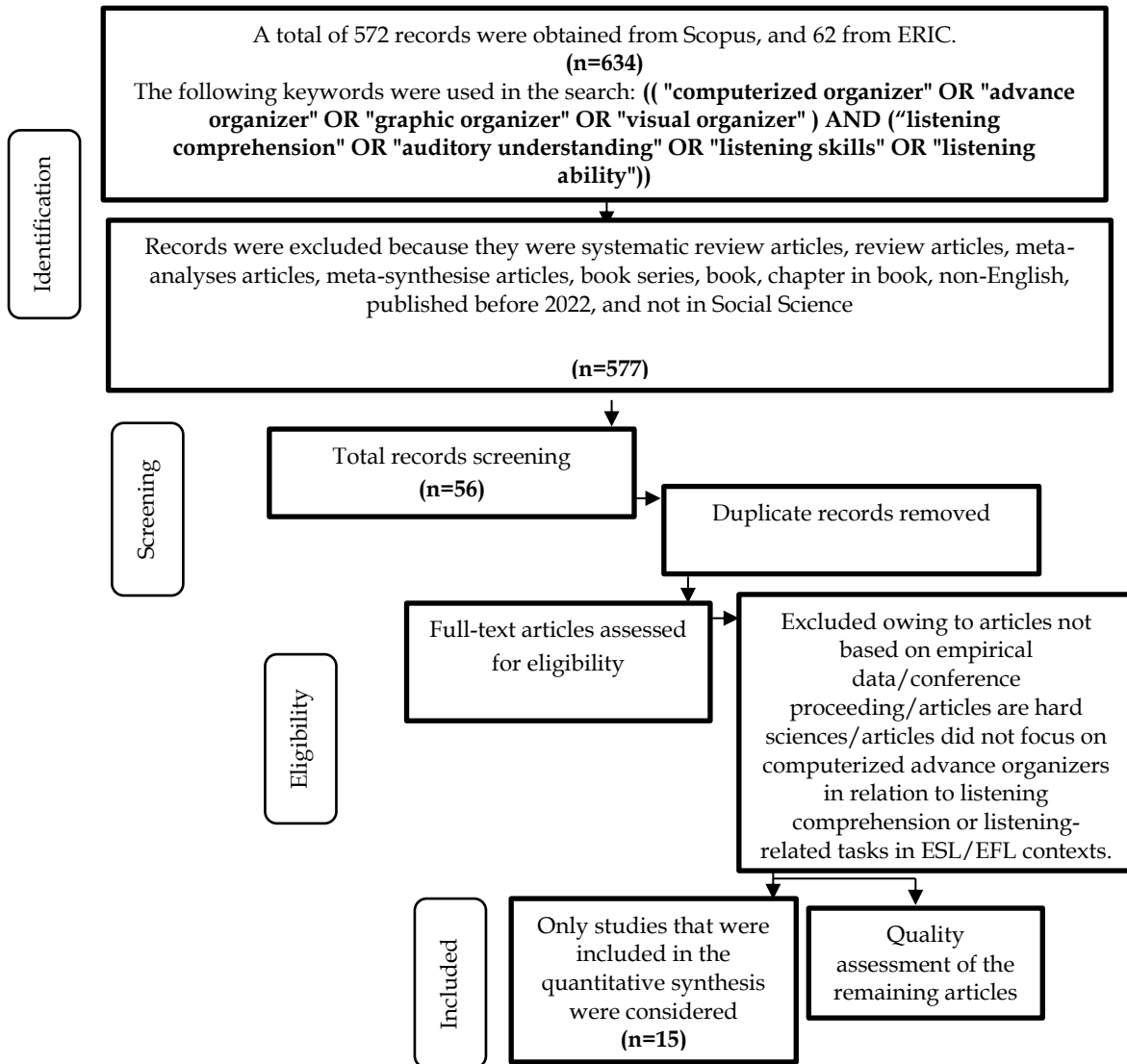
This review focused on studies that used quantitative research methodologies to give a structured and comparable analysis of impact. Although mixed methods and qualitative designs may be valuable, quantitative studies provide a more concrete foundation for evidence synthesis regarding the impact of CAOs on outcomes of listening comprehension.

### ***Language and Empirical Focus***

It is important to note that only empirical research articles in English were reviewed. English was chosen as the language of international academic publication, allowing a wider synthesis of the research contexts around the world. The need for empirical data meant that studies that presented theoretical arguments and/or commentary were excluded.

### ***Publication Timeline***

The review emphasized articles that were published in the last five years (2022 to the present) to include the recent advancements and technological implementations. This timeframe reflects current research trends in the rapidly evolving domain of technology-enhanced language learning.



**Figure 1: Flow Diagram of Research Selection Process Using PRISMA. Adapted from a Study by Moher et al. (2015)**

## Main Findings

Table 1 presents the 15 studies included in this scoping review based on the criteria stated above. The analysis revealed that the use of CAOs in ESL listening instruction provides substantial benefits for learners' comprehension and engagement in multimedia-supported learning contexts, directly addressing the research questions. From the analysis, CAOs are examined as instructional scaffolds that support learners' cognitive and metacognitive preparation for listening. Shamsi and Bozorgian (2024) added that CAOs enhance listening comprehension by guiding learners' engagement with multimodal input (Shamsi & Bozorgian, 2024). Specifically, research on CAOs encompasses multiple aspects of learner development that contribute to improved comprehension outcomes. CAO-based approaches are seen to facilitate the activation of background knowledge, the linkage of old knowledge with new information, and to help the learners focus on the relevant information before listening (Wang & Hsiao, 2022; Kajiura et al., 2025). These processes relate to better listening readiness and more focused listening engagement. CAOs are consistent in all the studies and highlight the

need for structured preparation and support learners in coping with the cognitive load and processing demands in listening (Wang & Hsiao, 2022; Kajiura et al., 2025). These strategies also help students to be engaged and motivated to learn in the context of multimedia learning environments through digital tools like multimedia input and interactive tools, and can be helpful in effective comprehension processes (Shamsi & Bozorgian, 2024; Lin & Wei, 2024). Moreover, visual tools like concept mapping help organize and present knowledge in multimodal contexts, which further aids learning processes (Tsai, 2025).

Themes and subthemes were also created from the scoping review to answer the second research question. Four major themes were identified, which can be broken down into several subthemes. Cognitive Activation (schema activation and vocabulary readiness) is the first theme, which means that learners are to be prepared with necessary linguistic and conceptual supports before listening (Wang & Hsiao, 2022; Kajiura et al., 2025). The second theme is Metacognitive Support (planning, monitoring, and evaluating), which refers to the support from instruction and learning strategies that help the learners manage their comprehension processes in the listening task (Shamsi & Bozorgian, 2024; Lin & Wei, 2024). The third theme is Affective Facilitation (motivation, confidence and reduced anxiety), which is related to supportive learning environments and interactive features that promote positive attitudes, motivation and engagement in listening tasks (Aldukhayel, 2023; Lin & Wei, 2024). The last theme is Technological Mediation (Multimedia Enhancement, Learner Control, and Interactivity), which emphasizes the role of digital aspects like multimedia input and visual cues in enhancing engagement and promoting comprehension, although their impact can differ depending on the type of task and learner factors (Tsai, 2025; Kwon & Yu, 2024). The summary of the included studies and mapping of themes and subthemes are outlined in Table 1 below.

**Table 1: Charting the Data of the 15 Articles**

Study	CAO / Design Feature	Key Outcome	Theme	Subtheme	Research Gap
<b>Kajiura et al. (2025), Japan</b>	Deferred multimodal input: transcript reading followed by normal-rate or fast-rate audio	Improved listening proficiency and adaptation to fast speech, especially with transcript support before fast-rate audio	Cognitive Activation	Sequential text-audio priming	Short 5-day intervention; tertiary EFL learners only; no metacognitive awareness measure; not tested in primary or virtual classrooms
<b>Wang &amp; Hsiao (2022), Taiwan</b>	Advance organizers before video: text aids only vs text aids + visual aids	Reduced cognitive load; stronger benefits for high-proficiency learners; positive perceptions of Advance Organizers (AOs)		AO modality & cognitive load reduction	Effects varied by proficiency; possible overload for lower-proficiency learners; no learner-control

					feature; not a virtual-classroom study
<b>Aldukhayel (2023), Saudi Arabia</b>	Social media comments are used as support while viewing YouTube videos	Better comprehension, reflected in more accurate and informed summaries; positive perceptions of comments as a help option		Lexical priming	Subtheme fit is weaker than the others because the study is about comment-based support, not true lexical priming; no standard listening test; no long-term or metacognitive data
<b>Tsai (2025), Taiwan</b>	Concept mapping in integrated listening-speaking tasks: expert-generated vs learner-generated maps	Improved multimodal communicative competence; learner-generated maps particularly supported organization and verbal-visual coherence		Learner-generated visual organization	Listening was not isolated as a standalone outcome; tertiary setting only; no primary or virtual context
<b>Shamsi &amp; Bozorgian (2024), Iran</b>	Multimedia input combined with explicit metacognitive instruction and collaborative listening	Improved listening comprehension and strategy use; learners reported positive views of multimedia-supported collaborative listening	Metacognitive Support	Strategy-based CAO framework	Small-scale qualitative/case-study design; difficult to isolate the independent effect of the scaffold from collaboration and metacognitive instructions (MI); limited generalisability
<b>Pujadas &amp; Muñoz (2024), Spain</b>	Captioned vs Uncaptioned television viewing across proficiency levels	Captions improved comprehension; the additive caption benefit was no longer significant at C2-level.		Adaptive support regulation	Subtheme is interpretive rather than explicit because learner control was not manipulated; university learners only;

					no younger learners; no direct metacognitive measure
<b>Kositthaivat (2025), Thailand</b>	Graphic organizers integrated with task-based learning	Significant improvement in listening and speaking proficiency, including Common European Framework of Reference for Languages(CEFR) progression	Affective Facilitation	Motivation through structured scaffolding	One-group pre-test/post-test design; mixed listening-speaking outcomes; no online/virtual comparison
<b>Lin &amp; Wei (2024), Macau/China context</b>	Social annotations on (Massive Open Online Course (MOOC) videos; learners could move among no visual aids, captions, and social annotation modes	Higher engagement and motivation; annotations supported comprehension checking and help-seeking		Socially mediated CAOs	Very small sample; MOOC-specific context; engagement emphasized more than direct performance gains
<b>Mulyani et al. (2022), Indonesia</b>	Bimodal English-Indonesian subtitles in video clips; perception-focused design	Learners generally reported positive perceptions, but poor subtitle design caused difficulty and frustration		Design quality & affect	No direct listening-comprehension outcome in this paper; emphasis on vocabulary support and learner perceptions
<b>Alkhalaf (2023), Saudi Arabia</b>	Academic lecture input across modes: listening, reading, reading-while-listening, viewing, and viewing with captions	Significant gains in incidental collocation learning, especially from reading, viewing, and viewing with captions	Technological Mediation	Multimodal input delivery	Collocation learning rather than listening comprehension was the main outcome; higher-education context only; no metacognitive measure

<b>Lo (2024), Macao SAR</b>	Dual-subtitled videos with immediate repetition, spaced repetition, or no repetition	Repetition improved vocabulary gains, with immediate repetition outperforming spaced repetition		Repetition & timing control	Vocabulary-focused rather than listening-focused; university EFL learners only; no strategy or metacognitive measure
<b>Helal et al. (2024), Saudi Arabia</b>	Word-based captions vs phonetic transcription captions	Both caption types improved stress pronunciation; no statistically significant difference between groups		Caption engineering	Pronunciation, not listening comprehension, was the main outcome; small sample; tertiary setting only
<b>Teng (2022), China</b>	Captioned vs non-captioned documentary viewing	Captioned viewing significantly improved incidental vocabulary learning; learner-related factors also influenced gains		Caption-supported input	Vocabulary learning was the main outcome; no metacognitive measure; future work should compare captioning types and mixed methods
<b>Teng (2025), China</b>	Bilingual-captioned videos for incidental learning of multiword expressions	Bilingual captioning facilitated Multiword Expression (MWE) learning; vocabulary knowledge and working memory significantly predicted gains		Dual-language redundancy	MWE learning rather than listening comprehension was measured; university preparatory students only; no direct metacognitive outcome
<b>Kwon &amp; Yu (2024), Korea</b>	Visual cues in a video-based listening test were examined through eye-tracking and stimulated recall	Mixed effects of visuals on attention and performance; some cues helped, but looking at speakers/distractors could relate to lower scores		Visual cue processing	Assessment study rather than classroom intervention; mixed findings; limited to a testing context

## ***Synthesis of Findings Based on Charted Studies***

A review of the charted studies indicates that CAOs and associated technology-mediated supports are not used in the same way in ESL/EFL environments. They are typically thought of as scaffolds to organize the way in which learners interact with oral input, but the pedagogical roles of these can differ according to the design and modality, as well as how they are integrated into the instruction. More significantly, the effectiveness of these supports in directly improving listening comprehension is inconsistent, with many studies focusing on the overall language outcomes instead of listening as a separate construct (Teng, 2022; Teng, 2025).

## ***Cognitive Activation Through Pre-Listening Structuring***

Research that falls under the heading of Cognitive Activation suggests that pre-listening structuring may be useful to ease the processing demands of the initial phase of listening comprehension in order to elicit learners' prior knowledge in the pre-listening phase to improve their initial understanding of the spoken text. For example, transcript-based sequencing has been proven to be beneficial in listening development as it allows students to combine information from the text with the information they hear. This helps with sound-meaning mapping and enhances listening skills (Kajiura et al., 2025). Likewise, a combined advance organizer of visual and textual information is found to facilitate comprehension, reduce cognitive load and offer structure scaffolding before listening tasks with a video (Wang & Hsiao, 2022). But the benefits are not equal for all learners; the effectiveness is dependent upon proficiency. Some learners have increased cognitive load due to the non-optimal alignment of the supports and their abilities. Besides pre-structured organizers, some texts, such as those in social media that contain explanations, summarization, and elaboration of the video content, have also been shown to help with comprehension (Aldukhayel, 2023).

However, this type of input tends to be more of an auxiliary and not a systematic pre-listening organizer, indicating significant conceptual differences between incidental help and systematic advance organizers. Even with these contributions, there are some drawbacks. One of the significant issues is that comprehension is assumed from output measures such as written summaries instead of verified listening assessments (Aldukhayel, 2023). These constrain the scope of identifying improvements as a result of listening processes, as per Aldukhayel (2023). In addition, previous research has mainly been conducted with secondary or tertiary students, and little empirical research is available on younger age groups. Lastly, cognitive activation is often thought of as a stand-alone preparatory stage, and there is not enough consideration of the more general metacognitive processes. Consequently, the idea of pre-listening support is often treated separately from the idea of listening, as if it were not part of a strategy-driven listening framework (Kajiura et al., 2025).

## ***Metacognitive Support and Strategy Regulation***

Although metacognitive support is relevant to listening, research on metacognitive support is still relatively underdeveloped as compared to cognitive activation. Evidence shows that CAOs can affect the understanding of learners and their regulation of the listening processes if they are integrated into the metacognitive teaching. Structured guidance in planning, monitoring, problem-solving, and evaluation has been found to benefit listening comprehension and strategic engagement, in particular (Shamsi & Bozorgian, 2024). However, the methodological

limitations of this evidence will be a limitation in the strength of this evidence. The research design of the present study was qualitative, and the case study method was used with a small sample ( $N = 20$ ). It is based on observation, learner logs and interviews as opposed to controlled experimental measures. Consequently, although gains are made in listening comprehension, the success is often closely linked to the support of collaboration, interaction and the use of metacognitive strategies together, and it is difficult to determine the independent effect of CAO-like supports or the multimedia. In addition, the study environment is one that is teacher-led and collaborative, with the learners being supported throughout the process by the teacher through the different stages (prediction, verification, evaluation, etc.) and not being in a fully autonomous or learner-controlled digital environment. This implies that development of metacognition occurs as a result of guided interaction and joint regulation and not as a result of self-regulation. Another limitation is the narrow participants' profile: the subjects of the study are upper-intermediate, female university-level EFL learners, which means that the results of the study cannot be generalized to other groups of subjects. Interestingly, there is no evidence for younger or primary level learners, suggesting that the use of support through metacognitive CAO in early ESL contexts is under-researched.

### ***Affective Facilitation and Learner Engagement***

The contribution of CAOs in relation to the affective aspects of learning is somewhat supported by empirical evidence, but to a limited extent. The study conducted by Lin and Wei (2024) revealed that social annotations can enhance learners' motivation and engagement in multimedia learning environments, as evidenced by their viewing time, pause count, and the number of comments and questions they made. Moreover, social annotations allow learners to track their comprehension by observing others' annotations and comparing them with their own, which encourages active participation in the learning process. It is worth emphasizing that the results are mostly observation-based and qualitative interview data, and not actual listening comprehension performance data. This separation does indicate a concern for the connection between engagement and learning outcomes. Although the learners' engagement and motivation are increased with social annotation, the study fails to prove that the presence of social annotation results in measurable gains in listening comprehension. Moreover, this study took place in an online learning environment, a Massive Open Online Course (MOOC) with a small number of students, and the results can only be generalised to formal classroom environments. In general, learners find social annotations helpful and interesting, but there is no clear indication that there is a correlation between affective responses (motivation and participation) and validated gains in understanding. This suggests that there is a disassociation between affective involvement and cognitive achievement, warranting more empirical studies to be conducted to directly measure the listening comprehension in the context of CAO-based interventions.

### ***Technological Mediation and Multimodal Design***

Technological mediation and multimodal design studies are indicative of a transformation towards the embedding of CAOs in a digitally mediated learning environment. In this strand, supports like captions, subtitles and multimodal input configurations serve to spread linguistic information between the auditory and the visual channels to access spoken input. Empirical evidence verifies that captions help to comprehend measurably, as they enhance the chance of giving a correct answer in listening tasks. However, there are differences in this advantage depending on the level of proficiency of the learners (Pujadas & Muñoz, 2024). In particular,

the captioned viewing has been proven to be very effective in improving comprehension results over the uncaptioned condition. However, this benefit is lost at higher proficiency, indicating that there is a point at which captioning is no longer a useful aid to understanding (Pujadas & Muñoz, 2024). Meanwhile, a significant number of studies in this area do not have listening comprehension as the dependent variable. Rather, they emphasize the learning of vocabulary and lexical development as learning outcomes. For example, captioned audiovisual input has been consistently linked to incidental vocabulary learning, such as word-form recognition, meaning recognition and multiword expressions learning (Teng, 2022, 2025). These results show that captions can be an effective way to help develop words, but they are indirect evidence of improvements in listening comprehension. This is important because vocabulary skills are not always transferred to improvements in real-time processing of spoken input, especially since listening is a task that is time- and cognitively demanding. One of the constraints is the fairly static nature of multimodal supports. While digital environments provide the learner with some control, for example, over playback and over caption switching, such affordances are not necessarily the subject of systematic study in empirical research. As such, CAOs are often implemented as static teaching aids, and not as flexible tools that can be flexibly managed by learners in a strategic way to suit their needs and proficiency. In general, technological mediation can increase the number of multimodal supports that are available to learners, but this does not automatically mean that there is increased engagement with listening processes. The evidence suggests that the main benefits of captions and multimodal input are in the areas of comprehension and vocabulary development, which are achieved through dual-channel processing and decreased decoding demands (Teng, 2022). Nevertheless, their efficacy depends on learner factors like proficiency and vocabulary knowledge (Pujadas & Muñoz, 2024). Cognitive and metacognitive aspects have to be explicitly included if the pedagogical potential of CAOs in the digitally mediated environment is to be fully exploited.

### ***Summary of Research Gaps and Future Directions***

To answer the third research question, the results of the studies examined are compiled and summarized to identify several important research gaps and directions for the design and pedagogical use of CAOs in the context of ESL listening. Firstly, there is a significant absence of empirical studies with primary-level ESL learners, especially in digital or virtual classrooms. For example, Shamsi and Bozorgian (2024) carried out their study with upper-intermediate university-level EFL learners aged 18-20, indicating a lack of younger learners in such studies. This mirrors a wider trend towards research studies of older learners and a lack of research into populations of primary learners. Second, listening comprehension is often studied in conjunction with other factors, such as vocabulary knowledge, interaction, or affordances of a multimedia text, instead of being studied as a standalone variable. The same study shows that listening development is not studied separately but in a combination of various factors such as collaboration, input of multimedia and metacognitive strategies. The research identified the significance of metacognitive processes, including planning, monitoring and evaluation, and also pointed out that little research has been conducted on the use of metacognitive strategies. This suggests a lack of knowledge about the exact mechanisms of such strategies for listening. Third, while a focus on metacognitive instruction is at the core of the intervention, there is not a strong emphasis on learner control or self-directed use of strategies in technology-mediated environments. The instructional design used in Shamsi and Bozorgian (2024) was teacher-led and teacher-structured learning with predefined stages (planning, monitoring, problem-solving, and evaluation), which did not allow for the learners' autonomy and personalization of the learning process. This implies that there is a lack of research on how learner-controlled CAOs

can be used to help in the process of self-regulated listening. These gaps, collectively, point to the need for systematically-designed, age-appropriate, digitally delivered and theoretically informed listening comprehension and metacognitive interventions based on CAO. Future research should move beyond examining general effectiveness to investigating how learners actively engage with metacognitive strategies and how instructional designs can better support self-regulated listening, particularly in underrepresented contexts such as primary ESL virtual classrooms.

## Discussion

This scoping review consolidates recent empirical evidence on the role of CAOs in ESL and EFL contexts. Taken together, the reviewed studies demonstrate that CAOs implemented through textual previews, captions, transcripts, graphic organizers, concept maps, and social or multimedia annotations play a facilitative role in supporting learners' comprehension processes and engagement in technology-mediated listening environments (Wang & Hsiao, 2022; Aldukhayel, 2023; Kajiura et al., 2025). A consistent pattern across the studies is that CAOs support learners' engagement with input and help manage the demands of multimodal listening materials. For instance, Wang and Hsiao (2022) determined that advance organizers presented before video viewing supported learners' comprehension, particularly among higher-proficiency learners. Similarly, Kajiura et al. (2025) demonstrated that transcript-based pre-listening input supported learners' engagement with spoken input and facilitated adaptation to speech conditions. These findings suggest that CAOs serve as preparatory supports that facilitate learners' engagement with listening tasks, contributing directly to improved listening comprehension.

In addition to being beyond comprehension, CAOs seem to have an impact on learners' attention allocation and engagement when completing listening tasks. Research in captioning and subtitling shows that textual supports can help learners to process spoken input, as they can give them more linguistic cues (Teng, 2022; Pujadas & Muñoz, 2024). However, the magnitude of these benefits depends on learners' attributes, such as proficiency and vocabulary. Pujadas and Muñoz (2024) found that captions aided comprehension, but there were differences in the level of proficiency. A number of studies also emphasize the function of CAOs in the listening process in promoting metacognitive involvement. Shamsi and Bozorgian (2024) proved that the use of multimedia to support collaborative listening and metacognitive teaching and learning of comprehension strategies, such as planning, monitoring and evaluating comprehension, was effective. While not all CAOs were clearly conceptualised as metacognitive tools, the integration of CAOs in the context of structured listening activities provided opportunities for students to engage strategically with listening activities. In a similar vein, Lin and Wei (2024) suggested that social annotations within a multimedia setting facilitated interaction and engagement with listening materials. From a teaching perspective, visual organizers like concept maps and graphic organizers were found to be useful CAOs for meaning construction and multimodal integration. Tsai (2025) found that concept mapping helped improve learners' listening comprehension by structuring the information and combining the verbal and visual information. Regarding research question 2, the results show that there are some thematic patterns that are common across the design and implementation of CAOs. This encompasses their function as pre-listening scaffolds, their multimodal and interactive characteristics, and the way they are used in digital spaces in conjunction with metacognitive and collaborative learning processes. Taken together, the results indicate that

CAOs can help to create instructional coherence by organizing information and aiding the organization of input by the learners in listening tasks.

Although the studies reviewed showed overall positive impacts of CAOs on listening comprehension and related language-learning outcomes, care must be taken when interpreting the results. Results are limited in their generalizability because of the differences in the participants, the level of the learners, instructional contexts, intervention designs, and outcome measures. Some studies showed differential effects between groups of learners, with higher proficient learners benefiting more than lower proficient learners. Furthermore, several researchers investigated more general language learning goals than listening comprehension, such as vocabulary learning, engagement of students, interaction in groups and metacognition. The results, therefore, should not be interpreted as a universal proof of the efficacy of CAOs in ESL/EFL listening but rather as evidence of the potential efficacy of CAOs in specific listening contexts. It is hence necessary to conduct future research to examine the effects of different learner characteristics, instructional conditions, and specific CAO designs on listening comprehension in various educational contexts.

### **Limitations and Recommendations**

One of the main drawbacks of this scoping review is the selection of Scopus and ERIC as the main databases for the selection of articles. Meanwhile, these databases provide an index of quality research in applied linguistics and educational technology. It is important to highlight that limiting the scope of the search to these sources might have resulted in missing some studies that would have been considered relevant in studies on audiovisual learning, captioning or multimodal scaffolding. As evidenced from the literature reviewed, the various terms used to refer to CAO-related constructs include text-based advance organizers, captions, visual cues, social annotations and graphic organizers (Wang & Hsiao, 2022; Aldukhayel, 2023; Kwon & Yu, 2024). Future reviews can thus profit from adding more databases to catch a wider variety of conceptual and methodological approaches to CAO-aligned research.

Another limitation is that the computerized instructional supports are continually changing. With the ongoing evolution of digital technologies, it is now hard to draw a clear line between advanced organizers and other types of multimedia scaffolding. Some of the studies included, therefore, may not actually be advance organizers but preparatory instructional supports that have similar purposes to advance organizers. This more general perspective allowed for a more thorough review of current practices in the classroom, but could have also added some heterogeneity to the evidence. Future research should develop clearer operational definitions of CAOs and make it clearer to differentiate them from other types of digital scaffolding, to be more conceptually consistent among studies. Further studies are also required to examine how the specific mechanisms of the various types of computerized advanced organizers affect listening comprehension in a variety of populations of learners and in different learning situations.

Another constraint is the level of development in previous studies. Many investigations are short-term in nature or quasi-experimental, which means that they are looking at short-term learning outcomes, not sustained learning outcomes. For example, research on text-based or multimodal advance organizers was mostly conducted with adult or tertiary students with short-term interventions (Wang & Hsiao, 2022; Kajiura et al., 2025; Kositchaivat, 2025). This restricts the scope of being able to assess the longer-term developmental effects of CAOs,

especially for younger ESL students in the school-based or virtual classroom setting. In methodological terms, most research is based on outcome measures like comprehension tests, vocabulary gains or learner perceptions, with few research studies investigating processes. While some recent studies have investigated the use of multimodality in input and visual cues in listening tasks (Kwon & Yu, 2024), these methods are not widely used in CAO research. Future research might incorporate process-based approaches like multimodal interaction analysis or learning analytics to gain deeper insights into the interaction between CAOs and learners' engagement with listening input. Lastly, there is limited integration between CAOs and metacognitive regulation in the literature. Research that includes collaborative or metacognitive scaffolding has shown positive results on listening outcomes (Shamsi & Bozorgian, 2024); however, few studies have singled out CAOs as a unique instructional component. Thus, longitudinal and multi-phase research designs are suggested to investigate the interaction of CAOs and metacognitive strategies in the long term and different instructional contexts.

## Implications

This scoping review provides a systematic compilation of the research on CAOs and digital scaffolds in ESL learning. In all the studies analyzed, the CAOs are mostly used as preparatory and structuring tools which facilitate the learners' engagement with the listening input through activation of prior knowledge and direction of attention (Wang & Hsiao, 2022; Kajiura et al., 2025). The results support the pedagogical importance of providing pre-listening and pre-viewing supports in a structured manner for technology-based learning environments. The review also points to the need to consider the congruence of CAO design with the characteristics of learners and the demands of the tasks. There is some evidence that text-based and caption-based supports can be effective or ineffective depending on the learners' levels and that the effectiveness is different across learners' groups (Wang & Hsiao, 2022; Pujadas & Muñoz, 2024). This highlights the importance of adaptive and differentiated CAO designs, as opposed to uniform implementation across contexts. Instructionally, the results suggest that CAOs are best suited for support in the context of instructional sequences that are coherent, and should not be used as stand-alone supports. Research on the use of visual organizers, like concept maps, shows that they can facilitate learners' information organization and interaction with multimodal input (Tsai, 2025). Likewise, the studies on social annotations and collaborative listening showed that scaffolds such as CAO can facilitate the engagement and interaction with listening materials (Aldukhayel, 2023; Lin & Wei, 2024; Shamsi & Bozorgian, 2024). In general, the implications suggest that principled CAO design, which takes into account differences among learners, the structure of the tasks, and instructional coherence, must be taken into account when incorporating digital organizers into ESL pedagogy.

## Conclusion

CAOs are important tools for listening and other language skills in digitally mediated ESL learning environments. While the available literature is scattered in the areas of captions, visual cues, social annotations and graphic organizers, this scoping review brings together evidence to show that CAOs can help learners to engage with the listening input and to facilitate comprehension. The studies reviewed reveal that CAOs like text preview, captions, transcripts, visual organizers and multimodal scaffolds can support listening outcomes and engagement, especially when considering the characteristics of the learners and the context of instruction (Wang & Hsiao, 2022; Kajiura et al., 2025; Tsai, 2025). Despite this, many of the studies have

been restricted to brief intervention periods, adult samples of learners and are more concerned with immediate rather than ongoing development. This review identifies and maps current research and identifies consistent gaps, thereby providing clear directions for future research. To deepen the understanding of how CAOs can be effectively designed and used, increasing the geographical scope of the databases, as well as introducing longitudinal and process-oriented methodologies and focusing on primary and virtual ESL contexts are required. In general, the review offers a systematic basis for further research in the field of CAOs in ESL listening instruction, with the goal of enhancing the theoretical and pedagogical integration of CAOs into listening instruction.

---

**Acknowledgements:** The first author wishes to express her heartfelt appreciation to the second author, her supervisor, for her invaluable guidance and insightful feedback throughout the preparation of this review. Sincere gratitude is also extended to Universiti Pendidikan Sultan Idris (UPSI) for providing the academic resources and institutional support that made this study possible.

**Funding Statement:** No funding was received for this research.

**Conflict of Interest Statement:** The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the Journal of Information System and Technology Management (JISTM).

**Ethics Statement:** This study did not involve any human participants, animals, or sensitive data requiring ethical approval. The authors confirm that the research was conducted in accordance with accepted academic integrity and ethical publishing standards.

**Author Contribution Statement:** Sangeetha Thiarajah was responsible for the conceptualization, literature synthesis, and drafting of the manuscript. Hema Vanita Kesevan contributed to supervision, critical revision, and provided expert guidance throughout the study. All authors read and approved the final version of the manuscript prior to submission.

---

## References

- Aldukhayel, D. (2023). The benefits of social media comments to L2 listening comprehension. *SAGE Open*, 13(2). <https://doi.org/10.1177/21582440231171320>
- Alkhalaf, M. (2023). Incidental learning of L2 collocations in an academic lecture: A multimedia theory perspective. *World Journal of English Language*, 13(6), 480-488. An eye-tracking study. *Language Testing*, 41(2), 412-432.
- Ausubel, D.(1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 51(5), 267-272.
- Helal, S. M., Mahdi, H., & Thabet, M. A. (2024). Captioning for clarity: Mastering stress with word-based vs. phonetic video transcriptions. *Frontiers in Linguistics*, 6(4), 6779. <https://doi.org/10.30564/fls.v6i4.6779>  
<https://doi.org/10.1177/02655322241239356> <https://doi.org/10.5430/wjel.v13n6p480>
- Kajiura, M., Smith, A., & Kinoshita, T. (2025). Deferred multimodal input enhances L2 listening and fast-speech adaptation: A predictive coding perspective.
- Kositchaivat, S. (2025). The effects of task-based learning and graphic organizer techniques on the developing student teacher professional competency in listening and speaking skills. *NECTEC Review*, 1(1). <https://doi.org/10.70730/NESC6564>
- Kwon, S. K., & Yu, G. (2024). The effect of viewing visual cues in a listening comprehension test on second language learners' test-taking process and performance:
- Lin, S., & Wei, W. (2024). Social annotations and second language viewers' engagement with multimedia learning resources in LMOOCs: A self-determination theory perspective. *Cogent Education*, 11(1), 2335715. <https://doi.org/10.1080/2331186X.2024.2335715>
- Lo, S. (2024). Vocabulary learning through viewing dual-subtitled videos: Immediate repetition versus spaced repetition as an enhancement strategy. *ReCALL*, 36(2), 123-140. <https://doi.org/10.1017/S0958344024000053>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4, Article 1. <https://doi.org/10.1186/2046-4053-4-1>
- Mulyani, n., Yusuf, Y. Q., Trisnawati, I. K., Syarfuni, Qamariah, H., & Wahyuni, S. (2022). Watch and learn: EFL students' perceptions of video clip subtitles for vocabulary instruction. *Pertanika Journal of Social Sciences and Humanities*, 30(S1), 1-18. <https://doi.org/10.47836/PJSSH.30.S1.01>
- Pujadas, G., & Muñoz, C. (2024). When to switch captions off? Exploring the effects of L2 proficiency and vocabulary knowledge on comprehension of captioned and uncaptioned TV. *Studies in Second Language Learning and Teaching*, 14(1), 45-68. <https://doi.org/10.14746/ssl.t.38036>
- Shamsi, E., & Bozorgian, H. (2024). Collaborative listening using multimedia through metacognitive instruction: A case study with less-skilled and more-skilled EFL learners. *Language Education and Technology*, 4(1), 45-62. <https://doi.org/10.1186/s40862-023-00248-8>
- System, 119,103852. <https://doi.org/10.1016/j.system.2025.103852>
- Teng, M. F. (2022). Incidental L2 vocabulary learning from viewing captioned videos: Effects of learner-related factors. *System*, 108, 102736. <https://doi.org/10.1016/j.system.2022.102736>

- Teng, M. F. (2025). Incidental learning of multiword expressions from bilingual captioned videos for language minority students. *Journal of Asia TEFL*, 22(1), 1201-1218. <https://doi.org/10.29140/jaltcall.v21n1.1201>
- Tsai, K.-J. (2025). Concept mapping for enhancing multimodal communication in an EFL classroom. *SAGE Open*, 15 (1). <https://doi.org/10.1177/21582440251368978>
- Wang, L., & Hsiao, H.-S. (2022). Effects of advance organizers on Taiwanese middle school students' English-video listening comprehension. *International Journal of Computer-Assisted Language Learning and Teaching*, 12 (3), 1-15.