Al-Muyassar: Journal of Arabic Education
Vol. 5, No. 3, 2026, pp. 1243-1261
E-ISSN 2829-095X
DOI: https://doi.org/10.31000/0wmm5b25
Received: June 23, 2026; Revised: August 27, 2026; Accepted: September 17, 2026
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Developing Arabic Listening Comprehension through a YouTube-Based Active Listening Model: A Design-Based Study in Postgraduate Arabic Language Education

تنمية فهم المسموع في اللغة العربية من خلال نموذج للاستماع النشط قائم على يوتيوب

دراسة قائمة على التصميم في تعليم اللغة العربية للدراسات العليا

Ach. Tabrani[1], Rizal Firdaus[2], Lalu Agus Pujiartha[3], Achmad Zuhrudin[4]
Shakhzod Akhmedov[5], Badar Al Mamari[6]
[123]Institut Muslim Cendikia Sukabumi, Indonesia
[4]UIN Walisongo Semarang, Indonesia
[5]The Samarkand State Institute of Foreign Languages, Uzbekistan
[6]Sultan Qaboos Institute, Oman
*Corresponding E-mail: ahmadtabrani16@gmail.com

ABSTRACT

This study aims to design and preliminarily implement a YouTube-based active listening model to improve Arabic listening comprehension among postgraduate students. The study employed a Design-Based Research (DBR) approach through two iterative design cycles involving analysis, design, implementation, evaluation, and redesign in the context of postgraduate Arabic language education. The participants were ten students from Institut Insan Cendekia in Sukabumi, Indonesia, enrolled in a course on the application of technology in Arabic language teaching. The model was structured around six interconnected actions (listen, predict, verify, listen again, reconstruct meaning, and reflect) embedded within a nine-stage instructional sequence. Each listening cycle was organized around a specific purpose that guided learners from global comprehension to analytical processing and metacognitive reflection. The study employed authentic or semi-authentic YouTube materials. The instructional intervention was conducted over eight weeks, comprising sixteen sessions of 90 minutes each. Arabic listening comprehension was assessed using a 20-item multiple-choice test. Content validity was established through expert judgment, while reliability was assessed using Cronbach's alpha. Data were collected through pre- and post-tests, classroom observations, learner responses, and reflective evidence. The mean listening comprehension score increased from 62 on the pre-test to 86 on the post-test, representing an absolute gain of 24 points and a relative improvement of 38.71%. The normalized gain (N-gain) was approximately 0.63, indicating a moderate level of improvement. This finding is interpreted as a preliminary indicator of learning improvement rather than inferential evidence of causal effectiveness or generalizability, given the small sample size and the absence of individual-level dispersion data. The study contributes a practical framework for integrating YouTube into Arabic listening instruction through active listening, metacognitive strategies, purposeful repeated listening, and meaning reconstruction.

Keywords: Listening comprehension; Active listening; YouTube.

مستخلص البحث

تهدف هذه الدراسة إلى تصميم نموذج للاستماع النشط قائم على يوتيوب وتنفيذه بصورة أولية لتطوير فهم المسموع باللغة العربية لدى طلبة الدراسات العليا .اعتمدت الدراسة توجها من البحث القائم على التصميم من خلال دورتين تصميميتين، شملت مراحل التحليل والتصميم وذلك في سياق تعليم اللغة العربية للدراسات العليا، بمشاركة عشرة طلاب من جامعة إنسان جنديكيا في سوكابومي في محاضرة تطبيق تكنولوجيا تعليم اللغة العربية. وبني النموذج حول ستة أفعال مترابطة هي :أصغ، وتوقع، وتحقق، وأعد الاستماع، وأعد بناء المعنى، وتأمل، مع تنظيم كل دورة

Citation:
Tabrani, A., Firdaus, R., Pujiartha, L. A., Zuhrudin, A., Akhmedov, S., & Al Mamari, B. (2026). Developing Arabic Listening Comprehension through a YouTube-Based Active Listening Model : A Design-Based Study in Postgraduate Arabic Language Education. *Al-Muyassar: Journal of Arabic Education, 5* (3), 1243–1261.

استماع حول غرض محدد ينتقل بالمتعلم من الفهم الكلي إلى المعالجة التحليلية والتأمل الميتامعرفي .واستخدمت الدراسة مواد يوتيوب أصيلة أو شبه أصيلة، وطُبّق التدخل التعليمي على مدى ثمانية أسابيع بواقع ست عشرة حصة، مدة كل حصة تسعون دقيقة. ولقياس فهم المسموع، استُخدم اختبار مكوّن من عشرين فقرة من نوع الاختيار من متعدد، وتم التحقق من صدق الأداة من خلال تحكيم الخبراء للتحقق من صدق المحتوى، ومن ثباتها باستخدام معامل ألفا كرونباخ. وجمعت البيانات من الاختبار القبلي والبعدي، والملاحظة الصفية، واستجابات المتعلمين، وأدلة التأمل .ارتفع متوسط الدرجات من ٦٢ في الاختبار القبلي إلى ٨٦ في الاختبار البعدي، أي بزيادة مطلقة قدرها ٢٤ درجة، وتحسن نسبي قدره ٣٨,٧١٪، وبلغ الكسب المعياري نحو ٠,٦٣، وتفسر هذه النتيجة بوصفها مؤشرا أوليا على تحسن التعلم، لا دليلا استدلاليا قابلا للتعميم، لصغر حجم العينة وعدم توافر بيانات التشتت الفردية .وتقدم الدراسة نموذجا تطبيقيا لدمج يوتيوب في تعليم الاستماع العربي على أساس الاستماع النشط والاستراتيجيات الميتامعرفية وإعادة بناء المعنى.

الكلمات المفتاحية :فهم المسموع؛ الاستماع النشط؛ يوتيوب.

INTRODUCTION

Listening comprehension is central to Arabic language learning, especially in academic settings that require learners to process connected discourse beyond the level of isolated words and sentences (Field, 2008; Rost, 2016). In academic and communicative situations, learners need to identify the main idea, extract key information and supporting details, link acoustic input to contextual cues, infer implicit meaning, and then reconstruct the content of the discourse coherently (Vandergrift, 2007). These demands become more complex for postgraduate students (Siegel, 2015). They face academic discourse genres such as lectures, discussions, presentations, interviews, and specialized dialogues, which vary in speech rate, vocabulary, discourse structure, and information density (Nation & Newton, 2020; Webb & Rodgers, 2009).

The difficulty of learning listening comprehension does not lie only in the nature of the acoustic input. It also depends on how learners direct the comprehension process while listening (Vandergrift, 2004). Effective instructional listening requires learners to set a purpose before listening, activate prior knowledge, and form expectations about the content (Graham, 2006). They then monitor their comprehension while listening, verify the hypotheses or information they reach, and re-listen selectively when a comprehension gap appears (Lynch, 2009). The process ends with meaning reconstruction and reflection on the strategies they used (Graham & Santos, 2015). From this perspective, active listening differs from receptive listening, which limits learners to following the audio and answering questions afterward (Brown & Lee, 2015). Active listening makes learners participants in building meaning and monitoring their own comprehension (Siegel, 2018).

This view suggests that the design of listening comprehension activities must move beyond presenting audio material as instructional input. Designers need to organize it within a sequence of tasks that guide comprehension before, during, and after listening. Digital audio and video materials, particularly content published on YouTube, expose learners to varied forms of spoken language in authentic or semi-authentic contexts. However, availability alone does not guarantee effective learning. The instructional value of digital content depends on how teachers select it, set listening purposes, design processing tasks, use re-listening, and provide appropriate support without replacing the listening process itself. This creates a need for a procedural model that integrates

digital content with active listening and metacognitive strategies, so that learners move systematically from global comprehension to analysis, verification, meaning reconstruction, and reflection.

In the digital environment, YouTube has become a broad source of audiovisual materials that can extend exposure to Arabic beyond the textbook (Asy'ari et al., 2025). Recent studies show that video content presents language in context, combines sound, image, and contextual cues, and offers opportunities for repetition and re-listening (Dizon, 2022; Vanderplank, 2016). Studies in Arabic language teaching also show that YouTube may support listening, pronunciation, vocabulary, and contextual comprehension (Mubarok et al., 2025). Yet the platform itself does not guarantee learning (Yuyun & Simamora, 2021). Learners may slip into passive viewing or rely excessively on subtitles and accompanying transcripts (Montero Perez et al., 2013). The selection of digital material matters because of the nature of instructional material itself. Good material should align with the objective, suit the learners' level, and allow segmentation and reuse across multiple tasks. This principle is consistent with the principles of materials development for language teaching (Tomlinson, 2018).

It is therefore important to distinguish between using YouTube as a video repository and using it as an instructional resource within a deliberate pedagogical design (Reinhardt, 2019). The instructional problem is not a shortage of materials (Godwin-Jones, 2020). The problem is how to turn audiovisual material into a sequence of listening tasks that require learners to predict, test their predictions, re-listen for a new purpose, verify their comprehension, and then reconstruct meaning and evaluate their strategy (Mayer, 2020; Sweller, 1988).

Design-based research offers a suitable framework for addressing this problem (Design-Based Research Collective, 2003). It starts from a real instructional need, designs an instructional solution, tests it in its natural setting, and collects evidence during implementation (Anderson & Shattuck, 2012; Barab & Squire, 2004). It then revises the design in light of the results and arrives at design principles that can inform similar contexts (Rodríguez, 2017; Wang, 2021; Yang & Lee, 2025).

A review of previous literature reveals growing interest in using YouTube and digital media in listening instruction, and in metacognitive strategies that help learners set a listening purpose, monitor their comprehension, and choose suitable ways to address comprehension difficulties (Dizon, 2022; Goh, 2008; Vandergrift & Tafaghodtari, 2010). However, the studies reviewed mostly treated these strands separately (Asy'ari et al., 2025). Some studies focused on listening as a strategic activity that requires understanding arguments, extracting information, and verifying meaning (Graham & Santos, 2015; Siegel, 2018). Other studies addressed the explicit teaching of metacognitive strategies (Cross, 2015; Lee, 2015). Neither strand offered a clear procedural model that integrates YouTube content with a complete cycle of active listening, metacognitive processing, and meaning reconstruction. Research on YouTube in Arabic language teaching has also not sufficiently explained, in the context of the present study, how to turn video content into graded listening tasks that move beyond word recognition to prediction, verification, re-listening, meaning reconstruction, and reflection (Pei &

Suwanthep, 2021).

Accordingly, this study addresses the need for a procedural instructional model that integrates YouTube, active listening, and metacognitive strategies in Arabic language teaching, and that specifies the model's components, task sequence, implementation, and evaluation for postgraduate students. The proposed model differs from the general use of YouTube as a source of audio material because it organizes learning through six interconnected actions: "listen, predict, verify, listen again, reconstruct meaning, and reflect." It also differs in context and target group, since it focuses on postgraduate students in Arabic language education. It further differs in how it examines effectiveness. It relies on design-based research through successive cycles of analysis, design, implementation, evaluation, and redesign, together with a listening comprehension test, classroom observation, learner responses, and reflective evidence. The study therefore does more than combine YouTube with metacognitive listening. It develops an integrated procedural model that specifies how YouTube content can guide learners from initial comprehension to monitoring, verification, meaning reconstruction, and reflection in postgraduate Arabic language education.

The study therefore addresses the following question: How can a YouTube-based active listening model be designed to help postgraduate students develop Arabic listening comprehension? The study aims to design the model, pilot it, analyze the listening practices it involves, and describe the change in mean performance between the pre-test and the post-test (Mackey & Gass, 2016). Its significance lies in combining active listening, metacognitive strategies, and design-based research in a single applicable model (Cohen et al., 2018; McKinley & Rose, 2020).

RESEARCH METHOD

The study adopted design-based research (DBR), which combines instructional design, testing in a real setting, formative evaluation, and redesign based on empirical evidence (McKenney & Reeves, 2019; Rodríguez, 2017). The study ran in two successive design cycles (Wang, 2021). In the first cycle, the researchers analyzed learners' needs and difficulties in listening comprehension, then built and tested an initial version of the YouTube-based active listening model. They collected data from classroom observation, learner performance, learner responses, reflective evidence, and a listening comprehension test (Yang & Lee, 2025). Based on the formative evaluation results, the researchers revised the instructions for the verification, re-listening, and meaning reconstruction tasks, and reorganized some YouTube materials to suit the learners' level.

In the second cycle, the researchers implemented the revised model and collected data again from the same sources to check whether the revisions worked and to refine the design. The results showed clearer sequencing of listening tasks, more purposeful use of re-listening, and more integrated meaning reconstruction among learners. These results led to limited final revisions in the instructions and activity sequence. The two cycles produced the final model and a set of preliminary design principles for using YouTube in active and metacognitive listening among postgraduate students.

The study took place in the postgraduate Arabic Language Education program at

Institut Muslim Cendikia in Sukabumi, Indonesia. Ten postgraduate students of Arabic language education took part. The researchers selected participants purposively because they represent the target group of the proposed instructional model and have a background in Arabic learning that allows them to engage in Arabic listening activities. The selection criteria were enrollment in the postgraduate Arabic language education program, willingness to take part in all research stages, ability to follow instructional materials in Arabic, and commitment to attend the intervention sessions and complete the listening and assessment tasks.

To establish actual listening proficiency, the study used the pre-test as a baseline for participants' performance before the model was applied. Their mean score was 62. The researchers also used classroom observation and learner responses to identify the difficulties participants faced while listening. Participants had prior experience with YouTube, but they had not used it for Arabic listening in a way structured by the active listening model proposed here. The researchers therefore designed the activities to move them from ordinary viewing and listening to guided listening based on prediction, verification, re-listening, meaning reconstruction, and reflection.

Because the sample was limited to ten students, the researchers chose them for their fit with the target instructional context, not for statistical representation of a population. The results therefore count as preliminary evidence tied to their instructional context, and they do not support statistical generalization to all postgraduate students. On ethical grounds, the researchers informed participants of the study's aims and procedures. Participants took part voluntarily after giving informed consent. The researchers kept their data confidential and did not reveal their identities when reporting results. They used the data for research purposes only and followed the institutional procedures at Institut Muslim Cendikia for conducting and participating in research.

The instructional material centered on YouTube videos suited to the learners' level, the topic, and the learning objective. Selection considered audio clarity, the possibility of dividing the material into segments, vocabulary suitability, a context that supports comprehension, and the possibility of re-listening (Dizon, 2022). Transcripts or subtitles served, when needed, as support for verification rather than a substitute for listening (Mubarok et al., 2025; Othman et al., 2025). The appeal of a video was not the only selection criterion. The criteria also included fit with the objective, potential for conversion into tasks, clarity of language, graded difficulty, and the possibility of replaying specific segments (Krashen, 1985). This approach is consistent with principles of language materials development that treat the fit between material, learners, and objectives as a core design requirement (Tomlinson, 2018).

The proposed model consists of nine interconnected classroom stages designed to organize active and metacognitive listening with YouTube content: (1) selecting the YouTube material and setting the listening purpose, (2) preparation and prediction, (3) first listening for global comprehension, (4) analytical listening, (5) collaborative verification, (6) re-listening with a new purpose, (7) meaning reconstruction, (8) metacognitive reflection, and (9) evaluation and the redesign decision (Vandergrift & Goh, 2012). These stages do not form rigid linear steps (Vandergrift & Cross, 2018). They

form an instructional cycle in which learners can return to an earlier stage when evaluation reveals a comprehension gap (Cross, 2015; Flavell, 1979; Lee, 2015).

The researchers collected data through a pre-test and a post-test of Arabic listening comprehension, along with classroom observation, learner responses, and reflective evidence (Cohen et al., 2018). The listening comprehension test contained 20 items and took 90 minutes. Its items measured comprehension of the main idea, extraction of explicit information, understanding of details, verification of meaning, inference from context, and meaning reconstruction (Field, 2008). The test items matched the listening comprehension learning objectives and the tasks of the YouTube-based active listening model, which ensured alignment between learning objectives, tasks, and the assessment instrument (Richards, 2015; Ur, 2018).

The researchers administered the pre-test before the instructional intervention and the post-test after it ended. They scored answers with a standard answer key, giving one point for each correct answer and zero for each incorrect answer, and then converted scores to a 100-point scale. They established content validity by reviewing the items against the listening learning objectives and the participants' level, and they revised item wording and instructions before administration. They assessed reliability with Cronbach's alpha, which reached 0.86 and indicates good internal consistency.

In addition to the test, classroom observation tracked learners' performance across the model's six actions, particularly prediction, verification, re-listening, meaning reconstruction, and reflection. The researchers also collected learner responses and reflective evidence to interpret the difficulties learners faced and the strategies they used while listening. The study thus combined quantitative and qualitative evidence in line with the nature of design-based research.

The results showed that the mean listening comprehension score rose from 62 on the pre-test to 86 on the post-test, an absolute gain of 24 points and a relative improvement of 38.71%, with a normalized gain of about 0.63. The sample was small, limited to ten participants, and individual scores and standard deviations were unavailable. For these reasons, the researchers did not run inferential tests or compute statistical effect sizes. The quantitative analysis covered only a descriptive comparison of the two means, the improvement rate, and the normalized gain, while the qualitative evidence informed the interpretation of the results.

Proposed Learning Design

The study proposes a learning design based on an active listening cycle, in which YouTube is not the "method" itself but a digital resource within a structured instructional scenario (Asy'ari et al., 2025). The design combines audiovisual exposure, prediction, repeated listening with changing purposes, collaborative verification, meaning reconstruction, and metacognitive reflection (Vandergrift & Goh, 2012). This structure draws on literature that highlights the role of metacognitive awareness and collaboration in improving listening comprehension, and on studies that show YouTube's potential to provide a rich linguistic and audiovisual context (Cross, 2015; Lee, 2015; Liu, 2020; Mubarok et al., 2025; Shabani & Heidarian, 2015). The structure also rests on the principle

that learners engage more when they have a clear task and responsibility for decisions during learning, rather than merely receiving material (Mercer & Dörnyei, 2020). Each stage therefore requires a specific cognitive or communicative action from the learner (Brown & Lee, 2015).

General Design Model

The proposed model rests on a sequential design framework that turns YouTube from a mere source of audio material into a structured learning environment for developing listening comprehension, active listening, and metacognitive awareness. The model has nine interconnected stages: YouTube, preparation and prediction, first listening, analytical listening, collaborative verification, re-listening with a new purpose, meaning reconstruction, metacognitive reflection, and evaluation. These stages draw on principles of auditory input processing, strategic listening, metacognitive learning, and collaborative learning. Learners move gradually from building expectations and global comprehension to processing details, verifying comprehension, reconstructing meaning, and reflecting on the learning process.

Stage 1: YouTube

The input is an authentic or semi-authentic audio or audiovisual clip from YouTube that suits the learners' level and the lesson objectives. The teacher selects the material and sets the objective and an appropriate length, while learners become familiar with the topic and context of the material. Moving to the next stage requires material linked to a clear learning objective and comprehensible at the learners' level. The expected output is ready material and a defined listening purpose.

Stage 2: Preparation and Prediction

The teacher presents the title, an image, or some contextual information without revealing the full content, then guides learners to activate prior knowledge and predict the content of the clip. Learners identify what they expect to hear and form initial hypotheses about the topic and key information. They move to the next stage after formulating clear predictions that they can check while listening. This stage activates prior knowledge and directs attention before learners receive the audio input.

Stage 3: First Listening

Learners listen to the clip for the first time without frequent pauses to grasp the main idea and the core message. The teacher directs attention to the global listening purpose, while learners form an initial understanding and compare the content with their predictions. The output of this stage is the identification of the main idea and essential information. Learners are ready to move on when they can express the general meaning of the clip, even in a preliminary form.

Stage 4: Analytical Listening

Learners listen to the material again and analyze details, relationships between ideas, important information, and meanings drawn from context. The teacher provides guided questions or tasks, while learners identify key information and relevant details and link them to the general meaning. The learner's ability to extract the required information and

explain the relationships within it shows whether the objective has been met.

Stage 5: Collaborative Verification

Learners work in pairs or small groups to compare their answers and predictions and discuss points where their comprehension differs. The teacher acts as a facilitator who guides the discussion without giving the answer directly. Learners compare their understanding with that of their peers and return to the audio evidence when needed. They move to the next stage when they can identify points of agreement and disagreement and justify their comprehension with evidence from the audio material.

Stage 6: Re-listening with a New Purpose

Learners do not re-listen to the clip merely for repetition. They re-listen to achieve a new purpose or to resolve a specific comprehension problem that emerged in earlier stages. Learners identify the segment that needs further processing and re-listen to it to verify information, interpret meaning, or resolve a conflict in comprehension. The teacher sets the task and directs attention, while learners take responsibility for locating the difficulty and choosing what to re-listen to. Achieving the specified purpose serves as the criterion for moving to meaning reconstruction.

Stage 7: Meaning Reconstruction

Learners reorganize the information they gathered from the clip and link it to context and prior knowledge to reach an integrated meaning. They can do this through summarizing, ordering ideas, answering a complex question, or paraphrasing the content of the clip in Arabic. The teacher helps organize the task and gives feedback, while learners build meaning independently or collaboratively. The objective is met when learners can present a coherent representation of the content that goes beyond recognizing isolated words.

Stage 8: Metacognitive Reflection

Learners reflect on their listening process by identifying what they understood, the difficulties they faced, the strategies they used, how well those strategies worked, and what they could change in a later listening task. The teacher asks reflective questions that help learners monitor and evaluate their learning. The output of this stage is clearer awareness of the learning process and of effective strategies for handling listening comprehension difficulties.

Stage 9: Evaluation

The teacher evaluates the achievement of learning objectives through tasks or questions that measure global comprehension, details, inference, and meaning reconstruction, and also draws on observation, learner responses, and reflective evidence. Evaluation does not stop at the final score. It also identifies strengths and difficulties and informs decisions on redesigning tasks or materials in the next cycle of design-based research.

The model thus rests on a sequential relationship between inputs, processes, and outputs. The input is YouTube material and a defined listening purpose. The processes

include preparation and prediction, graded listening, verification, re-listening, meaning reconstruction, and reflection. The expected outputs are improved listening comprehension, a greater ability to use listening strategies, and stronger awareness of one's own learning process. The stages are not rigid. Learners can return to an earlier stage when evaluation reveals a comprehension difficulty. The model therefore serves as a procedural framework that others can reapply, evaluate, and develop in similar contexts.

Detailed Lesson Scenario

Before listening: The teacher selects a clip of no more than 3–6 minutes at the initial stage and sets a clear linguistic and cognitive objective. The teacher then shows the title or a selected image without playing the audio and asks students to record their predictions (Vandergrift & Goh, 2012). This procedure draws on principles of task-based language teaching (Long, 2015). The teacher defines the outcome learners should reach before choosing the activity, then selects material that enables learners to complete the task and achieve the objective (Loewen, 2015).

First listening: The teacher plays the whole clip once. Students should not try to write down every word. The teacher asks them to answer global comprehension questions: What is the topic? Who is the speaker? What is the main issue? What is the speaker's general position?

Second listening: The teacher replays specific segments. The task now shifts to searching for information, vocabulary, or relationships between ideas. The teacher can pause the clip after a short segment when the goal is to analyze discourse structure or a precise meaning.

Collaborative verification: Students work in pairs or small groups and compare their answers. The teacher does not simply give the correct answer but asks: Which part did you hear that led you to choose this answer?

Re-listening with a new purpose: The teacher replays the clip with a different task. If the first goal was the main idea, the second goal becomes identifying the evidence, the speaker's position, or a causal relationship. Re-listening thus becomes reprocessing, not mere repetition.

Meaning reconstruction: Each student produces an oral or written summary of 3–5 sentences or paraphrases the main idea in accurate Arabic. Assessment here focuses on coherence of meaning rather than memorization of words.

Reflection: Each student answers three questions: What was the hardest part? Which strategy helped me? What will I do in the next listening task? The teacher uses the answers to adjust the design for the next lesson.

Roles of the Teacher, Learner, and Technology

This distribution of roles (Table 1) reflects a shift from a model in which instruction centers on presenting material to a model that centers on learner engagement with the task and on the teacher's role in facilitation, feedback, and organizing interaction (Brown & Lee, 2015). Learner engagement is a central factor in designing contemporary language

learning environments (Mercer & Dörnyei, 2020).

Table 1. Roles of the Teacher, Learner, and Technology in the Design

Component	Role in the Design
Teacher	Designs the task, directs attention, organizes verification, and facilitates reflection; not merely a video operator.
Learner	An active listener who forms a prediction, tests it, re-listens, justifies their comprehension, and reconstructs meaning.
YouTube	A source of audiovisual material and real-world context, not a substitute for the learning activity.
Transcript/Subtitles	Support for verification when needed, used after an attempt at comprehension rather than routinely before it.
Group interaction	A tool for revealing differences in interpretation and the gap between hearing and understanding.
Evaluation	Combines a comprehension test, observation, meaning reconstruction tasks, and reflection.

Design-Based Research Logic

The design works in a recurring cycle: analysis of students' needs → initial design → classroom implementation → collection of evidence from the test, observation, and reflection → identification of strengths and difficulties → revision of the video, task, or re-listening time → retesting. This cycle is consistent with design-based research, which links the instructional solution to its real context and uses implementation results to develop the design itself (Rodríguez, 2017; Wang, 2021; Yang & Lee, 2025).

RESULTS AND DISCUSSION

Implementation data from the two design cycles showed that the instructional value of YouTube content did not come from showing videos to learners in itself. It came from the way the researchers reorganized its use within a sequence of active listening tasks. In the first cycle, classroom observation, learner responses, and reflective evidence showed that learners needed greater clarity in some verification, re-listening, and meaning reconstruction tasks (OBS-C1; RESP-C1; REF-C1). Based on this evidence, the researchers revised the task instructions and reorganized some YouTube materials. Each re-listening then had a specific purpose linked to a comprehension problem that emerged in the previous stage.

In the preparation and prediction stage, the video title and context guided learners to form expectations about the topic before listening. The first listening stage then targeted global comprehension, and learners focused on identifying the main idea without having to understand all vocabulary and details from the start. Learners then moved to analytical listening to extract key information, details, and relationships between ideas. Classroom observations in the first cycle showed that some learners tended to treat listening as an attempt to understand words one after another. This

tendency required the teacher to redirect them toward the specific task purpose and contextual cues (OBS-C1). In the second cycle, observations and learner responses showed clearer task sequencing and more purposeful use of re-listening (OBS-C2; RESP-C2).

Re-listening with a new purpose was one of the components that changed markedly between the two cycles. In the first cycle, formative evaluation revealed a need to clarify why learners re-listen and what they should look for during re-listening (OBS-C1; REF-C1). The researchers therefore revised the task so that re-listening did not mean watching the whole clip again without a purpose. Instead, they tied it to a specific question, piece of information, or point left unresolved in the previous listening. In the second cycle, observations showed that learners became better able to direct repeated listening toward points of difficulty instead of re-listening at random (OBS-C2). Learner responses and reflective evidence, which focused on the difficulties learners faced and the strategies they used to address them, support this finding (RESP-C2; REF-C2).

The collaborative verification stage also showed that comparing answers among learners gave them an opportunity to review their comprehension and identify points of disagreement before moving to re-listening. The aim of collaborative work was not simply to reach a shared answer. It pushed learners to justify their comprehension and return to the audio evidence. In the meaning reconstruction stage, learners organized the information they had gathered and summarized or paraphrased the content of the material coherently. This shift from extracting information to reconstructing it made comprehension the result of an integrated process rather than a set of separate answers (OBS-C2; RESP-C2).

Metacognitive reflection evidence complements these findings. It helped learners review what they had understood, identify points of difficulty, and recognize the strategies that helped them reach meaning (Goh, 2008; Vandergrift & Tafaghodtari, 2010). The qualitative evidence therefore suggests that the improvement did not stem from YouTube as a standalone technology (Asy'ari et al., 2025; Mubarok et al., 2025). It stemmed from embedding YouTube within an instructional structure that gave each stage a specific function in building meaning (Godwin-Jones, 2020). This interpretation is consistent with studies that stress linking digital media to clear instructional objectives and tasks rather than treating technology as an end in itself (Dizon, 2022; Reinhardt, 2019; Richards, 2015).

The quantitative results lend preliminary support to this interpretation. The mean listening comprehension score of the ten students rose from 62 on the pre-test to 86 on the post-test, an absolute gain of 24 points, with a normalized gain (N-gain) of about 0.63. However, these results alone cannot show that YouTube was the direct causal factor in the improvement, because the intervention also included prediction, graded listening, collaborative verification, re-listening, meaning reconstruction, and metacognitive reflection. The main contribution of the study therefore lies in developing and refining a procedural model that specifies how to turn YouTube content into a sequence of active and metacognitive listening tasks that teachers can apply and redesign. It does not lie in demonstrating an effect of YouTube independent of the other model components.

Implementation also showed the importance of collaborative verification. Dialogue among learners revealed that some comprehension errors did not result from failing to hear the audio (Field, 2008). They resulted from inaccurate interpretation of cues or incorrect links between sentences (Lee, 2015). Verification was therefore not only a stage for correcting answers but also a way to expose the gap between hearing and understanding (Pei & Suwanthep, 2021).

Practice also showed that re-listening is more useful when the task changes. If the teacher replays the clip with the same question, re-listening may become mechanical repetition (Nation & Newton, 2020). When learners move from identifying the main idea to identifying the speaker's position, the evidence, or the relationship between ideas, re-listening becomes an opportunity for reprocessing (Vandergrift & Cross, 2018).

The meaning reconstruction stage stands out as an important link between listening and expression (Rost, 2016). Learners do not settle for scattered answers. They try to summarize or paraphrase what they understood in a coherent form (Lynch, 2009). Metacognitive reflection also helps learners identify the source of difficulty, such as speech rate, vocabulary, links between sentences, or overreliance on the accompanying transcript (Siegel, 2018).

Students' mean listening comprehension score rose from 62 on the pre-test to 86 on the post-test, an absolute gain of 24 points and a relative improvement of 38.71% over the pre-test mean. To estimate the size of the improvement relative to the possible distance between the pre-test mean and the maximum score, the researchers calculated the normalized gain (N-gain) with the following formula: N-gain = (post-test mean − pre-test mean) / (maximum score − pre-test mean), that is, (86 − 62) / (100 − 62) = 24/38 ≈ 0.63. Under the classification of N-gain values (Hake, 1998), this value falls in the moderate category and indicates a notable improvement in listening comprehension after the model was applied. However, readers should treat this result as preliminary exploratory evidence of improvement. The difference between the two means alone does not establish statistical significance or attribute the improvement causally to the model, especially given the small sample and the availability of means only, without individual scores or standard deviations. The quantitative analysis in this study therefore covers only the description and comparison of the two measurements, while classroom observation data, learner responses, and reflective evidence inform the interpretation of the observed change in student performance.

The findings agree with research showing that YouTube provides context-rich auditory input and increases opportunities for language exposure (Dizon, 2022; Mubarok et al., 2025; Shorouk, 2020; Yuyun & Simamora, 2021). They also agree with studies that highlight the role of metacognitive training in raising learners' awareness of listening strategies (Liu, 2020; Shabani & Heidarian, 2015). The present study, however, differs from these strands. It does not use video only as audio material, and it does not treat metacognitive strategies separately. It integrates both into a complete procedural cycle that gives each listening a specific instructional purpose, followed by verification, meaning reconstruction, and reflection. This contribution matters particularly for postgraduate students, because academic discourse requires understanding ideas,

relationships, and inferences, not merely recognizing vocabulary.

Theoretically, the study suggests that the effectiveness of digital content in listening instruction depends more on task structure and on how the teacher guides processing than on the medium alone. Practically, the model offers a sequence that teachers can transfer to other Arabic teaching contexts and adapt for lectures, interviews, seminars, and academic presentations. Future studies need to test these applications with larger samples and in broader contexts.

The results of implementing the model, together with classroom observations, learner responses, and reflection data, yielded a set of design principles. These principles go beyond using YouTube as a source of audio material and organize how digital content becomes a process of active and metacognitive listening. The results supported starting with global comprehension before moving to details, which allows learners to form an initial picture of the clip's content before focusing on specific information. They also showed the value of making prediction part of the listening process, because it prepares learners to compare their expectations with what they hear instead of listening passively. The activity sequence further showed the importance of changing the purpose of re-listening. Re-listening is not a repetition of the same clip but a means to address a specific comprehension problem or to search for new information. Collaborative verification gave learners an opportunity to compare their comprehension and justify their answers with audio evidence, while the accompanying transcript served as a support tool when needed, not a substitute for listening. The process ended with meaning reconstruction and metacognitive reflection, so that evaluation went beyond knowing the correct answer to reviewing how comprehension occurred and which strategies supported it.

The added scientific value of the study lies in bringing these principles together in an integrated procedural framework. The framework links YouTube audio material, prediction, graded listening, verification, re-listening with changing purposes, meaning reconstruction, and metacognitive reflection within a design cycle open to review and redesign. The contribution therefore goes beyond combining YouTube, active listening, and metacognitive strategies. It specifies how to order these components, what roles they play, and how learners move between them according to specific comprehension goals. Previous work treated listening as a strategic activity or used digital media to supply auditory input (Dizon, 2022; Mubarok et al., 2025). This study, by contrast, offers a procedural model that explains how to turn digital content into successive listening tasks that move learners from initial comprehension to verification, meaning reconstruction, and reflection (Rodríguez, 2017; Wang, 2021; Yang & Lee, 2025).

Nevertheless, the data of this study support some claims, while other claims represent a theoretical extension or a recommendation for future work. The rise in mean listening comprehension scores from 62 to 86, together with classroom evidence and learner responses, offers preliminary support for the model's suitability in the context studied. It does not, on its own, prove that all design principles represent causal relationships or that they generalize to other contexts. Accordingly, extending the model to short lectures, academic interviews, seminars, academic presentations, and specialized dialogues remains a recommendation for future research that requires empirical testing

in different contexts and samples. It is not a direct finding of the present study.

Several practical challenges need attention, including speech rate, multiple speakers, unfamiliar vocabulary, and dense visual effects (Azka et al., 2024). Subtitles or accompanying transcripts may also become a crutch if the teacher does not define their function (Montero Perez et al., 2013; Vanderplank, 2016). Video selection should therefore form part of the design, not a technical step separate from the lesson objectives (Dizon, 2022).

Derived Design Principles

The design starts from the listening purpose, not from the available video.

Learners move gradually from global comprehension to detailed analysis.

Each re-listening includes a new task with a clear purpose.

Prediction directs attention and builds hypotheses about meaning.

Collaborative verification reveals the gap between hearing and interpretation.

Transcripts or subtitles serve, when needed, as support for verification, not as a substitute for listening.

Meaning reconstruction and reflection form part of listening instruction.

Limitations of the Study

The sample is limited to ten students, so the quantitative result does not generalize to all postgraduate students.

The quantitative data rely on pre-test and post-test means. No individual data were available to compute a significance test or a more advanced effect size.

Judging whether the improvement lasts requires longer implementations and follow-up measurements.

Testing the model's validity requires trials in different institutions and levels and with varied types of Arabic content.

CONCLUSION

The study concludes that integrating YouTube into listening comprehension instruction adds more value when teachers place it within an active and structured instructional design than when they use it only as viewing material. The model follows a cycle that begins with preparation and prediction, moves to listening for global comprehension, then to analytical listening and collaborative verification, and then to re-listening with a new purpose, meaning reconstruction, and metacognitive reflection. Implementation with ten postgraduate students showed that mean performance rose from 62 to 86, with an absolute gain of 24 points and a normalized gain of about 0.63. This result remains a preliminary indicator because of the limited sample and quantitative data. The main value of the study lies in design principles that make YouTube a resource for active learning: selecting material based on the objective, grading comprehension, changing the function of re-listening, collaborative verification,

meaning reconstruction, and reflection. Future studies should expand the sample, collect individual data, run appropriate inferential tests, measure learning retention, compare the model with conventional listening instruction, and apply it to more specialized Arabic academic content.

REFERENCES

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher,* 41(1), 16–25. https://doi.org/10.3102/0013189X11428813

Asy'ari, H., Siraj, T., Rufaiqoh, E., & Yaqin, M. A. (2025). Computer aided Arabic language learning: A systematic review of technology integration and effectiveness. *Alsinatuna: Journal of Arabic Linguistics and Education,* 11(1), 1–19. https://doi.org/10.28918/alsinatuna.v11i1.12602

Azka, L., Hafsoh, U. N., & Ibrahim, F. M. A. (2024). The challenges of learning Arabic listening skills using YouTube as a learning tool. *Alsina: Journal of Arabic Studies,* 6(2), 229–246. https://doi.org/10.21580/alsina.6.2.23369

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences,* 13(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1

Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy* (4th ed.). Pearson Education.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.

Cross, J. (2015). Metacognition in L2 listening: Clarifying instructional theory and practice. *TESOL Quarterly,* 49(4), 883–892. https://doi.org/10.1002/tesq.258

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher,* 32(1), 5–8. https://doi.org/10.3102/0013189X032001005

Dizon, G. (2022). YouTube for second language learning: What does the research tell us? *Australian Journal of Applied Linguistics,* 5(1), 19–26. https://doi.org/10.29140/ajal.v5n1.636

Field, J. (2008). *Listening in the language classroom*. Cambridge University Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist,* 34(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Godwin-Jones, R. (2020). Building the porous classroom: An expanded model for blended language learning. *Language Learning & Technology,* 24(3), 1–18. http://hdl.handle.net/10125/44731

Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal,* 39(2), 188–213. https://doi.org/10.1177/0033688208092184

Graham, S. (2006). Listening comprehension: The learners' perspective. *System,* 34(2), 165–182. https://doi.org/10.1016/j.system.2005.11.001

Graham, S., & Santos, D. (2015). *Strategies for second language listening: Current scenarios and improved pedagogy*. Palgrave Macmillan.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics,* 66(1), 64–74. https://doi.org/10.1119/1.18809

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.

Lee, J. (2015). The impact of metacognitive instruction in blended learning on L2 listeners' metacognitive awareness and comprehension. *Multimedia-Assisted Language Learning,* 18(4), 70–98. https://doi.org/10.15702/mall.2015.18.4.70

Liu, Y. (2020). Effects of metacognitive strategy training on Chinese listening comprehension. *Languages,* 5(2), 21. https://doi.org/10.3390/languages5020021

Loewen, S. (2015). *Introduction to instructed second language acquisition*. Routledge.

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.

Lynch, T. (2009). *Teaching second language listening*. Oxford University Press.

Mackey, A., & Gass, S. M. (2016). *Second language research: Methodology and design* (2nd ed.). Routledge.

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press.

McKenney, S., & Reeves, T. C. (2019). *Conducting educational design research* (2nd ed.). Routledge.

McKinley, J., & Rose, H. (Eds.). (2020). *The Routledge handbook of research methods in applied linguistics*. Routledge.

Mercer, S., & Dörnyei, Z. (2020). *Engaging language learners in contemporary classrooms*. Cambridge University Press.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System,* 41(3), 720–739. https://doi.org/10.1016/j.system.2013.07.013

Mubarok, A., Tabrani, A., Marina, S., Alfarisi, U. M., Al Ayyubi, S., & Hably, R. (2025). The utilization of YouTube channels in teaching listening and speaking skills to non-native Arabic speakers. *Borneo Journal of Language and Education,* 5(2), 322–344. https://doi.org/10.21093/benjole.v5i2.11206

Nation, I. S. P., & Newton, J. M. (2020). *Teaching ESL/EFL listening and speaking* (2nd ed.). Routledge.

Othman, S., Arshad, A., Fauzi, A., & Mohd Zaidin, N. A. (2025). Exploring the integration of YouTube-based learning to improve Arabic learners' listening and speaking skills. *International Journal of Research and Innovation in Social Science,* 9, 4926–4936. https://doi.org/10.47772/IJRISS.2025.909000401

Pei, T., & Suwanthep, J. (2021). Development of L2 listening metacognitive awareness via online metacognitive listening practice. *International Journal of Distance Education Technologies,* 19(4). https://doi.org/10.4018/IJDET.286741

Reinhardt, J. (2019). Social media in second and foreign language teaching and learning: Blogs, wikis, and social networking. *Language Teaching,* 52(1), 1–39. https://doi.org/10.1017/S0261444818000356

Richards, J. C. (2015). *Key issues in language teaching*. Cambridge University Press.

Rodríguez, C. D. (2017). Design-based research. In C. A. Chapelle & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning*. Wiley. https://doi.org/10.1002/9781118914069.ch24

Rost, M. (2016). *Teaching and researching listening* (3rd ed.). Routledge.

Shabani, K., & Heidarian, M. (2015). The effect of metacognitive instruction on Iranian EFL learners' listening abilities and metacognitive awareness. *Journal of Education and Training,* 2(2). https://doi.org/10.5296/jet.v2i2.8033

Shorouk, A. (2020). The effect of using YouTube to increase the level of listening skills among non-native students of Arabic speakers in Malaysian universities. *Education Quarterly Reviews,* 3(2), 207–219. https://doi.org/10.31014/aior.1993.03.02.133

Siegel, J. (2015). Exploring L2 listening instruction: Theoretical, practical, and methodological issues. *System,* 52, 15–26.

Siegel, J. (2018). *Second language listening instruction: Interdisciplinary perspectives*. Routledge.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science,* 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Tomlinson, B. (Ed.). (2018). *Materials development in language teaching* (2nd ed.). Cambridge University Press.

Ur, P. (2018). *A course in language teaching: Practice and theory* (2nd ed.). Cambridge University Press.

Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics,* 24, 3–25. https://doi.org/10.1017/S0267190504000017

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching,* 40(3), 191–210. https://doi.org/10.1017/S0261444807004338

Vandergrift, L., & Cross, J. (2018). *Teaching L2 listening: Metacognition in action*. Routledge.

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning,* 60(2), 470–497.

https://doi.org/10.1111/j.1467-9922.2009.00559.x

Vanderplank, R. (2016). *Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning*. Palgrave Macmillan.

Wang, F. (2021). Design-based research on integrating learning technology tools into higher education classes to achieve active learning. *Computers & Education,* 166, 103935.

Webb, S., & Rodgers, M. P. H. (2009). Vocabulary demands of television programs. *Language Learning,* 59(2), 335–366. https://doi.org/10.1111/j.1467-9922.2009.00509.x

Yang, Z., & Lee, Y.-Y. (2025). Challenges and opportunities of design-based research in applied linguistics: Insights from a scoping review. *Research Methods in Applied Linguistics,* 4(1), 100178. https://doi.org/10.1016/j.rmal.2024.100178

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