Vol. 3 No. 4 (2026): Quartus
https://doi.org/10.69760/jales.2026004010

The Role of Authentic Audiovisual Materials in Developing EFL Speaking Fluency and Communicative Competence

Leyla Gulmammadova
Nakhchivan State University, Nakhchivan, Azerbaijan

Abstract

Speaking instruction in English as a foreign language (EFL) must help learners move from understanding language to using it in interaction. Authentic audiovisual materials can offer contextualized models of pronunciation, vocabulary, discourse, and non-verbal meaning, but exposure alone does not ensure speaking development. This integrative review examines how films, interviews, documentaries, online videos, and other audiovisual resources can support speaking fluency and communicative competence when paired with purposeful learner output. It asks: (1) through which pathways can audiovisual input support oral development; (2) which instructional tasks help learners transform viewing into spoken performance; and (3) what factors limit or shape the usefulness of these resources? The review synthesizes 22 recent peer-reviewed publications from 2022–2026 and eight foundational works. It identifies four recurring principles: select comprehensible and relevant material; segment viewing around a clear language goal; require learners to speak, interact, or record; and provide feedback linked to observable performance criteria. Recent studies on digital storytelling, dubbing, blended input, and video-mediated interaction show that structured production tasks can support aspects of oral performance, while findings across studies remain heterogeneous and context-dependent. The article offers a classroom sequence and assessment framework. It reports no new experiment or original learner data; its conclusions are a synthesis of published research.

Keywords: *authentic audiovisual materials; EFL; speaking fluency; communicative competence; video-based learning; digital storytelling*

How to cite: Gulmammadova, L. (2026). The role of authentic audiovisual materials in developing EFL speaking fluency and communicative competence. *Journal of Azerbaijan Language and Education Studies, 3*(4), 87-95. https://doi.org/10.69760/jales.2026004010

1. Introduction

Developing the ability to speak clearly and appropriately is a central objective of English language teaching. Speaking requires learners to retrieve words, organize ideas, pronounce them intelligibly, and respond to another person in real time. In many English as a foreign language (EFL) settings, however, meaningful exposure to English outside the classroom is limited. Learners may know grammatical rules and vocabulary but still hesitate in spontaneous communication because they have had few opportunities to listen to varied speech and use language for authentic purposes.

Authentic audiovisual materials—such as interviews, documentaries, public talks, films, podcasts with video, and online conversations—can bring contextualized speech into the classroom. In addition to words and grammar, learners can observe prosody, facial expression, gesture, turn-taking, and the

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

relationship between language and situation. Yet teachers should not assume that watching by itself develops speaking. Input must lead to retrieval, rehearsal, interaction, and feedback. Recent evidence on dubbing, digital storytelling, and video-supported communication suggests that it is often the activity surrounding a video, rather than the medium alone, that creates opportunities for oral practice (Jao et al., 2024; Huang, 2024; Wang & Lee, 2025).

The present article recasts the supplied manuscript as an integrative literature review and develops its central argument: audiovisual materials are most useful when they function as the starting point for purposeful speaking tasks. The review is guided by three questions.

RQ1. In what ways can authentic audiovisual input contribute to EFL learners' speaking fluency and communicative competence?

RQ2. Which instructional activities help learners convert viewing and listening into active spoken language use?

RQ3. What learner, material, and classroom factors should guide selection and assessment?

2. Review Approach

This article uses an integrative, focused review rather than a systematic review or meta-analysis. It brings together 22 recent peer-reviewed publications from 2022–2026 that address audiovisual input, video-based speaking practice, dubbing, digital storytelling, subtitling, or video-mediated interaction, together with eight foundational works on authenticity, communicative competence, comprehensible input, output, and multimedia learning. Sources were identified through scholarly publisher and journal records and DOI-linked publication pages, then retained for direct relevance to EFL or second-language oral development and the instructional use of audiovisual media.

The synthesis was organized thematically around four dimensions: (a) exposure to contextualized speech; (b) pronunciation, fluency, vocabulary, and discourse development; (c) learner engagement and interaction; and (d) instructional conditions and limitations. Because the source manuscript did not provide a reproducible database log, duplicate-screening record, or PRISMA flow, the review does not claim exhaustive coverage. It also does not pool effect sizes or treat results from different study designs as directly comparable. The evidence described below belongs to the cited studies, not to a new intervention conducted for this article.

3. Conceptual Foundations

3.1. Authenticity and contextualized language

Authenticity refers both to materials created for communication beyond the classroom and to the way classroom tasks invite meaningful language use. Breen (1985) and Gilmore (2007) caution, in different ways, that authenticity is not a property of a video alone: the task and the learners' engagement with it also matter. A film clip can be authentic as a cultural product yet unsuitable for a particular class if its speech is too fast, its references are unfamiliar, or no meaningful response follows. Teachers should therefore consider authenticity alongside relevance, comprehensibility, and instructional purpose.

Audiovisual materials can make language-context links visible. A learner can infer a speaker's intention from intonation, gesture, and conversational context, while noticing that the same expression may function differently across settings. Recent reviews of audiovisual input and film-based instruction emphasize this multimodal potential while also identifying vocabulary load, captions,

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

learner proficiency, and task design as conditions that affect comprehension (Montero Perez, 2022; Pavithra & Gandhimathi, 2024).

3.2. Communicative competence and interaction

Communicative competence includes grammatical resources as well as the ability to use language appropriately in social contexts (Hymes, 1972; Canale, 1983; Canale & Swain, 1980). For speaking lessons, this means assessing more than error-free sentences. Learners need to initiate and sustain exchanges, respond to others, clarify meaning, take turns, and adjust their language to purpose and audience. Authentic interviews and conversations can model these practices, but students must rehearse them in pair and group interaction. Online conversation projects likewise make interaction central by connecting EFL learners with interlocutors through video conferencing (Tran et al., 2024).

3.3. Input, output, and multimedia processing

Krashen's (1985) account of comprehensible input highlights the importance of language learners can understand while still finding challenging. Swain (1985) adds that producing language can prompt learners to notice gaps in what they can express. Applied to audiovisual instruction, these perspectives suggest a sequence: understand a segment, notice useful language, attempt a spoken response, and revise it after feedback. Multimedia learning theory also helps explain why combining image and sound can support meaning-making, while excessive information or unnecessary visual detail may compete for limited processing resources (Mayer, 2009). Research on media comparisons and on-screen text confirms that captions and subtitles can help some learners while imposing additional processing demands for others (Tarchi & Mason, 2023; Montero Perez, 2022).

4. Synthesis of Recent Evidence

4.1. Natural speech models support noticing, but viewing is not speaking

Authentic video exposes learners to connected speech, reductions, hesitation, discourse markers, varied accents, and speech delivered at natural rates. This exposure can help students notice how words are pronounced and how conversational turns are managed. The 2026 meta-analysis by Sutton and Webb synthesizes 56 experiments and examines several language outcomes, including speaking; its findings indicate that effects vary by video category and that evidence across outcomes should not be reduced to a single claim about fluency. The authors report stronger outcomes for educational and language-focused video categories than for entertainment-focused material, while the included studies differ in design and target skill (Sutton & Webb, 2026).

The distinction between input and output is important. A student may recognize an expression while viewing but fail to retrieve it in conversation. Research on blended input suggests that text-and-video conditions can support speaking development, but instruction still needs explicit speaking practice (Pu & Chang, 2025). Similarly, video-supported digital assignments can combine preparation, recording, and asynchronous interaction, creating practice beyond the class period (Hafour, 2024). These findings support the use of video as a source and prompt for language production, not as a replacement for it. Mobile-assisted language learning (MALL) can extend access to video and speaking practice beyond the classroom; Alisoy and Sadiqzade (2024) discuss both its educational affordances and implementation challenges. In this context, a mobile device is a delivery and recording tool, while learning still depends on the oral task, feedback, and subsequent interaction.

4.2. Speaking gains are associated with structured production tasks

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

Recent intervention studies examine task designs that require learners to produce language in response to audiovisual content. Digital storytelling integrates planning, scripting, recording, and oral presentation. Fu (2022) and Huang (2023) investigate its relationship to speaking proficiency and learner outcomes, while Tatlı et al. (2022) examine speaking and vocabulary development alongside cognitive load in an online EFL course. Robot-supported storytelling has also been studied as a way to increase interaction and narrative engagement (Liang & Hwang, 2023). Together, these studies suggest that learner-generated narration can create repeated opportunities to formulate and deliver spoken language.

Dubbing tasks make the connection between perception and production especially explicit: learners listen to a segment, rehearse timing and pronunciation, and record their own speech. Jao et al. (2024) report improvements in accuracy and fluency in a college-level dubbing project. Later work examines flow and speaking outcomes during video dubbing (Hwang et al., 2025), and app-supported dubbing across different proficiency levels (Wang & Lee, 2025). Reverse dubbing, in which learners translate and voice dialogue from an L1 clip into English, has also been investigated as a task for speaking proficiency and motivation (Huang, 2024). These interventions differ in learners, duration, measures, and implementation, so they support the pedagogical potential of structured practice without establishing one universally superior format.

Saber's (2025) study of authentic video materials with Iraqi undergraduate EFL learners provides a context-specific example of speaking instruction using video. Reviews of digital and film-based learning similarly identify discussion, role-play, retelling, and presentation as common ways to turn viewing into productive activity (Keleş et al., 2025; Pavithra & Gandhimathi, 2024). A meta-analysis of English-subtitled video studies across speaking, listening, and vocabulary also supports using captions as a scaffold while signaling that outcome-specific conclusions should be drawn carefully (Oktapiani et al., 2024). The practical implication is that post-viewing activity should be planned before the material is selected. A short game-based speaking round can provide one such activity. In a qualitative study of educational games in foreign-language classes, Ismayilli et al. (2025) report learners' perceptions of greater confidence and a less stressful setting for speaking. This context-specific, perception-based evidence supports trying low-pressure game prompts after viewing, but it does not establish that games guarantee measurable gains for all learners.

4.3. Communicative competence requires interaction and feedback

Speaking fluency is not simply a high rate of speech. It involves comprehensibility, connected ideas, appropriate timing, and the ability to maintain interaction. Digital storytelling studies link production tasks with speaking performance and willingness to communicate, while video-mediated conversation research emphasizes regular interaction with interlocutors as a route to meaningful listening and speaking practice (Huang, 2023; Tran et al., 2024). Clustered digital materials can also provide varied prompts for situational language practice when students collaborate and respond to one another (Salih & Omar, 2024).

Feedback helps learners convert repeated performance into improvement. Instructors can ask students to notice one feature at a time—such as pausing, intelligibility, use of target expressions, or response to a partner—rather than attempting to correct every error. Learner-generated video can support self-observation and reflection, while peer response can draw attention to comprehensibility and interaction. These practices should be supported by clear criteria and an opportunity to try again.

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

Video-mediated communication reviews also emphasize the role of task structure, interlocutor support, and access when organizing online speaking practice (Keleş et al., 2025).

4.4. Material selection and learner differences shape outcomes

The phrase "authentic video" covers a wide range of material, from short interviews to full-length films. A suitable clip should have a clear purpose, a manageable length, understandable audio, and language that is useful for the target group. Teachers can preview the clip, identify unfamiliar cultural references, prepare key vocabulary, and provide captions selectively. Subtitles may support access, but they can draw attention away from listening or encourage reading instead of processing speech; the choice should reflect the learning goal and learner proficiency (Montero Perez, 2022; Tarchi & Mason, 2023).

Learners also differ in background knowledge, confidence, listening proficiency, and familiarity with technology. A clip that encourages advanced learners to analyze stance and register may overwhelm beginners. Shorter segments, a vocabulary preview, repeated viewing, and optional captions can reduce barriers. At the same time, simplifying every text can deprive learners of exposure to natural speech. A gradual sequence from supported comprehension to independent response offers a practical balance. Recent classroom research on video dubbing and digital storytelling highlights the importance of proficiency, engagement, and scaffolding in interpreting outcomes (Wang & Lee, 2025; Tatlı et al., 2022). Emerging app-based speaking practice also points to the potential value of guided after-class feedback, although this evidence is adjacent to rather than direct evidence about authentic audiovisual input (Ma et al., 2025).

5. A Classroom Model for Video-Supported Speaking

The following sequence adapts the original manuscript's classroom model into a goal-based cycle. It can be used with a short interview, documentary extract, public talk, or scene from a film. Teachers should match the speech task to course level and allow alternatives when technology or captions are inaccessible.

Table 1. A goal-based sequence for audiovisual speaking lessons

Stage	Teacher and learner activity	Primary purpose
Before viewing	State the speaking outcome; activate background knowledge; preview essential vocabulary; ask learners to predict the situation.	Prepare comprehension and give viewing a clear purpose.
First viewing	Watch or listen for the main idea and identify speakers, setting, and communicative goal.	Build global comprehension without overloading learners.
Focused viewing	Replay a short segment; notice expressions, turn-taking, stress, intonation, or non-verbal cues. Use captions selectively.	Direct attention to language features relevant to the lesson.
Rehearsal	Shadow a short phrase, retell the segment, prepare a response, or record a first attempt.	Move from recognition to supported production.
Interaction and feedback	Discuss, role-play, dub, present, or play a low-stakes speaking game; give peer/teacher feedback and allow a second attempt.	Develop fluency, comprehensibility, interaction, and self-monitoring.

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

Assessment can use a brief analytic rubric covering task completion, comprehensibility, fluency, use of target expressions, and interaction. The rubric should reflect the lesson's stated objective: a pronunciation lesson should not be graded mainly on vocabulary range, and an interaction lesson should consider how learners respond to others. A brief pre-task and post-task performance sample may help teachers monitor progress, but any research claim based on such assessment requires a documented instrument, scoring process, and actual learner data.

6. Discussion

The review questions are answered through a consistent principle: audiovisual materials can supply contextualized models and communicative prompts, but speaking development depends on how learners use them. For RQ1, potential pathways include noticing pronunciation and discourse, connecting language with visual context, and acquiring expressions for interaction. For RQ2, the literature points to structured production tasks—retelling, discussion, role-play, dubbing, digital storytelling, and video-mediated conversation—as mechanisms that require retrieval and performance. For RQ3, proficiency, material difficulty, subtitle use, cognitive load, access, and feedback shape what learners can gain.

The evidence supports a measured conclusion rather than a blanket claim that authentic video automatically improves fluency. Meta-analyses and reviews include different languages, learners, video types, measures, and instructional designs. The 2026 synthesis by Sutton and Webb considers multiple L2 outcomes and notes category-level differences; individual classroom studies address narrower tasks and populations. Digital storytelling and dubbing research offers promising examples of active production, but transfer to spontaneous interaction and long-term retention requires further study. The outcomes of a well-designed task in one setting should not be generalized to all EFL learners.

Several practical conditions follow. First, teachers should select material after defining the oral outcome. Second, viewing should be segmented and paired with purposeful output. Third, feedback should be specific and followed by another performance opportunity. Finally, access and learner differences should be considered in advance. These principles preserve the original manuscript's emphasis on pronunciation, vocabulary, confidence, interaction, and appropriate material selection while grounding the claims in a clearer synthesis of recent research.

7. Limitations and Future Research

This focused review is limited by the lack of a registered protocol and reproducible database search record; it should not be described as a systematic review. The included sources use diverse designs and outcome measures, and the review did not calculate a pooled effect. Some studies examine speaking directly, whereas others focus on wider language learning, listening, engagement, or multimodal storytelling. Future studies should report participant characteristics, intervention length, material-selection rules, caption conditions, speaking measures, rater procedures, reliability evidence, and delayed outcomes. Research in underrepresented EFL contexts and comparisons across proficiency levels would clarify how the framework transfers to different classrooms.

The article presents no newly collected classroom data. Any future empirical submission based on the proposed classroom sequence should include only procedures actually conducted and results calculated

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

from the collected data. Sample sizes, scores, significance tests, and participant quotations must not be supplied as hypothetical findings.

8. Conclusion

Authentic audiovisual resources can enrich EFL speaking instruction by exposing learners to contextualized language and providing meaningful prompts for oral practice. Their value depends on purposeful selection and active use. A lesson should move learners from comprehension to noticing, rehearsal, interaction, feedback, and transfer. Films, interviews, documentaries, digital stories, and video-mediated conversations can support this sequence, but passive viewing alone cannot demonstrate communicative competence. This integrative review offers a practical classroom model and a cautious interpretation of current evidence. Its conclusions are literature-based and should be tested through transparent empirical research.

Declarations

Ethics Approval: Not applicable; this integrative review reports no research involving human participants.

Data Availability: Not applicable; no new dataset was generated or analysed.

Funding: This research received no external funding.

Conflict of Interest: The author declares no conflict of interest.

References

Alisoy, H., & Sadiqzade, Z. (2024). Mobile-assisted language learning (MALL): Revolutionizing language education. *Luminis Applied Science and Engineering, 1*(1), 60–72. https://doi.org/10.69760/lumin.202400002

Breen, M. P. (1985). Authenticity in the language classroom. *Applied Linguistics, 6*(1), 60–70.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 2–27). Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47.

Fu, J. S. (2022). Exploring the impacts of digital storytelling on English as a foreign language learners' speaking competence. *Journal of Research on Technology in Education, 54*(5), 679–694. https://doi.org/10.1080/15391523.2021.1911008

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching, 40*(2), 97–118.

Hafour, M. (2024). Interactive digital media assignments: Effects on EFL learners' overall and micro-level oral language skills. *Computer Assisted Language Learning, 37*(4), 986–1018. https://doi.org/10.1080/09588221.2022.2067180

Huang, H.-T. D. (2023). Examining the effect of digital storytelling on English speaking proficiency, willingness to communicate, and group cohesion. *TESOL Quarterly, 57*(1), 242–269. https://doi.org/10.1002/tesq.3147

Huang, H.-T. D. (2024). Investigating the effects of reverse dubbing tasks on English speaking proficiency and English learning motivation. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688241283153

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

Hwang, G.-J., Fathi, J., & Rahimi, M. (2025). Fostering EFL learners' speaking skills and flow experience with video-dubbing tasks: A flow theory perspective. *Journal of Computer Assisted Learning, 41*(2), e13120. https://doi.org/10.1111/jcal.13120

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.

Ismayilli, T. M., Mammadova, K. M., & Asadova, A. A. (2025). The impact of educational games on speaking skills in the foreign language teaching process. *Novitas-ROYAL (Research on Youth and Language), 19*(1), 229–240. https://doi.org/10.5281/zenodo.15228398

Jao, C.-Y., Yeh, H.-C., Huang, W.-R., & Chen, N.-S. (2024). Using video dubbing to foster college students' English-speaking ability. *Computer Assisted Language Learning, 37*(4), 585–607. https://doi.org/10.1080/09588221.2022.2049824

Keleş, N., Öztürk-Taş, H., Polat, H., Kurşun, E., Aydemir-Arslan, M., & Yıldız, M. (2025). The use of video-mediated communication tools for English language speaking practice: A systematic review. *Anadolu Journal of Educational Sciences International, 15*(4), 1629–1661. https://doi.org/10.18039/ajesi.1668148

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

Liang, J.-C., & Hwang, G.-J. (2023). A robot-based digital storytelling approach to enhancing EFL learners' multimodal storytelling ability and narrative engagement. *Computers & Education, 201*, Article 104827. https://doi.org/10.1016/j.compedu.2023.104827

Ma, M., Noordin, N., & Razali, A. B. (2025). Improving EFL speaking performance among undergraduate students with an AI-powered mobile app in after-class assignments: An empirical investigation. *Humanities and Social Sciences Communications, 12*, Article 370. https://doi.org/10.1057/s41599-025-04688-0

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Montero Perez, M. (2022). Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching, 55*(2), 163–192. https://doi.org/10.1017/S0261444821000501

Oktapiani, I., Ilham, I., Humaira, H., Bafadal, F., Adebayo, M., & Fasih, P. (2024). The effectiveness of English-subtitled videos in enhancing speaking, listening and vocabulary skills: A meta-analysis. *Journal of Languages and Language Teaching, 12*(1), 56–71. https://doi.org/10.33394/jollt.v12i1.9501

Pavithra, K., & Gandhimathi, S. N. S. (2024). A systematic review of empirical studies incorporating English movies as pedagogic aids in English language classroom. *Frontiers in Education, 9*, Article 1383977. https://doi.org/10.3389/feduc.2024.1383977

Pu, P., & Chang, D.-Y. S. (2025). Effects of different input modes on blended EFL speaking instruction: A quasi-experimental study. *Computer Assisted Language Learning, 38*(5–6), 1223–1248. https://doi.org/10.1080/09588221.2023.2273853

Saber, A. F. (2025). Examining the effectiveness of authentic video materials in developing speaking skills among undergraduate EFL Iraqi learners. *Journal of Tikrit University for Humanities, 32*(11, Part 2), 22–37. https://doi.org/10.25130/jtuh.32.11.2.2025.24

Salih, A. A., & Omar, L. I. (2024). Enhancing EFL learners' engagement in situational language skills through clustered digital materials. *Frontiers in Education, 9*, Article 1439104. https://doi.org/10.3389/feduc.2024.1439104

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**

Sutton, D., & Webb, S. (2026). The effects of audiovisual input on second language learning: A meta-analysis. *Studies in Second Language Acquisition, 48*(2), 571–595. https://doi.org/10.1017/S0272263126101612

Tarchi, C., & Mason, L. (2023). Learning across media in a second language. *European Journal of Psychology of Education, 38*, 1593–1618. https://doi.org/10.1007/s10212-022-00652-7

Tatlı, Z., Saylan, E., & Kokoç, M. (2022). Digital storytelling in an online EFL course: Influences on speaking, vocabulary, and cognitive load. *Participatory Educational Research, 9*(6), 89–112. https://doi.org/10.17275/per.22.130.9.6

Tran, N., Hoang, D. T. N., Gillespie, R., Yen, T. T. H., & Phung, H. (2024). Enhancing EFL learners' speaking and listening skills through authentic online conversations with video conferencing tools. *Innovation in Language Learning and Teaching*. Advance online publication. https://doi.org/10.1080/17501229.2024.2334809

Author	**Leyla Gulmammadova** Nakhchivan State University Nakhchivan, Azerbaijan Email: leylagulmemmedova1998@gmail.com ORCID: https://orcid.org/0009-0007-4358-501X

Received: 05.09.2026
Revised: 10.09.2026
Accepted: 22.11.2026
Published: 23.11.2026

This is an open access article under the Creative Commons Attribution 4.0 International License **ISSN 3078-6177**