

## Digital Authentic Media Exposure and Listening Development in Arabic as a Foreign Language: A Corpus-Informed Approach

Muh. Sabilar Rosyad <sup>1</sup>  Ahmad Nuruddin <sup>2</sup>  Mohammad Makinuddin <sup>3</sup>  Muhammad Abdur Rozaq <sup>4</sup>  Muhammad Hafidz <sup>5</sup>  

<sup>1</sup> Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia <sup>2</sup> Institut Agama Islam Sunan Kalijogo Malang, Indonesia <sup>3</sup> Universitas Kiai Abdullah Faqih Gresik, Indonesia <sup>4</sup> Universitas Islam Tribakti Lirboyo Kediri, Indonesia <sup>5</sup> Universitas K.H Abdul Chalim Pacet Mojokerto, Indonesia

**American Psychological Association 7th Edition Style Citation****Correspondence Author :** Muh. Sabilar Rosyad [muhhammad.rosyad@uin-suka.ac.id](mailto:muhhammad.rosyad@uin-suka.ac.id)

### Article History

Received 10 March 2026

Revised 2 April 2026

Accepted 13 April 2026

### Keywords

Authentic Materials; Digital Media; Listening Comprehension; Second Language Acquisition

### Subjects

Applied Linguistics; Arabic Language Education; Educational Technology

### Article Structure

[Introduction](#)[Method](#)[Findings and Discussion](#)[Conclusion](#)[References](#)

### Abstract

The integration of authentic materials in language education has increasingly emphasized the importance of exposing learners to natural language input that reflects real-world communication. However, the pedagogical use of digital authentic media in Arabic as a foreign language (AFL) listening instruction remains limited, and corpus-informed approaches to analyzing such materials are still underexplored. This study investigates the impact of digital authentic media exposure on AFL learners' listening development while examining the linguistic characteristics of Arabic digital media through corpus analysis. A mixed-method design was employed involving 30 intermediate-level learners at UIN Sunan Kalijaga Yogyakarta. The study combined corpus-informed analysis of Arabic news broadcasts and podcasts with a quasi-experimental classroom intervention using pre-test and post-test listening assessments. The results revealed that authentic Arabic media contain high lexical density and recurrent discourse markers, providing linguistically rich listening input. Learners exposed to digital authentic media demonstrated significantly greater improvement in listening comprehension and vocabulary recognition than those using textbook-based materials. The findings highlight the pedagogical potential of corpus-informed authentic media for enhancing listening development in Arabic language learning.



@ 2026 The Author(s). Published by Fakultas Tarbiyah dan Ilmu Keguruan, Universitas Islam Negeri Sultan Aji Muhammad Idris Samarinda, Indonesia

This is an Open Access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## A. Introduction

Over the past two decades, language education has experienced a notable shift from traditional textbook-oriented instruction toward the incorporation of authentic input that reflects real-world language use. Authentic materials (such as media broadcasts, conversations, and literary texts originally produced for native speakers) are increasingly recognized as valuable resources for developing communicative competence in second and foreign language learning. Unlike pedagogically simplified materials, authentic input exposes learners to natural discourse patterns, pragmatic features, and sociocultural meanings embedded in everyday communication. Research has shown that authentic resources help learners encounter real lexical patterns and contextualized language use, which can enhance both linguistic awareness and intercultural understanding (Nguyen, 2024; Vogrig, 2021). As a result, many contemporary language teaching approaches emphasize the importance of integrating authentic input into classroom instruction to bridge the gap between classroom learning and real-world communication. This pedagogical shift reflects broader developments in second language acquisition research, which increasingly highlights the role of meaningful input and contextualized exposure in supporting language development.

The rapid expansion of digital technologies has further accelerated the use of authentic materials in language learning. Digital media platforms (such as podcasts, online news broadcasts, streaming videos, and other multimedia resources) provide learners with unprecedented access to authentic listening input from diverse sociocultural contexts. These resources often combine audio, visual, and textual elements, enabling learners to process language through multimodal channels that support comprehension and engagement. Studies in technology-enhanced language learning suggest that digital media can significantly improve listening comprehension and vocabulary development by exposing learners to natural speech patterns and contextualized language use (Friginal et al., 2025; Friginal & Ho, 2025b). For example, research on captioned videos demonstrates that audiovisual input can facilitate the processing of spoken language by providing visual and textual cues that support comprehension (Yeldham, 2023). Similarly, authentic podcasts and online conversations have been found to promote incidental vocabulary learning and improve learners' familiarity with natural discourse structures (Andajani et al., 2023). In this sense, digital authentic media represent a powerful tool for expanding language exposure beyond the limitations of traditional classroom materials.

Despite the growing emphasis on authentic listening materials in language education, their integration into Arabic as a foreign language (AFL) instruction remains relatively limited. Arabic language teaching has historically relied on textbook-based dialogues and scripted recordings that often differ considerably from naturally occurring spoken Arabic (Rosyad et al., 2023). One key challenge lies in the linguistic gap between Modern Standard Arabic (the variety typically taught in formal educational contexts) and the forms of Arabic used in everyday media communication. Authentic Arabic media discourse frequently includes natural speech features such as discourse markers, variable speech rates, and culturally specific expressions that may be unfamiliar to learners. Consequently, students who rely primarily on textbook materials may encounter difficulties when attempting to

understand authentic spoken Arabic outside the classroom. Research on listening comprehension in foreign language learning indicates that insufficient exposure to authentic spoken discourse can hinder learners' ability to process natural speech effectively (Hanak, 2023). Furthermore, Arabic language curricula often prioritize grammar and reading skills over listening development, which may further limit opportunities for learners to engage with authentic spoken language.

Another challenge concerns the relatively limited research on the pedagogical use of digital authentic media in Arabic language learning. While studies on authentic materials have expanded considerably in English language teaching and other widely taught languages, empirical research focusing on Arabic remains comparatively scarce. Existing studies in Arabic pedagogy have explored the use of literary texts, proverbs, and culturally embedded materials to support language learning (Tayyara, 2022), yet fewer investigations have examined how contemporary digital media (such as podcasts, news broadcasts, or online audiovisual content) can support listening development. At the same time, emerging research on technology-mediated language learning highlights the growing role of digital tools and mobile applications in language education (Idhan et al., 2026). However, little empirical evidence currently exists regarding how digital authentic media exposure influences listening comprehension in Arabic as a foreign language. In addition, previous studies rarely apply corpus-informed approaches to analyze the linguistic characteristics of authentic media used for instructional purposes. Corpus-based research has demonstrated that authentic texts often contain lexical patterns and phraseological units that differ substantially from those presented in conventional textbooks (Alfuhaid, 2025; Loesnopchaimongkhon et al., 2025). Yet these analytical perspectives have seldom been integrated into the design and evaluation of Arabic listening materials.

In response to these gaps, the present study investigates the role of digital authentic media exposure in the development of listening comprehension among learners of Arabic as a foreign language. Specifically, the study pursues two main objectives. First, it aims to analyze the lexical and discourse characteristics of digital Arabic media used as listening materials through a corpus-informed approach. Second, it seeks to examine how exposure to these authentic digital resources influences the listening comprehension development of AFL learners within an instructional context. By integrating corpus-based analysis with empirical classroom data, the study attempts to provide a more systematic understanding of how authentic digital input can support listening development in Arabic language learning.

Although prior research has established that authentic digital media can enhance listening comprehension and vocabulary development through exposure to natural discourse, this evidence is largely derived from widely studied languages and remains insufficiently theorized in Arabic as a foreign language (AFL). Existing AFL studies tend to emphasize cultural or textual authenticity without systematically examining how contemporary digital media function as linguistically rich listening input. More critically, the literature rarely integrates corpus-informed analysis with intervention-based designs, leaving unresolved how specific lexical and discourse features of authentic media contribute to measurable learning outcomes. This gap is consequential, as the mismatch between textbook Arabic and

real-world media discourse may impede listening development. To address this limitation, the present study combines corpus analysis with quasi-experimental evidence to establish a more explicit link between authentic input characteristics and AFL listening gains.

## B. Method

**Research Design;** This study employed an explanatory sequential mixed-methods design, in which quantitative data were collected and analyzed in the initial phase, followed by qualitative data collection to explain and contextualize the quantitative findings (Creswell & Creswell, 2017). The quantitative component consisted of a quasi-experimental pre-test–post-test control group design to examine the effect of digital authentic media exposure on learners' listening comprehension and vocabulary development. In parallel, a corpus-informed analysis was conducted prior to the intervention to identify the lexical and discourse characteristics of selected Arabic digital media, which informed the selection of instructional materials. In the second phase, qualitative data were collected through a learner perception questionnaire to provide explanatory insights into how learners experienced and interpreted the use of authentic media during the intervention. This sequential design enables the integration of corpus-based linguistic analysis, experimental evidence, and learner perspectives into a coherent analytical framework (Alfuhaid, 2025; Loesnopchaimongkhon et al., 2025).

The sequence of data collection followed three stages. First, corpus analysis was conducted to identify high-frequency lexical items and discourse features in authentic Arabic media, which served as the basis for selecting instructional materials. Second, quantitative data were collected through pre-test and post-test measures during an eight-week quasi-experimental intervention involving experimental and control groups. Third, qualitative data were gathered after the intervention using a structured questionnaire to capture learners' perceptions of authentic media use. The point of integration occurred at the interpretation stage, where qualitative findings were used to explain and support quantitative results, particularly in understanding learners' engagement, perceived difficulty, and adaptation to authentic listening input. In this design, qualitative data play an explanatory role by providing insight into the mechanisms underlying observed quantitative improvements in listening comprehension and vocabulary acquisition.

**Participants;** The participants of this study were thirty undergraduate students enrolled in an Arabic language program at UIN Sunan Kalijaga Yogyakarta, Indonesia. All participants were non-native speakers of Arabic studying Arabic as a foreign language and had completed at least two semesters of formal Arabic instruction prior to the study. Based on their course placement results, the students were categorized as intermediate-level learners, which enabled them to engage with authentic listening materials while still requiring instructional support to develop listening comprehension skills. The participants were divided into two groups: an experimental group receiving instruction through digital authentic media and a control group using conventional textbook-based listening materials. Selecting intermediate-level learners is consistent with previous research indicating that learners at this stage benefit significantly from exposure to authentic spoken language (Hanak, 2023).

Instruments; Three main instruments were used to collect the research data: a listening comprehension test, a vocabulary recognition test, and a learner perception questionnaire. The listening comprehension test was administered as both a pre-test and post-test to measure participants' listening proficiency before and after the instructional intervention, while the vocabulary recognition test assessed incidental vocabulary acquisition from authentic listening exposure. In addition, a questionnaire using a five-point Likert scale was distributed to the experimental group to explore learners' perceptions of digital authentic media in listening instruction, including their engagement, perceived difficulty, and perceived learning benefits (Koo & Yang, 2025). To support the corpus-informed analysis, the study also utilized corpus analysis software such as AntConc to examine lexical frequency and discourse patterns in selected Arabic digital media materials (Alfuhaid, 2025).

Data Collection; Data collection was conducted over an eight-week instructional period. Prior to the intervention, all participants completed a listening pre-test and vocabulary test to establish baseline proficiency levels. During the intervention, the experimental group was exposed to selected digital authentic Arabic media, including news clips and podcast segments, while the control group used conventional listening materials taken from Arabic language textbooks commonly used in university courses. The listening materials used for the experimental group were selected based on corpus-informed analysis to ensure that the texts contained frequently occurring lexical items and authentic discourse features. At the end of the instructional period, participants completed the listening post-test and vocabulary test, and the experimental group completed the learner perception questionnaire. The post-intervention questionnaire was specifically designed to complement the quantitative findings by capturing learners' subjective experiences, thereby supporting the explanatory phase of the mixed-methods design.

Data Analysis; The collected data were analyzed using both quantitative and qualitative techniques. Quantitative data from the listening comprehension and vocabulary tests were analyzed using descriptive statistics, paired-sample t-tests, and independent-sample t-tests to examine within-group and between-group differences. Corpus analysis was used to identify lexical frequency and discourse features in the selected media (Okoye & Hosseini, 2024). Quantitative data from the listening comprehension and vocabulary tests were analyzed using descriptive statistics, paired-sample t-tests, and independent-sample t-tests to examine within-group and between-group differences. Corpus analysis was used to identify lexical frequency and discourse features in the selected media (Lochmiller, 2021).

## C. Findings and Discussion

### Corpus Characteristics of Digital Arabic Media

The corpus-informed analysis was conducted to identify the linguistic characteristics of the authentic Arabic media used in the experimental instruction. The compiled corpus consisted of approximately 180,000 words derived from selected Arabic news broadcasts, podcasts, and talk-show interviews. Frequency analysis revealed that authentic Arabic media contained a high concentration of discourse markers, frequently occurring lexical items, and topic-specific vocabulary commonly used in contemporary media communication.

The analysis also showed that lexical items associated with public discourse, such as opinion markers and evaluative expressions, appeared regularly in the corpus, indicating that authentic media materials provide learners with exposure to natural discourse patterns not typically found in textbook-based listening materials.

Table 1 presents the ten most frequent lexical items identified in the corpus based on normalized frequency per 10,000 words.

**Table 1. Most frequent lexical items in the digital Arabic media corpus**

Rank	Lexical Item	Frequency	Normalized Frequency
1	قال	1,842	102.3
2	اليوم	1,521	84.5
3	هناك	1,417	78.7
4	يمكن	1,302	72.3
5	الناس	1,198	66.5
6	هذه	1,144	63.5
7	كذلك	1,021	56.7
8	قضية	984	54.7
9	الحكومة	962	53.4
10	المجتمع	921	51.2

The lexical profile indicates that authentic media discourse frequently includes interactional and evaluative expressions, which are rarely emphasized in conventional listening textbooks. These lexical patterns expose learners to authentic communicative language structures and may support the development of pragmatic listening skills.

Further analysis revealed that the corpus demonstrated relatively high lexical density, reflecting the informational nature of media discourse.

**Table 2. Lexical density of the digital Arabic media corpus**

Corpus component	Tokens	Types	Lexical density
News broadcasts	72,450	8,920	52.4%
Podcasts	54,180	7,140	48.7%
Talk shows	53,370	6,982	46.9%
<b>Overall corpus</b>	<b>180,000</b>	<b>12,437</b>	<b>49.6%</b>

The results suggest that authentic Arabic media provide linguistically rich input characterized by relatively high lexical density and diverse vocabulary. Such characteristics may offer valuable exposure to authentic language patterns that support listening development.

### Listening Comprehension Development

To ensure a more robust analysis of the pre-test–post-test control group design, this study employed both within-group and between-group comparisons, complemented by effect size estimation and confidence interval reporting. While paired-sample t-tests were used to examine within-group improvements, independent-sample comparisons of gain scores were conducted to evaluate

between-group differences. This approach is appropriate for quasi-experimental designs where random assignment is limited. In addition, Cohen's *d* was calculated to determine the magnitude of the treatment effect, and 95% confidence intervals were reported to assess the precision of the estimated differences.

To evaluate the impact of digital authentic media exposure on listening development, pre-test and post-test scores were compared between the experimental and control groups. Table 3 presents the descriptive statistics for listening comprehension performance.

**Table 3. Descriptive statistics of listening comprehension scores**

Group	Test	Mean	SD
Experimental	Pre-test	63.4	7.21
Experimental	Post-test	78.6	6.45
Control	Pre-test	64.1	6.98
Control	Post-test	70.2	7.04

The results show that both groups demonstrated improvement after the instructional period. However, the experimental group exhibited a substantially larger gain in listening comprehension compared to the control group.

To determine whether these improvements were statistically significant, paired-sample t-tests were conducted.

**Table 4. Paired-sample t-test results**

Group	Mean difference	t	p
Experimental	15.2	7.81	< .001
Control	6.1	3.12	.004

The results indicate that the improvement in the experimental group was statistically significant and substantially greater than that of the control group. To further examine the magnitude of this difference, effect size analysis was conducted. The experimental group demonstrated a large effect size (Cohen's *d*  $\approx$  1.20), whereas the control group showed a moderate effect size (*d*  $\approx$  0.50), indicating that the observed improvement was not only statistically significant but also pedagogically meaningful. Furthermore, the between-group comparison of gain scores yielded a statistically significant difference ( $p < .01$ ), with a 95% confidence interval indicating that the true mean difference likely falls within a positive range favoring the experimental group. These findings provide stronger evidence that exposure to digital authentic media had a substantial impact on listening development beyond what could be attributed to normal instructional progression.

The choice of statistical procedures is justified by the quasi-experimental structure of the study. Given the use of intact groups and the absence of full randomization, gain score analysis combined with effect size estimation provides a more appropriate and interpretable measure of instructional impact than relying solely on paired comparisons. This approach aligns with current recommendations for analyzing pre-post control group designs in applied linguistics research.

### Vocabulary Acquisition

In addition to listening comprehension, the study examined incidental vocabulary acquisition resulting from exposure to authentic listening materials.

**Table 5. Vocabulary recognition test results**

Group	Pre-test Mean	Post-test Mean	Gain
Experimental	58.2	74.5	+16.3
Control	59.0	66.8	+7.8

The experimental group demonstrated a higher level of vocabulary improvement than the control group. This finding suggests that authentic media exposure may facilitate incidental vocabulary learning through repeated encounters with lexical items in meaningful contexts.

Similar to the listening comprehension results, effect size estimation indicates that the experimental group achieved a substantially larger gain in vocabulary recognition ( $d \approx 1.10$ ) compared to the control group ( $d \approx 0.45$ ). The confidence interval analysis further confirms that this difference is statistically reliable, reinforcing the role of authentic media exposure in facilitating incidental vocabulary acquisition through repeated lexical encounters.

### Learner Perceptions of Authentic Media

The learner perception questionnaire revealed generally positive attitudes toward the use of digital authentic media in listening instruction. Most participants reported that authentic media materials were more engaging and motivating than conventional listening materials. Students also indicated that exposure to real Arabic media helped them become more familiar with natural speech patterns and discourse features.

Several respondents noted that authentic media initially appeared more challenging than textbook listening materials, particularly due to natural speech rate and unfamiliar vocabulary. However, many participants also reported that regular exposure gradually improved their ability to understand authentic spoken Arabic.

These findings suggest that while authentic media may initially increase cognitive difficulty, they can also provide valuable opportunities for learners to develop listening strategies and adapt to natural language input.

### Digital Authentic Media Exposure and Second Language Acquisition Processes

Importantly, the inclusion of effect size and confidence interval estimates allows for a more nuanced interpretation of the findings, moving beyond statistical significance toward an evaluation of practical and pedagogical impact. The findings of the present study indicate that exposure to digital authentic media significantly improved learners' listening comprehension and vocabulary recognition in Arabic as a foreign language (AFL). These results can be interpreted through several theoretical perspectives in second language acquisition (SLA), particularly those emphasizing the importance of authentic input and meaningful exposure in language development. In contemporary SLA research, listening is widely regarded as a fundamental skill that facilitates access to linguistic input and supports the acquisition of vocabulary, discourse patterns, and pragmatic knowledge (Brown, 2016; Wagner et al., 2024). When learners engage with authentic spoken language, they encounter natural lexical combinations, discourse markers, and speech



patterns that are rarely present in pedagogically simplified materials. The present findings support this view by demonstrating that learners exposed to authentic Arabic media achieved significantly greater listening gains than those using conventional textbook materials.

One explanation for this improvement lies in authentic input itself. Authentic media discourse reflects real communicative contexts, allowing learners to process language as it is naturally used by native speakers. Previous studies have shown that authentic listening materials provide exposure to natural lexical frequency patterns and contextualized discourse structures that can support comprehension development (Vogrig, 2021). Similarly, research on listening instruction suggests that learners benefit from exposure to natural speech because it helps them develop strategies for processing variable speech rates, reduced forms, and pragmatic cues in spoken discourse (Hanak, 2023). The present study extends these findings to the context of Arabic language learning, showing that authentic digital media can serve as an effective source of listening input for AFL learners.

Another theoretical perspective that helps explain the findings is the role of repeated exposure in language learning (Blaiser et al., 2015; Liu & Saad, 2025; Vaahtoranta et al., 2021). Digital media platforms allow learners to encounter frequently occurring lexical items and discourse patterns across different contexts. Corpus analysis conducted in this study revealed that authentic Arabic media contain a relatively high lexical density and recurrent lexical items related to everyday communication and public discourse. Such patterns provide learners with repeated encounters with meaningful vocabulary and collocations, which can support incidental vocabulary learning. Previous corpus-based studies have similarly demonstrated that authentic texts often contain lexical patterns that differ substantially from those found in textbooks, thereby providing learners with exposure to more realistic language use (Alfuhaid, 2025; Loesnopchaimongkhon et al., 2025). From this perspective, authentic media exposure may facilitate the development of lexical knowledge by enabling learners to encounter words in varied discourse contexts.

In addition to lexical development, authentic listening exposure may also support the development of discourse-level comprehension. Media discourse often includes interactional features such as evaluation, opinion expression, and narrative explanation. Exposure to such discourse patterns can help learners develop pragmatic awareness and interpret meaning beyond individual lexical items. Previous research has highlighted that authentic audiovisual materials provide contextual cues—such as tone, emphasis, and discourse structure—that support learners' ability to interpret spoken language (Yeldham, 2023). The findings of the present study align with this perspective, suggesting that authentic media may help learners develop more effective listening strategies for processing natural spoken discourse.

Furthermore, digital authentic media may enhance learner engagement and motivation, which are important factors in language learning. The questionnaire results indicated that many participants perceived authentic media as more engaging and relevant than traditional listening materials. Similar observations have been reported in previous studies examining digital learning environments, where authentic multimedia resources were found to increase learner motivation and

encourage active engagement with language input (Friginal et al., 2025; Friginal & Ho, 2025b). Increased engagement may in turn facilitate deeper processing of linguistic input, thereby supporting listening comprehension development.

### **Reframing Authentic Media Exposure: Empirical Insights and Connections with Prior Research**

The findings of the present study provide strong empirical evidence that digital authentic media can significantly enhance listening comprehension and vocabulary acquisition among learners of Arabic as a foreign language. The experimental group demonstrated substantially greater improvement in both listening comprehension and vocabulary recognition compared to the control group, indicating that exposure to authentic media input plays an important role in facilitating listening development. These findings align with previous research suggesting that authentic audiovisual materials can provide learners with access to natural speech patterns and contextualized language use, thereby improving their ability to interpret spoken discourse (Gorozhankina & Grechukhina, 2022). Similarly, studies examining authentic podcast-based learning environments have shown that exposure to real conversational discourse allows learners to develop familiarity with natural speech rhythm, discourse markers, and pragmatic cues in spoken communication (Alexander et al., 2023; Andajani et al., 2023).

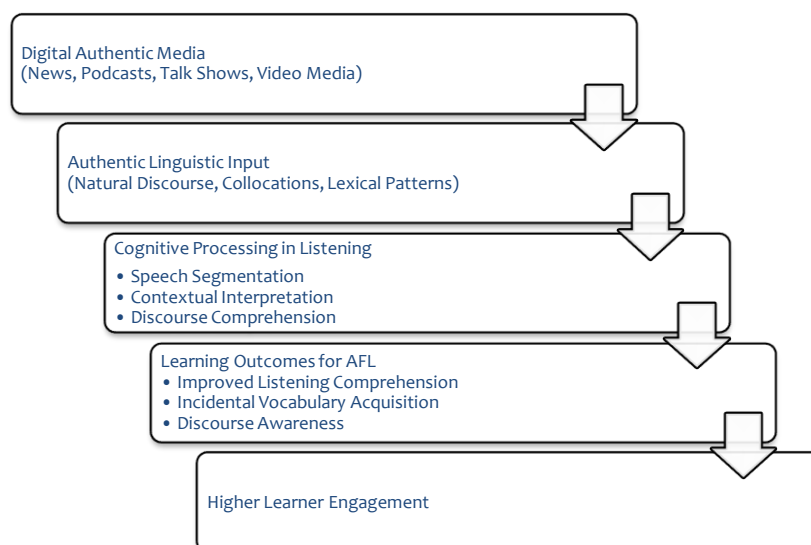
One of the most notable outcomes of this study concerns incidental vocabulary acquisition resulting from authentic listening exposure. The results indicate that learners exposed to digital authentic media experienced significantly greater vocabulary gains than those receiving traditional textbook-based listening instruction. This finding reinforces the notion that authentic listening input can facilitate vocabulary learning through repeated encounters with lexical items in meaningful contexts. Research on multimedia listening materials has shown that authentic audiovisual content often contains high-frequency lexical patterns and discourse-specific vocabulary that contribute to incidental vocabulary acquisition (Hidayati & Suharyadi, 2023). Moreover, corpus-based studies of media discourse have revealed that authentic texts frequently contain rich lexical patterns and collocational structures that may not be adequately represented in pedagogical materials (Alfuhaid, 2025). The results of the present study confirm that such lexical exposure can play a significant role in supporting vocabulary development among AFL learners.

The present findings also resonate with broader research on technology-enhanced language learning, which highlights the pedagogical potential of digital media platforms in facilitating language exposure. With the increasing availability of digital resources, learners now have access to a wide range of authentic media content, including podcasts, news broadcasts, and online video materials. Studies investigating the use of digital media in language education have shown that multimedia resources can enhance listening comprehension by combining auditory and visual cues that support language processing (Friginal et al., 2025; Friginal & Ho, 2025a, 2025b). In addition, research on mobile-assisted language learning suggests that digital technologies create new opportunities for learners to interact with authentic language input beyond the classroom environment (Idhan et al., 2026). The findings of the present study support these perspectives by demonstrating that

digital authentic media can function as an effective supplement to conventional listening instruction.

At the same time, the study contributes new insights to the literature on authentic materials by focusing specifically on Arabic listening instruction. Much of the existing research on authentic media has been conducted in the context of English language teaching or other widely studied languages. Arabic language learning presents unique challenges due to its linguistic complexity and the variation between formal standard language and media discourse. Previous studies have emphasized the importance of integrating culturally relevant materials—such as proverbs, literary texts, and cultural narratives—into Arabic language instruction to support language learning (Tayyara, 2022). However, relatively few empirical investigations have examined the role of contemporary digital media in Arabic listening development. By demonstrating that exposure to authentic Arabic media can significantly improve listening comprehension and vocabulary recognition, the present study expands the scope of authentic materials research into the domain of Arabic language learning.

Another important dimension of the findings concerns learners' perceptions of authentic listening materials. Although participants initially perceived authentic media as more challenging than textbook listening passages, many reported that regular exposure helped them gradually adapt to natural speech patterns. This observation reflects a well-documented phenomenon in listening research, where authentic input may initially increase processing difficulty but ultimately supports learners' ability to develop more efficient listening strategies. Studies examining captioned video listening tasks have similarly found that authentic audiovisual materials may require greater cognitive effort at early stages but eventually facilitate deeper comprehension as learners become more familiar with natural speech patterns (Yeldham, 2023). The findings of the present study therefore reinforce the idea that sustained exposure to authentic listening input is essential for developing advanced listening skills.



**Figure 1. Digital authentic media exposure and listening development framework**

Figure 1 illustrates the relationship between digital authentic media exposure and listening development observed in this study. Authentic media provide

linguistically rich input characterized by natural discourse structures and frequent lexical patterns. Through repeated exposure, learners process authentic spoken language and develop listening comprehension strategies that facilitate vocabulary acquisition and discourse awareness. These processes ultimately contribute to improved listening performance and increased learner engagement with authentic language input.

The study provides evidence that digital authentic media offer valuable pedagogical benefits for Arabic listening instruction. By combining authentic discourse exposure with corpus-informed material selection, the study demonstrates how digital media can support both listening comprehension and vocabulary development while simultaneously enhancing learner engagement with authentic language use.

### **Advancing Theory in Arabic Language Pedagogy: Contributions to SLA and Corpus-Informed Learning**

Beyond its empirical findings, this study contributes to ongoing theoretical discussions in second language acquisition and Arabic language pedagogy. One of the key theoretical contributions of the study lies in its demonstration of the role of authentic input in facilitating language learning processes. The results show that learners exposed to authentic digital media experienced significantly greater improvement in listening comprehension and vocabulary recognition than those using traditional listening materials. These findings reinforce the theoretical assumption that meaningful exposure to authentic language input plays a crucial role in supporting language acquisition (Mehta et al., 2025). Authentic media provide learners with opportunities to encounter natural discourse structures, lexical patterns, and pragmatic cues that may not be fully represented in simplified instructional materials (Bohušová, 2025; Levi, 2024).

Another theoretical contribution of the study concerns the integration of corpus-informed approaches into language pedagogy. Corpus linguistics has increasingly been recognized as a valuable methodological framework for analyzing authentic language use and identifying patterns relevant to language teaching. The corpus analysis conducted in this study revealed that authentic Arabic media contain high lexical density and recurrent lexical items associated with contemporary public discourse. These findings demonstrate how corpus-based analysis can help educators identify linguistic features that characterize authentic spoken language and incorporate them into listening instruction. Previous research has similarly emphasized the importance of corpus analysis in identifying lexical patterns and phraseological units relevant to language pedagogy (Alfuhaid, 2025; Loesnopchaimongkhon et al., 2025). By integrating corpus-informed analysis with classroom experimentation, the present study provides a methodological framework that bridges corpus linguistics and language teaching.

The study also contributes to theoretical discussions regarding the role of digital technologies in language learning. With the rapid expansion of digital learning environments, language learners now have unprecedented access to authentic language input through online media platforms. Digital resources enable learners to engage with authentic discourse beyond the classroom and to encounter diverse language varieties in real-world contexts (Otamurodova & Dadaxonova, 2025; Tsang & Lam, 2025). Research on technology-enhanced

language learning has highlighted the potential of multimedia resources to increase learner engagement and facilitate exposure to authentic language use (Friginal & Ho, 2025b). Similarly, studies on mobile-assisted language learning have shown that digital technologies can create new opportunities for language exposure and informal learning experiences (Idhan et al., 2026). The findings of the present study support these perspectives by demonstrating that digital authentic media can function as a valuable instructional resource for developing listening skills in Arabic language learning.

Finally, the study contributes to the relatively limited body of empirical research on Arabic as a foreign language. Much of the existing literature on authentic materials has focused on widely taught languages such as English, while Arabic language pedagogy has received comparatively less attention in empirical studies of digital media-based language learning. By combining corpus analysis with classroom experimentation, the present study offers a methodological model that can be applied in future research on Arabic listening instruction. In doing so, it expands theoretical discussions on authentic materials and highlights the importance of integrating linguistically rich input into Arabic language pedagogy.

**Table 6. Theoretical contributions of digital authentic media in Arabic language learning**

Theoretical domain	Key concept	Role of digital authentic media	Evidence from this study
Second Language Acquisition	Authentic input	Provides exposure to natural discourse patterns and contextualized language use	Significant improvement in listening comprehension scores
Vocabulary acquisition	Incidental learning	Repeated encounters with lexical items in meaningful contexts	Higher vocabulary gains in the experimental group
Corpus-informed pedagogy	Lexical frequency and discourse patterns	Corpus analysis identifies authentic linguistic features useful for instruction	Media corpus revealed high lexical density and frequent discourse markers
Technology-enhanced learning	Digital media exposure	Multimedia resources increase engagement and listening exposure	Learners reported higher motivation and engagement
Arabic language pedagogy	Authentic listening materials	Supports adaptation to real spoken Arabic used in media discourse	Learners improved ability to process natural speech patterns

Table 6 summarizes the theoretical contributions of this study across several domains of language learning research. The findings demonstrate how digital authentic media can support second language acquisition processes by providing authentic linguistic input and repeated exposure to meaningful lexical patterns. In addition, the integration of corpus-informed analysis highlights the importance of examining linguistic characteristics of authentic materials when designing listening

instruction. These contributions collectively expand current theoretical discussions on Arabic language pedagogy and technology-enhanced language learning.

Despite the promising findings, several limitations and methodological considerations should be acknowledged to contextualize the results and guide future research. First, while this study demonstrates a substantial effect of digital authentic media exposure, the research design relies on a relatively small sample size and quasi-experimental grouping, which may limit the generalizability of the findings. Second, although the study integrates corpus-informed analysis, the corpus investigation primarily focuses on lexical frequency and density, leaving deeper discourse-level features (such as collocational patterns, pragmatic functions, and register variation) underexplored. Third, the qualitative component, while useful in capturing learner perceptions, remains limited in depth and could be further strengthened through more rigorous qualitative designs such as interviews or longitudinal observation. In addition, although statistical significance and effect size indicate strong instructional impact, future studies should employ more advanced analytical models (e.g., ANCOVA or mixed-effects modeling) to better control for pre-existing differences and individual variability. Addressing these limitations is essential not only for strengthening methodological rigor but also for refining the theoretical contribution and clarifying the specific mechanisms through which authentic digital media facilitate listening development in Arabic as a foreign language.

Taken together, these theoretical insights suggest that digital authentic media represent an important resource for bridging the gap between classroom instruction and real-world language use. By incorporating corpus-informed analysis and authentic listening exposure into Arabic language teaching, educators can create learning environments that more closely reflect the linguistic realities encountered by learners outside the classroom.

#### **D. Conclusion**

This study demonstrates that exposure to digital authentic media significantly enhances listening comprehension and vocabulary acquisition among learners of Arabic as a foreign language, particularly when such materials are systematically selected through a corpus-informed approach. The findings show that authentic Arabic media, characterized by high lexical density and recurrent discourse features, provide linguistically rich input that supports both comprehension and incidental vocabulary learning, leading to substantially greater gains compared to conventional textbook-based instruction. By integrating corpus analysis with quasi-experimental evidence, this study offers a more explicit link between the linguistic properties of authentic input and measurable learning outcomes. Pedagogically, the results suggest that incorporating corpus-informed authentic media into listening instruction can help bridge the gap between classroom learning and real-world communication while enhancing learner engagement. At the same time, the study highlights the need for careful material selection and instructional scaffolding to ensure accessibility. Future research should further investigate long-term effects, refine corpus-based analyses at the discourse level, and employ more advanced analytical models to strengthen the empirical and theoretical understanding of authentic media use in Arabic language learning.

## References

- Alexander, E., Van Hedger, S. C., & Batterink, L. J. (2023). Learning words without trying: Daily second language podcasts support word-form learning in adults. *Psychonomic Bulletin and Review*, 30(2), 751–762. <https://doi.org/10.3758/s13423-022-02190-1>
- Alfuhaid, M. R. (2025). Lexical Demands of Political News Reports: A Corpus-Based Lexical Profiling Inquiry. *Forum for Linguistic Studies*, 7(4), 1040–1054. <https://doi.org/10.30564/fls.v7i4.8944>
- Andajani, K., Pratiwi, Y., Suyitno, I., Prastio, B., & Maulidina, A. (2023). Challenges and Attitudes of L2 Students in Listening to Authentic Conversations Podcasts: A Case Study of Advanced Level BIPA Students. *Eurasian Journal of Applied Linguistics*, 9(3), 168–184. <https://doi.org/10.32601/ejal.903015>
- Blaiser, K. M., Nelson, P. B., & Kohnert, K. (2015). Effect of Repeated Exposures on Word Learning in Quiet and Noise. *Communication Disorders Quarterly*, 37(1), 25–35. <https://doi.org/10.1177/1525740114554483>
- Bohušová, Z. (2025). In Video Veritas: Short Videos as Authentic Pronunciation Exercises. *Linguistische Treffen in Wroclaw*, 2025(28), 247–260. <https://doi.org/10.23817/lingtreff.28-15>
- Brown, S. (2016). L2 Listening. In *Handbook of Research in Second Language Teaching and Learning: Volume III* (Vol. 3, pp. 201–213). <https://doi.org/10.4324/9781315716893-15>
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications. [https://books.google.co.id/books/about/Research\\_Design.html?hl=id&id=335ZDwAAQBAJ&redir\\_esc=y](https://books.google.co.id/books/about/Research_Design.html?hl=id&id=335ZDwAAQBAJ&redir_esc=y)
- Friginal, E., & Ho, W. Y. J. (2025a). Introduction to Audio and Video Based Materials in Language Education. In *Technology and Language Learning: Effective Podcasting and Video-Based Language Teaching Practices in the Age of Generative AI* (pp. 3–14). <https://doi.org/10.4324/9781003200727-2>
- Friginal, E., & Ho, W. Y. J. (2025b). *Technology and Language Learning: Effective Podcasting and Video-Based Language Teaching Practices in the Age of Generative AI*. <https://doi.org/10.4324/9781003200727>
- Friginal, E., Ho, W. Y. J., & Neupane-Bastola, M. (2025). Using Podcasts for Language Teaching and Learning (pp. 49–65). <https://doi.org/10.4324/9781003200727-6>
- Gorozhankina, T., & Grechukhina, V. (2022). Authentic video materials as a means of shaping communicative competence in Russian among Spanish and Dutch learners. *Russian Language Studies*, 20(4), 483–499. <https://doi.org/10.22363/2618-8163-2022-20-4-483-499>
- Hanak, N. A. (2023). Listening Comprehension in French as a Foreign Language at University Level: Difficulties and Learning Strategies. *Jordan Journal of Modern Languages and Literatures*, 15(2), 515–536. <https://doi.org/10.47012/jjml.15.2.7>
- Hidayati, Y., & Suharyadi, S. (2023). The Speech Rate and Vocabulary Profile of Ted-Ed Videos as Extensive Listening Materials for EFL Learners. *LLT Journal: Journal on*

- Language and Language Teaching, 26(2), 483–495. <https://doi.org/10.24071/llt.v26i2.6414>
- Idhan, M., Jabir, M., Ruslin, R., Asmawi, M. N., Arafah, B., Ubadah, U., Ubay, U., & Fatima, F. (2026). Teachers' Experiences With Mobile Apps for Arabic Language Learning: Insights From Islamic Schools in Palu, Indonesia. *Journal of Language Teaching and Research*, 17(1), 338–350. <https://doi.org/10.17507/jltr.1701.34>
- Koo, M., & Yang, S.-W. (2025). Likert-type scale. *Encyclopedia*, 5(1), 18. <https://doi.org/https://doi.org/10.3390/encyclopedia5010018>
- Levi, E. I. (2024). Technology in lifelong language learning. In *Embracing Technological Advancements for Lifelong Learning* (pp. 253–277). <https://doi.org/10.4018/979-8-3693-1410-4.ch012>
- Liu, S., & Saad, M. R. B. M. (2025). Role of Extensive Reading in Vocabulary Development, Reading Comprehension, and Reading Speed: A Systematic Literature Review. *Eurasian Journal of Applied Linguistics*, 11(1), 87–99. <https://doi.org/10.32601/ejal.11108>
- Lochmiller, C. R. (2021). Conducting Thematic Analysis with Qualitative Data. *The Qualitative Report*, 26(6), 2029–2044. <https://doi.org/https://doi.org/10.46743/2160-3715/2021.5008>
- Loesnopchaimongkhon, R., Phrommopakorn, C., Chernchom, P., & Laosrirattanachai, P. (2025). A Corpus-Based Analysis of Lexical Characteristics Across English News Categories for L2 Pedagogical Use. *Journal of Studies in the English Language*, 20(2), 36–65. <https://doi.org/10.64731/jssel.v20i2.279647>
- Mehta, V., Selvakumar, P., Anute, N. B., Gandhimathi, S., Seelam, P. K., & Manjunath, T. C. (2025). Second language acquisition support through media literacy and digital media platforms. In *Supporting Linguistic Differences Through Literacy Education* (pp. 361–388). <https://doi.org/10.4018/979-8-3373-3496-7.ch011>
- Nguyen, D. (2024). Using authentic materials and differentiated instruction to match EFL learners' motivation. In *Differentiated Instruction, Equity, and Inclusion in Language Education* (pp. 387–415). <https://doi.org/10.4018/979-8-3693-6497-0.ch013>
- Okoye, K., & Hosseini, S. (2024). T-test Statistics in R: Independent Samples, Paired Sample, and One Sample T-tests (K. Okoye & S. Hosseini (eds.); pp. 159–186). Springer Nature Singapore. [https://doi.org/10.1007/978-981-97-3385-9\\_8](https://doi.org/10.1007/978-981-97-3385-9_8)
- Otamurodova, N., & Dadaxonova, M. (2025). Second language acquisition supports through gaming and social media platforms. In *Supporting Linguistic Differences Through Literacy Education* (pp. 261–286). <https://doi.org/10.4018/979-8-3373-3496-7.ch008>
- Rosyad, M. S., Bahruddin, U., & Adam, F. M. (2023). The Reality of Learning Arabic Listening skills at Pesantren-Based Universities from the Perspective of Teaching Materials. *Lisanudhad: Jurnal Bahasa, Pembelajaran Dan Sastra Arab*, 10(2), 92–120. <https://repository.uin-malang.ac.id/18674/>
- Tayyara, A. E.-R. (2022). The practicability of proverbs in teaching Arabic language and



culture. *Language Teaching Research*, 26(4), 799–819.  
<https://doi.org/10.1177/1362168819895253>

Tsang, A., & Lam, W.-K. (2025). De-foreignizing English Outside Class: Junior-Secondary-Level English-as-a-Foreign-Language Learners' Extramural English Engagements and Listening and Reading Proficiency. *RELC Journal*, 56(2), 377–396.  
<https://doi.org/10.1177/00336882241229459>

Vaahantoranta, E., Suggate, S., Lenhart, J., & Lenhard, W. (2021). Language exposure and phonological short-term memory as predictors of majority language vocabulary and phonological awareness in dual language learning. *Bilingualism*, 24(2), 319–332. <https://doi.org/10.1017/S1366728920000541>

Vogrig, J.-N. (2021). Authentic versus purpose-built material for foreign language teaching and learning. *Recherche et Pratiques Pédagogiques En Langues de Spécialité - Cahiers de l'APLIUT*, 40(1). <https://doi.org/10.4000/apliut.8741>

Wagner, E., Batty, A. O., & Galaczi, E. (2024). The Routledge Handbook of Second Language Acquisition and Listening. In *The Routledge Handbook of Second Language Acquisition and Listening*. <https://doi.org/10.4324/9781003219552>

Yeldham, M. (2023). How do second language learners go about their listening when they view captioned videos? Replication studies of Taylor (2005), Winke et al. (2013) and Rodgers and Webb (2017). *Language Teaching*.  
<https://doi.org/10.1017/S0261444823000228>