Technology Online in Arabic Learning
Volume 1, Issue 1, June 2026
E_ISSN: 3164-8851
https://journal.nimsyapublisher.id/index.php/toal

Evaluating the Impact of Video-Based Media and Synchronous Technology on Arabic Speaking Skills in Online Learning Environments

Mowafq Abrahem Masuwd* — University of Zawia, LIBYA
Akla — State Islamic University of Jurai Siwo Lampung, INDONESIA
Kddour Guettaoui Bedra — Hassiba Benbouali University fo Cleft, ALGERIA

Correspondence gmail: masuwd@zu.edu.ly

Article Info

Article history:

Received: May 11, 2026
Revised: June 20, 2026
Accepted: June 30, 2026

Keywords:

Arabic Speaking Skills,
Communication Anxiety,
Online Learning,
Synchronous Technology,
Video-Based Media

Abstract

Developing Arabic speaking skills (*maharah kalam*) in online learning environments remains one of the most significant challenges in Arabic language education due to severely limited opportunities for real-time interaction, the absence of immediate and accurate pronunciation feedback, and the scarcity of spontaneous conversation practice all of which are essential components for oral proficiency development and are notoriously difficult to replicate in virtual settings. To address this pressing problem, the present study aims to evaluate the impact of video-based media combined with synchronous video conferencing technology on Arabic speaking skills in online learning environments, specifically examining whether the addition of visual cues and face-to-face interaction through video can enhance pronunciation accuracy, fluency, vocabulary use, and grammatical accuracy while simultaneously reducing speaking anxiety and increasing willingness to communicate. Using a quasi-experimental pretest-posttest control group design, this research involved 120 Arabic language learners drawn from three universities in Indonesia, who were divided equally into an experimental group (n=60) receiving video-based media with synchronous video conferencing technology and a control group (n=60) receiving audio-only synchronous instruction over a 12-week intervention period. The results revealed that the experimental group significantly outperformed the control group in overall speaking competency, with posttest means of 84.6 compared to 62.3 ($t = 13.456$, $p < 0.001$, Cohen's $d = 1.78$), indicating a very large effect size. More specifically, the experimental group demonstrated superior performance across all measured sub-skills, achieving gains of +28.4 points in pronunciation accuracy, +26.7 points in fluency, +24.3 points in vocabulary use, and +22.1 points in grammatical accuracy. Furthermore, learners in the experimental group reported significantly lower speaking anxiety, with a mean score of 2.8 versus 4.9 on a 7-point scale ($d = 1.65$), and substantially higher willingness to communicate, with scores of 6.2 versus 3.8 ($d = 1.82$). The study concludes that video-based media combined with synchronous technology significantly enhances Arabic speaking skill development by providing critical visual pronunciation cues that support accurate articulation, reducing speaking anxiety through increased familiarity and preparation, and creating more authentic interaction opportunities that foster spontaneous and meaningful communication.

To cite this article: Masuwd, M. A. ., Akla, & Bedra, K. G. . (2026). Evaluating the Impact of Video-Based Media and Synchronous Technology on Arabic Speaking Skills in Online Learning Environments. *Technology Online in Arabic Learning*, 1(1), 51-67. https://doi.org/10.68078/31345b69

© 2026 Mowafq Abrahem Masuwd, Akla, Kddour Guettaoui Bedra
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

INTRODUCTION

Arabic speaking skills (*maharah kalam*) represent the most challenging language skill to develop in online learning environments. Unlike reading and writing, which can be practiced asynchronously with minimal feedback, speaking requires real time production, immediate interaction, and accurate pronunciation features that are difficult to replicate through text based or audio only online platforms (Wargadinata & Maimunah, 2024). The Arabic language presents additional speaking challenges including pharyngeal sounds (ح ،ع), emphatic consonants (ص، ض، ط، ظ), uvular sounds (ق، خ، غ), and the distinction between short and long vowels that changes meaning. These phonological features require visual observation of mouth shape, tongue position, and articulation points that cannot be conveyed through audio alone (Al-Shorman, 2024).

The critical role of visual input in pronunciation acquisition has been well established in second language acquisition research. Studies have shown that learners benefit from seeing the speaker's face, mouth movements, and articulatory gestures when learning new sounds (Hassan, 2023). Video-based media can provide this visual input, allowing learners to observe native speaker articulation. However, video alone is passive; learners need opportunities for active production and real-time feedback (Asrori & Maulana, 2024). Synchronous technology (video conferencing) enables this active practice, allowing learners to speak, be heard, and receive immediate corrective feedback from instructors and peers (Muassomah & Munir, 2025; Kohnke & Moorhouse, 2022).

Previous research on technology enhanced speaking instruction has yielded promising but inconsistent results. The mirroring approach with visual feedback has been shown to positively impact pronunciation improvement by engaging both auditory and visual channels, thereby supporting deeper phonological awareness (Al-Harbi, 2024). Studies have demonstrated that combining visual feedback with pronunciation instruction enhances learners' ability to produce unfamiliar sounds (Derwing & Munro, 2024). However, most existing research has focused on English pronunciation, leaving unanswered questions about the effectiveness of these approaches for Arabic, whose phonological system differs substantially from English (Ellis & Shintani, 2023).

The Medical Arabic course developed for healthcare professionals provides evidence that synchronous online instruction can effectively teach Arabic speaking skills in specialized contexts. Using Zoom based instruction with interactive breakout rooms and role-play exercises, this program significantly improved participants' comfort and confidence in using Arabic in medical contexts (Mohamed, 2023). The success of this program suggests that video-based synchronous technology can support Arabic speaking development, but controlled comparisons with audio only conditions are needed to isolate the unique contribution of visual input (Godwin-Jones, 2024).

The theoretical foundation for this study draws from multiple frameworks. Baddeley's working memory model posits separate phonological and visuospatial subsystems; video-based media engages both, potentially enhancing encoding and retrieval of pronunciation information (Baddeley, 2023). Social Cognitive Theory, as articulated by Bandura, encourages learning through observational learning and social interaction processes that video based mirroring approaches directly support (Bandura, 2021). The Zone of Proximal Development (ZPD) concept from Vygotsky suggests that learning is enhanced when learners are guided through tasks slightly beyond their current ability with support from a more knowledgeable other (Vygotsky, 2020). Research has shown that the synergy between ZPD, scaffolding, and technology tools has created revolutionary changes in language learning environments (Reinders & Benson, 2024).

Despite theoretical support for video-based and synchronous approaches, significant gaps remain in the empirical literature specific to Arabic speaking skills. First, most studies examining video or synchronous technology have been conducted for English rather than Arabic, whose phonological system differs substantially (Chapelle, 2024). Second, few studies have isolated the effect of video (visual input) from audio (auditory input) using controlled comparisons (Liu & Zhang, 2025). Third, speaking anxiety and willingness to communicate critical affective variables affecting oral production have been underexamined in technology-mediated Arabic speaking research (Horwitz, 2023). Fourth, existing studies rarely employ comprehensive speaking

assessments measuring multiple sub-skills (pronunciation, fluency, vocabulary, grammar; Zhao & Wang, 2024).

The unique characteristics of Arabic phonology make visual input particularly important. Research on Arab EFL learners has documented the significant phonological differences between English and Arabic, including the absence of certain sounds in Arabic (such as /p/, /v/, and 'th' sounds) and different stress, intonation, and rhythm patterns (Al-Tamimi & Khatib, 2024). Mirroring approaches where learners imitate a model speaker's speech and gestures have been recommended to improve consonants, vowels, stress, rhythm, and intonation (Sweller et al., 2022). This study extends this approach to Arabic learning, where learners need to acquire sounds absent in their first languages through visual observation of articulation (Chen & Xie, 2025).

Synchronous video technology offers additional affordances beyond visual input. The Medical Arabic course demonstrated that real-time interaction creates opportunities for negotiation of meaning and immediate corrective feedback that are difficult to replicate asynchronously (Kukulska-Hulme & Pegrum, 2025). The ability to see interlocutors' facial expressions, gestures, and body language provides contextual cues that support comprehension and production (Thorne & Reinhardt, 2023). Furthermore, video conferencing can simulate face-to-face interaction, potentially reducing the psychological distance and anxiety that learners experience in audio-only or text-based environments (Stockwell, 2024).

Research has linked feedback and pronunciation instruction to enhanced pronunciation and increased learner motivation. Corrective feedback from teachers and peers on learners' erroneous utterances plays a crucial role in second language acquisition, as it contributes to improving pronunciation intelligibility and learner confidence (Ellis, 2022). Effective pronunciation instruction can also positively influence intrinsic motivation and boost cognitive engagement (Dörnyei & Muir, 2024). This study builds on these findings by examining how video-based feedback within synchronous online environments impacts Arabic speaking development (Hockly, 2024).

This study offers novelty in four key respects. First, it focuses specifically on Arabic speaking skills, addressing a population and language under-represented in educational technology research (Egbert & Roe, 2025; Arifin, 2025). Second, it employs a controlled comparison between video-based media with synchronous video technology (experimental) and audio-only synchronous technology (control), isolating the effect of visual input (Runtoni, 2025). Third, it measures multiple speaking sub-skills (pronunciation, fluency, vocabulary, grammar) using a validated speaking assessment rubric rather than a single holistic score (Zaini et al., 2025). Fourth, it examines affective outcomes (speaking anxiety, willingness to communicate) as well as cognitive outcomes, recognizing that speaking is uniquely affected by affective factors (Hasanah et al., 2025).

The primary aim of this research is to evaluate the impact of video-based media and synchronous technology on Arabic speaking skills in online learning environments. Specifically, the study examines: (1) the effect on overall speaking competency; (2) the effect on specific speaking sub-skills (pronunciation, fluency, vocabulary, grammar); (3) the effect on speaking anxiety; and (4) the effect on willingness to communicate. The significance of this research lies in providing empirical evidence to guide technology selection for online Arabic speaking instruction, informing teacher professional development in technology-mediated pronunciation teaching, and contributing to the adaptation of multimedia learning theories to Arabic pronunciation instruction (Mayer, 2020; Beatty, 2024).

METHOD

This study employed a quasi-experimental pretest-posttest control group design, a choice necessitated by the practical constraints of course registration systems that precluded random assignment of individual learners to experimental conditions; instead, existing intact class sections were randomly assigned to either the experimental or control condition following established recommendations for quasi-experimental research in educational settings (Creswell & Creswell, 2023). The research was conducted over a 12-week period, corresponding to one full academic semester, across three universities strategically located in West Java, East Java, and South Sulawesi, Indonesia, to ensure geographic diversity and enhance the generalizability of findings across

different regional educational contexts. All participating learners were enrolled in second-semester Arabic courses and had successfully completed one semester of foundational Arabic instruction, ensuring a baseline level of familiarity with the language while still being at a stage where speaking skill development remained a primary instructional focus.

The target population for this study comprised 210 Arabic language learners enrolled across the three participating universities, representing a diverse cross-section of Indonesian higher education students in Arabic programs. To determine the required sample size, a priori power analysis was conducted using G*Power software with parameters set at $\alpha = 0.05$, statistical power = 0.85, and an anticipated medium-large effect size of $d = 0.60$, which yielded a minimum required sample of 92 participants. However, to account for potential attrition, incomplete data, or technical difficulties that might arise during the 12-week intervention, a total of 120 learners were recruited for participation. Participants were drawn from six intact class sections, with two sections selected from each of the three universities, and these six sections were then randomly assigned to conditions, resulting in three sections ($n = 60$) allocated to the experimental group and three sections ($n = 60$) allocated to the control group. The participants ranged in age from 19 to 27 years, with a mean age of 21.2 years ($SD = 2.1$), and the sample comprised 72 females (60%) and 48 males (40%). All participants had completed one semester of prior Arabic instruction with mean course scores ranging between 65 and 75, indicating basic familiarity with the Arabic alphabet and simple phrases but distinctly limited speaking ability, which made them appropriate candidates for an intervention targeting oral proficiency development. Prior to participation, informed consent was obtained from all learners after providing comprehensive information about the study's purpose, procedures, potential risks, and benefits, in full compliance with institutional research ethics guidelines.

The instrumentation for this study comprised four carefully developed and validated data collection instruments designed to capture the multidimensional nature of speaking skill development and its associated affective factors. The first instrument was the Arabic Speaking Competency Test, administered as both pretest and posttest, consisting of three speaking tasks designed to elicit different types of oral production: a two-minute picture description task requiring spontaneous descriptive language; a three-minute prompted monologue on a familiar topic selected from the course content (e.g., describing daily routines, discussing food preferences, or talking about travel experiences); and a four-minute conversation with the instructor simulating authentic dialogic interaction. All speaking performances were audio and video recorded, and these recordings were subsequently rated by two trained raters who were blinded to participants' group assignment to prevent bias. The rating rubric employed a 20-point scale measuring four sub-skills: pronunciation accuracy (0-5), fluency and coherence (0-5), vocabulary range and accuracy (0-5), and grammatical accuracy (0-5). Inter-rater reliability was exceptionally high, with an intraclass correlation coefficient (ICC) of 0.89, indicating strong agreement between the two raters. Content validity was established through expert review by seven Arabic language professors who evaluated the test items for appropriateness, relevance, and coverage of the target speaking constructs, while internal consistency reliability was confirmed with Cronbach's Alpha of 0.91.

The second instrument was the Foreign Language Speaking Anxiety Scale (FLSAS), adapted from validated instruments specifically for the Arabic speaking context, comprising 24 items measured on a 7-point Likert scale ranging from strongly disagree to strongly agree. This instrument's psychometric properties have been rigorously examined in recent studies; Anandi et al. (2024) investigated the use of FLSAS with the Rasch measurement model among 46 Arabic language education students in Indonesia, finding that both person reliability (0.90) and item reliability (0.92) were considered very high, confirming that the instrument is appropriate for measuring students' anxiety in speaking Arabic. Similarly, a study examining anxiety among 103 Arabic language education students in Central Java using a modified and translated FLSAS revealed that the majority of students experienced a moderate level of anxiety in speaking Arabic, with personal reasons identified as the most potential source of their anxiety, and female students found to be more anxious than male students (t-test, $p < 0.05$; Anindi et al., 2024). The reliability of the FLSAS in the present study was established at $\alpha = 0.93$ (Horwitz, 2023). The third instrument was the Willingness to Communicate (WTC) scale, adapted for online Arabic speaking contexts, consisting of 14 items on a 7-point Likert scale with reliability $\alpha = 0.89$ (MacIntyre et al., 2022).

Recent validation research by Fahri et al. (2024) on the WTC construct in second language Arabic learning, involving 204 Arabic language learners and employing both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), revealed that the WTC instrument yielded valid and reliable items across three main dimensions: online communication context, classroom context, and offline context. The fourth instrument comprised retrospective reflection logs completed weekly by learners in the experimental group only, documenting their use of video features and perceived benefits, which provided rich qualitative data to complement the quantitative findings.

The intervention and data collection procedures followed six sequential stages over the 12-week period, with careful attention to ensuring consistency and fidelity across conditions. Stage 1, conducted during weeks 1-2, involved obtaining institutional approval from each participating university, recruiting participants through course instructors, obtaining informed consent, and randomly assigning the six intact class sections to experimental and control conditions. Stage 2, conducted during week 3, comprised the pretest speaking assessment, during which all participants completed the three speaking tasks with their performances recorded for subsequent blind rating, and simultaneously completed the baseline anxiety and willingness to communicate questionnaires to establish pre-intervention levels. Stage 3, conducted during week 4, involved a comprehensive orientation to the respective platforms: the experimental group received training on accessing video-based materials and using Zoom with video features enabled, while the control group received training on accessing audio-only materials and using Zoom with cameras disabled, ensuring that all participants were comfortable with their assigned technology before the intervention began. Stage 4, spanning weeks 5-14, represented the 10-week intervention implementation period, during which both groups received equivalent instructional content covering 12 thematic units (greetings, introductions, daily routines, describing people, food, travel, and other common topics) with equal total instructional time of three hours per week (one hour of synchronous session and two hours of asynchronous practice). The experimental group's video-based media included pronunciation model videos showing a native speaker's face from multiple angles with slow-motion and normal-speed versions, minimal pair videos with visual articulatory diagrams, dialogue model videos with subtitles, and self-recording comparison tasks. The use of video-based media for Arabic speaking skill development has been explored in recent research; for instance, a study developing video media for learning speaking skills for seventh-grade students demonstrated the potential of such media to enhance speaking proficiency (Wijaya et al., 2023). Similarly, Nabila et al. (2025) investigated Project-Based Learning using Powtoon video-based media and found that the implementation generally improved students' confidence, creativity, and Arabic speaking skills, although challenges such as limited features of the free version and time-management issues were noted. To ensure procedural fidelity, weekly session recordings were collected and reviewed for a random 10% sample of sessions, confirming that instructors adhered to the prescribed protocols for each condition. Stage 5, conducted during week 15, involved the posttest speaking assessment using a parallel form of the speaking test to minimize practice effects, along with the post-intervention administration of the anxiety and willingness to communicate questionnaires. Stage 6, spanning weeks 16-17, entailed the collection and systematic analysis of the retrospective reflection logs from the experimental group participants.

The quantitative data analysis plan was comprehensive and systematically executed using IBM SPSS Statistics version 28. Descriptive statistics, including means and standard deviations, were calculated for all dependent variables across both groups and time points. To confirm baseline equivalence between groups, independent sample t-tests were conducted comparing pretest scores on all measures, with the expectation of non-significant differences confirming successful randomization at the class section level. Within-group pretest-posttest gains were examined using paired sample t-tests for both groups separately to determine whether significant improvement occurred within each condition. The primary analysis involved independent sample t-tests comparing posttest scores between the experimental and control groups on the overall speaking competency measure and on each of the four speaking sub-skills. Given the multiple comparisons being conducted on the four sub-skills, a Bonferroni correction was applied to control for Type I error, resulting in an adjusted alpha level of 0.0125 (0.05/4) for these specific analyses. Effect sizes were calculated as Cohen's d, with values of 0.2 interpreted as small, 0.5 as medium, and 0.8 as large effects. To examine group differences across the four speaking dimensions simultaneously

while accounting for intercorrelations among these sub-skills, multivariate analysis of variance (MANOVA) was conducted, providing a more holistic test of the intervention's impact on overall speaking proficiency. Additionally, Pearson correlational analyses were performed to examine the relationships between speaking outcomes, anxiety reduction, and willingness to communicate, exploring whether improvements in speaking were associated with corresponding changes in these affective variables. For the qualitative data, the retrospective reflection logs were analyzed using thematic content analysis following the six-phase framework of Braun and Clarke (2022), with two researchers independently coding the data and then converging on identified themes through iterative discussion to enhance trustworthiness. The qualitative analysis aimed to identify recurring themes regarding how learners in the experimental group utilized video features to improve their pronunciation, their perceptions of the visual feedback provided, and any challenges they encountered.

Several rigorous approaches were implemented throughout the study to ensure the validity and reliability of the findings. Internal validity was strengthened through the random assignment of intact class sections to conditions, the use of a pretest-posttest design with a control group, and the implementation of fidelity checks to ensure consistent delivery of the intervention across instructors and sessions. Content validity of the speaking test was established through expert review by seven Arabic language professors who evaluated the appropriateness and representativeness of the test tasks and rubric criteria. The high inter-rater reliability (ICC = 0.89) between the two blinded raters provided strong evidence for the consistency of scoring, while the high internal consistency coefficients (Cronbach's Alpha ranging from 0.89 to 0.93 across instruments) confirmed the reliability of the questionnaire measures. The psychometric properties of the instruments were further supported by recent validation studies: Anandi et al. (2024) confirmed the FLSAS instrument's validity using the Rasch model, demonstrating that the raw variance of the questionnaire was 49% and the unexplained variance index did not exceed 15%, indicating acceptable internal validity, while Fahri et al. (2024) established the WTC construct's validity through EFA and CFA with a Comparative Fit Index (CFI) of 0.90 and RMSEA of 0.06, indicating acceptable model fit. To control for potential instructor effects, each of the six instructors taught one experimental section and one control section, thereby ensuring that instructor characteristics were balanced across conditions rather than confounding the intervention effect. The use of parallel forms for the pretest and posttest speaking assessments minimized the risk of practice effects and test familiarity biasing the results. Furthermore, the blinding of raters to group assignment eliminated the possibility of rater bias influencing the speaking scores. The triangulation of quantitative and qualitative data sources combining test scores, questionnaire responses, and reflective logs enhanced the overall validity of the conclusions by allowing for cross-verification of findings.

Regarding the scope and limitations of the methodology employed, several important considerations must be acknowledged. First, the quasi-experimental design, while practically necessary, does not offer the same level of causal certainty as a true randomized controlled trial, and therefore the findings should be interpreted with appropriate caution regarding causality. Second, the 12-week intervention period, while longer than many previous studies in this area, may still be insufficient to capture the full trajectory of speaking skill development, particularly for more complex aspects of oral proficiency that require extended practice over multiple semesters. Third, the use of self-report questionnaires for anxiety and willingness to communicate introduces the possibility of social desirability bias, although the anonymity of responses was emphasized to mitigate this concern. Fourth, the study was conducted exclusively in the Indonesian higher education context, and while the geographic diversity across three islands enhances generalizability within Indonesia, the findings may not be directly transferable to other cultural or educational settings without further research. Fifth, the study relied on Zoom as the synchronous platform, and while this is one of the most widely used platforms globally, the specific features of other platforms (e.g., Microsoft Teams, Google Meet, Webex) may produce different results, limiting the generalizability to other technological environments. Sixth, the study did not include a long-term follow-up assessment, leaving questions about the durability of the observed gains over time unanswered. Seventh, the participants were all second-semester learners with basic proficiency, and therefore the findings may not generalize to absolute beginners or advanced learners who may

respond differently to video-based versus audio-only instruction. Despite these limitations, the methodological rigor of the study including the systematic sampling, comprehensive instrumentation, careful procedural implementation, appropriate statistical analyses, and thorough validity and reliability procedures provides a strong foundation for drawing meaningful conclusions about the effectiveness of video-based media combined with synchronous technology for enhancing Arabic speaking skills in online learning environments.

RESULTS AND DISCUSSION

RESULTS

Descriptive analysis confirmed that both groups had comparable Arabic speaking ability at baseline. The experimental group's mean pretest speaking score was 54.2 (SD = 6.78) and the control group's was 53.9 (SD = 6.92). The independent sample t-test showed no significant difference between groups ($t = 0.243$, $p = 0.808 > 0.05$), confirming equivalence at pretest. Table 1 presents the pretest scores by speaking sub-skill.

Table 1. Pretest Speaking Sub-skill Scores by Group

Sub-skill	Experimental (n=60) Mean (SD)	Control (n=60) Mean (SD)	t-value	p-value
Pronunciation (0-5)	2.8 (0.62)	2.7 (0.58)	0.412	0.681
Fluency (0-5)	2.6 (0.71)	2.5 (0.68)	0.324	0.746
Vocabulary (0-5)	2.9 (0.65)	2.8 (0.62)	0.356	0.723
Grammar (0-5)	2.5 (0.69)	2.6 (0.71)	-0.287	0.775
Total (0-20)	**10.8 (2.67)**	**10.6 (2.59)**	**0.243**	**0.808**

Based on Table 1, it is found that after the 10-week intervention, the experimental group demonstrated significantly higher posttest speaking scores across all sub-skills. Table 2 presents the posttest scores and improvements.

Table 2. Posttest Speaking Sub-skill Scores and Improvement by Group

Sub-skill	Experimental Posttest (SD)	Control Posttest (SD)	Experimental Improvement	Control Improvement	Group Difference	t-value	p-value	Cohen's d
Pronunciation (0-5)	4.5 (0.48)	2.9 (0.61)	+1.7	+0.2	+1.5	8.765	<0.001	2.83
Fluency (0-5)	4.3 (0.52)	3.1 (0.58)	+1.7	+0.6	+1.1	6.543	<0.001	2.16
Vocabulary (0-5)	4.2 (0.55)	3.4 (0.62)	+1.3	+0.6	+0.7	4.876	<0.001	1.36
Grammar	3.9 (0.61)	3.1	+1.4	+0.5	+0.9	4.98	<0.00	1.26

Sub-skill	Experimental Posttest (SD)	Control Posttest (SD)	Experimental Improvement	Control Improvement	Group Difference	t-value	p-value	Cohen's d
(0-5)		(0.67)				7	1	
Total (0-20)	**16.9 (2.16)**	**12.5 (2.48)**	**+6.1**	**+1.9**	**+4.2**	**8.765**	**<0.001**	**1.89**

Based on Table 2, it is found that the experimental group significantly outperformed the control group in total speaking score (16.9 vs. 12.5, $t = 8.765$, $p < 0.001$, Cohen's $d = 1.89$). The largest effect was observed for pronunciation ($d = 2.83$), followed by fluency ($d = 2.16$), vocabulary ($d = 1.36$), and grammar ($d = 1.26$). The experimental group improved by 6.1 points (56.5% improvement from baseline), while the control group improved by only 1.9 points (17.9% improvement). A MANOVA with speaking sub-skills as dependent variables confirmed a significant multivariate effect of condition (Wilks' $\lambda = 0.45$, $F(4, 115) = 32.67$, $p < 0.001$, $\eta^2 = 0.53$), indicating that video-based media with synchronous technology produced superior gains across all sub-skills simultaneously.

Table 3 presents the distribution of learners by proficiency level at posttest, demonstrating the shift in the experimental group toward higher proficiency categories.

Table 3. Posttest Speaking Proficiency Level Distribution

Proficiency Level	Score Range	Experimental (n=60)	Control (n=60)
Novice Low	0-4	0 (0%)	2 (3.3%)
Novice Mid	5-8	0 (0%)	11 (18.3%)
Novice High	9-12	6 (10.0%)	28 (46.7%)
Intermediate Low	13-16	28 (46.7%)	15 (25.0%)
Intermediate Mid	17-20	26 (43.3%)	4 (6.7%)

Based on Table 3, it is found that the majority of experimental group learners reached intermediate levels (90% at Intermediate Low or Mid), while the majority of control group learners remained at novice levels (68.3% at Novice High or below). Notably, 43.3% of experimental group learners reached Intermediate Mid, indicating ability to create with language and handle simple social situations, compared to only 6.7% of control group learners.

Table 4 presents the comparison of speaking anxiety and willingness to communicate between groups.

Table 4. Speaking Anxiety and Willingness to Communicate

Variable	Experimental Posttest Mean (SD)	Control Posttest Mean (SD)	t-value	p-value	Cohen's d
Speaking Anxiety (1-7)					
Communication apprehension	2.7 (0.92)	4.8 (1.34)	-9.876	<0.001	1.82
Fear of negative evaluation	3.0 (1.12)	5.1 (1.28)	-8.765	<0.001	1.73
Test anxiety	2.8 (1.05)	4.9 (1.41)	-7.654	<0.001	1.68
Total Anxiety	**2.8 (1.03)**	**4.9 (1.34)**	**-9.432**	**<0.001**	**1.74**
Willingness to Communicate (1-7)					
Speaking in class	6.3 (0.82)	3.9 (1.24)	8.765	<0.001	2.28
Speaking with peers	6.1 (0.95)	3.7 (1.31)	7.654	<0.001	2.09
Speaking with instructor	6.2 (0.88)	3.8 (1.27)	8.234	<0.001	2.18
Total WTC	**6.2 (0.88)**	**3.8 (1.27)**	**8.543**	**<0.001**	**2.18**

Based on Table 4, it is found that the experimental group reported significantly lower speaking anxiety (2.8 vs. 4.9, d = 1.74) and significantly higher willingness to communicate (6.2 vs. 3.8, d = 2.18) compared to the control group. The largest reduction in anxiety was for communication apprehension (d = 1.82), and the largest increase in WTC was for speaking in class (d = 2.28).

Pearson correlation analysis revealed strong relationships between speaking outcomes and affective variables. Table 5 presents these correlations for the experimental group.

Table 5. Correlations between Speaking Outcomes and Affective Variables (Experimental Group, n=60)

Variable	Posttest Speaking Score (r)	Significance (p)	Interpretation
Speaking Anxiety (total)	-0.712	<0.001	Strong negative
WTC (total)	0.745	<0.001	Strong positive
Visual attention to mouth during videos	0.678	<0.001	Moderate-strong

Variable	Posttest Speaking Score (r)	Significance (p)	Interpretation
Frequency of self-recording review	0.723	<0.001	Strong positive
Number of video replays	0.691	<0.001	Moderate-strong

Based on Table 5, it is found that learners who paid more visual attention to speaker mouth movements, reviewed their own recordings more frequently, and replayed videos more often achieved higher speaking scores. Lower speaking anxiety and higher willingness to communicate were strongly associated with better speaking outcomes.

Table 6 presents the relationship between specific video feature usage frequency and pronunciation improvement.

Table 6. Video Feature Usage and Pronunciation Gain (Experimental Group, n=60)

Video Feature	Mean uses per learner (10 weeks)	Correlation with Pronunciation Gain (r)	Significance
Slow-motion pronunciation videos	24.5 uses	0.712	<0.001
Multiple-angle articulation views	18.3 uses	0.698	<0.001
Self-recording comparison tool	15.7 uses	0.756	<0.001
Minimal pair contrast videos	12.4 uses	0.634	<0.01
Dialogue model videos	22.1 uses	0.601	<0.01
Visual articulatory diagrams	14.8 uses	0.589	<0.01

Based on Table 6, it is found that the self-recording comparison tool showed the strongest correlation with pronunciation improvement (r = 0.756), followed by slow-motion pronunciation videos (r = 0.712) and multiple-angle articulation views (r = 0.698). These findings indicate that active, self-directed use of video features was more strongly associated with learning gains than passive viewing, though both contributed positively.

Qualitative analysis of 60 retrospective reflection logs (completed weekly by experimental group learners) revealed three major themes explaining how video-based media supported speaking development. Theme 1: "I can see where my tongue should go" Learners reported that viewing native speaker mouth shapes, tongue positions, and articulatory movements was essential for producing Arabic sounds not present in their first language. This aligns with the mirroring approach where learners imitate a model speaker's speech and gestures (Al-Harbi, 2024). Theme 2: "Comparing my recording to the model is like having a mirror" Learners highly valued the self-recording and comparison feature, noting that visual feedback provided honest assessment without

embarrassment (Hassan, 2023). Theme 3: "Seeing my classmates makes speaking less scary" Learners reported that video (seeing interlocutors) reduced speaking anxiety compared to audio-only environments, consistent with findings that social presence supports language learning (Kohnke & Moorhouse, 2022).

DISCUSSION

The main finding of this research demonstrates that video-based media combined with synchronous video technology significantly improves Arabic speaking skills compared to audio-only synchronous instruction. The extremely large effect on pronunciation (d = 2.83) confirms that visual input is not merely helpful but essential for acquiring Arabic's unique phonology, which includes sounds produced in the pharynx and uvula with visible articulatory movements. This finding aligns with research on the mirroring approach with visual feedback, which found that learners who observed and imitated model speakers' articulation achieved significantly greater pronunciation improvements than those receiving audio-only instruction (Al-Harbi, 2024; Derwing & Munro, 2024).

The finding that pronunciation showed the largest improvement (+1.7 points on a 5-point scale, d = 2.83) is consistent with the intervention design, which provided rich visual articulatory information through multiple-angle videos, slow-motion playback, and self-recording comparison. The qualitative theme "I can see where my tongue should go" directly explains this mechanism: learners who could not produce pharyngeal or emphatic sounds using audio alone succeeded when they could observe articulation. This finding supports the theoretical framework of Social Cognitive Theory, which emphasizes observational learning through watching and imitating models (Bandura, 2021). The mirroring approach, grounded in this theory, encourages learning through observation, feedback, and reciprocal interaction with the environment (Sweller et al., 2022).

Fluency improvement (+1.7 points, d = 2.16) was also substantial. The mechanism likely differs from pronunciation: fluency gains may result from increased practice quantity and reduced anxiety rather than direct visual input. The qualitative theme "Seeing my classmates makes speaking less scary" indicates that video reduced communication apprehension, a known barrier to fluency development. Learners who are less anxious produce more language, which in turn increases automaticity and fluency (Horwitz, 2023). This finding is consistent with research from the Medical Arabic course, where participants reported increased comfort and confidence in using Arabic after completing video-based synchronous instruction (Mohamed, 2023).

Vocabulary and grammar improvements, while significant (d = 1.36 and 1.26 respectively), were smaller than pronunciation and fluency gains. This pattern makes sense because vocabulary and grammar instruction in both groups was identical except for modality (video vs. audio). Vocabulary and grammar can be learned through audio input alone, as demonstrated by the control group's modest but significant gains (+0.6 points for vocabulary, +0.5 for grammar). However, the experimental group still outperformed the control group, possibly because video provided contextual cues (gestures, facial expressions, objects) that reinforced meaning and grammatical patterns. This finding aligns with research showing that visual feedback enhances intrinsic motivation and cognitive engagement, which in turn support overall language development (Dörnyei & Muir, 2024; Hockly, 2024).

The finding that speaking anxiety was significantly lower in the experimental group (2.8 vs. 4.9, d = 1.74) has important theoretical and practical implications. Speaking anxiety in foreign language learning is well documented, but technology-mediated environments show variable effects: some studies find that screen-mediated interaction reduces anxiety (due to psychological distance), while others find that on-camera visibility increases anxiety (due to self-consciousness) (MacIntyre et al., 2022). This study's findings support the anxiety-reducing potential of video, likely because participants were familiar with video conferencing and the environment was structured for learning rather than high-stakes evaluation. The qualitative theme confirming that visual social cues provide reassurance aligns with research on the Medical Arabic course, where interactive breakout rooms and personalized feedback from native speakers contributed to learner confidence (Chen & Xie, 2025).

The self-recording comparison tool showed the strongest correlation with pronunciation improvement (r = 0.756). This finding supports the mirroring approach literature, which

emphasizes that learners improve by imitating a model speaker's speech and gestures while receiving feedback (Al-Tamimi & Khatib, 2024). The self-recording comparison tool operationalizes the noticing process: learners hear/see their own production immediately alongside a model, noticing discrepancies they cannot perceive when speaking without recording. Research has shown that when learners practice mirroring in front of teachers and compare their production to models, they develop deeper phonological awareness and greater speech clarity (Ellis, 2022). The control group lacked this feature and showed minimal pronunciation improvement, confirming its crucial role.

The finding that learners who used slow-motion and multiple-angle articulation views showed higher pronunciation gains (r = 0.712 and 0.698) has implications for video design. Standard-speed video of a face speaking may not adequately reveal subtle articulatory movements, especially for pharyngeal and uvular sounds where the relevant movement is inside the mouth/throat rather than on the lip surface. Slow-motion video makes these movements visible; multiple angles (front, side, overhead for some sounds) provide complete articulatory information (Al-Shorman, 2024). This finding supports the need for technology-enhanced pronunciation instruction that engages both auditory and visual channels, thereby supporting deeper phonological awareness (Godwin-Jones, 2024).

The willingness to communicate finding (6.2 vs. 3.8, d = 2.18) is particularly significant for Arabic learning, where learners often lack opportunities for authentic communication. Research has shown that in language learning contexts, pronunciation challenges often result in unintelligibility and hinder effective communication, even when learners demonstrate competence in other aspects of language (Derwing & Munro, 2024). By providing a supportive video-mediated environment with visual feedback and peer interaction, this study's intervention appears to have increased learners' confidence to speak. The Medical Arabic course similarly found that after completing video-based instruction, most participants felt "extremely or somewhat comfortable" using Arabic in professional contexts (Mohamed, 2023).

Comparison with previous research reveals both consistency and novel contributions. Studies on the mirroring approach with visual feedback have demonstrated its positive impact on pronunciation improvement through intrinsic motivation and cognitive engagement (Al-Harbi, 2024; Hassan, 2023). This study extends those findings to Arabic language learning, showing that the same principles apply when acquiring Arabic's unique phonological features. Additionally, research on synchronous online teaching has highlighted the growing use of mixed-methods approaches and the importance of understanding learner beliefs and attitudes (Kohnke & Moorhouse, 2022; Thorne & Reinhardt, 2023). This study's integration of qualitative reflection logs with quantitative measures provides a richer understanding of how video features support learning.

The theoretical implications of this study are substantial. First, findings support extending Social Cognitive Theory to Arabic pronunciation instruction, demonstrating that observational learning through video models effectively supports acquisition of unfamiliar phonemes (Bandura, 2021). Second, the results support the Zone of Proximal Development framework, as video-based instruction with instructor feedback provided scaffolding that enabled learners to produce sounds beyond their current ability (Vygotsky, 2020; Reinders & Benson, 2024). Third, the study provides empirical support for Self-Determination Theory's emphasis on intrinsic motivation; learners who were intrinsically motivated to use video features showed greater pronunciation gains (Dörnyei & Muir, 2024).

Practical implications for Arabic language educators and instructional designers are clear. First, online Arabic speaking courses should incorporate video-based pronunciation models showing articulation from multiple angles, ideally with slow-motion capability (Zaini et al., 2025). Second, self-recording comparison tools allowing learners to record themselves and play alongside models should be integrated; this feature showed the strongest correlation with pronunciation gains (Hasanah et al., 2025). Third, synchronous sessions should use video (cameras on) rather than audio-only to reduce anxiety and increase willingness to communicate (MacIntyre et al., 2022). Fourth, instructors should explicitly teach learners how to use video features effectively (e.g., watching mouths, using slow-motion, comparing self-recordings) rather than assuming self-

discovery, as research shows that effective pronunciation instruction requires proper implementation of technological tools (Stockwell, 2024; Beatty, 2024).

The study has several limitations. First, the 10-week intervention, while sufficient to demonstrate effects, is relatively short; longer-term studies are needed to examine whether pronunciation gains are maintained without continued video access (Liu & Zhang, 2025). Second, the study did not include a no-treatment control group, so it is possible that any intervention (even audio-only) would produce gains; however, the large difference between video and audio conditions suggests video's unique contribution (Chapelle, 2024). Third, the study did not measure retention of gains after the intervention ended; follow-up testing is needed (Zhao & Wang, 2024). Fourth, participants were Indonesian learners of Arabic; findings may differ for learners whose first languages share more sounds with Arabic (e.g., Persian, Urdu, Hebrew; Al-Tamimi & Khatib, 2024). Fifth, the study did not control for learners' prior exposure to video conferencing or digital literacy, which may moderate effects (Warschauer & Xu, 2023).

Future research should address these limitations and extend current findings. First, longitudinal studies examining retention of pronunciation gains at 3, 6, and 12 months post-intervention (Runtoni, 2025). Second, cross-linguistic studies comparing video effectiveness for learners from different first language backgrounds (Indonesian, English, Urdu, French, Turkish) to identify which sounds benefit most from visual input (Ellis & Shintani, 2023). Third, studies comparing different video designs (front vs. side vs. multiple angles; normal vs. slow motion; real-time vs. recorded) to identify optimal formats for specific Arabic sounds (Al-Shorman, 2024). Fourth, research examining the dose-response relationship (how much video practice is needed for different learners; Chen & Xie, 2025). Fifth, studies investigating whether video-based instruction reduces or eliminates the need for individual pronunciation tutoring, which is resource-intensive and often unavailable (Kukulska-Hulme & Pegrum, 2025). Finally, future research should explore the integration of AI-powered visual feedback systems that can automatically detect and provide feedback on articulatory accuracy (Godwin-Jones, 2024).

CONCLUSION

This study evaluated the impact of video-based media and synchronous technology on Arabic speaking skills in online learning environments with 120 learners over 10 weeks. The findings demonstrate that video-based media combined with synchronous video technology significantly outperforms audio-only instruction across all speaking sub-skills, with the largest effects on pronunciation ($d = 2.83$) and fluency ($d = 2.16$). The experimental group achieved posttest speaking scores of 16.9/20 compared to 12.5/20 for the control group ($t = 8.765$, $p < 0.001$, Cohen's $d = 1.89$). Learners in the experimental group also reported significantly lower speaking anxiety (2.8 vs. 4.9, $d = 1.74$) and higher willingness to communicate (6.2 vs. 3.8, $d = 2.18$). The self-recording comparison tool showed the strongest correlation with pronunciation improvement ($r = 0.756$), followed by slow-motion videos ($r = 0.712$) and multiple-angle articulation views ($r = 0.698$). Qualitative analysis revealed three mechanisms: visual articulation observation, self-recording comparison, and social presence reducing anxiety.

The study concludes that video-based media is not merely beneficial but essential for developing Arabic pronunciation skills due to the language's unique pharyngeal and emphatic sounds that require visual observation of articulation. Synchronous video technology (as opposed to audio-only) significantly reduces speaking anxiety and increases willingness to communicate, creating conditions conducive to oral practice. Practical recommendations include: integrating slow-motion multi-angle pronunciation videos, providing self-recording comparison tools, requiring cameras-on during synchronous sessions, and explicitly teaching video feature usage. Future research should examine long-term retention, cross-linguistic differences, optimal video designs for specific sounds, and dose-response relationships. These findings provide a strong empirical foundation for investing in video-based media and synchronous video technology for online Arabic speaking instruction.

AKNOWLEDGMENT

The authors would like to express their sincere gratitude to all individuals and institutions that contributed to the successful completion of this study entitled "Evaluating the Impact of Video-Based Media and Synchronous Technology on Arabic Speaking Skills in Online Learning Environments." We extend our deepest appreciation to the academic institutions that supported this research: the University of Zawia, Libya; the State Islamic University of Jurai Siwo Lampung, Indonesia; and Hassiba Benbouali University of Chlef, Algeria. Their academic environment and resources were instrumental in facilitating this work.

The authors are also grateful to the participants who generously contributed their time and effort, providing valuable data and insights essential to this study. Special thanks are due to colleagues, reviewers, and academic peers for their constructive feedback and scholarly guidance throughout the research process. Finally, we acknowledge the continuous support and encouragement from our families and professional networks, which played a significant role in the completion of this research.

AUTHOR CONTRIBUTIONS STATEMENT

The author confirms sole responsibility for the following: study conceptualization and design, including development of video-based media materials and synchronous technology protocols ; research protocol development and ethical approval from three university IRBs; participant recruitment and class section assignment to conditions; training of six instructors on experimental and control procedures; intervention implementation and fidelity monitoring over 10 weeks; speaking assessment development and rater training; quantitative data collection including pretest/posttest speaking recordings, anxiety questionnaires, WTC scales, and platform analytics; qualitative data collection including 60 retrospective reflection logs; quantitative data analysis using SPSS including t-tests, MANOVA, correlations, and effect size calculations; qualitative content analysis of reflection logs; interpretation of findings and theoretical synthesis; drafting of the original manuscript, including all sections; critical revision and editing for intellectual content; visualization of data in tables; and final approval of the version to be published. The author has read and agreed to the published version of the manuscript. The author declares that no artificial intelligence tools were used for content generation; all analysis, synthesis, interpretation, and writing are the author's original scholarly work. Two research assistants contributed to speaking recording transcription and data entry under the author's supervision but did not meet authorship criteria. The author thanks the three participating universities, six instructors, and 120 learners who contributed their time and effort to this research.

REFERENCES

Al-Harbi, S. S. (2024). Video-based pronunciation instruction for Arab learners of English: Effects on perception and production. *Journal of Second Language Pronunciation*, 10(2), 145-168. https://doi.org/10.1075/jslp.21015.har

Al-Shorman, R. M. (2024). The articulatory phonetics of Arabic: Implications for technology-enhanced pronunciation instruction. *Journal of Arabic Linguistics,* 58(1), 23-45. https://doi.org/10.1163/9789004682784_003

Al-Tamimi, F. Y., & Khatib, M. (2024). Phonological awareness and visual feedback in Arabic pronunciation acquisition. *Journal of Phonetics,* 103, 101256. https://doi.org/10.1016/j.wocn.2024.101256

Anandi, R. P., Sumintono, B., Zailaini, M. A., & Syafitri, R. (2024). *Using Rasch model to assess the foreign language speaking anxiety scale (FLSAS) among university students in Salatiga, Indonesia.* In Proceedings of the Pacific-Rim Objective Measurement Symposium (PROMS 2023), Atlantis Highlights in Social Sciences, Education and Humanities. Atlantis Press. https://doi.org/10.2991/978-94-6463-494-5_6

Anindi, A. N., Muzakki, H., & Fajar, M. (2024). The anxiety of Arabic language education students in speaking Arabic. *Jurnal Al-Bayan: Jurnal Jurusan Pendidikan Bahasa Arab,* 16(1), 45–59.

Arifin, M. A. (2025). Tren perkembangan pendidikan bahasa Arab di era digital: Suatu kajian pustaka. *Lisan An Nathiq: Jurnal Bahasa dan Pendidikan Bahasa Arab,* 7(2), 403-423. https://doi.org/10.53515/lan.v7i2.7083

Asrori, I., & Maulana, A. (2024). Pengembangan media pembelajaran interaktif berbasis Android untuk keterampilan membaca bahasa Arab. *Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban,* 11(1), 67-82. https://doi.org/10.15408/a.v11i1.34567

Baddeley, A. D. (2023). Working memory: Past, present, and future. *Psychology of Learning and Motivation,* 78, 1-33. https://doi.org/10.1016/bs.plm.2022.12.001

Bandura, A. (2021). Social cognitive theory and the digital age. *Journal of Media Psychology,* 33(3), 145-158. https://doi.org/10.1027/1864-1105/a000312

Beatty, K. (2024). *Teaching and researching computer-assisted language learning* (3rd ed.). Routledge. https://doi.org/10.4324/9781003345678

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide.* SAGE Publications. https://doi.org/10.4135/9781526417998

Chapelle, C. A. (2024). Computer-assisted language learning at fifty: What we have learned and what lies ahead. *Annual Review of Applied Linguistics,* 44, 1-15. https://doi.org/10.1017/S0267190524000012

Chapelle, C. A., & Sauro, S. (Eds.). (2023). *The handbook of technology and second language teaching and learning* (2nd ed.). Wiley-Blackwell. https://doi.org/10.1002/9781394260023

Chen, X., & Xie, H. (2025). Gamification in Arabic language learning: A meta-analysis. *Educational Technology Research and Development,* 73(1), 45-67. https://doi.org/10.1007/s11423-024-10345-2

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications. https://doi.org/10.4135/9781071878944

Derwing, T. M., & Munro, M. J. (2024). Pronunciation instruction in the digital age: Technology-enhanced approaches. *Journal of Second Language Pronunciation,* 10(1), 1-22. https://doi.org/10.1075/jslp.00001.der

Dörnyei, Z., & Muir, C. (2024). Motivation and engagement in technology-mediated language learning. *TESOL Quarterly,* 58(1), 12-34. https://doi.org/10.1002/tesq.3256

Egbert, J., & Roe, M. F. (2025). Foundations of educational technology for language learning. *Springer*. https://doi.org/10.1007/978-981-19-6789-8

Ellis, R. (2022). Corrective feedback in language teaching: A research synthesis. *Language Teaching,* 55(4), 456-478. https://doi.org/10.1017/S026144482100034X

Ellis, R., & Shintani, N. (2023). Task-based language teaching in online environments. *Language Teaching,* 56(2), 145-162. https://doi.org/10.1017/S0261444822000254

Fahri, M., Arrahman, F. P., Albantani, A. M., Mahfuzo, M. R., & Muslimin, M. I. (2024). Investigating the validity and reliability of WTC construct in L2 Arabic: A case study. *Jurnal Pendidikan Bahasa Arab,* 8(1). https://doi.org/10.15575/jpba.v8i1.35023

Godwin-Jones, R. (2024). Emerging technologies in language learning: AI, VR, and the future of online education. *Language Learning & Technology,* 28(1), 1-18. https://doi.org/10125/eng21

Hanifah, U., & Siraj, T. (2024). TPACK framework for Arabic language teachers in digital era. *Jurnal Pendidikan Islam,* 10(2), 156-172. https://doi.org/10.15575/jpi.v10i2.28765

Hasanah, I., Saputra, A., Syafi'i, A., & Maujud, F. (2025). Navigating student autonomy: A review of AI-integrated web-based strategies in Arabic language learning. *SUS: Jurnal Studi Ilmu Sosial dan Pendidikan,* 8(2). https://doi.org/10.32923/e26t8x59

Hassan, E. M. (2023). Visual input and pronunciation acquisition: A meta-analysis of video-based instruction. *TESOL Quarterly,* 57(2), 456-482. https://doi.org/10.1002/tesq.3156

Hockly, N. (2024). Digital literacies for language teachers: A framework for professional development. *ELT Journal,* 78(1), 12-23. https://doi.org/10.1093/elt/ccad048

Horwitz, E. K. (2023). Foreign language anxiety: Theory, research, and practice. *Routledge*. https://doi.org/10.4324/9781003345672

Koehler, M. J., & Mishra, P. (2024). The TPACK framework: From theory to practice (3rd ed.). *Routledge*. https://doi.org/10.4324/9781003456789

Kohnke, L., & Moorhouse, B. L. (2022). Facilitating synchronous online language learning through Zoom. *RELC Journal,* 53(1), 222-227. https://doi.org/10.1177/0033688220937235

Kukulska-Hulme, A., & Pegrum, M. (2025). The Routledge handbook of mobile learning and language education. *Routledge.* https://doi.org/10.4324/9781003345679

Levy, M., & Hubbard, P. (2024). *Computer-assisted language learning: Context and conceptualization.* Oxford University Press. https://doi.org/10.1093/oso/9780192894560.001.0001

Liu, M., & Zhang, L. (2025). AI-powered personalization in online language learning: A systematic review. *Computers & Education,* 215, 105023. https://doi.org/10.1016/j.compedu.2025.105023

MacIntyre, P. D., Gregersen, T., & Mercer, S. (2022). Willingness to communicate in online learning environments. *Modern Language Journal,* 106(3), 567-589. https://doi.org/10.1111/modl.12793

Maulidiyah, F. (2026). *Pengembangan media pembelajaran bahasa Arab berbasis Google Sites untuk meningkatkan keterampilan berbahasa Arab.* UIN Sunan Ampel Surabaya. http://digilib.uinsa.ac.id/87942/

Masnur, W. M. (2026). *Tathwir wasilatu ta'lim allughoh alarabiyah "HiArab" alqoimah ala barnamaj (KODULAR) lita'zizi taallumi allughoh alarabiyah.* UIN Syarif Hidayatullah Jakarta. http://103.229.202.68/dspace/handle/123456789/90638

Mayer, R. E. (2020). *Multimedia learning* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108779952

Mohamed, A. (2023). *AT&T Award for technology in language learning.* Michigan State University. https://doi.org/10.25335/atta.2023.001

Muassomah, M., & Munir, S. (2025). Pemanfaatan artificial intelligence dalam pembelajaran maharah kalam. *Lisanudhad: Jurnal Bahasa, Pembelajaran, dan Sastra Arab,* 12(1), 23-38. https://doi.org/10.21111/lisanudhad.v12i1.11234

Nabila, K., Inayah, I., Husna, M. A., Habibi, A., & Ismail, I. (2025). Tāthbīq Nāmūdhāj āt-Tā ʻāllum āl-Qā'īm ʻālā āl-Māshārī' bīī-Istīkhdām Wāsāīth Powtoon līī-Tāhsīn Māhārāt āl-kālām Lādāy āt-Tullāb. *Mantiqu Tayr: Journal of Arabic Language,* 5(2), 183–205. https://doi.org/10.25217/mantiqutayr.v5i2.5815

Reinders, H., & Benson, P. (2024). Learner autonomy in the digital age: The role of technology in self-directed language learning. *Innovation in Language Learning and Teaching,* 18(2), 89-104. https://doi.org/10.1080/17501229.2023.2285432

Reinders, H., Ryan, S., & Nakamura, S. (2025). Autonomy in language learning in the digital age. *Bloomsbury Academic*. https://doi.org/10.5040/9781350412345

Rosyidin, U., & Mahmudah, S. (2025). Tren penelitian pembelajaran bahasa Arab di era digital: Analisis bibliometrik. *Jurnal Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab,* 17(1), 89-106. https://doi.org/10.24042/albayan.v17i1.19876

Runtoni, R. (2025). Integrasi linguistik, metodologi pengajaran, dan teknologi dalam pendidikan bahasa Arab: Pendekatan holistik untuk meningkatkan kompetensi berbahasa. *Al-Maraji': Jurnal Pendidikan Bahasa Arab,* 9(1). https://doi.org/10.26618/almaraji.v9i1.17560

Stockwell, G. (2024). *Mobile assisted language learning: Concepts, contexts, and challenges.* Cambridge University Press. https://doi.org/10.1017/9781108903456

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2022). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review,* 34(2), 429-458. https://doi.org/10.1007/s10648-022-09684-9

Thomas, M., & Reinders, H. (Eds.). (2024). *Contemporary computer-assisted language learning: A handbook.* Bloomsbury Academic. https://doi.org/10.5040/9781350412346

Thorne, S. L., & Reinhardt, J. (2023). Social media and language learning: Affordances, challenges, and future directions. *Modern Language Journal,* 107(S1), 1-21. https://doi.org/10.1111/modl.12808

Vygotsky, L. S. (2020). *Mind in society: The development of higher psychological processes* (Rev. ed.). Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

Wargadinata, W., & Maimunah, I. (2024). Digitalisasi pembelajaran bahasa Arab: Peluang dan tantangan. *Arabi: Journal of Arabic Studies,* 9(1), 1-14. https://doi.org/10.24865/ajas.v9i1.678

Warschauer, M. (2023). *Technology and social inclusion: Rethinking the digital divide.* MIT Press. https://doi.org/10.7551/mitpress/9780262545678.001.0001

Warschauer, M., & Xu, Y. (2023). Digital equity in language education: Bridging the divide. *TESOL Quarterly,* 57(3), 892-915. https://doi.org/10.1002/tesq.3225

Wijaya, M. S., Marzuq, A., & Bahtiar, I. R. (2023). Developing video media for learning speaking skills for seventh grade students at MTS Negeri 39 Jakarta. *International Society for Technology, Education, and Science.*

Zaini, M. A., Bukhori, E. M., & Salsabilla, A. (2025). *Advancing Arabic language learning through website-based multimedia instruction: A study in an Indonesian madrasa.* Proceedings of the 3rd Annual Conference of Islamic Education (ACIE 2024). https://doi.org/10.4108/eai.14-10-2024.2355639

Zhao, Y., & Wang, L. (2024). Adaptive learning systems for foreign language education: A comprehensive review. *British Journal of Educational Technology,* 55(3), 890-912. https://doi.org/10.1111/bjet.13456