Jurnal Pustaka Indonesia (JPI)
2023, Vol. 3, No. 1, 32 –42
https://doi.org/10.62159/jpi.v3i1.765

OPEN ACCESS Research Article

Listening and Vocabulary Gains via Arabic Song Media: A Two-Cycle Classroom Study in Grade X

Muhammad Sufian[1*], Rimadhona[2], Guntur Cahaya Kesuma[3], Umi Hijriyah[4], Erlina[5]
[12345] Universitas Islam Negeri Raden Intan Lampung, Lampung, Indonesia

ABSTRACT

This article reports a classroom study examining whether Arabic song media can improve Grade-X students' listening skills and vocabulary mastery. Drawing on a descriptive qualitative design with two iterative cycles (Plan Act Observe Reflect), the study was implemented in one intact class (n = 29) at a private senior high school. Data sources comprised observation fieldnotes, lesson artefacts (tests, rubrics), and teacher student feedback across cycles. Baseline assessment showed only 7 students meeting the minimum criterion in listening and 12 in vocabulary. After Cycle I, the number of students meeting the criterion rose to 19 (≈65%) for listening and 17 (≈59%) for vocabulary; after Cycle II, gains reached 26 (≈90%) and 25 (≈86%), respectively. Pedagogically, improvements were associated with explicit objectives, focused listening tasks, and iterative feedback; practically, song media increased engagement and supported repeated, meaningful exposure to target lexis. The study concludes that Arabic songs, when embedded in clear task designs and continuous feedback, can deliver substantial short-term gains in listening and vocabulary for Grade-X learners. Limitations include the single-class scope and lack of a control group; implications and future directions are outlined.

ARTICLE HISTORY
Received 30 Januari 2025
Revised 10 Maret 2025
Accepted 3 April 2025

KEYWORDS
Arabic songs; listening comprehension; vocabulary mastery; classroom study

PUBLISHER'S NOTE
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike (CC BY 4.0) license

CORRESPONDING AUTHOR
Muhammad Sufian, Universitas Islam Negeri Raden Intan Lampung, Lampung, Indonesia. Email: iyansufian1998@gmail.com

Introduction

Listening is widely recognized as the foundational receptive skill in second and foreign language learning, shaping learners' lexical growth, phonological attunement, and discourse awareness long before productive skills fully emerge (and, crucially, continuing to constrain them thereafter). Within contemporary second-language listening pedagogy, process-oriented approaches emphasize metacognitive awareness, strategy training, and iterative cycles of prediction verification using meaningful input (rather than mere answer-checking), a stance articulated programmatically by Vandergrift and Goh and widely adopted in subsequent instructional designs (Goh & Vandergrift, 2021). In parallel, research on vocabulary learning has clarified that listening can support both incidental and deliberate word learning when exposure is sufficiently frequent and salient, especially under conditions of repeated encounters and guided attention to form meaning mappings (Feng & Webb, 2020b; Nie et al., 2022).

Against this backdrop, song-based listening has attracted sustained interest as a potentially high-leverage classroom tool. Songs naturally bundle several ingredients known to facilitate

learning: (1) rich, repeated input; (2) prosodic scaffolds (rhythm, stress, melody) that aid segmentation and rehearsal; (3) strong affective engagement; and (4) cultural narratives that support meaning-making. Systematic reviews and classroom studies converge on the finding that, under pedagogically structured use, songs can improve vocabulary recognition and retention as well as aspects of listening comprehension although effect sizes vary with learner level, task design, and the way teachers exploit lyrics before/during/after listening (A. C. S. Chang, 2024; Hamilton et al., 2024; Tilwani et al., 2022). From a cognitive perspective, melodies and metrical regularities recruit verbal musical working memory resources, strengthening rehearsal in the phonological loop and promoting chunking; converging behavioral and neurocognitive evidence suggests partially shared mechanisms for verbal and tonal memory that may explain the "song advantage" reported in several experiments (Schulze & Koelsch, 2012; Williamson et al., 2010). Complementary theoretical lenses Dual Coding Theory and the Cognitive Theory of Multimedia Learning predict benefits when verbal material is paired with nonverbal auditory structure and presented with principled redundancy and coherence, provided extraneous load is minimized and attention remains on the target language (Loor, 2025; Suárez et al., 2016).

Empirically, experimental and quasi-experimental studies across L2 contexts show that listening to songs can yield measurable gains in word-form recognition, form meaning connection, and collocation learning, with repetition and task-embedded retrieval practice amplifying effects (Nie et al., 2022; Zhang, 2022). In secondary-school settings, the "time on task" advantage of songs students' willingness to replay outside class appears to sustain exposure far beyond a single lesson, a finding echoed in recent school-based studies using English songs for Indonesian tenth-graders (Nandita et al., 2025). Importantly, song-mediated input can also lower anxiety and enhance enjoyment, thus potentially reducing the affective filter and supporting sustained listening effort classic mechanisms anticipated by Krashen's socio-affective model and elaborated in contemporary work on classroom emotion and engagement (Murad & Othman, 2023). That said, cautionary reviews note that "background music" during study can sometimes hinder complex learning if it competes with linguistic processing a reminder that, pedagogically, songs must be the object of listening (with scaffolded tasks), not mere background sound (Lehmann & Seufert, 2017).

Within Arabic as a foreign language (AFL), the case for song-based pedagogy is compelling but comparatively under-documented relative to English and other L2s. A notable study on American learners of Arabic found that songs improved the memorability of vocabulary and formulaic expressions, suggesting that melody and prosody help stabilize mappings between Arabic phonology/orthography and meaning an especially valuable scaffold given Arabic's non-Latin script and rich morphology (Moufarrej & Salameh, 2019). Emerging studies in Indonesian higher-education and school contexts report positive learner responses and improvements in Arabic vocabulary and listening, while also cataloging practical constraints (song selection, classroom management, and alignment with assessment) (Hanifah, 2021; Khudriyah, 2022). In short, the international evidence base supports carefully designed song use; the Indonesian AFL literature, though growing, still lacks detailed, classroom-anchored frameworks at the upper-secondary level that connect song-based listening to measurable vocabulary outcomes.

The local curricular context further motivates this inquiry. In Indonesian senior high schools general and faith-based alike Arabic listening and vocabulary (mufradāt) are core components of communicative competence, yet teachers frequently report bottlenecks: limited authentic audio,

uneven exposure to spoken varieties, and student disengagement when listening is treated as a test-like activity rather than a meaning-oriented practice. A process-driven approach that embeds songs as structured listening texts combined with pre-listening lexical activation, guided lyric decoding, and post-listening retrieval tasks aligns with current listening pedagogy and offers a feasible route to increase input quantity/quality without sacrificing classroom manageability (Goh & Vandergrift, 2021). At the same time, any classroom innovation must respect school norms, age appropriateness, and religious sensitivities in material selection concerns repeatedly raised in practitioner reports on Arabic song use and addressed through principled criteria (theme, speed, diction clarity, cultural fit, copyright) (Siraj et al., 2023).

Cognitively, songs may assist Indonesian learners of Arabic through several nested mechanisms. First, rhythm and melody provide temporal scaffolds that help parse syllabic patterns, stress, and intonation, easing segmentation in a language where phonotactics and prosody differ markedly from Indonesian/English. Second, the coupling of melodic contours with lexical items supports richer encoding (verbal + musical), consistent with dual-route memory accounts and with research showing interference/synergy between verbal and musical working memory systems (Fennell et al., 2021; Schulze & Koelsch, 2012). Third, songs naturally afford spaced and interleaved exposure, as learners tend to replay tracks voluntarily; such repetition is a powerful driver of incidental vocabulary uptake, especially for spoken-form recognition and collocation salience (Tilwani et al., 2022; Zhang, 2022). Fourth, the inherently affective dimension of music can broaden attention and persistence, thereby increasing the likelihood of noticing and form meaning consolidation during repeated listens (Murad & Othman, 2023). Together, these mechanisms predict benefits for both listening comprehension processes (prediction, inference, monitoring) and lexical development (breadth and depth).

Nevertheless, the literature also signals boundary conditions. Gains are attenuated when lyrics are too fast, lexically dense, or culturally opaque; when classroom tasks devolve into karaoke-style singing without focused listening; or when assessment remains misaligned with instructional targets (e.g., testing isolated word lists but teaching only via global lyric appreciation). Recent syntheses emphasize that explicit planning across the pre-while-post listening arc previewing key lexis, setting listening purposes, exploiting lyric gaps/annotation, and orchestrating retrieval practice mediates outcomes more than the mere presence of songs. Additionally, while enthusiasm is a virtue, background-music studies warn against adding music as a concurrent stimulus during unrelated tasks, given possible attentional competition; in language classrooms, educators should treat songs as the primary text, not ambient sound.

The present study is situated in a typical Indonesian private Islamic high school (SMA Al-Azhar 3 Bandarlampung) where Arabic is a required subject and where teachers seek practical, evidence-based ways to elevate listening practice and vocabulary acquisition for Grade X learners. Building on the foregoing, our aims are twofold: first, to determine whether the use of Arabic songs as audio media can measurably improve students' listening skills and mastery of target mufradāt; second, to document enabling and constraining factors in implementation (material selection, task design, time management, classroom management, learner affect) that condition success in this context. By marrying classroom-level design to theory metacognitive listening instruction, incidental vocabulary learning through repeated aural exposure, and dual-coding/multimedia principles we intend to contribute a fine-grained account that is methodologically transparent and realistically

adoptable by school practitioners (Feng & Webb, 2020a; Suárez et al., 2016; Vandergrift & Goh, 2012).

Our contribution is threefold. Conceptually, we extend song-based listening beyond "motivational add-on" toward a principled design for Arabic in secondary education, specifying how pre-listening lexical activation, lyric-focused while-listening tasks, and post-listening retrieval cycles can be orchestrated to maximize both comprehension and vocabulary outcomes. Empirically, we add to a still-limited AFL evidence base by reporting outcomes with Grade X learners in Indonesia and by triangulating artifacts (listening tasks, lyric worksheets), teacher feedback, and student performance. Methodologically, we provide implementation heuristics and selection criteria for Arabic songs (tempo, enunciation, lexical targetability, cultural appropriateness) and map them to task types aligned with recognized listening pedagogy (Hamilton et al., 2024).

Finally, we acknowledge potential risks and ethical/procedural considerations copyright and licensing for classroom use; alignment with school codes of conduct; and the necessity of assessing learning against curricular targets rather than generic enjoyment. To mitigate these, we adopt a transparent selection protocol, embed tasks that make vocabulary uptake visible (e.g., spoken-form recognition and form meaning mapping checks), and treat songs as one strand in a balanced listening curriculum that also includes non-musical texts. In doing so, we respond to calls in recent reviews to move beyond "songs as fun" toward songs as structured input a stance consistent with both the science of listening instruction and the cognitive science of memory.

Methods

This study adopted a descriptive qualitative design embedded in a two-cycle classroom implementation following a Plan Act Observe Reflect sequence typical of critical participatory action research (Kemmis et al., 2014). Conducted in one intact Grade-X class (n = 29) at SMA Al-Azhar 3 Bandarlampung, the work was situated in a naturalistic setting to capture processes as they unfolded in real time (Creswell & Poth, 2018; Erlina et al., 2025; Lincoln & Guba, 1985). Arabic song media were curated against explicit criteria clear enunciation, moderate tempo, recurring target lexis, cultural/ethical fit and then integrated into lesson plans with stated objectives and success criteria (KKM). Each lesson followed a stable arc: pre-listening (activate prior knowledge; preview 6 10 target words; set listening purpose), while-listening (2 3 plays with graduated tasks from gist to selective gap-fill and verification), and post-listening (brief retrieval practice, pronunciation focus, and short pair exchanges using target vocabulary). Between cycles, small-scale design refinements (time-boxing, simplified prompts, sentence frames, tighter lexical load with more spaced repetition) were made based on observations and debriefs, consistent with iterative improvement logic in classroom inquiry (Afifurrahman et al., 2025; Patton, 2015).

Data were gathered through participant observation (structured fieldnotes), student artefacts (lyric worksheets and class tasks), and listening/vocabulary assessments at three points (baseline, end-Cycle I, end-Cycle II). Analysis combined descriptive statistics (counts/percentages meeting KKM per time point) with qualitative constant comparison to relate observed behaviors and artefacts to outcomes (Glaser & Strauss, 1967), organized within the Miles Huberman Saldaña interactive flow of data reduction, data display, and conclusion drawing/verification (Miles M. B. & Saldana, 2014). Trustworthiness was enhanced via triangulation across sources, methods, and time, a simple audit

trail of plans/debrief notes, and brief member checks with the teacher. School approval was secured; parental notification followed school procedure; student data were anonymized; and songs were used strictly for educational purposes in line with cultural ethical guidelines.

Results and Discussion

Effectiveness of Arabic Songs as a Medium for Listening Skills and Vocabulary Mastery

The data reveal a consistent and sustained pattern of improvement across two key indicators—listening comprehension and vocabulary mastery—over three measurement points (pre-cycle, end of Cycle I, and end of Cycle II). This "stepwise increase" pattern indicates that songs are effective when positioned as structured listening texts (pre while post design), rather than as mere background entertainment. At the aggregate level, the class target of ≥85% of students achieving the mastery threshold (KKM) was reached by the end of Cycle II for both skills, with adequate safety margins (listening 89.7%; vocabulary 86.2%). This is pedagogically significant: surpassing the threshold helps buffer against daily performance fluctuations, reducing the risk of falling back below the KKM benchmark.

Quantitatively, listening mastery rose from 24.1% to 65.5% and than 89.7%, equivalent to 19 additional students achieving mastery compared to baseline (+65.5 percentage points, pp). Vocabulary mastery increased from 41.4% to 58.6% and than 86.2%, equivalent to 13 additional students (+44.8 pp). In other words, listening demonstrated both higher absolute gains and greater relative growth than vocabulary. Cognitively, this makes sense: stabilization of sound recognition is typically achieved more quickly through rhythmic melodic repetition, whereas deeper lexical enrichment requires more frequent retrieval practice and exposure across varied contexts.

The sharp leap from pre-cycle to Cycle I (listening: +41.4 pp; vocabulary: +17.2 pp) reflects the initial effect of explicitly structured task design: pre-listening (schema activation + lexical preview), while-listening (2 3 directed playbacks: gist, detail, verification), and post-listening (brief retrieval, pronunciation drills, mini-sentence production). The repetition of choruses/refrains reinforced sound stabilization and lexical marking, accelerating recognition of key words.

From Cycle I to Cycle II, micro-refinements (tighter time-boxing, simplified gap-fill prompts, use of sentence frames to encourage production, reduced lexical density per lesson, and enhanced spaced repetition) contributed to further improvement: listening +24.1 pp, vocabulary +27.6 pp. Interestingly, vocabulary "caught up" during this stage due to more explicit form meaning verification tasks and reinforced recall (matching, short dictations, expressing target phrases in mini-dialogues). This pattern suggests that inter-cycle refinement is not optional but rather a critical lever for stabilizing achievement beyond the ≥85% threshold.

Table 1. Mastery of Listening & Vocabulary per Phase (n = 29)

Phase	Listening Mastery (n, %)	Vocabulary Mastery (n, %)	Class Target (≥85%)
Pre-cycle	7 (24.1%)	12 (41.4%)	Not achieved
Cycle I	19 (65.5%)	17 (58.6%)	Not achieved
Cycle II	26 (89.7%)	25 (86.2%)	Achieved (both)

Table 2. Change in Mastery Across Phases (pp = percentage points)

Skill	Δ Pre → C1 (n; pp)	Δ C1 → C2 (n; pp)	Δ Pre → C2 (n; pp)
Listening	+12; +41.4 pp	+7; +24.1 pp	+19; +65.5 pp
Vocabulary	+5; +17.2 pp	+8; +27.6 pp	+13; +44.8 pp

Table 2 shows incremental mastery across the two skills. For listening, mastery increased by +12 students (+41.4 pp) from Pre to Cycle I, by +7 (+24.1 pp) from Cycle I to Cycle II, and by +19 (+65.5 pp) overall. For vocabulary, mastery rose by +5 students (+17.2 pp) in Cycle I, +8 (+27.6 pp) in Cycle II, and +13 (+44.8 pp) overall. Both skills showed consistent improvement, with listening gains outpacing vocabulary.

Relative improvement ratios further confirm the trend: listening leaped ≈2.72× (+172%) from baseline to Cycle I and ≈3.72× (+272%) by Cycle II, while vocabulary grew ≈1.42× (+42%) in Cycle I and ≈2.08× (+108%) by Cycle II. This highlights the early advantage of songs for stabilizing sound recognition, while vocabulary required a second cycle emphasizing spaced repetition and explicit retrieval practice. Nonetheless, by the end of Cycle II, both skills had more than doubled their baseline performance, underscoring the effectiveness of the structured pre while post design.

Table 3. Relative Improvement Ratios vs Baseline (n = 29)

Skill	Mastery Pre-cycle	Cycle I	Ratio vs Baseline
Listening	0.241	0.655	2.72× (~+172%)
Vocabulary	0.414	0.586	1.42× (~+42%)

Table 3 shows that listening accelerated most strongly from the first cycle, reflecting the robust effect of rhythmic melodic repetition on sound processing. Vocabulary rose more slowly in Cycle I but "caught up" in Cycle II, consistent with the addition of explicit form meaning verification and sentence-frame supports. By Cycle II, both skills exceeded twice their baseline performance, confirming that songs are effective when treated as structured texts with scaffolded tasks.

To close residual vocabulary gaps, ongoing retrieval practice and spaced repetition are essential in subsequent sessions. The evidence thus indicates that songs provide a stronger initial "head-start" in listening, while vocabulary accelerates in later cycles as retrieval and production are emphasized. Pedagogically, this underscores the importance of aligning assessments with instructional targets: listening tests should measure practiced objectives (gist/detail, target-word detection), while lexical tests should evaluate form meaning mappings of words explicitly taught through songs.

Overall, the evidence demonstrates that Arabic songs can effectively elevate both listening and vocabulary mastery to surpass class targets within two cycles, provided that: (a) songs are carefully curated (clear articulation, moderate tempo, lexical repetition), (b) pre while post implementation is disciplined with scaffolded tasks and rapid feedback, and (c) inter-cycle refinements reduce lexical load while enhancing retrieval quality. Listening benefits earlier due to rhythm melody reinforcement of sound forms, whereas vocabulary requires explicit verification and

recall for comparable mastery levels. Following this blueprint, other schools and teachers can replicate results with high likelihood of success, while remaining mindful of design limitations and the need to test long-term retention.

Discussion

The sharp gains in listening and vocabulary mastery confirm that songs are effective when positioned as structured listening texts rather than mere background fillers. The key lies in goal-oriented teaching that incorporates lexical pre-listening (schema activation + preview of target words/phrases), scaffolded tasks (gist → detail → verification), and rapid feedback to sustain focus and momentum. Cognitively, rhythmic melodic repetition stabilizes phonological representations and facilitates sound meaning mapping, while the affective pull of music boosts attention and persistence in listening. Experimental evidence supports this claim: singing has been shown to outperform spoken repetition in short-term acquisition of foreign-language phrase pairs; simple, repetitive melodic patterns are most beneficial for verbatim recall, whereas constantly changing melodies may disrupt focus (Ludke et al., 2014; Wallace, 1998). Listening yielded both absolute and relative gains greater than vocabulary, indicating that sound patterns tend to stabilize faster than the expansion of semantic/collocational networks, which require more intensive retrieval practice and varied usage contexts.

The pre while post approach used here structured 2 3 playbacks moving from gist to detail to verification aligns with pedagogical models of *listening-as-process*, which emphasize metacognitive awareness (planning, monitoring, evaluating) rather than merely answering comprehension questions. The literature warns that a product-oriented obsession with "correct answers" can cause teachers to overlook training in process (strategies, attention to prosodic cues, stepwise form meaning processing). Process- and metacognition-oriented models offer stronger scaffolds for L2 listening development both inside and outside classrooms, consistent with the measurable cycle-to-cycle gains observed in this study (Collins, 2022; Field, 2009; Goh & Vandergrift, 2021; Wagner et al., 2024).

The vocabulary gains that "caught up" in Cycle II suggest that after sound stabilization, success hinges on frequent, brief retrieval practices (quick matching, micro dictations, use of target phrases in sentences) combined with spaced repetition across lessons. Cognitive research shows that retrieval practice outperforms elaborative study even on reasoning-demanding items because the act of recall itself strengthens memory traces (the testing effect). Likewise, the spacing effect consistently enhances long-term retention when exposures are distributed optimally. These principles were operationalized through Cycle II refinements: reducing lexical load, providing sentence frames, and distributing repetitions to make recall more robust (Karpicke & Blunt, 2011).

Mastery measured against the minimum competency standard (KKM) is meaningful only if it is isomorphic with what was taught: listening items must assess gist and detail skills explicitly practiced, while vocabulary items must test the form meaning pairs actually embedded in the songs. This principle of alignment resonates with L2 vocabulary research that stresses selecting sufficiently frequent targets, repeated processing, and meaningful use across tasks all of which increase the likelihood of words moving from passive recognition to active ownership (Nation, 2001).

Music provides affective leverage (engaging, enjoyable), yet literature warns that affect translates into achievement only when directed by well-structured tasks. Experiments comparing

singing with spoken repetition show advantages for the singing condition in phrase recall; however, classic studies also emphasize that simplicity and repetition in musical patterns are crucial, since overly variable melodies can disrupt lyric processing. Practical implications for classroom replication include curating songs with clear articulation, moderate tempo, repeated refrains, and manageable lexical density; using 2 3 playbacks with distinct goals (gist/detail/verification); and reinforcing recall through post-listening retrieval tasks (L. Chang & Ding, 2021; Ludke et al., 2014).

The concise and focused pre while post framework also addresses common large-class problems such as attention dispersion and fatigue. The *listening-as-process* approach encourages teachers to teach *how* to listen (recognizing prosodic cues, parsing phrases, anticipating rhyme/rhythm patterns), not merely to test comprehension (Amanvermez Incirkus, 2025). Within the song context, rhythm and rhyme serve as natural boundary markers that hold attention, while structured tasks prevent music from becoming a distractor. This explains why the largest leap occurred immediately after the structured design was implemented (Pre to Cycle I), followed by additional gains from micro-refinements (Cycle I to II)(Field, 2009).

First, position songs as the main text, not as background audio; each playback requires explicit goals and small deliverables. Second, align assessment tasks with instruction (gist/detail items for listening, recycled word/phrase lists for vocabulary) (MacPhail et al., 2023; Yang & Song, 2023). Third, schedule retrieval practices that are short but frequent, and provide feedback highlighting specific next steps (syllable pronunciation, sandhi marking, collocation choices). These principles are consistent with high-impact feedback meta-analyses and robust evidence on retrieval and spacing effects from learning psychology.

Methodologically, the single-class design without a control group limits strict causal inference; some improvements may stem from task familiarity. The narrow song corpus deliberately chosen for repetition means generalization to other genres/tempi requires further testing. Moreover, reliance on categorical mastery rates (KKM) rather than continuous scores and the absence of delayed post-tests preclude conclusions about long-term retention. Future research should therefore include comparison groups, determine optimal dosage (number of cycles/sessions), test genre/tempo variation, and use delayed post-tests to assess durability all while maintaining task design principles rooted in *listening-as-process* and retrieval reinforcement.

Based on the results of the analysis, it can be concluded that (1) Proven effectiveness: mastery rates surged, with the class target (≥85%) achieved for both listening and vocabulary. (2) Design "recipe": combining curated songs, scaffolded tasks, formative feedback, and language scaffolding (sentence frames) was the main driver of improvement. (3) Affective cognitive bridge: musical engagement transformed into measurable gains when bound by structured tasks and isomorphic assessment. (4) Meaningful iteration: small inter-cycle refinements produced tangible improvements in outcome stability, consistent with evidence on feedback, retrieval practice, and spacing in educational and SLA research.

Conclusion

This study confirms that Arabic songs, when employed as structured listening texts, substantially improve students' listening comprehension and vocabulary mastery. Over two instructional cycles, mastery rates rose consistently, with listening reaching 89.7% and vocabulary 86.2%, surpassing the class target of ≥85%. The sharper gains in listening highlight the immediate effect of rhythmic–

melodic repetition in stabilizing phonological recognition, while vocabulary growth accelerated in the second cycle through explicit retrieval practice, sentence frames, and spaced repetition. These findings demonstrate that songs function most effectively not as background entertainment but as carefully curated instructional texts within pre–while–post listening frameworks, supported by graded tasks and rapid feedback.

Beyond classroom outcomes, this research contributes a replicable blueprint for integrating music into language pedagogy: selecting songs with clear articulation and lexical repetition, embedding them in structured task cycles, and refining implementation across iterations. While the single-class design and limited song corpus constrain generalizability, the evidence underscores the potential of music to bridge affective engagement and cognitive achievement. Future studies should expand to multiple schools, explore varied song genres, and include delayed post-tests to assess retention. Taken together, these findings position Arabic songs as a powerful, evidence-based medium for aligning enjoyment with measurable learning gains in second language acquisition.

References

Afifurrahman, M., Erlina, E., Ghazi, F., Koderi, K., & Sufian, M. (2025). Efektivitas penerapan model induktif kata bergambar dan investigasi kelompok terhadap hasil belajar kitabah siswa sekolah dasar. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 10*(01), 231–247. https://doi.org/10.23969/jp.v10i01.23507

Amanvermez Incirkus, F. (2025). How do listening lessons based on a holistic approach effect Turkish pre-service teachers' awareness of metacognitive listening strategies? *SAGE Open, 15*(1), 21582440251330040. https://doi.org/10.1177/21582440251330042

Chang, A. C. S. (2024). The effects of varying practice modes on L2 learners' vocabulary retention through songs: Listening, singing, and oral reading. *Tesl-Ej, 28*(3), 1–18. https://doi.org/10.55593/ej.28111a9

Chang, L., & Ding, Y. (2021). Comparing the effects of different post-listening output tasks on second language incidental vocabulary acquisition: Revisiting the involvement load hypothesis. *Current Journal of Applied Linguistics, 44*(4), 506–519. https://doi.org/10.1515/CJAL-2021-0032

Collins, H. K. (2022). When listening is spoken. *Current Opinion in Psychology, 47*, 101402. https://doi.org/10.1016/j.copsyc.2022.101402

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage.

Erlina, E., Koderi, K., & Sufian, M. (2025). Designing a gender-responsive Qira'ah learning module: Bridging equality and inclusivity in Islamic higher education. *Jurnal Ilmiah Islam Futura, 25*(1), 239–262.

Feng, Y., & Webb, S. (2020). Learning vocabulary through reading, listening, and viewing: Which mode of input is most effective? *Studies in Second Language Acquisition, 42*(3), 499–523. https://doi.org/10.1017/S0272263119000494

Feng, Y., & Webb, S. (2020b). Learning vocabulary through reading, listening, and viewing. *Studies in Second Language Acquisition, 42*(3), 499–523. https://doi.org/10.1017/S0272263119000494

Fennell, A. M., Bugos, J. A., Payne, B. R., & Schotter, E. R. (2021). Music is similar to language in terms of working memory interference. *Psychonomic Bulletin & Review, 28*(2), 512–525. https://doi.org/10.3758/s13423-020-01833-5

Field, J. (2009). *Listening in the language classroom*. Cambridge University Press. https://doi.org/10.1017/CBO9780511575945

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine. https://books.google.co.id/books?id=oUxEAQAAIAAJ

Goh, C. C. M., & Vandergrift, L. (2021). *Teaching and learning second language listening*. Routledge. https://doi.org/10.4324/9780429287749

Hamilton, C., Schulz, J., Chalmers, H., & Murphy, V. A. (2024). Investigating the substantive linguistic effects of using songs for teaching second or foreign languages to preschool, primary and secondary school learners: A systematic review of intervention research. *System, 124*, 103350.

https://doi.org/10.1016/j.system.2024.103350
Hanifah, U. (2021). Fun Arabic learning through songs media. *Tanwir Arabiyyah: Arabic As Foreign Language Journal, 1*(2), 73–82. https://doi.org/10.31869/aflj.v1i2.2873
Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner*. Springer Singapore. https://doi.org/10.1007/978-981-4560-67-2
Khudriyah, K. (2022). The effectiveness of songs as media to teach listening at SMA Primaganda. *Ilmuna: Jurnal Studi Pendidikan Agama Islam, 4*(1), 12–28. https://doi.org/10.54437/ilmuna.v4i1.314
Lehmann, J. A. M., & Seufert, T. (2017). The influence of background music on learning in the light of different theoretical perspectives and the role of working memory capacity. *Frontiers in Psychology, 8*(OCT), 1–11. https://doi.org/10.3389/fpsyg.2017.01902
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.
Loor, K. M. M. (2025). Songs as pedagogical tools for improving listening skills in ESL learners: A systematic literature review. *Sinergia Academica*, 723–739.
Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. *Memory and Cognition, 42*(1), 41–52. https://doi.org/10.3758/s13421-013-0342-5
MacPhail, A., Tannehill, D., Leirhaug, P. E., & Borghouts, L. (2023). Promoting instructional alignment in physical education teacher education. *Physical Education and Sport Pedagogy, 28*(2), 153–164. https://doi.org/10.1080/17408989.2021.1958177
Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.
Moufarrej, G., & Salameh, C. (2019). The effects of songs on vocabulary retention in foreign language acquisition. *Al-'Arabiyya, 52*, 101–124. https://www.jstor.org/stable/48600457
Murad, T., & Othman, J. (2023). Arab EFL students' vocabulary enrichment through listening to songs from the English language. *International Journal of Advanced Education and Research, 8*(1), 68–75.
Nandita, D., Thamrin, N. S., Anggreni, A., & Nadruna. (2025). Improving students' vocabulary mastery through listening English songs of SMA Negeri 1 Torue. *Journal of English Language and Education, 10*(4), 994–999.
Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
Nie, K., Fu, J., Rehman, H., & Zaigham, G. H. K. (2022). An empirical study of the effects of incidental vocabulary learning through listening to songs. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.891146
Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). Sage.
Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences, 1252*(1), 229–236. https://doi.org/10.1111/j.1749-6632.2012.06447.x
Siraj, T., Sofiyah, E. M., Rahmawati, A., Fifilyana, A., Febriyantika, A. E., Husnah, B., & Ampel, U. I. N. S. (2023). Penggunaan media lagu anak-anak dalam meningkatkan kemampuan menghafal kosakata bahasa Arab konsisten di semua tingkatan meramalkan keuntungan masa depan. *Ihtimam: Jurnal Pendidikan Bahasa Arab, 6*, 54–64.
Suárez, L., Elangovan, S., & Au, A. (2016). Cross-sectional study on the relationship between music training and working memory in adults. *Australian Journal of Psychology, 68*(1), 38–46. https://doi.org/10.1111/ajpy.12087
Tilwani, S. A., Amini MosaAbadi, F., Shafiee, S., & Azizi, Z. (2022). Effects of songs on implicit vocabulary learning: Spoken-form recognition, form-meaning connection, and collocation recognition of Iranian English as a foreign language learners. *Frontiers in Education, 7*, 797344. https://doi.org/10.3389/feduc.2022.797344
Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.
Wagner, E., Batty, A. O., & Galaczi, E. (2024). *The Routledge handbook of second language acquisition and listening*. Taylor & Francis. https://books.google.co.id/books?id=dcMOEQAAQBAJ
Wallace, J. M. (1998). *Teaching vocabulary*. Education Book.
Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition, 38*(2), 163–175. https://doi.org/10.3758/MC.38.2.163
Yang, Y., & Song, X. (2023). Multidecadal variation of the Earth's inner-core rotation. *Nature Geoscience, 16*(2),

182–187. https://doi.org/10.1038/s41561-022-01112-z
Zhang, P. (2022). How does repetition affect vocabulary learning through listening to the teacher's explicit instruction? The moderating role of listening proficiency and preexisting vocabulary knowledge. *Language Teaching Research.* https://doi.org/10.1177/13621688221140521