

A Multisensory Conceptual Model for Early Arabic Vocabulary Learning in Raudhatul Athfal Settings

Husnul Hafidhoh* Mubarak Fatahillah

Sekolah Tinggi Agama Islam Kuningan, Indonesia

Email: husnulhafidhoh@gmail.com¹* mubarakfatahillah@gmail.com²

Abstract

Early Arabic vocabulary learning in Raudhatul Athfal is often reduced to the rote repetition of word lists, although young children retain words more durably through meaningful, repeated, and multisensory experience. This article aims to formulate a conceptual instructional model for early Arabic vocabulary learning, with RA Raudlotul Hidayah as the intended implementation context. The study used an integrative literature review of 25 sources published between 2020 and 2025, identified through Scopus, ERIC, Google Scholar, and Garuda and screened against explicit inclusion and exclusion criteria. Each source was coded with five analytical questions on meaning, repetition, child response, media use, and observation, and the codes were synthesized thematically. The resulting model was appraised against four criteria: developmental appropriateness, linguistic accuracy, feasibility in short sessions, and observable learning evidence. The synthesis produced five instructional principles, namely concrete meaning, coordinated sensory input, movement based response, spaced repetition, and observational assessment, which were organized into the cyclical See–Hear–Move–Say–Use model. Children meet a concrete referent, hear an accurate Arabic model, respond through movement, attempt oral production, and use the word in a routine or game. The outputs also include an evidence to model matrix and a three level observation rubric. Digital media is positioned as a supporting resource, with safeguards for screen exposure, linguistic verification, and inclusion. Because no classroom data were collected, the model remains a proposal whose effectiveness must be tested through a pilot study.

Keywords: arabic vocabulary; early childhood; multisensory learning; raudhatul athfal; total physical response

INTRODUCTION

Vocabulary is the most visible entry point for introducing a new language to young children, but vocabulary learning should not be reduced to repeating lists. Children develop meaning through action, social interaction, objects, pictures, routines, and emotional engagement. Teaching therefore needs to connect each Arabic word with an experience that children can see, hear, touch, move toward, or use in a simple communicative situation. The problem is global in scope. Foreign language programmes for young children are expanding, and a systematic review of early childhood education and care found that such programmes can foster foreign language development without harming the first or majority language; however, the effects depend on input quantity, language policy, and teacher strategies, and children's wellbeing is protected only when programmes are play based and language policy is not too strict (Thieme et al., 2022).

Vocabulary research also shows that frequency alone is insufficient. A meta-analysis of 26 studies (N = 1,918) reported only a medium relationship between the number of encounters and incidental word learning ($r = .34$), with the effect moderated by the support available to learners (Uchihara et al., 2019). At the same time, digital exposure is rising faster than guidance can contain it: a global meta-analysis of 63 studies with 89,163 children found that only 24.7% of children younger than two years and 35.6% of children aged two to five years met screen time guidelines (McArthur et al., 2022). These data indicate that early vocabulary programmes

need deliberate design of repetition, sensory support, and media use rather than more exposure alone.

Young language learners have limited attention spans and rely strongly on contextual support. Waddington, (2023) explains that language teaching for children must connect linguistic input with meaningful activity. Thieme et al., (2022) similarly emphasize the importance of play, repetition, interaction, and carefully adjusted teacher language. Arabic learning in Raudhatul Athfal usually introduces greetings, classroom objects, colours, numbers, body parts, family members, animals, and simple Islamic expressions. The selection is appropriate when the vocabulary is concrete, useful, pronounceable, and connected to children's daily lives. An early childhood Arabic strategy should avoid overloading children with formal grammar or too many unrelated words in one meeting.

In Indonesia, Arabic is commonly introduced in Raudhatul Athfal and other Islamic early childhood and lower primary institutions, but the available evidence is mostly small scale and practice oriented. Case studies describe strategies that combine memorization, songs, games, and picture media (Marni & Rahmi, 2025; Zulfida et al., 2024), while teachers increasingly design their own digital media with tools such as Canva (Warmansyah et al., 2023). A community service programme at an Indonesian PAUD reported a 53% increase in children's recognition of Arabic vocabulary after play based cycles, yet it also recommended a broader vocabulary scope, longer duration, and parental collaboration, which indicates that gains depend on how activities are sequenced and sustained (Manik et al., 2024). What is still missing is a shared instructional structure that teachers can use to plan, sequence, and observe vocabulary learning across sessions, including at RA Raudlotul Hidayah, where Arabic vocabulary is part of the routine programme but no documented instructional model is yet available.

Multisensory learning provides a useful foundation because it coordinates visual, auditory, motor, and verbal experience. Dual coding theory, originally proposed by Allan Paivio, holds that information is easier to remember when verbal and nonverbal representations support one another, and Mayer, (2021) builds on this principle in the cognitive theory of multimedia learning. Multimedia learning also shows that pictures and words can support understanding when unnecessary information is controlled (Mayer, 2021). Experimental evidence with eight year old learners shows that foreign words learned with pictures or self performed gestures were remembered better for several months after learning than words learned without such enrichment (Andrä et al., 2020).

Movement is particularly relevant in early childhood. Total physical response, the method originally developed by James Asher, connects language input with bodily action and reduces pressure for immediate verbal production. Children can demonstrate comprehension by pointing, touching, moving, sorting, or acting before they are expected to say the word independently (Zulfida et al., 2024). A mixed studies review of embodied approaches in language education identifies movement, gesture, and drama as recurrent vehicles for vocabulary learning, particularly among beginners (Jusslin et al., 2022), and a classroom experiment with primary school children found that foreign vocabulary learned through task relevant physical activity outperformed a sedentary condition (Schmidt et al., 2019).

Digital resources can provide clear pictures, audio repetition, and short stories. However, young children still need adult mediation, social interaction, and physical play. Digital tools

should support the learning sequence instead of replacing teacher-child conversation or hands-on activity. Policy-oriented analysis warns that technology in early childhood must be evaluated in relation to quality, access, safety, and developmental appropriateness (OECD, 2023). For RA Raudlotul Hidayah, a conceptual model is useful because it can guide lesson preparation without pretending that classroom outcomes have already been measured. The model can later be tested through observation, vocabulary performance tasks, teacher reflection, and documentation of children's participation.

A common instructional problem is the confusion between repetition and meaningful learning. Repeating a word many times may strengthen sound familiarity, but repetition becomes more effective when the child encounters the word through different actions and contexts. The same word can be heard in a song, selected in a game, used in a routine, and recognized in a story. Variation protects children from mechanical chanting without comprehension. Teacher talk also requires careful design. Instructions should be short, supported by gesture, and delivered at a pace that allows processing.

The teacher does not need to translate every word immediately when an object or action makes meaning visible. At the same time, the teacher should avoid turning the session into continuous performance in which children only watch. Comprehensible input remains important, but young children require input that is connected to visible objects, actions, routines, and emotional security. The principle, originally formulated by Stephen Krashen in the Input Hypothesis, holds that acquisition is supported when learners understand messages slightly beyond their current competence, and four decades of subsequent research still assign communicatively embedded input a central role in acquisition (Lichtman & VanPatten, 2021).

Tasks should require children to do something meaningful with language rather than repeat isolated items without purpose. Task based language teaching organizes attention around meaning while creating opportunities for form to be noticed (Ellis et al., 2019). Previous studies have therefore examined parts of the problem separately rather than together. Systematic reviews describe the conditions under which early foreign language programmes work (Thieme et al., 2022), how digital technologies support young children's language and literacy (Liu et al., 2024), and how adult guidance during play supports learning (Skene et al., 2022), while Indonesian studies document local strategies for Arabic vocabulary in single settings (Marni & Rahmi, 2025; Zulfida et al., 2024).

Experimental work on gestures and pictures demonstrates memory benefits (Andrä et al., 2020; Schmidt et al., 2019), but it was conducted with primary school learners of European languages and does not specify how teachers should sequence activities, balance physical and digital media, or document progress without written tests. None of these strands offers an integrated and observable instructional cycle for Arabic vocabulary in Raudhatul Athfal; this is the research gap addressed here.

The gap is urgent for three reasons. Arabic vocabulary is already part of the daily programme of many Raudhatul Athfal institutions, so instructional decisions are being made now, often without a shared framework. Rising screen exposure among young children McArthur et al., (2022) increases the risk that ready made videos and slides replace interaction. Moreover, early childhood quality depends on intentional interaction and observation (Children, 2020; OECD, 2021), which teachers can deliver only with practical planning tools.

The novelty of this article does not lie in any single technique, since multisensory input, total physical response, play, and repetition are each well established. It lies in integrating them into one repeatable cycle in which every stage is explicitly derived from the reviewed evidence, carries a defined observable indicator, and is adapted to the constraints of Arabic instruction in Raudhatul Athfal, namely short sessions, verified pronunciation, low cost materials, and observation without written tests. The model is therefore a design proposal rather than a claim that one method is universally superior.

Accordingly, this article aims to formulate a conceptual instructional model for early Arabic vocabulary learning through an integrative literature review. It addresses three research questions: (1) how multisensory principles can be translated into a short and practical Arabic vocabulary sequence for young children; (2) how digital and physical media should be balanced within that sequence; and (3) how teachers can observe vocabulary learning without formal written tests.

The study offers theoretical and practical benefits. Theoretically, it connects comprehensible input, dual coding, embodied learning, and guided play within one instructional cycle for a language and age group that remain underexplored. Practically, it gives Raudhatul Athfal teachers, including those at RA Raudlotul Hidayah, a planning cycle, an evidence to model matrix, and an observation rubric that require no specialist equipment. For institutions and teacher educators, the model provides a basis for professional development and for a subsequent pilot study that can test it with actual classroom data.

METHOD

The article used an integrative literature review. This approach allows theoretical, empirical, and professional sources to be combined in order to produce an instructional model. It was chosen because the aim was to combine heterogeneous evidence into a new conceptual framework rather than to estimate a pooled effect, which is the main purpose of an integrative review (Snyder, 2019). The review focused on young language learners, vocabulary learning, multisensory representation, movement-based instruction, multimedia, digital storytelling, and early childhood assessment. The final analytical corpus comprised 25 sources published between 2020 and 2025, selected through the protocol described below.

Sources were identified in Scopus, ERIC, Google Scholar, and Garuda (the Indonesian national research index) for publications from January 2020 to June 2025 in English or Indonesian. Search strings combined three concept blocks with the Boolean operator AND: (a) young learners, early childhood, preschool, kindergarten, or Raudhatul Athfal; (b) vocabulary, Arabic, or foreign language; and (c) multisensory, total physical response, gesture, play, digital storytelling, multimedia, or assessment. Policy and position documents were retrieved directly from the OECD, UNESCO, and NAEYC repositories, and the reference lists of included reviews were searched by hand. The protocol and the eligibility criteria are summarized in Table 1.

Table 1. Review Protocol and Eligibility Criteria

Component	Specification
Review type	Integrative literature review aimed at building a conceptual instructional model (Snyder, 2019)

Component	Specification
Sources searched	Scopus, ERIC, Google Scholar, Garuda; OECD, UNESCO, and NAEYC repositories; backward citation searching
Period and language	January 2020 to June 2025; English and Indonesian
Search blocks	(a) young learners OR early childhood OR preschool OR kindergarten OR Raudhatul Athfal AND (b) vocabulary OR Arabic OR foreign language AND (c) multisensory OR total physical response OR gesture OR play OR digital storytelling OR multimedia OR assessment
Inclusion criteria	(1) Concerns children aged about 3 to 8 years or their teachers; (2) addresses vocabulary, language, or literacy learning, or media and assessment in early childhood education; (3) peer reviewed article, scholarly book, or official policy or position document; (4) full text accessible
Exclusion criteria	(1) Adult or secondary learners only; (2) no pedagogical implication for early language learning; (3) opinion pieces without scholarly grounding; (4) duplicates or inaccessible full texts
Screening	Title and abstract screening, full text eligibility check, and credibility appraisal by source type, conducted independently by both authors
Extraction	Structured form recording source type, context, age group, and answers to the five analytical questions

Screening proceeded in three stages. Titles and abstracts were first checked against the inclusion and exclusion criteria, the full texts of potentially relevant sources were then assessed for eligibility, and the remaining sources were appraised for credibility according to their type (empirical study, review, theoretical work, or policy document). Both authors screened independently and resolved disagreements through discussion. The procedure follows the transparency logic of PRISMA 2020, adapted to a non statistical integrative review (Page et al., 2021), and produced a final corpus of 25 sources (Figure 1; Table 2).

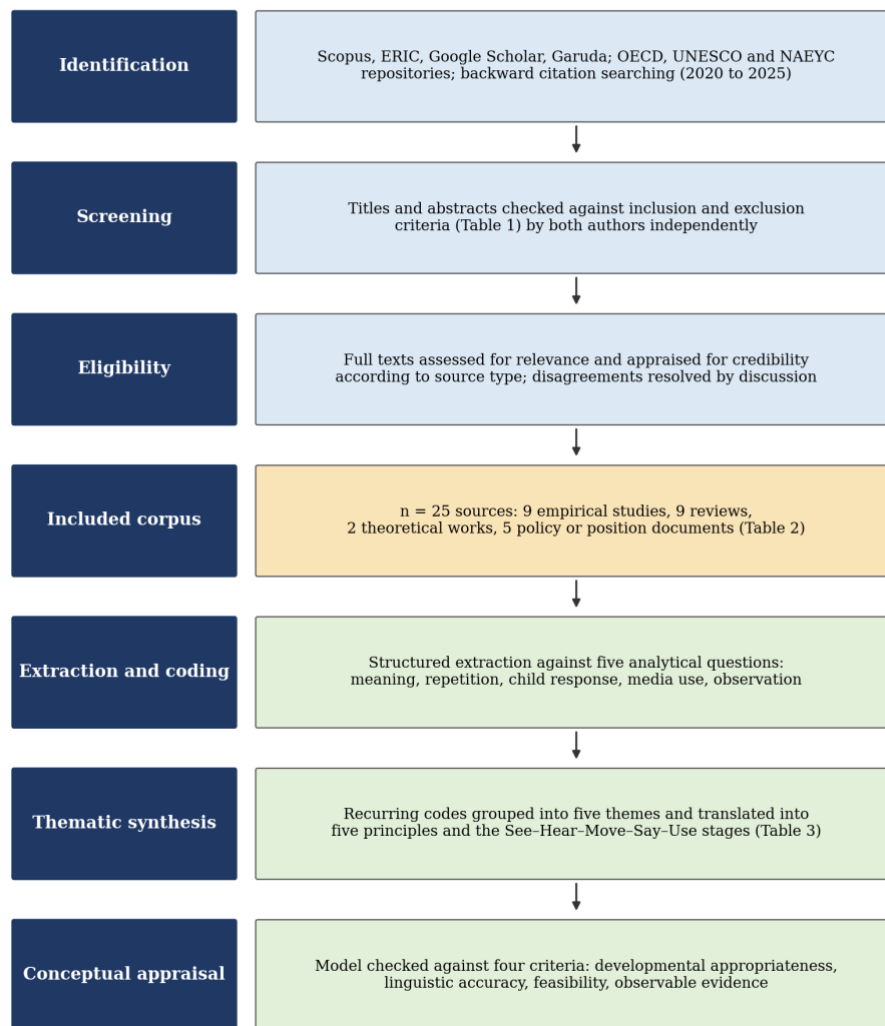


Figure 1. Source Identification, Screening, and Synthesis Procedure

Table 2. Composition of the Analytical Corpus by Source Type

Source type	n	Sources in the analytical corpus
Empirical studies	9	Hanifah et al., (2024); Manik et al., (2024); Marni & Rahmi, (2025); Misbahuddin & Setiyawan, (2023); Rifqiyah et al., (2025); Roche et al., (2024); Waddington, (2023); Warmansyah et al., (2023); Zulfida et al., (2024)
Systematic, narrative, and meta-analytic reviews	9	Astuti, (2025); Finders et al., (2023); Liu et al., (2024); Papavlasopoulou et al., (2026); Purnama et al., (2022); Ranta et al., (2023); Skene et al., (2022); Thieme et al., (2022); Torres et al.,(2021)
Theoretical and framework works	2	Falloon, (2020); Mayer, (2021)
Policy and position documents	5	Children, (2020); OECD, (2021); UNESCO, (2023)
Total	25	

The selected sources were analysed according to five questions: how meaning is introduced, how input is repeated, how children respond, how media is used, and how learning

is observed. Data were extracted with a structured form that recorded, for each source, its type, context, age group, and answer to each question. The extracted statements were coded inductively, and codes that recurred across at least two source types were grouped into themes. Each theme was then translated into an instructional principle, and the principles were ordered according to the developmental logic of moving from comprehension to production, which produced the five stages of the model (Table 3).

The final model was evaluated conceptually against four criteria: developmental appropriateness, linguistic accuracy, feasibility in short sessions, and the availability of observable learning evidence. Developmental appropriateness meant consistency with responsive, play based practice for children aged about four to six years; linguistic accuracy meant that every Arabic item and audio model can be verified by a competent speaker; feasibility meant that the cycle fits a ten to fifteen minute segment with low cost materials; and observability meant that every stage yields behaviour that a teacher can record while teaching. This study involved no human participants and collected no primary data. RA Raudlotul Hidayah is named only as the intended implementation context, and no claim is made that the model has produced measured effects there.

Foundational theories, namely comprehensible input, dual coding, total physical response, and task based learning, were not coded as corpus items. They were used, through recent authoritative syntheses that credit their originators, to interpret the coded themes (Andrä et al., 2020; Ellis et al., 2019; Jusslin et al., 2022; Lichtman & VanPatten, 2021). Benchmark data on vocabulary repetition and screen exposure McArthur et al., (2022); Uchihara et al., (2019) were likewise used only to frame the problem. The model was also examined for classroom practicality. Each stage had to fit within a short early childhood session, allow participation without requiring reading, and produce evidence that a teacher could observe while teaching. Complex testing procedures were excluded because they would be inconsistent with the proposed setting.

RESULT AND DISCUSSION

Results of the Literature Synthesis

The coding of the 25 sources produced five themes, each answering one analytical question. First, meaning is introduced most effectively through concrete referents such as objects, pictures, and actions that make the word meaning visible before production is expected. Second, input becomes useful when short and accurate models are repeated across varied activities and spaced across sessions rather than chanted in long lists. Third, children show comprehension first through action and play, and adult guidance is most productive when it is responsive and gradually reduced. Fourth, digital media supports language learning when it is embedded in intentional, adult mediated practice and checked for accuracy, whereas unmediated screen use adds little. Fifth, learning in early childhood is best evidenced through observable behaviour during interaction rather than written tests. Table 3 links each theme to its sources, the principle derived from it, and the corresponding stage of the model.

Table 3. Evidence to Model Matrix

Analytical question	Synthesized finding	Principle	Stage	Key corpus sources
How is meaning introduced?	Concrete referents anchor meaning before speech is required	Concrete meaning	See	Astuti, (2025); Marni & Rahmi, (2025); Mayer, (2021)
How is input repeated?	Short accurate models, varied and spaced across sessions	Coordinated sensory input; spaced repetition	Hear	Misbahuddin & Setiyawan, (2023); Purnama et al., (2022); Zulfida et al., (2024)
How do children respond?	Action and play show comprehension; guidance is responsive and faded	Movement based response	Move, Say	Roche et al., (2024); Skene et al., (2022); Torres et al., (2021); Waddington, (2023)
How is media used?	Media helps when embedded, mediated, and verified; screen time is limited	Digital media as support	Hear, Use	Liu et al., (2024); OECD, (2023); UNESCO, (2023); Warmansyah et al., (2023)
How is learning observed?	Behaviour during interaction is the main evidence; teacher competence matters	Observational assessment	All stages	Children, (2020); Finders et al., (2023); OECD, (2021); Ranta et al., (2023)

Model Development

The five principles were organized into the See–Hear–Move–Say–Use model. The sequence begins with comprehension and gradually moves toward oral production. It is intentionally cyclical rather than linear because children need to revisit the same vocabulary across several meetings and contexts (Figure 2).

See introduces a concrete referent through a real object, puppet, large picture, or classroom event. Hear provides a short and accurate Arabic model. Move allows children to show understanding through action. Say invites oral production without forcing every child to respond at the same time. Use places the word inside a greeting, game, sorting activity, song, or simple classroom routine.



Comprehension first, then production; the cycle is repeated with the same words in later meetings

Figure 2. The See–Hear–Move–Say–Use Instructional Cycle

Table 4. The See–Hear–Move–Say–Use Instructional Model

Stage	Teacher Action	Child Response	Evidence of Learning
See	Show one clear object, image, or action.	Attend, point, compare, or identify.	Visual attention and correct selection.
Hear	Model the Arabic word slowly and naturally.	Listen and discriminate the sound.	Response to the correct word.
Move	Give an action-based instruction.	Touch, bring, sort, imitate, or move.	Comprehension without verbal pressure.
Say	Invite individual or group production.	Repeat, complete, or name.	Approximate or accurate oral production.
Use	Place the word in a game, song, story, or routine.	Use the word for a simple purpose.	Spontaneous or prompted contextual use.

Discussion

Vocabulary selection determines the feasibility of the model. Each session should introduce a small set of related words, preferably three to five items. Zulfida et al., (2024) show that vocabulary learning requires repeated encounters and varied use. A larger list may create the appearance of coverage but reduce meaningful repetition. The teacher’s language model must be verified before instruction. Pictures should not be prepared before the meaning, script, pronunciation, and contextual use are checked. Audio can be recorded by a competent speaker and embedded in a presentation, but the teacher should still lead live interaction. Digital storytelling can extend the context when stories remain short and visually clear (Purnama et al., 2022).

A sample colour lesson can begin with three coloured objects. The teacher shows each object, models the Arabic word, and asks children to hold or move toward the corresponding colour. Children then repeat the words through a rhythmic chant and use them in a sorting game. The final routine may ask each child to choose one object while saying or completing the target word. Assessment should remain observational. Teachers can record whether a child attends to the model, selects the correct object, follows an action instruction, attempts the word, and uses it during play. These indicators are more suitable than written tests. The same checklist can be used across several sessions to distinguish temporary imitation from stable understanding.

Differentiation is necessary because children vary in language exposure, speech development, confidence, and sensory needs. A child who does not speak can still demonstrate comprehension through pointing or movement. A child who learns quickly can use the word in a short phrase. Inclusive design therefore provides more than one way to participate. Digital media should be limited to moments when it adds a function that physical materials cannot provide efficiently, such as replaying accurate audio, showing a short animated sequence, or documenting children's work. Teachers should avoid rapid visual transitions, crowded slides, and prolonged passive viewing. The educational goal is coordinated experience, not screen-based entertainment.

The model can be organized within a ten- to fifteen-minute focused sequence embedded in a longer play period. See and Hear may require two minutes, Move can occupy three to four minutes, Say can occur in brief turns, and Use can continue through a game or routine. The timing remains flexible because children's engagement and classroom transitions vary. Repetition should be planned across three dimensions: immediate repetition within the session, delayed repetition in later meetings, and contextual repetition in different activities. Zulfida et al., (2024) show that repeated encounters contribute to vocabulary learning when learners attend to meaning and use. A weekly plan can therefore recycle earlier words while introducing only a small number of new targets.

Teacher prompts should gradually reduce support. At first, the teacher may model the full word while showing the object. Later, the teacher can pause before the final sound, offer two choices, or ask a child to retrieve the object. This progression reflects scaffolding because support changes in response to the child's demonstrated understanding. The model can also support family participation without transferring the burden of formal teaching to parents. A simple home card may show one picture, provide verified audio through a link, and suggest a one-minute routine. The activity should remain optional, playful, and free from performance pressure. Families should be informed that accurate modeling and positive interaction are more important than the number of repetitions.

Observation records can use three performance levels: not yet observed, emerging with support, and demonstrated independently. The teacher records one or two indicators per session to keep documentation manageable. Over time, the record can show whether a child only imitates the sound, understands the meaning, or uses the word in context. The model should not be applied rigidly to children with different communication profiles. Some children may respond through gaze, gesture, picture selection, or assisted communication. The central question is whether the child demonstrates meaningful participation, not whether every child produces identical speech at the same moment.

Potential implementation risks include excessive vocabulary, overstimulating media, inaccurate audio, and the assumption that enthusiastic participation equals learning. These risks can be mitigated through limited targets, language verification, simple visual design, repeated observation, and reflection after each thematic cycle. With respect to the first research question, the cycle differs from existing approaches in what it combines rather than in any single component. Total physical response emphasizes comprehension through action but gives limited guidance on production and contextual use; multimedia principles specify how pictures and words should be combined but not how children respond bodily (Mayer, 2021); play based pedagogy protects agency but leaves the sequencing of language input to teacher judgement (Roche et al., 2024; Waddington, 2023); and Indonesian Arabic studies report useful songs, games, pictures, and literacy play without a common sequence across sessions (Manik et al., 2024; Marni & Rahmi, 2025; Rifqiyah et al., 2025).

The See–Hear–Move–Say–Use cycle places these elements in a developmental order, from referent to input, action, production, and use, and attaches an observable indicator to each stage. Gesture and movement are included only when they represent word meaning, because physical and digital play is most valuable when it supports purposeful movement and social engagement (Torres et al., 2021), and embodied vocabulary gains are associated with task relevant rather than incidental movement (Schmidt et al., 2019). Low cost objects, reusable picture cards, and teacher recorded audio keep the cycle feasible for RA Raudlotul Hidayah without continuous internet access.

Regarding the second research question, the reviewed evidence converges on a mediated rather than screen centred role for digital media. Digital technologies support young children's language and literacy when they are embedded in intentional practices (Liu et al., 2024), digital storytelling works when stories are short and visually clear (Purnama et al., 2022), and audiovisual Arabic input for young children requires adult mediation, repetition, and clear links to interaction (Misbahuddin & Setiyawan, 2023). Simple design tools can help teachers produce materials, but those materials still require checks for linguistic accuracy and instructional fit (Warmansyah et al., 2023).

Policy sources add that technology must be judged by quality, inclusion, and human centred purpose rather than novelty (OECD, 2023; UNESCO, 2023), which is especially relevant because most young children already exceed screen time guidelines (McArthur et al., 2022). In the model, therefore, media appears mainly in the Hear stage, for verified audio, and in the Use stage, for brief stories, while See, Move, and Say rely on objects, bodies, and interaction. For the third research question, observation without written tests is consistent with the view that early childhood quality rests on meaningful interactions and responsive pedagogy (Children, (2020); OECD, (2021) and on language environments in which children participate actively (Finders et al., 2023).

The three level rubric in Table 5 operationalizes this view by recording, at each stage, whether a behaviour is not yet observed, emerging with support, or demonstrated independently. Adult guidance during play improves learning when it is responsive and gradually withdrawn (Skene et al., 2022), which justifies reducing prompts as children move from Hear to Use, and materials should invite use and respond to local needs rather than merely display vocabulary (Astuti, 2025).

Table 5. Observation Rubric for the See–Hear–Move–Say–Use Cycle

Stage	Not yet observed	Emerging with support	Demonstrated independently
See	Does not attend to the referent	Looks at or points to the referent after a prompt	Selects the correct referent among two or three options
Hear	No response to the spoken word	Responds to the word after repetition or gesture	Responds correctly to the word at its first model
Move	Does not follow the action instruction	Follows the instruction by imitating peers	Performs the correct action from the word alone
Say	No oral attempt	Approximates the word with a model or first sound cue	Names the item clearly without a model
Use	Word not used in the activity	Uses the word when prompted in a game or routine	Uses the word spontaneously in play or routine

Implementation depends on teacher competence. Digital competence involves pedagogical, ethical, and evaluative judgement rather than software operation alone (Falloon, 2020), and early childhood pedagogical competence integrates planning, interaction, inclusion, assessment, and reflection (Ranta et al., 2023). Teacher education studies show that such competence grows through guided design, practice, and reflection (Hanifah et al., 2024; Papavlasopoulou et al., 2024), and professional learning for digital education should be connected to pedagogy, collaboration, and school strategy (OECD, 2025).

For RA Raudlotul Hidayah, this implies pronunciation rehearsal, joint preparation of low cost materials, and peer review of recorded trial lessons before classroom use. These arguments establish the plausibility of the model, not its effectiveness. The corpus is limited to 25 sources, most Indonesian studies are small scale, and the experimental evidence on gestures and pictures comes from older children and European languages (Andrä et al., 2020; Schmidt et al., 2019). The model should therefore be treated as a conceptual framework that requires empirical validation before any claim about learning outcomes can be made.

CONCLUSION

The See–Hear–Move–Say–Use model provides a practical multisensory sequence for early Arabic vocabulary learning by integrating concrete meaning, accurate oral input, movement, oral production, contextual use, and observational assessment. The five-stage cycle can be implemented within ten to fifteen minutes, with digital media limited mainly to verified audio and brief stories, while objects, movement, and interaction remain central. Learning is assessed through a three-level behavioural rubric rather than written tests. For RA Raudlotul Hidayah, the model can serve as a practical framework for lesson planning and teacher development. However, the model remains a conceptual instructional proposal because no empirical classroom validation has been conducted. Future research should therefore test its effectiveness through classroom implementation, vocabulary assessment, observation, delayed recall, and teacher interviews to provide empirical evidence of its feasibility and effectiveness.

REFERENCES

- Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & Von Kriegstein, K. (2020). Learning Foreign Language Vocabulary With Gestures And Pictures Enhances Vocabulary Memory For Several Months Post-Learning In Eight-Year-Old School Children. *Educational Psychology Review*, 32(3), 815–850. <https://doi.org/10.1007/S10648-020-09527-Z>
- Astuti, W. (2025). A Systematic Literature Review On Language Development Strategies In Early Childhood Education: Effective Pedagogical Approaches. *Ejournal.Upi.Eduw Astuti, E Yafiemimbar Pendidikan*, 2025•*Ejournal.Upi.Edu*, 10(2), 126–138. <https://doi.org/10.17509/Mimbardik.V10i2.82960>
- Children, N. A. For The E. Of Y. (2020). *Developmentally Appropriate Practice: A Position Statement*. National Association For The Education Of Young Children. <https://www.naeyc.org/resources/position-statements/dap/principles>
- Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). *Task-Based Language Teaching: Theory And Practice*. Cambridge University Press. <https://doi.org/10.1017/9781108643689>
- Falloon, G. (2020). From Digital Literacy To Digital Competence: The Teacher Digital Competency (Tdc) Framework. *Educational Technology Research And Development*, 68(5), 2449–2472. <https://doi.org/10.1007/S11423-020-09767-4>
- Finders, J., Wilson, E., & Duncan, R. (2023). Early Childhood Education Language Environments: Considerations For Research And Practice. *Frontiers In Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1202819/full>
- Hanifah, U., Binti Adam, Z., Faizin, M., Miftakhul Jannah, I., & Hanafi, Y. (2024). Accelerating The Digitalisation Of Learning Post-Covid-19 Era To Improve The Pedagogical Competence Of Pre-Service Arabic Teachers. *Taylor & Francisu Hanifah, Z Binti Adam, M Faizin, I Miftakhul Jannah, Y Hanaficogent Education*, 2024•*Taylor & Francis*, 11(1). <https://doi.org/10.1080/2331186x.2024.2413241>
- Jusslin, S., Korpinen, K., Lilja, N., Martin, R., Lehtinen-Schnabel, J., & Anttila, E. (2022). Embodied Learning And Teaching Approaches In Language Education: A Mixed Studies Review. *Educational Research Review*, 37, 100480. <https://doi.org/10.1016/J.Edurev.2022.100480>
- Lichtman, K., & Vanpatten, B. (2021). Was Krashen Right? Forty Years Later. *Foreign Language Annals*, 54(2), 283–305. <https://doi.org/10.1111/Flan.12552>
- Liu, S., Reynolds, B. L., Thomas, N., & Soyoof, A. (2024). The Use Of Digital Technologies To Develop Young Children's Language And Literacy Skills: A Systematic Review. *Journals.Sagepub.Coms Liu, Bl Reynolds, N Thomas, A Soyoofsage Open*, 2024•*Journals.Sagepub.Com*, 14(1). <https://doi.org/10.1177/21582440241230850>
- Manik, W., Hasibuan, M. S. M., Karo, T. K., & Dongoran, I. (2024). Introduction Of Arabic Vocabulary To Early Childhood Through A Community Service Programme At Paud An-Nur Serba Jadi Iv, Sunggal District, Deli Serdang. *Al-Arkhabii: Jurnal Pengabdian Masyarakat*, 4(4), 45–53. https://doi.org/10.51590/Jpm_Assunnah.V4i4.748
- Marni, M., & Rahmi, R. (2025). Arabic Learning Strategy For Early Childhood In Raudhatul

- Athfal Harapan Bunda. *Tanwir Arabiyyah: Arabic As Foreign Language Journal*, 5(1), 65–78. <https://doi.org/10.31869/Aflj.V5i1.6671>
- Mayer, R. E. (2021). *Multimedia Learning (3rd Ed.)*. Cambridge University Press.
- Mcarthur, B. A., Volkova, V., Tomopoulos, S., & Madigan, S. (2022). Global Prevalence Of Meeting Screen Time Guidelines Among Children 5 Years And Younger: A Systematic Review And Meta-Analysis. *Jama Pediatrics*, 176(4), 373–383. <https://doi.org/10.1001/Jamapediatrics.2021.6386>
- Misbahuddin, M., & Setiyawan, A. (2023). Youtube Media On Arabic Language Acquisition For Four-Year-Old Children. *Asalibuna*, 7(2). <https://doi.org/10.30762/Asalibuna.V7i02.2393>
- Oecd. (2021). *Education Policy Outlook 2021: Shaping Responsive And Resilient Education In A Changing World*.
- Oecd. (2023). *Teachers And School Leaders As Lifelong Learners*. Oecd Publishing.
- Oecd. (2025). *Preparing Teachers For Digital Education: Continuing Professional Learning On Digital Skills And Pedagogies (Oecd Education Policy Perspectives No. 122)*. Oecd Publishing. <https://doi.org/10.1787/Af442d7a-En>
- Page, M. J., Mckenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S., ... Moher, D. (2021). The Prisma 2020 Statement: An Updated Guideline For Reporting Systematic Reviews. *Bmj.Commj Page, Je Mckenzie, Pm Bossuyt, I Boutron, Tc Hoffmann, Cd Mulrow, L Shamseerbmj, 2021•Bmj.Com, 372*. <https://doi.org/10.1136/Bmj.N71>
- Papavlasopoulou, S., Undheim, M., Meaney, T., & Esmaeeli, S. (2024). Early Childhood Pre-Service Teachers' Preparation For Using Technology With Children: A Systematic Literature Review. *European Journal Of Teacher Education*, 49(2). <https://doi.org/10.1080/02619768.2024.2341935>
- Papavlasopoulou, S., Undheim, M., Meaney, T., & Esmaeeli, S. (2026). Early Childhood Pre-Service Teachers' Preparation For Using Technology With Children: A Systematic Literature Review. *European Journal Of Teacher Education*, 49(2), 461–478. <https://doi.org/10.1080/02619768.2024.2341935>
- Purnama, S., Ulfah, M., Ramadani, L., Rahmatullah, B., & Ahmad, I. F. (2022). Digital Storytelling Trends In Early Childhood Education In Indonesia: A Systematic Literature Review. *Jpud - Jurnal Pendidikan Usia Dini*, 16(1), 17–31. <https://doi.org/10.21009/Jpud.161.02>
- Ranta, S., Kangas, J., Harju-Luukkainen, H., Ukkonen-Mikkola, T., Neitola, M., Kinosh, J., Sajaniemi, N., & Kuusisto, A. (2023). Teachers' Pedagogical Competence In Finnish Early Childhood Education—A Narrative Literature Review. *Mdpi.Coms Ranta, J Kangas, H Harju-Luukkainen, T Ukkonen-Mikkola, M Neitola, J Kinosheducation Sciences, 2023•Mdpi.Com, 13(8)*. <https://doi.org/10.3390/Educsci13080791>
- Rifqiyah, N., Arifin, I., & Sobri, A. Y. (2025). Enhancing Early Childhood Language Skills Through Literacy Play Activities: A Classroom Action Research Study. *Muallimun: Jurnal Kajian Pendidikan Dan Keguruan*, 5(2), 221–234. <https://doi.org/10.23971/Muallimun.V5i2.10783>

- Roche, A., Carroll, C., & O'malley, M. P. (2024). An Exploration Of Pre-School Teacher's Experiences Of Using Play-Based Approaches To Develop Language And Literacy Skills Within Their Classrooms Of Children Aged. *Journals.Sagepub.Coma Roche, C Carroll, Mp O'malleychild Language Teaching And Therapy*, 2024•*Journals.Sagepub.Com*, 40(3), 278–296. <https://doi.org/10.1177/02656590241276687>
- Schmidt, M., Benzing, V., Wallman-Jones, A., Mavilidi, M.-F., Lubans, D. R., & Paas, F. (2019). Embodied Learning In The Classroom: Effects On Primary School Children's Attention And Foreign Language Vocabulary Learning. *Psychology Of Sport And Exercise*, 43, 45–54. <https://doi.org/10.1016/j.psychsport.2018.12.017>
- Skene, K., O'farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., & Ramchandani, P. G. (2022). Can Guidance During Play Enhance Children's Learning And Development In Educational Contexts? A Systematic Review And Meta-Analysis. *Academic.Oup.Comk Skene, Cm O'farrelly, Em Byrne, N Kirby, Ec Stevens, Pg Ramchandanichild Development*, 2022•*Academic.Oup.Com*, 93(4), 1162–1180. <https://doi.org/10.1111/Cdev.13730>
- Snyder, H. (2019). Literature Review As A Research Methodology: An Overview And Guidelines. *Journal Of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Thieme, A. M. M. M., Hanekamp, K., Andringa, S., Verhagen, J., & Kuiken, F. (2022). The Effects Of Foreign Language Programmes In Early Childhood Education And Care: A Systematic Review. *Taylor & Francisammm Thieme, K Hanekamp, S Andringa, J Verhagen, F Kuikenlanguage, Culture And Curriculum*, 2022•*Taylor & Francis*, 35(3), 334–351. <https://doi.org/10.1080/07908318.2021.1984498>
- Torres, P. E., Ulrich, P. I. N., Cucuiat, V., Cukurova, M., Fercovic De La Presa, M., Luckin, R., Carr, A., Dylan, T., Durrant, A., Vines, J., & Lawson, S. (2021). A Systematic Review Of Physical-Digital Play Technology And Developmentally Relevant Child Behaviour. *International Journal Of Child-Computer Interaction*, 30, 100323. <https://doi.org/10.1016/j.ijcci.2021.100323>
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The Effects Of Repetition On Incidental Vocabulary Learning: A Meta-Analysis Of Correlational Studies. *Language Learning*, 69(3), 559–599. <https://doi.org/10.1111/Lang.12343>
- Unesco. (2023). *Global Education Monitoring Report 2023: Technology In Education—A Tool On Whose Terms?* Unesco. <https://doi.org/10.54676/Uzqv8501>
- Waddington, J. (2023). Early Foreign Language Education: Play As A Site For Child Agency. *Taylor & Francisj Waddingtonjournal Of Multilingual And Multicultural Development*, 2024•*Taylor & Francis*. <https://doi.org/10.1080/01434632.2023.2290095>
- Warmansyah, J., Negeri, U. I., Yunus, M., Nurlaila, B., Islam, U., Mahmud, N., Batusangkar, Y., Tinggi, S., Islam, A., Pariangan, A., & Fitri, A. D. (2023). The Utilization Of Canva Application In Supporting Early Childhood Arabic Language Learning: Teachers'perceptions. *Pdfs.Semanticscholar.Orgj Warmansyah, N Nurlaila, A Mudinillah, Ad Fitrilughawiyah: Journal Of Arabic Education And Linguistics*, 2023•*Pdfs.Semanticscholar.Org*, 5(2), 141–154. <https://doi.org/10.31958/Lughawiyah.V5i2.10781>
- Zulfida, S., Rafli, Z., Murtadho, F., Nabighoh, M. I.-A., & 2024, Undefined. (2024). Arabic

Vocabulary Learning Strategies In Early Childhood: A Case Study At An Integrated Islamic Elementary School. *Researchgate.Nets* Zulfida, Z Rafli, F Murtadho, Ms Islaman Nabighoh, 2024•*Researchgate.Net*, 26(2), 269–286. <https://doi.org/10.32332/An-Nabighoh.V26i2.269-286>