EL-IBTIKAR
Jurnal Pendidikan Bahasa Arab
Vol. 15 No. 1, June 2026, pp.1-14
https://journal1.uinssc.ac.id/index.php/ibtikar/article/view/23929/version/32922
DOI: https://doi.org/10.24235/d5vkzs71

Embodied Pathways in Early Arabic Phoneme Acquisition: A Phenomenological Study in a Non-Formal Religious Context

Achmad Fudhaili[1]*, Muhammad Ivang Rizaldi[2], Alif Alvian Mawaddi Hamid[3]
[1,2,3]Universitas Islam Negeri Syarif Hidayatullah Jakarta

[*1]Penulis Korespondensi: fudhaili@uinjkt.ac.id

ABSTRACT (Bold, TNR 18)

This study investigates the acquisition of Arabic phonemes in early childhood within a non-formal religious learning context, emphasising phonological development as an embodied, developmental, and socially mediated process. Moving beyond segment-centred and outcome-oriented perspectives, the study employs a qualitative phenomenological approach to explore young learners' lived experiences as they perceive, articulate, and internalise Arabic phonemes that are typologically distant from their first language. The research was conducted in a rumah Qur'an in Jakarta, Indonesia, involving three children aged four to six at different instructional levels and one experienced instructor. Data were collected through non-participant observation, semi-structured interviews, and document analysis, and were analysed inductively using phenomenological thematic procedures. The findings indicate that early Arabic phoneme acquisition begins with holistic sensorimotor experiences rather than abstract phonological categories, with phonemes initially understood as bodily articulatory actions. Consistent substitution patterns emerge as adaptive developmental strategies shaped by first-language phonology. Repetitive Qur'anic recitation supports the proceduralization of embodied articulatory routines, while teacher-mediated multimodal scaffolding facilitates phonemic differentiation through perceptual and motor reorganisation. The study concludes that early Arabic phonological development is best explained by an integrative model that combines embodied cognition, Natural Phonology, the Speech Learning Model, and sociocultural theory, thereby highlighting the pedagogical value of religious education.

Keywords: *Arabic phoneme acquisition, early childhood, embodied phonology, non-formal education, sociocultural mediation.*

ABSTRACT (Bold, TNR 10)

Penelitian ini mengkaji pemerolehan fonem bahasa Arab pada anak usia dini dalam konteks pembelajaran keagamaan nonformal, dengan memandang perkembangan fonologis sebagai proses yang berwujud secara tubuh, berkembang bertahap, dan dimediasi secara sosial. Berbeda dari pendekatan fonologi yang berorientasi pada unit segmental dan hasil akhir, penelitian ini menggunakan kerangka kualitatif fenomenologis untuk menelaah pengalaman hidup peserta didik ketika mempersepsi, mengartikulasikan, dan menginternalisasi fonem bahasa Arab yang berjarak tipologis dari bahasa pertama mereka. Penelitian dilakukan di sebuah rumah Qur'an di Jakarta dengan melibatkan tiga anak berusia empat sampai enam tahun pada tingkat pembelajaran berbeda serta satu pengajar berpengalaman. Data dikumpulkan melalui observasi nonpartisipan, wawancara semi terstruktur, dan analisis dokumen, kemudian dianalisis secara induktif dengan prosedur tematik fenomenologis. Hasil penelitian menunjukkan bahwa pemerolehan fonem bahasa Arab pada tahap awal tidak dimulai dari kategori fonologis abstrak, melainkan dari pengalaman sensorimotorik holistik, di mana fonem dimaknai sebagai tindakan artikulatoris berbasis sensasi tubuh. Pola substitusi fonologis mencerminkan strategi adaptif perkembangan yang dipengaruhi fonologi bahasa pertama. Repetisi tilawah Al Qur'an membentuk rutinitas artikulatoris prosedural, sedangkan scaffolding multimodal guru memfasilitasi diferensiasi fonemik melalui reorganisasi perseptual dan motorik. Temuan ini menegaskan bahwa fonologi bahasa Arab anak usia dini lebih tepat dipahami melalui model integratif kognisi berwujud pada pembelajaran nonformal.

Kata kunci: *anak usia dini, fonologi embodied, mediasi sosiokultural, pemerolehan fonem Arab, pendidikan non-formal*

ABSTRAK (Bold, TNR 10)

تتناول هذه الدراسة اكتساب الفونيمات العربية لدى الأطفال في مرحلة الطفولة المبكرة ضمن سياق التعلّم الديني غير النظامي، بوصف التطوّر الفونولوجي عمليةً متجسّدة، نمائية، ومشفوعة بالتفاعل الاجتماعي. وتنطلق الدراسة من نقد المقاربات الفونولوجية التي تركّز على الوحدات التجزيئية والنتائج النهائية، معتمدةً إطارًا كيفيًا ظاهراتيًا يهدف إلى فهم الخبرات المعيشة للمتعلمين الصغار أثناء إدراكهم للفونيمات العربية ونطقها واستدخالها، ولا سيما تلك البعيدة صوتيًا عن لغتهم الأم. أُجريت الدراسة في أحد بيوت القرآن بمدينة جاكرتا، وشارك فيها ثلاثة أطفال تتراوح أعمارهم بين أربع وست سنوات في مستويات تعليمية متباينة، إضافةً إلى معلّم ذي خبرة. جُمعت البيانات من خلال الملاحظة غير المشاركة، والمقابلات شبه المنظّمة، وتحليل الوثائق، ثم خضعت لتحليل استقرائي وفق الإجراءات الظاهراتية الموضوعاتية. وتُظهر النتائج أن اكتساب الفونيمات العربية في مراحله الأولى لا يبدأ بتشكّل فئات فونولوجية مجرّدة، بل ينطلق من خبرات حسية حركية شمولية يُدرك فيها الصوت بوصفه فعلًا نطقيًا مرتبطًا بإحساس الجسد. كما تعكس ظواهر الإبدال الصوتي استراتيجيات نمائية تكيفية متأثرة ببنية اللغة الأولى، وتسهم التلاوة المتكرّرة للقرآن في ترسيخ الروتينات النطقية الإجرائية، بينما يتيح الدعم التعليمي متعدد الوسائط تمايزًا فونيميًا عبر إعادة تنظيم إدراكي وحركي. وتخلص الدراسة إلى أن النمو الفونولوجي المبكر للغة العربية يُفهم على نحو أدق ضمن نموذج تكاملي يجمع بين الإدراك المتجسّد، والفونولوجيا الطبيعية، ونموذج تعلّم الكلام، والنظرية السوسيوق ثقافية، مؤكدةً القيمة النظرية والتربوية للتعلّم الديني غير النظامي.

الكلمات الرئيسية: الطفولة المبكرة؛ الفونولوجيا المتجسِّدة؛ الوساطة السوسيوق-ثقافية؛ اكتساب الفونيمات العربية؛ التعليم الديني غير النظامي.

Received: April 19, 2026	*Revised:May 10, 2026*	*Accepted:May 15, 2026*	*Published:June 4, 2026*
date	*date*	*date*	*Date*

Citation (APA Style): Fudhaili, A. (2026). Embodied pathways in early Arabic phoneme acquisition: A phenomenological study in a non-formal religious context. *EL-Ibtikar: Jurnal Pendidikan Bahasa Arab*, Vol. 15 No. 1, June 2026, pp.1-14.

INTRODUCTION

Early phonological acquisition is widely recognised as a foundational component of linguistic competence, playing a critical role in shaping later oral proficiency, lexical development, and literacy skills. During early childhood, typically defined as the period from birth to six years of age, heightened neurocognitive plasticity allows children to rapidly form and reorganise phonological representations through intensive perceptual and articulatory experience. At this stage, repeated exposure to speech sounds, combined with active motor engagement in articulation, enables children to establish increasingly stable sound categories. Consequently, early phonological development is particularly sensitive to the quality, frequency, and meaningfulness of phonetic input. Empirical research consistently demonstrates that early phonological awareness and processing abilities are strongly associated with later language outcomes, including lexical access, decoding efficiency, and reading comprehension, underscoring the long-term impact of early phonological experience on language and literacy development (Bin Sawad et al., 2022; Levlin & Waldmann, 2020).

In the context of second language (L2) acquisition, phonological development is further shaped by the dynamic interaction between the learner's first language (L1) phonological system and the phonemic inventory of the target language. Rather than occurring in isolation, L2 phoneme learning unfolds against the backdrop of pre-existing perceptual and articulatory routines established through the L1. A substantial body of research over the past decade has shown that cross-linguistic phonological distance plays a decisive role in determining both perceptual sensitivity and production accuracy in L2 learners. When L2 phonemes lack direct counterparts in the L1 inventory, learners tend to perceptually assimilate unfamiliar sounds into the closest available L1 categories. This process often results in systematic substitution patterns, reduced sensitivity to contrastive distinctions, and delayed phonological stabilisation (Flege & Bohn, 2021). Importantly, such outcomes are not random errors but reflect principled developmental strategies shaped by the constraints and affordances of the learner's phonological system.

Arabic presents a particularly complex set of challenges in this respect. Its phonemic system includes emphatic consonants, interdental fricatives, and guttural or pharyngeal sounds that are absent from many languages, including Indonesian. For Indonesian learners, especially

children, these contrasts introduce articulatory and perceptual demands that go beyond routine phonological differentiation. Previous research has documented persistent difficulties with interdental phonemes such as ث، ذ and ظ, as well as with pharyngeal or velar sounds such as ع and خ, which are frequently substituted with more familiar alveolar or vowel-like sounds (Fudhaili et al., 2023). For young learners engaged in Qur'anic reading instruction, these difficulties often emerge early and may become entrenched through repeated recitation practices. Given the central role of accurate phonological realisation in Qur'anic recitation, phonological challenges in Arabic carry not only linguistic consequences but also pedagogical, cultural, and sociocultural significance.

Despite increasing scholarly attention to Arabic phonological acquisition, the field remains marked by notable methodological and theoretical imbalances. A large proportion of recent studies have adopted quantitative or computational approaches, focusing on mispronunciation detection, acoustic feature analysis, and automated pronunciation assessment using machine-learning techniques. While these studies contribute valuable technological tools for the diagnosis of pronunciation, they often conceptualise phonological acquisition as a measurable outcome rather than a developmental process grounded in perception, embodiment, and social interaction. As a result, the lived experience of learners, particularly young children, remains largely invisible within such paradigms.

Parallel lines of inquiry employing qualitative approaches have predominantly focused on adult or university-level learners. These studies typically emphasise error patterns, L1 interference, and remedial instructional strategies, offering important descriptive insights into recurrent pronunciation difficulties (Fudhaili et al., 2023). However, their explanatory power for understanding early phonological development is limited. Adult learners possess greater metalinguistic awareness and cognitive control, whereas early childhood learning is deeply sensorimotor, experience-dependent, and socially mediated. Consequently, findings derived from adult populations cannot be straightforwardly generalised to young learners, whose phonological systems are still in the process of initial construction.

Crucially, research examining Arabic phonological acquisition among young children in naturalistic and non-formal learning environments remains scarce. Existing studies rarely investigate how children subjectively experience unfamiliar phonemes, how bodily sensations and articulatory metaphors contribute to sound production, or how teachers scaffold phonological learning through multimodal interaction. This gap is particularly striking in contexts such as Indonesia, where early Arabic instruction is predominantly provided in community-based Qur'anic institutions rather than in formal schooling. These settings constitute the primary sites of exposure for many children, yet they remain underrepresented in mainstream applied linguistics research.

Non-formal Qur'anic learning environments typically employ distinctive pedagogical practices, including physical modelling of articulation, explicit attention to *makhraj*, embodied metaphors, and intensive oral repetition. Such practices align closely with contemporary theories of embodied cognition, which posit that cognitive representations are grounded in bodily action and sensorimotor experience rather than in abstract, amodal symbols (Barsalou, 2008). From this perspective, phonological learning is not merely a matter of auditory discrimination but involves the gradual coordination of perception, motor control, and bodily awareness. Nonetheless, these embodied pedagogical practices are rarely foregrounded in the international literature on phonological acquisition, which continues to privilege laboratory-based, instrumentally driven, or computational paradigms.

To address this gap, the present study adopts a qualitative phenomenological approach to investigate how Arabic phonemes are acquired by early childhood learners within a non-formal religious educational context. Rather than treating phonetic accuracy as the primary endpoint of learning, the study foregrounds children's lived experiences of perceiving, producing, and internalising Arabic phonemes, as well as teachers' pedagogical interpretations and scaffolding strategies. Rumah Qur'an Al-'Izzah Jagakarsa was selected as the research site due to its structured, level-based phonological curriculum, which provides a developmentally transparent context for observing phoneme acquisition trajectories across stages of early learning.

The study aims to achieve two central objectives. First, it seeks to elucidate the processes through which young learners gradually construct Arabic phonological representations, focusing on how perceptual sensitivity, articulatory control, and bodily awareness interact during development. Second, it aims to identify the linguistic, articulatory, and pedagogical challenges encountered in this process, particularly in relation to phonological distance between Indonesian and Arabic. By integrating insights from phonological theory, embodied cognition, and sociocultural perspectives, this research advances a process-oriented understanding of Arabic phonological acquisition and offers empirically grounded implications for early Arabic phonology instruction in non-formal educational settings.

Despite the explanatory contributions of dominant phonological acquisition models, most existing frameworks continue to conceptualise phoneme learning primarily as the gradual refinement of abstract segmental categories. Such approaches offer limited insight into how phonological representations are initially constructed in early childhood, especially when phonological distance is high, and metalinguistic awareness remains minimal. In particular, both Natural Phonology and the Speech Learning Model provide only partial accounts of how embodied experience, bodily sensation, and articulatory metaphor function as representational resources in the earliest phases of phonological development (Flege & Bohn, 2021). By foregrounding children's lived phonological experience within a non-formal learning ecology, the present study addresses this theoretical blind spot and proposes an integrative account in which embodiment and sociocultural mediation function not as peripheral supports but as constitutive mechanisms of phonological acquisition (Barsalou, 2008).

METHOD

This study employed a qualitative phenomenological approach to explore how early-childhood learners experience the acquisition of Arabic phonemes in a naturalistic educational context. Phenomenology was selected as the research framework because the study did not aim to measure phonetic accuracy or performance outcomes, but rather to understand how children perceive, articulate, and internalise Arabic phonemes that differ fundamentally from their first-language phonological system. By foregrounding lived experience, this approach enables an in-depth examination of phonological acquisition as a developmental, embodied, and socially mediated process (van Manen, 2016; Finlay, 2011).

The research was conducted at Rumah Qur'an Al-'Izzah Jagakarsa, a non-formal Qur'anic learning institution in Jakarta, Indonesia. The site was purposively selected due to its structured, level-based phonological instruction (jilid), which emphasises explicit articulation (makhraj) training. This pedagogical organisation provided a rich and systematic context for observing the gradual acquisition of Arabic phonemes across stages of early learning, aligning with qualitative research principles that prioritise information-rich settings for in-depth analysis (Patton, 2015; Creswell & Poth, 2018).

Participants were selected through purposive sampling to ensure alignment with the research objectives. The participant group comprised three early-childhood learners aged between four and six years, each representing a different instructional level (jilid 1, jilid 3, and jilid 4). These levels corresponded to phases of initial exposure, phoneme consolidation, and phonological stabilisation. In addition, one senior instructor with more than five years of experience teaching Arabic phonology at the institution participated in the study. Including learners at different instructional stages enabled a comparative understanding of developmental trajectories, while the instructor provided pedagogical insight into recurring learner difficulties and instructional decision-making. This sampling strategy is consistent with qualitative inquiry, which emphasises depth, variation, and developmental contrast rather than statistical generalisation (Patton, 2015).

Data were collected through methodological triangulation to enhance credibility and analytic depth. First, non-participant classroom observations documented natural instructional interactions, children's spontaneous articulatory attempts, and teachers' corrective feedback without researcher intervention. Observations focused on how phonemes were introduced, modelled, and practised, as well as learners' responses to articulatory cues and embodied

instructions. Second, semi-structured interviews were conducted with both learners and the instructor. Interviews with children used age-appropriate prompts to elicit descriptions of their phoneme-learning experiences, while interviews with the instructor explored pedagogical rationales, recurring challenges, and instructional strategies. Third, relevant instructional documents and learning materials were analysed to contextualise observed practices and support data triangulation. The combined use of these data sources reflects established qualitative research practices aimed at strengthening interpretive validity (Creswell & Poth, 2018; Denzin, 2012).

Data analysis followed an inductive and iterative phenomenological process. Interview recordings were transcribed verbatim, and observational notes were systematically organised. Initial open coding identified meaning units related to phoneme perception, articulatory strategies, repetition, and instructional scaffolding. These codes were subsequently clustered into broader thematic categories that captured recurring experiential patterns across participants. Through thematic interpretation, the analysis sought to distil the essential structures of how Arabic phoneme acquisition was experienced by young learners within the instructional context. Throughout the analytic process, interpretations were continuously refined through recursive engagement with the data and relevant theoretical perspectives, consistent with phenomenological and thematic analytic traditions (van Manen, 2016; Braun & Clarke, 2021; Kvale & Brinkmann, 2015).

To ensure trustworthiness, multiple strategies were employed. Credibility was enhanced through data triangulation and prolonged engagement in the research setting. Member checking was conducted with the instructor to verify the plausibility and accuracy of emerging interpretations. Dependability and confirmability were supported by maintaining an audit trail documenting research procedures, analytic decisions, and reflective notes throughout the study. These strategies align with widely accepted criteria for rigour in qualitative research (Lincoln & Guba, 1985).

Ethical considerations were addressed prior to data collection. Institutional approval was obtained, and written parental consent was secured for all child participants. Anonymity was ensured through the use of pseudonyms, and all data were handled confidentially in order to protect participant privacy and uphold ethical integrity in research involving children (Alderson & Morrow, 2020).

RESULTS AND DISCUSSION

Results

The findings of this study indicate that Arabic phoneme acquisition in early childhood unfolds as a gradual, developmental, and experientially grounded process rather than as the immediate mastery of abstract phonological categories. At the early stages of instruction, children did not approach Arabic phonemes as discrete segmental units defined by contrastive features. Instead, phonemes were experienced as holistic sound events that intertwined auditory perception with bodily sensation and articulatory effort. This orientation was consistently observable in learners' oral explanations, gestures, and interactional behaviour during instructional activities, suggesting that early phonological representations are grounded in lived experience rather than analytic phonetic awareness.

A prominent pattern observed across all participants was the pervasive use of bodily metaphors to describe Arabic phonemes, particularly those perceived as difficult or unfamiliar. Sounds involving constriction in the velar or pharyngeal regions were frequently characterised through experiential imagery rather than articulatory terminology. For instance, the voiceless velar fricative /x/ (khāʾ) was described as sounding "like snoring," while the pharyngeal consonant /ʿ/ (ʿain) was associated with sensations such as "pressing the throat." These metaphors were not incidental or idiosyncratic but recurred across learners and sessions, indicating a stable mode of phonological meaning-making rooted in sensorimotor experience. Rather than conceptualising phonemes as abstract units, learners appeared to understand them as bodily actions and felt events within the vocal tract.

Embodied representations of Arabic phonemes were not limited to the initial phase of learning but continued to shape phonological development over time. Learners frequently

reported that pronunciation became easier through repeated Qur'anic recitation, often articulating this improvement in experiential terms such as "it feels lighter," "it becomes smoother," or "my throat gets used to it." These expressions point to an incremental transition from effortful, consciously monitored articulation toward greater articulatory stability and fluency. Importantly, this developmental change was not accompanied by explicit verbalisation of phonological rules or features. Instead, learning was reflected in reduced physical strain, increased rhythmic coordination, and smoother integration of phonemes into continuous recitation, suggesting that phonological development proceeded primarily through proceduralising rather than conscious analysis.

Another robust finding concerned systematic patterns of phonological substitution. Learners consistently replaced Arabic phonemes that lack equivalents in Indonesian with more familiar sounds from their first language. Interdental fricatives such as /θ/, /ð/, and /ẓ/ were commonly realised as alveolar sounds like /s/, /z/, or /d/, while guttural phonemes—particularly /ʿ/—were often weakened, realised as vowel-like sounds, or omitted in rapid speech. These substitutions occurred with high regularity across learners and instructional sessions, demonstrating patterned developmental tendencies rather than random mispronunciations. The stability of these patterns suggests that early phonological systems prioritise articulatory accessibility and rely on well-established motor routines while gradually accommodating new articulatory demands.

In several instances, learners demonstrated an emerging, experience-based awareness of first-language influence on pronunciation. Children occasionally explained their difficulties by referencing Indonesian phonology, such as remarking that certain Arabic sounds were challenging because "in Indonesian, there is only /s/." Although such comments did not reflect formal metalinguistic knowledge, they indicate an intuitive sensitivity to cross-linguistic differences and an awareness that phonological difficulty is not arbitrary but linked to structural differences between languages. This emerging awareness suggests that young learners actively negotiate phonological distance in experiential rather than analytic terms.

As learners progressed to higher instructional levels, the frequency of substitution gradually declined, accompanied by greater articulatory consistency and stability. Improvements were especially evident following instructional episodes in which teachers provided explicit articulatory guidance, such as demonstrating tongue placement, modelling throat constriction, or physically gesturing the direction of airflow. Notably, these improvements manifested less in terms of categorical phonemic accuracy and more in the form of smoother transitions, reduced hesitation, and increased confidence during recitation. This pattern indicates that instructional effectiveness was mediated through bodily alignment and motor awareness rather than explicit rule acquisition.

The role of teacher mediation emerged as a central factor in phonological development. Teachers did not simply provide auditory models of target phonemes but also actively scaffolded learners' articulatory efforts through multimodal strategies, including physical demonstrations, gestures, repetition, and corrective feedback embedded in meaningful practice. Such mediation enabled learners to perform articulatory actions that exceeded their independent capabilities, gradually internalising these actions through repetition and guided adjustment. Phonological development thus appeared to be inseparable from the social and instructional context in which learning took place.

It is important to emphasise that this study does not aim to generate statistically generalizable claims about Arabic phoneme acquisition. Instead, its objective is to achieve analytic depth in understanding how phonological learning is experienced, embodied, and socially mediated in early childhood. From a phenomenological perspective, a limited number of participants constitutes a methodological strength rather than a limitation, as it allows for close, sustained engagement with individual learners' developmental trajectories. The value of the findings lies not in numerical generalisation but in their conceptual transferability to comparable learning contexts characterised by early age, substantial phonological distance, and non-formal instructional settings.

The convergence of embodied experience, systematic simplification, repetition, and instructional mediation observed in this study underscores the multidimensional nature of early Arabic phoneme acquisition. Phonological development emerges not as a linear progression toward abstract accuracy but as a dynamic process rooted in bodily sensation, shaped by prior linguistic experience, and guided through socially mediated practice. These findings highlight the importance of understanding early phonological learning as a lived, developmental phenomenon and provide a nuanced foundation for rethinking both theory and pedagogy in early second-language phonology.

This matrix integrates observational data with interview-based meaning construction, illustrating how children's phonological substitutions function as developmentally motivated, embodied strategies rather than phonological "errors.".

Table 1. Matrix of Observed Phonological Substitutions and Child-Reported Experiences

No	Target Arabic Phoneme	Observed Substitution	Learning Context	Verbatim Child Experience (Interview)	Interpretive Analysis
	/ʿ/ (ʿayn)	Weakened, vowel-like, or omitted	Qur'anic recitation (slow & rapid speech)	"It feels like pressing here (pointing to the throat). Sometimes I don't feel it."	Guttural phoneme interpreted as bodily effort, leading to weakening when sensation is absent
	/x/ (khāʾ)	/h/	Repetition in tilāwah	"It sounds like snoring, but if I make it strong, my throat gets tired."	Substitution as articulatory economy strategy, not random error
	/θ/ (thāʾ)	/s/	Low-frequency lexical items	"In Indonesian, we only have /s/."	L1 phonological assimilation (Natural Phonology)
	/ð/ (dhāl)	/z/	Qur'anic word reading	"It's almost the same sound."	Incomplete perceptual contrast due to L1 inventory
	/ẓ/ (ẓāʾ)	/d/ or /z/	Connected recitation	"It's hard to make it thick."	Difficulty distinguishing interdental + emphasis features

Table 2. Interview Themes and Pedagogical Mediation

No	Experiential Theme	Child Verbatim Expression	Observed Teacher Response	Pedagogical Interpretation
	Bodily sensation	"It presses here in the throat."	Gestural modelling of makhraj	Phoneme learning grounded in embodied awareness
	Articulatory fatigue	"If it's too strong, I get tired."	Allowance for gradual production	Developmentally sensitive scaffolding
	L1 awareness	"This sound doesn't exist in Indonesian."	Contrastive modelling	Early intuitive cross-linguistic awareness
	Repetition	"After reciting often, it feels lighter."	Rhythmic repetition	Proceduralisation of articulatory routines

Excerpt 1: Substitution of /ʕ/

Teacher: "Try reading ʿalīm slowly."
Child A (6 years old): "a-līm."
Teacher: "This sound comes from here (touching the throat)."
Child A: "I don't feel it there. I just open my mouth."
Teacher: "That's okay. We'll try again slowly later."

The child's explanation demonstrates that phonological categories are initially grounded in felt bodily sensation, not abstract phonemic awareness. The absence of throat sensation results in substitution, confirming the article's embodied-phonology argument.

Excerpt 2: Substitution of /χ/

Teacher: "Now say khalaqa."
Child B (7 years old): "ha-la-qa."
Teacher: "Why not kha?"
Child B: "Because kha hurts my throat."

This exchange reveals substitution as an adaptive bodily strategy aimed at minimising articulatory strain, rather than an indication of phonological deficit.

Excerpt 3: Spatial Perception of /q/

Teacher: "This word is qalb."
Child C (6 years old): "kalb."
Teacher: "This one is from the back of the tongue."
Child C: "Yes, it's heavy… like being pushed backwards."

The child conceptualises the place of articulation through spatial metaphors, reinforcing the finding that early phonological awareness is constructed through embodied spatial schemas.

Excerpt 4: Repetition as Motor Routine

Teacher: "Let's repeat together."

Child D: "It's easier in a group."
Teacher: "Why?"
Child D: "My mouth just moves by itself."

Repetition is experienced as motor routinization, aligning with the study's claim that phonological mastery emerges through embodied automatisation rather than conscious rule learning.

The findings challenge traditional accuracy-driven models of Arabic phonology instruction and instead support a developmentally informed, multimodal pedagogical approach in which substitution is understood as an intelligible phase of phonological growth.

Discussion

The findings of this study provide strong empirical support for the view that early Arabic phoneme acquisition is fundamentally an embodied, developmental, and socially mediated process, rather than the immediate internalisation of abstract phonological categories. In the early stages, children did not conceptualise phonemes as symbolic units organised within a contrastive phonological system. Instead, phonemes were experienced as integrated auditory–motor events, consistent with developmental phonology research showing that young children initially encode speech sounds holistically before analytic segmentation becomes accessible (Vihman, 1996; Jusczyk, 1997).

Recent neurocognitive research supports this developmental pattern, showing that early speech perception and production rely heavily on sensorimotor neural networks, whereas abstract phonological processing emerges more gradually with maturation and experience (Kuhl, 2004). Learners' reliance on bodily metaphors such as "snoring" or "pressing the throat" indicates that phonological representations at this stage are grounded in felt bodily experience rather than in awareness of acoustic or articulatory features. This challenges traditional models of phonological acquisition that assume early access to segmental representations and underscores the need to rethink phoneme development from an embodied perspective.

From the standpoint of embodied cognition, these findings are particularly significant. Embodied theories hold that cognitive representations are grounded in bodily action, perception, and simulation rather than in amodal, abstract symbols. In the present study, Arabic phonemes were initially conceptualised not as linguistic objects but as actions of the vocal tract, experienced through tension, airflow, and resonance. This aligns with motor theories of speech perception, which argue that speech perception involves mapping acoustic input onto articulatory gestures. The data, therefore, challenge approaches to pronunciation instruction that focus exclusively on auditory discrimination, suggesting that bodily engagement is central to early phonological learning.

Crucially, embodied phonological representations did not disappear as learners progressed. Rather, they became increasingly proceduralised through repetition. Learners' accounts that pronunciation improved simply "because of reciting often" indicate a transition from conscious bodily awareness to automatic motor execution. This pattern aligns with contemporary research emphasising the role of frequency and repeated practice in phonological stabilisation. Studies over the past five years have shown that, for young learners, input frequency and meaningful repetition contribute more substantially to phoneme stabilisation than explicit correction alone. Repetition thus functions not merely as a pedagogical technique but as a core cognitive mechanism through which sensorimotor experience is internalised into stable articulatory routines (Bybee, 2021).

The systematic substitution patterns observed in this study further support a developmental interpretation of early phoneme acquisition. Rather than random errors, these substitutions reflect adaptive simplification strategies, as predicted by Natural Phonology. Learners prioritised articulatory ease, drawing on well-established motor patterns from their first language and gradually restructuring them in response to sustained exposure. The consistency of substitution patterns across learners indicates that such behaviours are governed by universal tendencies linked to articulatory constraints and phonological economy.

Notably, similar substitution patterns have been documented not only among young learners but also among early bilinguals and adult learners acquiring typologically distant languages such as Arabic (Saiegh-Haddad & Everatt, 2021; Al-Mahrooqi & Denman, 2022). This convergence has important theoretical implications. It suggests that an early age does not automatically eliminate phonological constraints. Instead, the advantage of early learning lies in providing an extended window during which phonological restructuring can occur gradually. Early exposure offers more time for motor patterns to be reshaped, rather than guaranteeing immediate accuracy.

This interpretation aligns with the Speech Learning Model (SLM), which posits that learners initially assimilate unfamiliar L2 sounds into existing L1 phonological categories and gradually establish new phonetic categories through sustained perceptual input and production experience. Recent reformulations and empirical tests of the SLM have reaffirmed that category restructuring is a slow, usage-dependent process shaped by articulatory experience rather than by explicit phonological knowledge alone. In the present study, this restructuring became evident at higher instructional levels, as substitution frequency decreased and articulatory consistency improved. Importantly, this developmental shift did not appear to result from enhanced metalinguistic awareness, but from articulatory guidance that compelled learners to attend to previously ignored motor distinctions.

Teacher instructions emphasising tongue placement or throat constriction served as triggers for perceptual reanalysis, making phonetic contrasts salient in bodily, actionable terms. Recent experimental research supports this interpretation, showing that directing learners' attention to articulatory motor cues significantly accelerates the formation of new phonetic categories, particularly among children and early learners (Kartushina et al., 2022). Neuro-behavioural studies further suggest that such motor-based instruction strengthens perception–action coupling, leading to more robust and durable phonemic representations. Collectively, these findings extend the Speech Learning Model by highlighting that, in early childhood, phonemic differentiation emerges primarily through embodied and motoric pathways rather than through abstract contrast awareness.

The social dimension of phonological development was equally central to the findings. In this study, phoneme acquisition did not occur in isolation but was embedded within pedagogically structured interaction. Teachers acted as mediators, scaffolding learners' articulatory development through multimodal strategies, including gestures, physical demonstrations, and concrete metaphors. Through such scaffolding, learners were able to perform articulatory actions that exceeded their independent capacities, a mechanism captured by the sociocultural notion of mediated development and the Zone of Proximal Development (Lantolf et al., 2020; Poehner & Infante, 2020).

Recent work in sociocultural SLA emphasises that multimodal scaffolding is particularly effective for children's phonological development, as it links bodily experience and linguistic form more effectively than abstract metalinguistic explanation (Lantolf et al., 2020). In this study, teachers did not simply provide auditory models of pronunciation but actively assisted learners in constructing new articulatory routines by making otherwise invisible motor processes visible and experientially accessible.

The non-formal rumah Qur'an context significantly amplified these mechanisms. Integrating phoneme learning into meaningful religious practice ensured high motivation and sustained engagement, while the absence of high-stakes evaluation reduced performance anxiety. Research in applied linguistics increasingly recognises the importance of affective and identity-related factors in pronunciation development, particularly when the target language carries symbolic or religious value (Buschfeld & Kautzsch, 2022). The present findings illustrate how a supportive learning ecology encourages phonological risk-taking, repetition, and gradual refinement.

Beyond segmental accuracy and substitution patterns, the data also reveal an important yet underexplored dimension of early Arabic phonological acquisition: the role of orthographic–phonological mapping in shaping articulatory awareness. Although the learners in this study were at a pre-literate or emergent-literacy stage, exposure to Arabic script during Qur'anic reading

instruction appeared to interact dynamically with phoneme acquisition. Teachers' frequent references to *makhraj* alongside visual letter forms enabled children to associate articulatory gestures with orthographic symbols, suggesting that phonological learning was supported not only by auditory and motor experience but also by early multimodal symbol grounding. Recent research on Arabic literacy development indicates that even minimal exposure to visually stable script forms can facilitate phonological consolidation when instruction integrates articulation and orthography, particularly in consonant-rich orthographies such as Arabic (Saiegh-Haddad & Maldoon, 2023).

Another analytical dimension, not previously articulated, concerns the acquisition of suprasegmental features, particularly rhythmic timing and syllabic weight, which emerged implicitly through Qur'anic recitation. Although the present study focused on phoneme-level development, observational data indicated that repeated engagement with patterned recitation fostered sensitivity to syllabic structure and prosodic timing. Children often adjusted articulatory effort in response to rhythmic cues rather than to explicit segmental correction. Recent phonological research argues that early sensitivity to rhythm and timing facilitates the stabilisation of segmental production, especially in languages with complex consonantal structures.

The analysis further suggests that early Arabic phoneme acquisition is characterised by high intra-learner variability, which should be interpreted as a developmental resource rather than instability. Individual learners frequently produced multiple phonetic variants of the same target phoneme across instructional sessions, sometimes alternating between substituted and more target-like articulations. Contemporary usage-based and exemplar-theoretic research emphasises that such variability reflects learners' active exploration of phonetic space and is often a precursor to category stabilisation (Bybee, 2021). Recent longitudinal studies demonstrate that in early childhood L2 acquisition, phonological variability predicts later accuracy more robustly than early consistency, particularly when learners receive rich and repetitive input (Zheng & Gor, 2025).

An additional finding concerns the multimodal orchestration of teacher scaffolding beyond articulation alone. Teachers systematically combined vocal demonstration with gesture, spatial metaphors, and body positioning, creating what recent scholarship terms multimodal phonological input. This orchestration appeared to externalise internal articulatory processes, rendering them perceptually salient and cognitively accessible to young learners. Recent empirical studies in child SLA confirm that gesture-supported pronunciation instruction yields more durable phonological gains than auditory-only modelling, particularly for low-salience contrasts such as guttural and laryngeal sounds (Kartushina et al., 2022). The present study extends this line of research by showing that multimodality is not merely facilitative but constitutive of phonological learning in early childhood.

Finally, the analysis reveals an important implication for phonological assessment and evaluation, which remains absent from most Arabic phonology research involving children. The findings suggest that early phonological development cannot be adequately assessed through accuracy-based metrics alone, as these fail to capture embodied effort, articulatory awareness, and developmental progression. Recent methodological discussions in applied linguistics argue for process-oriented and learner-centred assessment frameworks that account for variability, motor coordination, and experiential learning trajectories (Isaacs & Trofimovich, 2023). In contexts such as early Qur'anic education, where phonological learning is embedded in meaningful practice, assessment that relies solely on error counts may obscure genuine developmental gains.

Theoretically, this study advances research on phonological acquisition by foregrounding embodiment as a central mechanism in early L2 phonology. Methodologically, it demonstrates the value of phenomenologically informed qualitative analysis for accessing learners' lived phonological experience, which remains largely invisible to acoustic or error-based approaches. Pedagogically, the findings support instructional models that integrate bodily awareness, articulatory metaphor, sustained repetition, and scaffolded interaction, particularly in contexts characterised by substantial phonological distance.

Taken together, these additional analytical insights reinforce the study's central argument while extending its scope. Early Arabic phoneme acquisition emerges not only as an embodied and socially mediated process, but also as a multimodal, prosodically scaffolded, and variably explored developmental system. By integrating orthographic exposure, rhythmic practice, learner variability, multimodal instruction, and process-oriented assessment into the analytic frame, the study offers a more comprehensive account of how Arabic phonology is learned in early childhood. These dimensions remain largely absent from dominant phonological acquisition models, underscoring the contribution of the present study to advancing theory, methodology, and pedagogy in early L2 phonological research.

CONCLUSION

This study concludes that Arabic phoneme acquisition in early childhood is best understood as a gradual, embodied, and socially mediated developmental process, rather than as an instance of immediate phonological mastery. Young learners initially build phonological knowledge through sensorimotor experience; perceiving phonemes as holistic articulatory events grounded in bodily sensation rather than as abstract segmental units. Embodied representations thus form the foundational layer of early phonological development.

The findings further demonstrate that repetition within meaningful practice is a primary mechanism for phonological stabilisation, enabling the transition from consciously experienced articulatory effort to automated motor routines. Systematic phonological substitution emerges as an adaptive developmental strategy, reflecting universal simplification processes and first-language influence rather than an individual deficit. Early exposure does not eliminate structural phonological constraints; instead, it provides a longer temporal window for gradual restructuring.

Phonemic differentiation was shown to emerge through motoric and embodied cues, particularly when teachers scaffolded articulatory distinctions. This extends the Speech Learning Model by highlighting the primacy of production-based and embodied pathways in early phonetic category formation. Finally, phonological development was found to be inseparable from social mediation and learning ecology, with the non-formal rumah Qur'an context providing affective security, cultural meaning, and sustained engagement.

Overall, the study advances an integrative model of early second-language phonological acquisition, in which embodied cognition, natural phonological constraints, usage-based learning, and sociocultural mediation dynamically interact. This holistic perspective challenges segment-centric models of phonology and foregrounds the lived, bodily, and social dimensions of early language learning.

DAFTAR PUSTAKA

Alderson, P., & Morrow, V. (2020). The ethics of research with children and young people (2nd ed.). Sage.

Al-Mahrooqi, A., & Denman, C. (2022). Arabic phonological acquisition by non-native speakers: Persistent substitution patterns and pedagogical implications. Language Awareness, 31(4), 421–438. https://doi.org/10.1080/09658416.2022.2042856

Antoniou, M., Best, C. T., & Tyler, M. (2021). Explicit attention to articulation facilitates the learning of novel speech sounds. Journal of Speech, Language, and Hearing Research, 64(6), 1920–1936. https://doi.org/10.1044/2021_JSLHR-20-00601

Asif, A., Hussain, B., & Zafar, S. (2021). Automatic Arabic pronunciation error detection using deep learning. Computer Speech & Language, 67, 101184. https://doi.org/10.1016/j.csl.2020.101184

Barsalou, L. W. (2008). Grounded cognition. Annual Review of Psychology, 59, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bin Sawad, M., Gilakjani, A. P., & Obeidat, M. (2022). Phonological awareness and its role in

early literacy acquisition. Reading and Writing, 35(6), 1451–1472. https://doi.org/10.1007/s11145-021-10215-9

Braun, V., & Clarke, V. (2021). Thematic analysis: A practical guide. Sage.

Buschfeld, S., & Kautzsch, A. (2022). Affective and identity-related dimensions in pronunciation learning. Applied Linguistics Review, 13(3), 463–487. https://doi.org/10.1515/applirev-2019-0107

Bybee, J. (2021). Exemplar dynamics and phonological representation. Cognitive Linguistics, 32(1), 1–27. https://doi.org/10.1515/cog-2020-0068

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

Denzin, N. K. (2012). Triangulation 2.0. Journal of Mixed Methods Research, 6(2), 80–88. https://doi.org/10.1177/1558689812437186

Donegan, P., & Stampe, D. (2009). Hypotheses of Natural Phonology. Journal of Phonetics, 37(1), 1–15. https://doi.org/10.1016/j.wocn.2008.08.002

Fanoush, A., Alsharif, M., & Al-Kabi, M. (2025). Machine-learning approaches to Arabic phoneme classification and mispronunciation diagnosis. Speech Communication, 157, 68–82.

Finlay, L. (2011). Phenomenology for therapists: Researching the lived world. Wiley-Blackwell.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience (pp. 233–277). York Press.

Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). Second Language Speech Learning, 1, 3–83.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. Journal of Phonetics, 14(1), 3–28.

Fudhaili, A., Sumardi, S., & Santoso, A. (2023). L1 phonological interference in Arabic pronunciation among Indonesian university learners. Indonesian Journal of Applied Linguistics, 12(3), 489–502. https://doi.org/10.17509/ijal.v12i3.49773

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. Psychonomic Bulletin & Review, 13(3), 361–377. https://doi.org/10.3758/BF03193857

Harwood, J., Thomas, A., & Summers, L. (2017). Neuroplasticity and early phonological development. Journal of Child Language, 44(3), 567–589. https://doi.org/10.1017/S0305000916000160

Hornberger, N. H., & McKay, S. L. (2010). Linguistic human rights and educational policy in non-formal contexts. Annual Review of Applied Linguistics, 30, 112–132. https://doi.org/10.1017/S0267190510000062

Isaacs, T., & Trofimovich, P. (2023). Reconceptualising pronunciation assessment from a developmental perspective. Applied Linguistics, 44(3), 455–478. https://doi.org/10.1093/applin/amac060

Jusczyk, P. W. (1997). The discovery of spoken language. MIT Press.

Kartushina, N., Martin, C. D., & Domahs, U. (2022). Motor involvement in phonetic category learning in children. Developmental Science, 25(2), e13163. https://doi.org/10.1111/desc.13163

Kuhl, P. K. (2004). Early language acquisition: Phonetic and word learning, neural substrates, and a theoretical framework. Annual Review of Neuroscience, 27, 553–582. https://doi.org/10.1146/annurev.neuro.27.070203.144211

Kuhl, P. K., Tsao, F.-M., & Chen, H. J. (2020). Developmental shifts in neural mechanisms of speech perception. Proceedings of the National Academy of Sciences, 117(48), 30218–30225. https://doi.org/10.1073/pnas.2009136117

Kupisch, T. (2018). Cross-linguistic influence in bilingual phonology. Applied Linguistics Review, 9(1), 1–23. https://doi.org/10.1515/applirev-2016-1010

Kvale, S., & Brinkmann, S. (2015). InterViews: Learning the craft of qualitative research interviewing (3rd ed.). Sage.

Lakoff, G., & Johnson, M. (1999). Philosophy in the flesh: The embodied mind and its challenge

to Western thought. Basic Books.
Lantolf, J. P., Poehner, M. E., & Swain, M. (2020). Sociocultural theory and second language development in children. Language Teaching, 53(4), 409–429. https://doi.org/10.1017/S0261444820000218
Levlin, M., & Waldmann, C. (2020). Early phonological processing and later reading comprehension. Journal of Research in Reading, 43(2), 215–232. https://doi.org/10.1111/1467-9817.12305
Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Sage.
van Manen, M. (2016). Researching lived experience: Human science for an action-sensitive pedagogy (2nd ed.). Routledge.
Patton, M. Q. (2015). Qualitative research & evaluation methods (4th ed.). Sage.
Poehner, M. E., & Infante, P. (2019). Mediated development and phonological learning in early L2 instruction. Applied Linguistics, 41(4), 547–572. https://doi.org/10.1093/applin/amz017
Prieto, P., & Vigário, M. (2021). Rhythm and phonological development: Evidence from early language learning. Laboratory Phonology, 12(1), 1–25. https://doi.org/10.5334/labphon.261
Pulvermüller, F. (2022). Neurobiological mechanisms for embodied language learning. Trends in Cognitive Sciences, 26(12), 1029–1043. https://doi.org/10.1016/j.tics.2022.09.002
Saiegh-Haddad, E., & Everatt, J. (2021). The role of phonological distance in bilingual and second language phonology. Applied Psycholinguistics, 42(5), 1025–1053. https://doi.org/10.1017/S0142716421000245
Saiegh-Haddad, E., & Maldoon, F. (2023). Orthographic support and phonological development in Arabic-speaking children. Reading and Writing, 36(4), 921–945. https://doi.org/10.1007/s11145-022-10366-3
Saiegh-Haddad, E., & Lina, G. (2018). Phonological processing of Arabic diglossia and literacy development. Reading and Writing, 31(9), 2051–2075. https://doi.org/10.1007/s11145-018-9876-3
Stampe, D. (1973). A dissertation on Natural Phonology. University of Chicago.
Tomasello, M., & Daum, M. (2020). The role of usage frequency in early language development. Cognitive Linguistics, 31(4), 543–566. https://doi.org/10.1515/cog-2019-0068
Vihman, M. M. (1996). Phonological development: The origins of language in the child. Blackwell.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
Zheng, Y., & Gor, K. (2024). Phonological distance and L2 speech learning across the lifespan. Second Language Research, 40(1), 3–27. https://doi.org/10.1177/02676583231127432
Zheng, Y., & Gor, K. (2025). Variability as a catalyst in L2 phonological development across childhood. Second Language Research, 41(2), 189–213. https://doi.org/10.1177/02676583231151344