

A Phonological Model of Arabic Speech Acquisition in Malay Children (Model Fonologi Pemerolehan Bahasa Arab dalam Kalangan Kanak-kanak Melayu)

MUHAMMAD RIDHA ALI HUDDIN^{1*}, SHAHIDI A HAMID², MD SHAFIIN SHUKOR³
& SITI MARYAM AJ JALILAH ROSLAN¹

¹ Faculty of Languages and Linguistics, Universiti Malaya,
50603 Lembah Pantai, Kuala Lumpur, Malaysia

² Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor, Malaysia

³ UKM Pakarunding Sdn. Bhd., Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor, Malaysia

Received: 27 November 2025 / Accepted: 9 February 2026

ABSTRACT

The acquisition of Arabic sounds by second language learners is crucial in language learning. This study examines the acquisition of selected Arabic consonants among Malay primary school children, particularly focusing on the most difficult Arabic sounds, thus, addressing critical knowledge gaps in existing research. Prior studies predominantly focus on separate issues either in sound perception or production, yet in the context of English, neglecting the integration of both language acquisition processes. Unlike previous research, this study proposes a reciprocal model which interplay of both Arabic sound perception and production. Using perception and production tests involving 30 children, the study reveals a mismatch between sound perception and articulation accuracy. The results demonstrated patterns of substitution, transfer, and developmental errors, providing insight into the interaction between learners' L1 influence and their developing L2 phonological competence. Findings also reveal that sounds without counterparts in Malay, such as emphatic consonants and uvulars, are the most problematic, often leading to systematic substitution with phonetically proximate sounds. Perceptual difficulties were found to correlate strongly with production errors, highlighting the importance of auditory discrimination in successful phonological acquisition. Drawing on the Perceptual Assimilation Model (PAM), Speech Learning Model (SLM), and Source-Filter Theory (SFT), the proposed Arabic Speech Phonological Model (M-ASP) integrates mind consciousness and airflow control to explain and enhance Arabic pronunciation in Malay children. The study underscores the need for integrative teaching approaches that address both the auditory and articulatory dimensions of Arabic phonology, thereby supporting more effective and sustainable learning outcomes for L2 learners of Arabic. The model extends existing phonological theories by embedding cognitive and physiological processes into a recursive speech chain. The findings offer a theoretically grounded, data-driven framework that supports Arabic pronunciation instruction in multilingual and Islamic educational contexts.

Keywords: Arabic phonemes; phonology; speech perception; perception-production model; Malay learners

ABSTRAK

Pemerolehan bunyi bahasa Arab oleh pelajar bahasa kedua adalah penting dalam pembelajaran bahasa. Kajian ini meneliti pemerolehan konsonan Arab terpilih dalam kalangan kanak-kanak sekolah rendah Melayu, dengan penekanan khusus pada bunyi Arab yang paling sukar, sekaligus menangani jurang pengetahuan kritikal dalam kajian sedia ada. Kajian-kajian terdahulu kebanyakannya menumpukan kepada isu berasingan sama ada dalam persepsi bunyi atau pengeluaran bunyi, terutamanya dalam konteks bahasa Inggeris, tanpa menekankan integrasi kedua-dua proses pemerolehan bahasa. Berbeza dengan kajian terdahulu, kajian ini mencadangkan model timbal balik yang menggabungkan interaksi antara persepsi bunyi Arab dan pengeluarannya. Menggunakan ujian persepsi dan pengeluaran melibatkan 30 kanak-kanak, kajian ini mendedahkan ketidaksesuaian antara persepsi bunyi dan ketepatan artikulasi. Hasil kajian menunjukkan corak penggantian

* Corresponding author: mardha67@um.edu.my

(substitution), pemindahan (transfer), dan kesilapan perkembangan, memberikan pemahaman mengenai interaksi antara pengaruh L1 pelajar dan kompetensi fonologi L2 mereka yang sedang berkembang. Kajian ini juga menunjukkan bahawa bunyi yang tiada padanan dalam bahasa Melayu, seperti konsonan tegas (emphatic) dan uvular, merupakan yang paling bermasalah, selalunya membawa kepada penggantian sistematik dengan bunyi yang hampir secara fonetik. Kesukaran persepsi didapati berkorelasi kuat dengan kesilapan pengeluaran, menekankan kepentingan diskriminasi auditori dalam pemerolehan fonologi yang berjaya. Berdasarkan Model Perceptual Assimilation (PAM), Model Pembelajaran Ucapan (SLM), dan Teori Sumber-Filtrasi (SFT), Model Fonologi Ucapan Arab yang dicadangkan (M-ASP) menggabungkan kesedaran minda dan kawalan aliran udara untuk menerangkan dan meningkatkan sebutan bahasa Arab dalam kalangan kanak-kanak Melayu. Kajian ini menekankan keperluan pendekatan pengajaran integratif yang merangkumi kedua-dua dimensi auditori dan artikulatori fonologi Arab, seterusnya menyokong hasil pembelajaran yang lebih efektif dan lestari bagi pelajar L2 bahasa Arab. Model ini memperluas teori fonologi sedia ada dengan memasukkan proses kognitif dan fisiologi ke dalam rangkaian ucapan berulang. Hasil kajian menyediakan kerangka berasaskan teori dan data yang menyokong pengajaran sebutan bahasa Arab dalam konteks pendidikan berbilang bahasa dan pendidikan Islam.

Kata Kunci: Fonem Arab; fonologi; persepsi ucapan; model persepsi-penghasilan; pelajar Melayu

INTRODUCTION

Speech perception can be understood as a complex cognitive process intricately linked to brain development and literacy acquisition (Heald & Nusbaum 2014). To frame this development, this study draws upon Piaget's Cognitive Developmental Theory, which postulates that children construct knowledge through environmental interaction and communication. This occurs through adaptation process, where new information is either integrated into existing mental structures (assimilation) or forces those structures to modify (accommodation). These structures or schemas then form the foundation for organizing knowledge, including language.

Complementing this, Frith's Theory of Reading Acquisition (1986) provides a specific model for literacy, outlining three progressive stages: the logographic (recognizing symbols), the alphabetic (linking letters to sounds), and the orthographic (internalizing spelling patterns for automatic reading). Central to navigating these stages is phonological processing which is the ability to use sound knowledge to process language (Nagaraja et al., 2025). This capability is crucial for speech perception (Défossez et al., 2023), and difficulties here are a primary factor in perception challenges. Phonological processing comprises three key skills: phonological awareness (manipulating sounds in words), phonological memory (storing and retrieving sound sequences), and rapid automatic naming (efficiently recoding symbols into sounds) (Wagner & Torgesen, 1987). Research confirms these skills are distinct from general cognitive ability and are vital for early literacy.

Among these, phonological awareness is particularly critical as a metalinguistic skill, requiring conscious reflection on language structure. It involves manipulating sounds at the syllable, rhyme, and phoneme levels. Effective speech perception, enabled by these skills, is fundamental for interpreting spoken language, understanding context, and shaping communicative dynamics. The relationship between speech perception and production is complex; while often seen as interdependent "sides of the same coin," evidence also suggests they can sometimes operate independently (Baese-Berk et al. 2024; McGettigan & Tremblay, 2018; Arjmandi & Roozbeh Behroozmand, 2024).

This study investigates this relationship in the context of second language acquisition, specifically focusing on Malay-speaking primary school children learning Arabic sounds. Despite regular exposure to Arabic through daily worship, it is unclear if their perception of these sounds aligns with their production proficiency. This research aims to explore the role of sound perception in this process.

LITERATURE REVIEW

Speech production is a complex process that involves sound articulation, grammar formulation, and conveying meaning. It is also a highly intricate sensorimotor task that requires careful coordination and processing across multiple areas of the brain (Kearney & Guenther, 2019). Difficulty in pronouncing L2 or foreign language sounds usually occurs with phoneme sounds that are different from those of the native language (Flege, 1995; Gilakjani & Ahmadi, 2011; Kartushina et al., 2016). Previous studies have also stated that errors occur with these sounds when pronounced with letter sounds that are almost the same in the native language (Jabar, 2004; Kamaruzzaman & Nik Rahimi 2000). Typically, the forms of sound errors that occur are the conversion of phoneme sounds to phoneme sounds that are almost the same or even sounds that are not the same as L1 or L2. However, different forms of sound errors may exist among Malay speakers. Previous studies have shown that there are three forms of pronunciation errors in Arabic sounds made by Malay children, such as changes in letter morphemes, changes in letter properties, and changes in letter morphemes and letter properties (Kamaruzzaman & Nik Rahimi 2000). The characteristics of sound pronunciation describe the properties of the sound and the way the phoneme sound is pronounced, whether with nasality, thickness, inclination, aspiration, inhalation, exhalation, etc.

However, Hamid (2010) found that the same phonemes between L1 and L2 also cause pronunciation difficulties for speakers. This situation can be seen in one example of phoneme sounds in Arabic that are also found in Malay, such as the sound /h/. This phoneme sound seems easy because it exists in the Malay language sound system. However, in reality, it can cause difficulties for Kelantan Malay speakers because the pronunciation of Kelantan Malay speakers tends to inhale air (inhalation) when producing this phoneme sound (Huddin et al., 2017), thus producing the sound /h/.

Therefore, the importance of sound perception among Malay children is not just knowledge in discriminating the sounds heard, rather when the existing L2 phonological knowledge is combined with cognitive awareness (Nasrullah, 2025), it will also affect the level of sound production ability. Returning to the studies earlier, most of them do not touch on the forms and characteristics that underlie the problem of pronunciation difficulties among Malay children. Previous studies on Arabic phonetics only show the letters that are difficult for students and Malay Muslims to pronounce (Azidan 2004; Issraq 2004; Issraq 2017; Nik Rahimi et al. 2010), rather than unraveling and justifying the phonetic characteristics that are the cause of the problem phenomenon. Therefore, studies on the characteristics and properties of Arabic sounds produced by Malay children in reciting the Quran need to be conducted in more depth so that the influence of sound perception and other factors can be seen more clearly.

In particular, many studies conducted to overcome the pronunciation difficulties of Arabic letter sounds among Malays have found that the pronunciation difficulties of Arabic sounds among Malays usually occur with sounds that are not in their mother tongue. Pronunciation errors for pharyngealized sounds such as [s^h], [d^h], [t^h], and [ð^h] are common errors made by the Malay Muslim community. Radzi & Hamid (2022) found that Malays experience pronunciation problems, especially with the letters /ħ/, /ʕ/, /q/, /t^h/, /d^h/, /x/ and /s^h/. The findings of the study clearly show that most of the problems occur with pharyngealized consonants (Binasfour & Setter 2018) which are also found in Surah al-Fātiḥah.

Studies on Arabic pharyngealized and non-pharyngealized sounds have been conducted by several researchers (Radzi, 2019, 2020; Alwabari 2013; Binasfour & Setter 2018; Nik Rahimi et al. 2010), which focus on the similarities and differences of phonemes between L1 and Arabic as L2 (Azidan 2004; Kamaruzzaman & Nik Rahimi 2000). Based on previous studies conducted on Arabic pharyngealization (Azidan 2004; Nik Rahimi et al. 2010; Majdan & Shahidi 2022; Shehata 2024), it was found that native Malay speakers are weak in mastering Arabic pharyngealized sounds. However, the researchers found that pharyngealized sounds are not the main factor that makes it difficult for non-native Arabic speakers. On the contrary, this study highlights the level of ability of non-native speakers, namely Malay children, from the aspect of perception and sound production together. Previous studies have focused more on the aspect of sound production alone without investigating the aspect of sound perception (see Issraq, 2004, 2017; Radzi et al., 2015, 2019, 2020, 2022; Jabar, 2004; Nik Rahimi, 2010; Ramli et al., 2016).

To contribute more towards the studies on phonological speech perception and production, as well as to develop a phonological model for the most difficult produced Arabic sounds by Malay children, we have aimed this study to address the following research questions:

- 1) Which Arabic phonemes are the most difficult in phonological perception among Malay children?
- 2) To what extent does the phonological perception have a one-to-one correlation with the produced one?
- 3) How can these be addressed through a predictive Arabic phonological model?

METHODOLOGY

PARTICIPANTS

In this study, the sample size consists of 30 subjects aged between 10 and 12 years. In the context of phonetic production tests involving voice recordings, a sample of six respondents for each gender is adequate (Ladefoged 2003). Based on several acoustic studies conducted, Hamid (2010) selected 20 subjects for the perception test and 24 subjects for the sound production test. Shirley (2014) included 20 subjects solely for the sound production test, while Binasfour & Setter (2018) established a study sample of 46 subjects (all female) from three different native language backgrounds (Urdu, Mandarin, and Tagalog).

A total of 30 Malay children (18 boys and 12 girls) aged between 10 and 12 years were selected as subjects for this study for the purposes of the perception and sound production tests. Consent forms were distributed to the parents to ask for their permission in allowing their children as study subjects prior the beginning of the data collection procedure. The average age of the subjects is 10.9 years. All subjects are speakers of the KL/Selangor dialect, attend national schools, and receive religious education at private religious schools in Putrajaya. Each subject was informed by the researcher a few days before the actual testing session. The subjects were given a general introduction about the study, the research procedures, the materials used during the tests, and how to respond to those materials. A specific date was set at the end of that meeting.

INSTRUMENTS

The sound perception test in this study is designed and implemented following the procedures and guidelines from the ABX Discrimination Task. The ABX Test design is adapted and developed according to the accessibility of the study. This test aims to identify the level of Malay children's ability to respond to the Arabic sounds presented. Additionally, this test is used to identify which Arabic sounds are most challenging for Malay children. The target sounds were embedded at the final position of each phoneme. The list of words for the sound perception test can be seen in the table below:

TABLE 1. List of Phoneme that Consist of the Target Sound

	Choice of Answers			Target Sound
1	A. فَلَح [falax]	B. فَح [falah]	C. فَه [falah]	[x]/[خ]
2	A. حَدِيد [hadid]	B. حَدِيْد [had ^h id ^h]		[d ^h]/[ض]
3	A. بَاس [bās]	B. بَاس [bās ^h]	C. بَد [bað]	[s]/[س]
4	A. سَرَاه [syarah]	B. سَرَح [syarah]	C. سَرَخ [syarax]	[h]/[ه]
5	A. سَمَاء [samaʔ]	B. سَمَع [samaʕ]	C. سَمَك [samak]	[ʕ]/[ع]
6	A. فَالَك [falak]	B. فَلَح [falax]	C. فَلَوق [falaq]	[q]/[ق]
7	A. فَسَوق [fasaq]	B. فَسَح [fasax]	C. فَسَح [fasah]	[x]/[خ]
8	A. بَز [baz]	B. بَد [bað]	C. بَاس [bās]	[z]/[ز]
9	A. حَيْد [hajð]	B. حَيَس [hajs]	C. حَيْث [hajθ]	[ð]/[ذ]
10	A. عَكَاذ [ukkað]	B. عَكَاز [ukkaz]	C. عَكَاظ [ukkað ^h]	[ð ^h]/[ظ]
11	A. شَوَك [ʃawʔ]	B. شَوَق [ʃawq]	C. شَوء [ʃawʔ]	[k]/[ك]
12	A. شَوَك [ʃawʔ]	B. شَوَق [ʃawq]	C. شَوخ [ʃawx]	[q]/[ق]
13	A. حَدِيْد [hadid]	B. حَدِيْب [had ^h id ^h]	C. حَدِيْب [had ^h id]	[d]/[د]
14	A. بَد [bað]	B. بَاس [bās]	C. بَاس [bās ^h]	[s ^h]/[ص]
15	A. هَمَس [hamas ^h]	B. هَمَس [hamas]		[s]/[س]
16	A. سَرَح [syarah]	B. سَرَاه [syarah]	C. سَرَخ [syarax]	[h]/[ح]
17	A. سَمَاء [samaʔ]	B. سَمَع [samaʕ]	C. سَمَك [samak]	[ʔ]/[ء]
18	A. فَالَك [falak]	B. فَلَح [falax]	C. فَلَوق [falaq]	[x]/[خ]
19	A. فَسَوق [fasaq]	B. فَسَح [fasax]	C. فَسَح [fasah]	[h]/[ح]
20	A. بَز [baz]	B. بَد [bað]	C. بَاس [bās]	[ð]/[ذ]
21	A. حَيَز [hiz]	B. حَيَس [his]	C. حَيِص [his ^h]	[s]/[س]
22	A. عَكَاذ [ukkāð]	B. عَكَاز [ukkāz]	C. عَكَاظ [ukkāz ^h]	[z]/[ز]
23	A. سَقَاظ [saqat ^h]	B. سَكَاظ [sakat]	C. سَقَاظ [saqat]	[ʔ]/[ظ]
24	A. أَصْوَات [as ^h wāt]	B. أَصْوَد [aswad]	C. أَصْوَض [aswad ^h]	[d]/[د]
25	A. أَسَاظ [asaz]	B. أَسَاث [asaθ]	C. أَسَاظ [asas]	[s]/[س]
26	A. فَظ [faz ^h]	B. فَذ [fað]	C. فَز [faz]	[ð ^h]/[ظ]
27	A. حَيْد [hajð]	B. حَيْث [hajθ]	C. حَيَز [hajz]	[ð]/[ذ]
28	A. إِسْمَح [ismah]	B. إِسْمَع [ismaʕ]	C. إِسْمَك [ismak]	[ʕ]/[ع]
29	A. حَرَس [haras]	B. حَرَاذ [harað]	C. حَرَاظ [haraz]	[z]/[ز]
30	A. فَرَح [farah]	B. فَرَخ [farax]	C. فَرَاظ [faray]	[x]/[خ]
31	A. سَمَاء [samaʔ]	B. سَمَع [samaʕ]	C. سَمَك [samak]	[ʔ]/[ء]
32	A. مُرُوق [murūr]	B. مُرُوع [murūʕ]		[r]/[ر]
33	A. سَمَاء [samaʔ]	B. سَمَع [samaʕ]	C. سَمَك [samak]	[k]/[ك]
34	A. سَقَاظ [saqat ^h]	B. سَقَاظ [saqat]		[t]/[ت]
35	A. أَصْوَات [as ^h wāt]	B. أَصْوَد [as ^h wad]	C. أَصْوَض [as ^h wad ^h]	[t]/[ت]
36	A. أَثَاظ [aθās]	B. أَثَاث [aθāθ]		[θ]/[ث]
37	A. فَظ [fað ^h]	B. فَذ [fað]	C. فَز [faz]	[ð]/[ذ]
38	A. حَيْد [hajð]	B. حَيْث [hajθ]	C. حَيَز [hajz]	[θ]/[ث]
39	A. إِسْمَح [ismah]	B. إِسْمَع [ismaʕ]	C. إِسْمَك [ismak]	[h]/[ح]
40	A. حَرَس [haras]	B. حَرَاذ [harað]	C. حَرَاظ [haraz]	[s]/[س]
41	A. فَرَح [farah]	B. فَرَخ [farax]	C. فَرَاظ [faray]	[y]/[غ]
42	A. سَمَاء [samaʔ]	B. سَمَع [samaʕ]	C. سَمَك [samak]	[k]/[ك]
43	A. مُرُوق [murūr]	B. مُرُوع [murūʕ]		[y]/[غ]
44	A. سَمَاء [samaʔ]	B. سَمَع [samaʕ]	C. سَمَك [samak]	[ʕ]/[ع]
45	A. فَه [falah]	B. فَح [falax]	C. فَلَح [falax]	[h]/[ه]

Data were collected through two methods: a sound perception test (multiple-choice) and a sound production test (recorded pronunciation).

DATA COLLECTION PROCEDURES

Data collection comprised two principal tasks: a perception test and a production test. In the perception task, participants discriminated between minimal pairs of Arabic words presented auditorily. A key methodological strength was the use of a native Arabic speaker to record the stimuli, ensuring phonetic authenticity. However, ecological validity may have been compromised by presenting each target sound five times consecutively, a procedure that could artificially enhance discrimination performance by reducing phonological memory load. Furthermore, the forced-choice paradigm administered in a controlled classroom setting, while practical, may not fully approximate natural listening conditions. For the production task, participants pronounced a list of 28 Arabic words within a soundproof booth. A significant strength here was the triple recording of each lexical token to control for nervousness or sporadic errors, alongside the inclusion of filler words to promote stable articulation, thereby enhancing the reliability of the production data. Crucially, the study's core analytical power derives from its within-subjects design, wherein the same cohort of children completed both the perception and production tasks, enabling a direct intra-individual comparison of competencies across these two domains.

DATA ANALYSIS

The analysis strategy was appropriately tailored to the research questions but remains primarily descriptive. For the perception test, descriptive statistics (percentages, means, standard deviations) was adopted to quantify accuracy, broken down by linguistic factors like place and manner of articulation. This will effectively identify which sounds are most problematic. The production data will be analyzed using spectrographic analysis with Praat software. This is a major methodological strength, as it provides an objective, acoustic measure of production quality that goes beyond subjective human judgment. The spectrograms will offer concrete evidence of how the children's articulations of Arabic sounds align with or deviate from native-like production.

RESULTS AND FINDINGS

Overall, this study has successfully built a model for learning the pronunciation of Arabic fricative and plosive sounds uttered by Malay children. The construction of this pronunciation learning model framework is based on the theory and conceptual framework that was built at the beginning of the study. By taking into account the results and findings of the study for descriptive analysis and acoustic analysis as well as Pearson correlation analysis, a model for learning the pronunciation of Arabic sounds was produced through a multiple linear regression model analysis. The research gaps highlighted in this model framework are the control of airflow based on the Filter-Source Theory as well as two elements that are absorbed when applying the Filter-Source Theory during the sound production process, namely mind consciousness and L2 phonological knowledge. Although the Arabic pronunciation model was produced through a multiple linear regression model analysis, the researcher took into account support from related L2 pronunciation learning theories such as the Filter-Source Theory (SFT), the Perceptual Assimilation Model (PAM), the Speech Learning Model (SLM) and the Speech Chain Process Theory.

Based on the Source-Filter Theory, students need to go through two sound filtering processes in the larynx and vocal tract. Resonance is the tendency of one of the four cavities to

vibrate with the largest amplitude in response to a frequency close to its natural frequency. Speech sound is the response that comes from the vocal tract system where a sound source is introduced into and filtered by the resonant characteristics of the vocal tract. SFT explains how the vocal tract filters speech sound including the source function, transfer function and output function. There are 4 different cavities in the vocal tract; laryngeal, pharyngeal, oral and nasal. Each cavity has its own valve. There are also 4 valves; labial, lingual, velopharyngeal and laryngeal. The labial valve manipulates the air flow involving the use of the lips and teeth. The lingual valve is a constriction of the tongue. Both the labial and lingual valves are in the oral cavity. Therefore, the sound produced depends on the source and the filter, which is the vocal tract. To apply Filter-Source Theory, teachers can also construct a series of harmonics representing the glottal source. Calculate the formant frequency for the vocal tract tube using the 1/4 wavelength resonator formula (open-closed tube). Sketch the filter function for the vocal tract. Pass the glottal source through the vocal tract filter to observe the speech output spectrum.

Through the review of the knowledge vocabulary carried out by the researcher, there are four theories that support the element of mental consciousness in sound production, namely Higher-order Theories, Global Workspace Theories, Re-entry and Predictive Processing Theories and Integrated Information Theory (Seth & Bayne 2022). These two elements are very important in influencing the quality of the speech sound produced. Without these two elements, a non-native speaker would simply pronounce the L2 sound in a passing manner without the existence of a visualization or shadowing process in the brain to produce a certain sound. The speech sound produced is in fact a representation of the imagination in the brain. If there is such an imagination process, then the speaker has filled one of the research gaps that have not been stated in any model of learning Arabic sound pronunciation in Malaysia.

There are basically three levels in the processing of speech sound perception (Brian, 2008; Aldridge, 1978; Kollmeier et al., 2008). At the auditory level, humans receive speech piece by piece. The speech is then perceived in terms of its acoustic characteristics. Concepts such as area of articulation, manner of articulation, distinctive features, and VOT (Voice Onset Time: the time between the release of air for the pronunciation of a consonant and the vibration of the vocal cords for the following vowel sound) are very useful here because it is these things that separate one sound from another. The sounds in speech are stored in the listener's auditory memory.

Next, at the phonetic level, the sounds are then identified. In our brain process, we see whether the sound is consonant, voiced, nasal and others. Likewise, the environment of the sound determines whether the sound is followed by a vowel or a consonant. If it is a vowel, the details of the vowel such as front vowel, back vowel, high vowel, low vowel and so on. If the speech is 'see', then the sound [s] is analyzed in the mind first and the determination of the sound is done by observing aspects such as the place of articulation, the manner of articulation, and the characteristics of its differences. Then the VOT is also examined due to its importance in determining the time when the vibration of the vocal cords occurs. These sound segments are then stored in phonetic memory. Auditory memory and phonetic memory have differences in terms of storing allophonic variations. Auditory memory stores all the allophonic variations found in the sound while phonetic memory only stores phonemic characteristics. For example, when the sound [b] from the word '/baba/' is heard, what is stored in auditory memory is not the phoneme /b/ and not just the place of articulation, the manner of articulation, and the features of its differences, but the influence of the sound [a] that follows it. Therefore, the sound [b] is also followed by lip-rounding. In phonetic memory, things like this are no longer needed because as long as the sound

is perceived as the sound [b], the details are no longer important. In other words, whether the sound [b] is followed by lip-rounding or not, the sound [b] must also be heard.

Another mental analysis is to see how the sounds are arranged because of the order of sounds that decide a word. The difference in the order of the sounds [a], [k], and [n] can form different words. When the sound [k] is heard first, then [a] and [n], then the sound [kan] is heard, similarly if the sound [n] is heard first, then the sound [nak] is heard. At the phonological level, phonological rules are applied by the mind to a series of sounds heard to determine whether the sounds follow the phonotactic rules of our language.

Based on the Speech Chain process by Hayward (2000), the process of producing speech sounds includes three stages, starting from the linguistic stage (brain), then moving to the physiological stage (muscles and nerves), then moving to the acoustic stage (target sound produced). However, the researcher formulated a slight improvement in terms of the speech process cycle by revealing the research gap in this built model. This Arabic pronunciation learning model reveals that there are six stages in the process of producing speech sounds, in which the first stage initially from linguistic level (brain). It then shift to physiological level (oral) but at the third stage (second linguistic level), the cycle must be linked to the linguistic stage again, which is the brain, before moving to the fourth stage, which is the Filter-Source Theory (airflow control and air filtration). The fifth stage in this process (target sound produced) is also the source for PAM (perception test), thus making these cycles mutually dependent on each other. The framework of the Arabic pronunciation learning model built by the researcher in this study can be seen more clearly in the Figure 1:

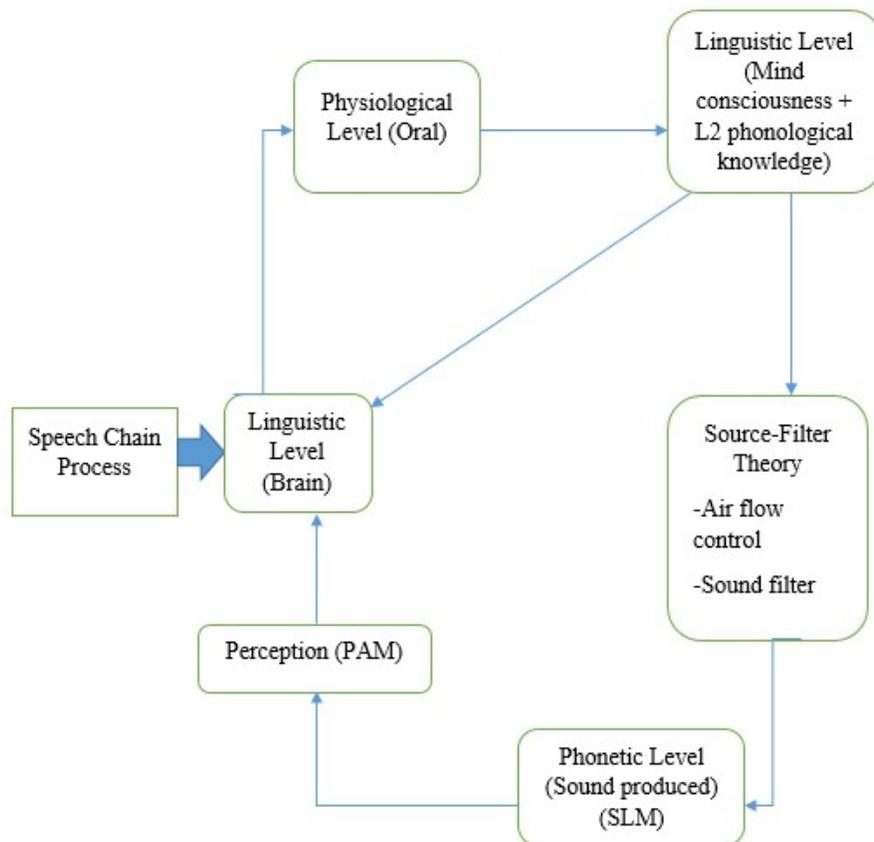


FIGURE 1. Speech Chain Process (Adapted from Hayward, 2000)

Based on this model framework, learning Arabic pronunciation starts from understanding the PAM, SLM and SFT theories. In short, the main thing that must be understood is the sound system of the native language and L2. According to Best et al. (1988), through PAM, L2 sounds that are the same as the L1 sound category will facilitate the acquisition of L2. However, Flege (1995) holds the opposite view, namely that L1 sounds that resemble the L2 sound category actually cause difficulties in the acquisition of L2. If observed in more detail, this study is more inclined towards the theory and view of Flege (1995), where the sounds of the target segment of Arabic that are at a low level of ability are sounds that have phonetic categories in L1. The smaller the gap in sound differences between L1 and L2, the more difficult it is to discriminate for non-native speakers. The small gap in differences causes confusion in the perception of non-native Arabic speakers.

Next, SLM is also related to the Source-Filter Theory (SFT), which is its influence on the aspect of sound production. Each process of sound production is not contradictory as described in the Filter-Source Theory. According to SFT, the air flow coming from the lungs induces tissue vibrations in the vocal folds (i.e., two small muscle folds located in the larynx) and produces the "source" sound. Turbulent air flow is also created in the glottis or in the vocal tract to produce a noisy sound source. The spectral structure of this source sound is shaped by the "filter" of the vocal tract. Through the filtering process, frequency components corresponding to the resonance of the vocal tract are amplified, while other frequency components are attenuated. The source sound mainly characterizes the vocal pitch (i.e., the fundamental frequency), while the filter forms the timbre. SFT provides a very accurate description of the production of ordinary sounds and has been successfully applied to speech analysis, synthesis, and processing. Separate control of the source (phonation) and the filter (articulation) provides advantages for acoustic communication, especially for human language, which requires the expression of multiple phonemes realized by flexible maneuvers of the vocal tract configuration (Tokuda, 2021).

However, SFT cannot stand alone without the support of mental awareness and L2 phonological knowledge. This means that throughout the process of controlling the airflow in the vocal tract, students must focus their full attention (mental awareness) and apply existing L2 phonological knowledge to realize good sound production. Mental awareness includes the visualization process for each sound representation in the brain. This is supported by Higher-Order Theories of Consciousness. These two elements also directly affect SFT. At the same time, the element of mental awareness also affects the level of student perception (PAM), so the student's perception also affects their L2 phonological knowledge. It clearly shows the existence of such a cycle in the process of producing accurate Arabic sounds.

In short, two elements, namely mental awareness and L2 phonological knowledge, are applied during the application of SFT, and then applied to the perception process (PAM). PAM and SLM also have a two-way cycle. Then, this process chain improves the Speech Chain Process Theory (Hayward, 2000) by explaining that there are several stages that need to be placed in the process of learning sound production.

The speech chain process, which has enhanced the speech chain process model by Hayward (2000), was the foundation and further extended to the Arabic pronunciation model that was produced in this study. Specifically, the researcher produced an Arabic pronunciation model by taking into account the independent variables that contributed to the achievement of the study subject's sound production test. By using a linear regression analysis model, the researcher produced an Arabic pronunciation model through quantitative analysis since the existing data was also quantitative.

Through the analysis of the multiple linear regression model carried out, a pronunciation model was produced based on the independent variables that influence the dependent variables. Among the independent variables involved are social aspects, namely age, gender, duration of Arabic and Quranic learning, duration of Quranic learning hours outside of school, duration of Quranic extra classes at school, family environment and background, and family economy. Meanwhile, the dependent variable is the Frication Duration (FD) which can reflect the level of Arabic sound pronunciation among the study subjects. For the process of conducting the multiple linear regression model analysis, all Frication Duration value data for 30 study subjects for three sounds that were concluded to be problematic were converted to a scale of 1-10 to represent the level of accuracy of the values produced by the study subjects.

Arabic Speech Phonological Model (M-ASP) was produced specifically for three problematic sounds in the perception and production of sounds among Malay children, namely the sounds [ð], [z] and [h]. Meanwhile, the sound [ʔ] was removed from the list of problematic sounds after spectrograph analysis was conducted, revealing that the mean value of the Voice Onset Time for this sound [ʔ] in the sound production test was quite good, namely 28.83ms. If look closely, this mean value is only slightly higher than the standard mean value for the Voice Onset Time for the sound [ʔ] in previous knowledge vocabulary studies, which is 28ms (Saif & Ziyad, 2021). In the process of analyzing the linear regression model, the data involved should be analyzed in a correlation manner first. The tables below display the results of the correlation of sound production for the three sounds that are problematic in sound production with independent variables.

TABLE 2. Correlation of sound production /ð/ and independent variable

Variables	/ð/	AQ Class	PF2	L	TBBA	TBAQDS	ML	KTAQLS
/ð/	1	0.261	0.025	0.214	0.270	0.414*	0.204	0.083
Quranic Class	0.261	1	0.594**	0.210	0.137	0.111	0.200	0.210
PF2	0.025	0.594**	1	0.000	0.041	0.325	0.000	0.236
Boy	0.214	0.210	0.000	1	0.459*	0.184	0.000	0.333
TBBA	0.270	0.137	0.041	0.459*	1	0.194	0.047	0.306
TBAQDS	0.414*	0.111	0.325	0.184	0.194	1	0.015	0.282
Multilingual	0.204	0.200	0.000	0.000	0.047	0.015	1	0.272
KTAQLS	0.083	0.210	0.236	0.333	0.306	0.282	0.272	1

TABLE 3. Correlation of sound production /z/ and independent variable

Variables	/z/	AQ Class	PF2	L	TBBA	TBAQDS	ML	KTAQLS
/z/	1	0.152	0.295	0.245	0.133	0.284	0.255	0.216
Quranic Class	0.152	1	0.506**	0.039	0.027	0.263	0.097	0.171
PF2	0.295	0.506**	1	0.210	0.137	0.111	0.200	0.210
Boy	0.245	0.039	0.210	1	0.459*	0.184	0.000	0.333
TBBA	0.133	0.027	0.137	0.459*	1	0.194	0.047	0.306
TBAQDS	0.284	0.263	0.111	0.184	0.194	1	0.015	0.282
Multilingual	0.255	0.097	0.200	0.000	0.047	0.015	1	0.272
KTAQLS	0.216	0.171	0.210	0.333	0.306	0.282	0.272	1

TABLE 4. Correlation of sound production /h/ and independent variable

Variables	/h/	AQ Class	PF2	L	TBBA	TBAQDS	ML	KTAQLS
/h/	1	0.362*	0.222	0.093	0.021	0.170	0.493**	0.238
Quranic Class	0.362*	1	0.908**	0.000	0.041	0.325	0.000	0.236
PF2	0.222	0.908**	1	0.039	0.027	0.263	0.097	0.171
Boy	0.093	0.000	0.039	1	0.459*	0.184	0.000	0.333
TBBA	0.021	0.041	0.027	0.459*	1	0.194	0.047	0.306
TBAQDS	0.170	0.325	0.263	0.184	0.194	1	0.015	0.282
Multilingual	0.493**	0.000	0.097	0.000	0.047	0.015	1	0.272
KTAQLS	0.238	0.236	0.171	0.333	0.306	0.282	0.272	1

TABLE 5. Correlation sound production /ð/ and independent variables

Variables	Unstandardized Coefficients		Standardized Coefficients		t-test	Significance	Collinearity Statistics	
	Beta	Ralar Piawai	Beta				Tolerance	VIF
Fixed Variable	2.860	1.843			1.552	0.135		
Mind Consciousness	0.784	0.384	0.400		2.044	0.053	0.557	1.794
Prior Knowledge of /ð/	0.302	0.392	0.152		0.771	0.449	0.544	1.838
Boy	1.518	0.824	0.406		1.841	0.079	0.438	2.285
Learning Duration of AL	1.063	0.444	0.496		2.397	0.025	0.497	2.011
Extra Class AQ (outside)	0.195	0.058	0.567		3.375	0.003	0.754	1.327
Multilingual	1.841	0.632	0.482		2.912	0.008	0.776	1.289
Extra Class AQ (inside)	0.382	0.138	0.613		2.767	0.011	0.433	2.308
R Square					0.532			
Adjusted R Square					0.383			
ANOVA Test F					3.568			
ANOVA Significance					0.010			

VIF= Variance Inflation Factor (VIF)

Table 5 shows the relationship between [ð] and the independent variables. This independent variable explains the factors that influence or improve the pronunciation of the sound [ð]. The ANOVA test shows that the F value is 3.568 and the significant value is 0.010. The ANOVA test results show that the model built is suitable for the available data. Then, the R square value is 0.532 showing that the independent variable can explain the dependent variable by 53.2%. The VIF test shows that there is no overlap between the independent variables. This is evidenced by the VIF value below 10. Next, the relationship between mind consciousness and the pronunciation of the sound [ð] shows a positive and significant relationship (Beta = 0.784, significant = 0.053). This result indicates that mental awareness is important to the pronunciation of the sound [ð]. The relationship between L2 phonological knowledge and the pronunciation of the sound [ð] shows a positive and insignificant relationship (Beta = 0.302, significant = 0.449). This indicates that L2 phonological knowledge needs to be given more emphasis in an effort to improve the level of pronunciation of the sound [ð]. Next, the relationship between the duration of learning Arabic and the pronunciation of the sound [z] shows a positive and significant relationship (Beta = 0.636, significant = 0.115). This result indicates that the duration of learning Arabic is important to the pronunciation of the sound [z]. The relationship between the duration of learning the Quran outside of school and the pronunciation of the sound [z] shows a positive and significant relationship (Beta = 0.122, significant = 0.021). This indicates that the duration of learning Arabic is important to the pronunciation of the sound [z]. Furthermore, the relationship between multilanguage and the pronunciation of the sound [ð] shows a positive and significant relationship

(Beta = 1.553, significant = 0.024). This result indicates that the multilanguage factor is important to the pronunciation of the sound [ð]. Meanwhile, the relationship between the duration of additional Quranic classes at school and the pronunciation of the sound [ð] also shows a positive and significant relationship (Beta = 0.389, significant = 0.012). This result indicates that the factor of the duration of additional Quranic classes in schools is also important for the pronunciation of the sound [ð].

TABLE 6. Correlation of sound production /z/ and independent variable

Variables	Unstandardized Coefficients		Standardized Coefficients	t-test	Significance	Collinearity Statistics	
	Beta	Ralar Piawai	Beta			Tolerance	VIF
Malar	4.400	1.592		2.763	0.011		
KM zal	0.064	0.337	0.036	0.189	0.852	0.654	1.529
PF Zal	0.555	0.302	0.354	1.838	0.080	0.641	1.560
Lelaki	0.809	0.689	0.271	1.176	0.252	0.447	2.237
TB BA	0.561	0.374	0.327	1.500	0.148	0.499	2.005
KTAQLS	0.132	0.048	0.481	2.738	0.012	0.768	1.302
ML	1.583	0.540	0.519	2.934	0.008	0.758	1.319
KTAQDS	0.306	0.116	0.615	2.651	0.015	0.441	2.267
R Square				0.478			
Adjusted R Square				0.312			
ANOVA				2.876			
Test F				0.027			
Significance							

VIF= Variance Inflation Factor (VIF)

Table 6 shows the relationship between [z] and independent variables. This independent variable explains the factors that influence or improve the pronunciation of the sound [z]. The ANOVA test shows that the F value is 2.876 and the significant value is 0.027. The ANOVA test results show that the model built is in accordance with the available data. Then, the R square value is 0.478 showing that the independent variable can explain the dependent variable by 47.8%. The VIF test shows that there is no overlap between the independent variables. This is evidenced by the VIF value below 10. Next, the relationship between mental awareness and the pronunciation of the sound [z] shows a positive and insignificant relationship (Beta = 0.064, significant = 0.852). This result indicates that mental awareness needs to be given more emphasis in an effort to improve the level of pronunciation of the sound [z]. The relationship between L2 phonological knowledge and pronunciation of the sound [z] shows a positive and insignificant relationship (Beta = 0.555, significant = 0.080). This indicates that L2 phonological knowledge also needs to be given more emphasis in an effort to improve the level of pronunciation of the sound [z]. Next, the relationship between the duration of learning Arabic and pronunciation of the sound [z] shows a positive and insignificant relationship (Beta = 0.561, significant = 0.148). This result indicates that the duration of learning Arabic needs to be emphasized in an effort to improve the level of pronunciation of the sound [z]. The relationship between the duration of learning the Quran outside of school and pronunciation of the sound [z] shows a positive and significant relationship (Beta = 0.132, significant = 0.012). This indicates that the duration of learning Arabic is important to the pronunciation of the sound [z]. Furthermore, the relationship between multilanguage and pronunciation of the sound [z] shows a positive and significant relationship (Beta = 1.583, significant = 0.008). This result indicates that the multilanguage factor is important to the pronunciation of the sound [z]. Meanwhile, the relationship between the duration of additional

Quranic classes at school and the pronunciation of the sound [z] also showed a positive and significant relationship (Beta = 0.306, significant = 0.015). This result indicates that the factor of the duration of additional Quranic classes at school is also important to the pronunciation of the sound [z].

TABLE 7. Correlation of sound production /h/ and independent variable

Variables	Unstandardized Coefficients		Standardized Coefficients	t-test	Significance	Collinearity Statistics	
	Beta	Ralar Piawai	Beta			Tolerance	VIF
Malar	7.123	1.043		6.827	0.000	0.156	6.419
KM zal	0.863	0.416	0.757	2.078	0.050	0.163	6.141
PF Zal	0.445	0.454	0.349	0.980	0.338	0.444	2.253
Boy	0.669	0.464	0.311	1.441	0.164	0.496	2.017
TB BA	0.158	0.252	0.128	0.628	0.537	0.754	1.326
KTAQLS	0.085	0.033	0.430	2.596	0.016	0.799	1.252
ML	0.813	0.353	0.370	2.303	0.031	0.430	2.327
KTAQDS	0.101	0.079	0.282	1.286	0.212	0.156	6.419
R Square				0.545			
Adjusted R Square				0.401			
ANOVA	F Test			3.769			
	Significance			0.008			

VIF= Variance Inflation Factor (VIF)

Table 7 shows the relationship between [h] and independent variables. This independent variable explains the factors that influence or improve the pronunciation of the sound [h]. The ANOVA test shows that the F value is 3.769 and the significant value is 0.008. The ANOVA test results show that the model built is in accordance with the available data. Then, the R square value is 0.545 showing that the independent variable can explain the dependent variable by 54.5%. The VIF test shows that there is no overlap between the independent variables. This is evidenced by the VIF value below 10. Next, the relationship between mental awareness and the pronunciation of the sound [h] shows a positive and significant relationship (Beta = 0.863, significant = 0.050). This result indicates that mental awareness is important to the level of pronunciation of the sound [h]. The relationship between L2 phonological knowledge and the pronunciation of the sound [h] shows a positive and insignificant relationship (Beta = 0.445, significant = 0.338). This indicates that L2 phonological knowledge also needs to be given more emphasis in an effort to improve the level of pronunciation of the sound [h]. Next, the relationship between the duration of learning Arabic and the pronunciation of the sound [h] shows a positive and insignificant relationship (Beta = 0.158, significant = 0.537). This result indicates that the duration of learning Arabic needs to be given more attention in an effort to improve the level of pronunciation of the sound [h]. The relationship between the duration of learning the Quran outside of school and the pronunciation of the sound [h] shows a positive and significant relationship (Beta = 0.085, significant = 0.016). This indicates that the duration of learning Arabic is important to the pronunciation of the sound [h]. Next, the relationship between multilanguage and the pronunciation of the sound [h] shows a positive and significant relationship (Beta = 0.813, significant = 0.031). This result indicates that the multilanguage factor is important to the pronunciation of the sound [h]. Meanwhile, the relationship between the duration of additional Quranic classes at school and the pronunciation of the sound [h] also shows a positive and insignificant relationship (Beta = 0.101, significant = 0.212). This result indicates that the factor of the duration of additional Quranic classes at school

also needs to be given more attention in an effort to improve the level of pronunciation of the sound [h].

It can be concluded that these five independent variables are important factors that influence the level of pronunciation of the three Arabic sounds that were found to be problematic in the sound production test conducted in this study. As a result, a pronunciation model for the three Arabic sounds, namely the sounds [ð], [z] and [h], was designed. The pronunciation model of the Arabic sound can be seen in Figure 2 as follows:

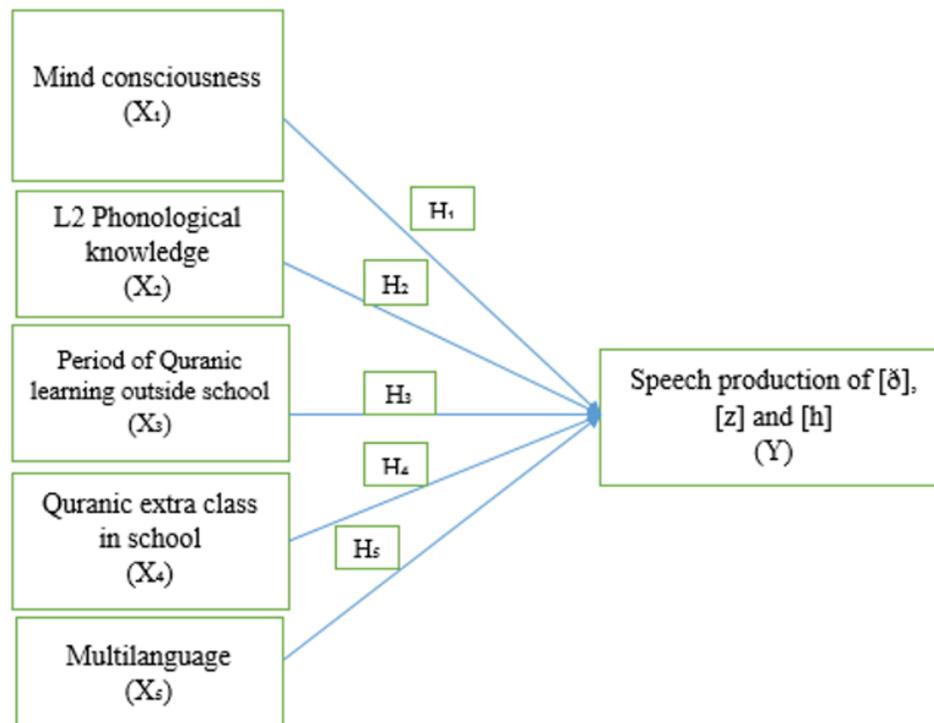


FIGURE 2. The pronunciation model of the Arabic sound

*X₁-X₅: Independent variable

*Y: Dependent variable

*H₁-H₅: Hypothesis impact X on Y

DISCUSSION

The development of the Arabic Speech Phonological Model (M-ASP) produced in this study integrates key phonetic theories, including the Filter-Source Theory (SFT), the Perceptual Assimilation Model (PAM), and the Speech Learning Model (SLM). This model emphasizes the critical roles of two cognitive-linguistic elements; mind consciousness and L2 phonological knowledge. These factors influence both the perception and production of Arabic sounds among Malay primary school children, especially for phonemes absent or variably realized in their L1. The improved model contributes significantly to existing literature by filling gaps in the understanding of perception–production mismatch in L2 Arabic phonology. Unlike prior studies that often isolated production, this study provides an integrated view by empirically analyzing how

perception proficiency aligns or misaligns with actual articulation. This dual analysis enhances the pedagogical framework for teaching Arabic as a second language. Importantly, the model extends the theoretical framework of Hayward's Speech Chain Process (2000) by introducing two additional stages; the recursive involvement of linguistic processing after physiological activation and the role of airflow control as a filtering mechanism. This nuanced layering illustrates how Arabic phoneme acquisition is not a linear chain but a cyclical-cognitive loop, where perception feedback actively refines future production. Further, the inclusion of mental awareness is aligned with current cognitive neuroscience models of speech, notably the Higher-Order Theories of Consciousness (Carruthers & Gennaro, 2023). These perspectives suggest that conscious monitoring and visualization of articulatory targets enhance phonetic accuracy. From a practical standpoint, this study offers a scalable model for Arabic phoneme instruction in multilingual educational settings, particularly in Malaysia. The model is especially pertinent in reinforcing Arabic literacy in Islamic education, where accurate recitation (e.g., Qur'an recitation) is essential. The findings support curriculum development, teacher training, and the formulation of instructional interventions targeting difficult Arabic sounds, particularly [ð], [z], and [h]. In essence, this model contributes a novel, theory-informed, data-driven framework that advances our understanding of L2 phonological acquisition and provides a solid foundation for empirical applications in Arabic language education.

CONCLUSION

This study depicts a comprehensive investigation into the perceptual and articulatory challenges faced by Malay primary school children in acquiring Arabic phonemes. The identification of specific sounds that are poorly produced despite being perceived well underscores the difficulty of L2 phonological acquisition. By a rigorous analysis, the proposed Arabic Speech Phonological Model (M-ASP) addresses this gap by incorporating cognitive awareness and phonetic principles rooted in contemporary speech science. The model's theoretical contributions lie in its synthesis of PAM, SLM, and SFT with additional aerodynamics of speech and cognitive elements namely airflow control and mental visualization (mind consciousness). A fairly significant of its pedagogical relevance, thus offering relevant guidance for educators, curriculum designers, and language therapists working with non-native Arabic learners. By focusing on the most difficult phonemes and mapping them within a structured instructional model, this study paves the way for enhanced Arabic pronunciation instructional methods. Eventually, the research not solely enhances the scholarly discourse on L2 phonology but also bridges theory and practice in a culturally and linguistically meaningful context particularly in non-native (Malay) settings.

REFERENCES

- Aldridge, J. W. (1978). Levels of processing in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 4(1), 164–177. <https://doi.org/10.1037/0096-1523.4.1.164>
- Alwabari, S. (2013). *Non-native Production of Arabic Pharyngeal and Pharyngealized Consonants* (Carleton University), (Tesis Sarjana). Retrieved from <https://curve.carleton.ca/system/files/etd/3ce1d8e6-b4d9-4eb2-930b-058e405c41c8/>
- Antoniou, Mark, Michael D. Tyler, and Catherine T. Best. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics*, 40(4), 582-594. <https://doi.org/10.1016/j.wocn.2012.05.005>

- Arjmandi, M. K., & Behroozmand, R. (2024). On the interplay between speech perception and production: insights from research and theories. *Frontiers in Neuroscience*, 18. <https://doi.org/10.3389/fnins.2024.1347614>
- Baese-Berk, M. M., Kapnoula, E. C., & Samuel, A. G. (2024). "The Relationship of Speech Perception and Speech Production: It's Complicated." *Psychonomic Bulletin & Review*, September. <https://doi.org/10.3758/s13423-024-02561-w>.
- Baker, W. (2005). The interaction of the bilingual's two phonetic systems: Differences in early and late Korean-English bilinguals. Paper presented at the Proceedings of the 4th International Symposium on Bilingualism.
- Best, C. T., Gerald W. McRoberts, and Nomathemba M. Sithole. (1988). Examination of Perceptual Reorganization for Non-native Speech Contrasts: Zulu Click Discrimination by English-speaking Adults and Infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 345-360. <https://doi.org/10.1037/0096-1523.14.3.345>
- Best, C. T. (1995). A Direct Realist View of Cross-language Speech Perception. In Strange, W. (Ed.). *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Maryland: New York Press.
- Binasfour, H., & Setter, J. (2018). The Influence of Vowel Context on the Perception and Production of Emphatic Sounds among L2 Learners of Arabic. *Language Studies Working Papers*, 9, 14-25.
- Binasfour, H. S. M. (2018). *Investigating the Perception and Production of the Arabic Pharyngealised Sounds by L2 Learners of Arabic* PhD diss., University of Reading.
- Brian C.J Moore; Basic auditory processes involved in the analysis of speech sounds. *Philos Trans R Soc Lond B Biol Sci* 12 March 2008; 363 (1493), 947–963. <https://doi.org/10.1098/rstb.2007.2152>
- Carruthers, P., & Gennaro, R. "Higher-Order Theories of Consciousness", *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/fall2023/entries/consciousness-higher/>
- Défossez, A., Caucheteux, C., Rapin, J., Kabeli, O., & King, J. R. (2023). Decoding speech perception from non-invasive brain recordings. *Nature Machine Intelligence*. <https://doi.org/10.1038/s42256-023-00714-5>
- Flege, J. E. (1995). Second Language Speech Learning Theory, Findings and Problems. In Strange, W. (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 233-277). Maryland: New York Press.
- Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, 36, 67–81. <https://doi.org/10.1007/BF02648022>
- Gillon, G. T. (2004). Phonological awareness: from research to practice. New York: Guilford Press.
- Hamid, S. A. (2010). *An acoustic and perceptual analysis of the phonetic properties of Malay English in comparison to those of Malay*. PhD diss., University of Newcastle upon Tyne.
- Hayward, K. (2000). *Experimental Phonetics*. England: Perarson Education Ltd.
- Heald, S. L., & Nusbaum, H. C. (2014). Speech perception as an active cognitive process. *Front Syst Neurosci*, 8, 1-15. <https://doi.org/10.3389/fnsys.2014.00035>
- Horslunda, C. S., Aarhoj, A. E., & Ocke.-Schwen B. (2015). Perceptual assimilation and identification of English consonants by native speakers of Danish. Paper presented at the to appear in: Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, UK.
- Huddin, M. R. A., Ali, S. S. A., & Yeob, S. J. (2017). *Fa'iliyyah al-Lahjat al-Malayuwiyyah fi Mustawa al-Kifayah al-Nutqiyah li-annas al-Arabiyy: Lahjata Wilayah Kelantan wa Qodah ammuzajain*. IIUM Press. ISBN: 978-967-418-816-0
- Issraq Ramli (2004). *Pengaruh Unsur Fonologi Dialek Melayu Sabah dalam Sebutan Bahasa Arab*. (Disertasi Sarjana). Diakses dari <http://studentsrepo.um.edu.my/>
- Issraq Ramli & Mohd Zaki Abd Rahman. (2017). Aplikasi Teori Mahjub terhadap Pembaikan Sebutan Bunyi Bahasa Arab dalam Kalangan Pelajar Sabah. *Jurnal Pusat Penataran Ilmu & Bahasa*. Universiti Malaysia Sabah, Malaysia. 25.
- Jabar, M. A. A. (2004). Gangguan Bunyi Melayu dalam Sebutan Arab: Satu Analisis Ringkas. *Pertanika Journal of Social Sciences & Humanities*, 12(2), 101-110. ISSN: 0128-7702
- Kearney, E., & Guenther, F. H. (2019). Articulating: the neural mechanisms of speech production. *Language, Cognition and Neuroscience*, 34(9), 1214–1229. <https://doi.org/10.1080/23273798.2019.1589541>
- Kollmeier, B., Brand, T., Meyer, B. (2008). Perception of Speech and Sound. In: Benesty, J., Sondhi, M.M., Huang, Y.A. (eds) Springer Handbook of Speech Processing. Springer Handbooks. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-49127-9_4
- Korpala, P., & Mikolaj, S. (2020). The perception of native vs. nonnative Danish speech: Bent and Bradlow's matched interlanguage speech intelligibility benefit revisited. *Scandinavian Philology*, 18, 284–296. <https://doi.org/10.21638/11701/spbu21.2020.204>

- Ladefoged, P. (2003). *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell Publishing.
- Lai, Yi-hsiu. (2010). English vowel discrimination and assimilation by Chinese-speaking learners of English. *Concentric: Studies in Linguistics*, 36(2), 157-182.
- McGettigan, C & Tremblay, P. (2018). Links between perception and production examining the roles of motor and premotor cortices in understanding speech. *The Oxford Handbook of Psycholinguistics*, 306-332.
- McGuire, G. (2010). A brief primer on experimental designs for speech perception research. Laboratory Report, 77
- Nagaraja, J. S., Hamid, B. A., & Maamor, N. (2025). Phonological Development in Hearing-Impaired Malay-Speaking Children. *GEMA Online Journal of Language Studies*, 25(1), 33-55. <https://doi.org/10.17576/gema-2025-2501-03>
- Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition* 16(4), 343-352. <https://doi.org/10.3758/BF03197045>
- Nasrullah, R. (2025). Working Memory and Second Language Acquisition: Effects on Vocabulary, Grammar, Reading, and Writing. *GEMA Online Journal of Language Studies*, 25(1), 102-118. <https://doi.org/10.17576/gema-2025-2501-06>
- Nik Muhammad Rahimi, Harun Baharudin, Ghazali Yusri, Kamarul Syukri Mat Teh & Mohamed Amin Embi. Pembelajaran Konsonan Arab mengikut Pelat Bahasa Melayu. *GEMA Online Journal of Language Studies*, 10(3), 1-14.
- Perron, M., Vuong, V., Grassi, M. W., Imran, A., & Alain, C. (2024). Engagement of the speech motor system in challenging speech perception: Activation likelihood estimation meta-analyses. *Human Brain Mapping*, 1-26. <https://doi.org/10.1002/hbm.70023>
- Piaget, J. (1962). The stages of the intellectual development of the child. *Bulletin of the Menninger Clinic*, 26(3), 120-128.
- Radzi, M. P., & Hamid, S. A. (2022). Model Pemerian Ciri-Ciri Akustik Bunyi Plosif dan Frikatif Arab Yang Diujarkan Oleh Penutur Natif Melayu. *Global Journal of al-Thaqafah*, 12(2), 104-117. ISSN: 2232-0474 I E-ISSN: 2232.0482
- Radzi, M. P., Hamid, S. A., & Aman, R. (2019). Ciri-ciri Akustik Bunyi Faringealisasi Arab oleh Penutur Natif Melayu: Penelitian Berasaskan Perbezaan Jantina. *Jurnal Melayu*, 18(1), 13-28.
- Radzi, M. P., & Hamid, S. A. (2020). Saya Ujar /t/, /t̤/ dan /t̥/ Secara Akustiknya Berbeza. *Global Journal of al-Thaqafah*, 10(2), 78-83.
- Ramli, N. A., Che Radiah Mezah, C. R., & Thai, Y. N. (2016). Penguasaan Pelajar Melayu Terhadap Tekanan Suara Menyebut Perkataan Arab dari Sudut Intensiti. *Jurnal Kemanusiaan*, 25(1), 110-123.
- Saif Mohammed al-Tai & Ziyad Rakan Kasim. (2021). An Acoustic Study of the Production of Iraqi Arabic Stop Consonants. *Adab Al-Rafidayn*, 86, 67-97.
- Saito, K., Tran, M., Suzukida, Y., Sun, H., Magne, V., & Ilkan, M. (2019). How do L2 listeners perceive the comprehensibility of foreign-accented speech? Roles of L1 profiles, L2 proficiency, age, experience, familiarity and metacognition. *Studies in Second Language Acquisition*, 41, 1133-1149. <https://doi.org/10.1017/S0272263119000226>
- Seth, A.K., Bayne, T. Theories of consciousness. *Nat Rev Neurosci*, 23, 439-452 (2022). <https://doi.org/10.1038/s41583-022-00587-4>
- Shehata, A. (2024). Arabic speech intelligibility: Perception of spoken Arabic by native and non-native speakers. *Language Teaching Research* 1-16. <https://doi.org/10.1177/13621688241231628>
- Sun, L., & Vincent J. van Heuven. (2007). Perceptual assimilation of English vowels by Chinese listeners: Can native-language interference be predicted? *Linguistics in the Netherlands*, 24(1), 150-161. <https://doi.org/10.1075/avt.24.15sun>
- Tokuda, I. (2021, November 29). The Source-Filter Theory of Speech. Oxford Research Encyclopedia of Linguistics. Retrieved 2nd Sept. 2025, from <https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-894>
- Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, 71(1), 4-21. <https://doi.org/10.1159/000356237>
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101(2), 192-212. <https://doi.org/10.1037/0033-2909.101.2.192>

ABOUT THE AUTHORS

Muhammad Ridha Ali Huddin (Ph.D) is a senior lecturer in the Faculty of Languages and Linguistics, Universiti Malaya. His expertise lies in phonetics and phonology, psycholinguistics and language acquisition.

Shahidi A. Hamid (Ph.D) is an Associate Professor at the Faculty of Social Science and Humanities, National University of Malaysia. His expertise lies in experimental phonetics, second language learning and Malay dialectology.

Md Shafiin Shukor (Ph.D) is a researcher in UKM Pakarunding. His expertise lies in the statistics and economics.

Siti Maryam Aj Jalilah Roslan is an Arabic Teacher in Sungai Ramal Dalam Primary School, Ministry of Education Malaysia. Her expertise lies in Arabic pedagogy.