Journal of Modern Languages
Vol 35(1) 2025
https://doi.org/10.22452/jml.vol35no1.2

Acoustic Analysis of Malay /r/ in Connected Speech: A Preliminary Study on Tap and Trill Variants

W. A Wan Aslynn[1]
wanaslynn@iium.edu.my
Department of Audiology and Speech-Language Pathology
Kulliyyah of Allied Health Sciences
International Islamic University Malaysia
25200 Kuantan, Pahang, Malaysia
https://orcid.org/0000-0003-4952-3972

Muhammad Roslan
uwaiscreative@gmail.com
Department of Audiology and Speech-Language Pathology,
Kulliyyah of Allied Health Sciences
International Islamic University Malaysia
25200 Kuantan, Pahang, Malaysia
https://orcid.org/0009-0003-7931-1484

Nur Hanisah Tukiran
hanisahtukiran@iium.edu.my
Department of Audiology and Speech-Language Pathology,
Kulliyyah of Allied Health Sciences,
International Islamic University Malaysia,
25200 Kuantan, Pahang, Malaysia
https://orcid.org/0000-0003-0863-5409

Received: 2 March 2025 Accepted: 12 June 2025

Abstract

Differences in regional dialects in the production of consonants and vowels can influence the pronunciation of Standard Malay. This study aims to acoustically characterise the pronunciation of the Malay phoneme /r/, a consonant known to exhibit phonetic variation.

[1] Corresponding author

The data analysed in this study contained both tap [ɾ] and trill [r] realisations of /r/. A total of 74 tokens containing the /r/ sound were extracted from audio recordings of four Standard Malay-speaking participants. Based on acoustic analysis, it was found that the /r/ phoneme was consistently realised as either a tap or trill in initial and medial positions, while it was frequently omitted in word-final position. These findings have clinical relevance for speech-language pathology, particularly in differentiating between typical dialectal variation and disordered speech.

Keywords: Acoustic Analysis, Malay Consonants, Standard Malay, Tap, Trill

1. Introduction

Pronunciation plays a crucial role in effective communication, and variations in phonetic features among speakers often reflect the influence of regional dialects and second-language learning. In Malaysia, Standard Malay serves as the national language and is used widely in formal settings. However, the linguistic diversity within the country has led to variations in how Malay phonemes are pronounced. Among these, the phoneme /r/ often exhibits notable variability, influenced by regional dialects and sociolinguistic factors (Schmittauer, 2024).

Phonetic studies have demonstrated that regional dialects exert a significant influence on the pronunciation of vowels and consonants in Standard Malay (Shahidi & Rahim, 2010) and as noted by Idris Aman (2000), language not only reflects sociocultural identity but also holds the power to influence societal perceptions and relationships. The phoneme /r/ is particularly susceptible to such variation, with realisations including the tap [ɾ], the trill [r], the approximant [ɹ], the voiced velar fricative [ɣ], the voiced uvular fricative [ʁ], and the voiced pharyngeal fricative [ʕ]. While regional dialectal influences are recognised as a significant factor in phonetic variation, there is a lack of comprehensive acoustic studies that objectively characterise the /r/ phoneme in connected speech. Understanding the acoustic properties of /r/, such as its first formant (F1) values and duration, can provide valuable insights into its articulation and variability. This is crucial for addressing variability in Malay pronunciation in various settings including educational, professional, and clinical settings.

Understanding phonetic variations, in this case, the realisation of /r/, may also influence speech intelligibility. Different realisation of this phoneme could reduce intelligibility, especially for listeners unfamiliar with regional or idiosyncratic speech patterns. For instance, the omission

of /r/ in word-final position, commonly observed among southern Standard Malay speakers, may result in perceptual ambiguity or miscommunication if the listener relies on phonemic cues for word recognition (Shahidi & Rahim, 2010). These variations, while acceptable from a sociolinguistic perspective, can pose challenges in cross-dialectal interactions or when clear enunciation is expected, such as in media, education, or public speaking (Abu Bakar, 2019). Furthermore, for second language learners or individuals undergoing speech therapy, inconsistent /r/ production may hinder phonological acquisition or reinforce non-standard articulatory patterns, complicating intervention strategies (Gurevich, 2021). By acoustically characterising the /r/ phoneme, particularly through measurable parameters like F1 and duration, clinicians and educators can be better equipped to differentiate between dialectal variation and potential articulation disorders, thereby supporting more accurate diagnosis and effective language instruction.

While previous literature has acknowledged the presence of multiple articulatory variants of /r/, including approximants and fricatives, there remains a lack of systematic acoustic characterisation particularly in naturalistic speech contexts. In the present study, we examine the production of the tap [ɾ] and the trill [r] in the connected speech of Malay speakers. Although other variants of /r/ have been observed in broader dialectal contexts, this analysis focuses specifically on these two realisations as captured in our recordings. The main objective of the study was to acoustically characterise the realisations of tap [ɾ] and the trill [r] by Standard Malay speakers.

Despite the acknowledgement of phonetic variation in Malay /r/ realisations, much of the discussion in existing literature remains impressionistic or sociolinguistic in focus, with limited empirical grounding in acoustic data. This presents a methodological gap, as objectively documenting such variation through acoustic analysis offers a more precise understanding of how these sounds are physically realised in natural speech. Empirical data not only enable researchers to describe the phoneme more accurately but also provide a foundation for comparing articulatory behaviour across dialects, speakers, and communicative contexts.

Moreover, the absence of detailed acoustic benchmarks makes it difficult for clinicians, educators, and language technologists to discern whether a given /r/ production falls within the expected range of standard variation or reflects an atypical articulation. In clinical settings, particularly speech-language pathology, this distinction is critical for differentiating dialectal influence from disordered speech patterns. Similarly, in language instruction, teachers benefit from

a clearer understanding of how phonemes like /r/ are commonly produced, especially when assisting learners from diverse dialectal backgrounds. In speech technology, such as automatic speech recognition systems, acoustic characterisation of phonetic variants is vital for enhancing system accuracy in recognising regionally inflected or non-standard forms of Malay.

By narrowing the focus to the tap [ɾ] and trill [r] variants—both of which are attested in connected speech among Standard Malay speakers—this study aims to fill a specific empirical gap. These two realisations are of particular interest due to their articulatory and acoustic salience. Through a detailed acoustic analysis of these sounds, this study contributes to the broader effort to map the phonetic landscape of Standard Malay in a systematic, data-driven manner. Accordingly, the following research questions are addressed:

1. What are the acoustic characteristics of the tap [ɾ] and trill [r] realisations of the phoneme /r/ as produced by Standard Malay speakers?
2. In which word environments do the tap [ɾ] and trill [r] tend to occur?

2. Literature Review

This section begins by exploring Standard Malay pronunciation, with particular attention to /r/ produced in different Malay dialects. It then examines three key areas that underpin the current study: cross-linguistic perspectives of /r/, the acoustic differences between the tap [ɾ] and trill [r], and the phonological realisation of /r/ in Malay. Together, these discussions provide the theoretical and empirical foundation for understanding how /r/ is produced and varies in connected speech in Standard Malay.

2.1 Standard Malay

Malay is the national and official language of Malaysia. It is the main language of intergovernmental communication (Abu Bakar, 2019) and is widely used for inter- and intra-ethnic communication. However, there are both geographical and social Malay dialects throughout Malaysia including two contact Malay varieties, Baba Malay and Chetti Malay (Austin & Pillai, 2020). Malay pronunciation is basically divided into two variations in Peninsular Malaysia: south- and north-centred (Shahidi & Rahim, 2010). In general, both have different realisations of /a/ and /r/ in the word-final position. For the north-centred Malay variants, orthographic *r* in a word-final

position is sometimes realised as a glottal stop /ʔ/ for example the word *besar* (‘big’) is produced as / bəsaʔ/. The examples for the two varieties are shown in Table 1.

Table 1. Realisations of Word Final *a* and *r* in Malay

Words	South-centred Malay	North-centred Malay
rasa (‘taste’)	/rasə/	/rasa/
apa (‘what’)	/apə/	/apa/
besar (‘big’)	/ bəsa /	/ bəsa/ or / bəsaʔ/

2.2 The Phoneme /r/

This section examines three key areas relevant to the present study: cross-linguistic perspectives of /r/, the acoustic differences between [ɾ] and [r], and the phonological realisations of /r/ in Malay.

2.2.1 *Cross-Linguistic Perspectives on /r/*

The phoneme /r/ is one of the most typologically variable consonants across the world’s languages, with considerable diversity in its articulatory and acoustic realisations. Ladefoged and Maddieson (1996) identified a range of articulations including taps, trills, approximants, and various fricatives, with languages differing significantly in which variants are used and how they are distributed. For instance, Spanish contrasts the alveolar tap [ɾ] and trill [r] phonemically, with minimal pairs such as *pero* (‘but’) and *perro* (‘dog’). In contrast, English typically uses a postalveolar approximant [ɹ], which does not contrast phonemically with other rhotic types but varies somewhat in realisation across dialects (Zharkova, 2016). In Tamil and other South Indian languages, /r/ is often realised as a retroflex approximant or tap, while in French, the uvular fricative [ʁ] predominates. The Czech language provides another example, where the phoneme /ř/ is realised as a raised alveolar non-sonorant trill, which is both acoustically and articulatorily complex (Howson, Komova, & Gick, 2014). This cross-linguistic diversity reflects not only the articulatory flexibility of /r/ but also their high susceptibility to phonetic variation due to sociolinguistic and aerodynamic factors.

Understanding these cross-linguistic patterns provides a useful comparative backdrop for analysing Malay /r/ realisations, particularly as Malay does not exhibit phonemic contrast between tap and trill but shows considerable variation in their contextual usage. Including such typological insights strengthens situating the present study within the broader field of acoustic phonetics and

supports the relevance of documenting rhotic variation in under-researched languages like Standard Malay.

2.2.2 *Acoustic Differences between the Tap and Trill*

The tap [ɾ] and the trill [r] are two commonly observed realisations of the /r/ phoneme across a range of languages, including Standard Malay. Although they share the same alveolar place of articulation, their production mechanisms differ significantly in both articulatory and acoustic terms, which has implications for their identification and analysis in connected speech.

From an articulatory perspective, a tap is characterised by a single, rapid contact of the tongue tip against the alveolar ridge, with no sustained airflow or vibratory pattern (Ladefoged & Maddieson, 1996). In contrast, a trill involves multiple rapid contacts of the tongue tip with the alveolar ridge, set in motion by a sustained airstream. This articulatory distinction results in different acoustic profiles, particularly in the domains of duration and formant behaviour.

Duration is a key distinguishing factor where the tap is a brief segment, typically lasting between 20 and 40 milliseconds, and appears as a single closure or burst on the waveform and spectrogram (Ladefoged & Maddieson, 1996). In contrast, the trill, due to its repeated closures, exhibits a longer segment duration, often exceeding 100 milliseconds, depending on speech rate and individual speaker physiology (Dhananjaya, Yegnanarayana, & Bhaskararao, 2012). This temporal contrast makes duration a valuable cue for differentiating between the two variants in acoustic analyses.

Another critical acoustic feature is the F1, which reflects changes in vocal tract configuration and glottal aperture. The production of trills often results in greater F1 fluctuation, due to repeated constrictions and expansions of the vocal tract during tongue vibration (Howson et al., 2014). Additionally, Dankovicová (1999) notes that the airflow variability associated with trills can be perceived as increased breathiness, which in turn affects F1 amplitude and contour. In contrast, tap production tends to result in a more stable F1 trajectory as the articulatory gesture is brief and not aerodynamically sustained.

Spectrographically, taps appear as a single transient closure, often accompanied by a brief silence or burst, while trills are visualised as a sequence of periodic vertical striations, each corresponding to a tongue-tip contact (Dhananjaya et al., 2012). These differences in spectrographic patterning support the use of acoustic tools to reliably distinguish between tap and

trill variants in empirical speech data. The examples of these sounds on the spectrograms are illustrated in Figures 1a and 1b.

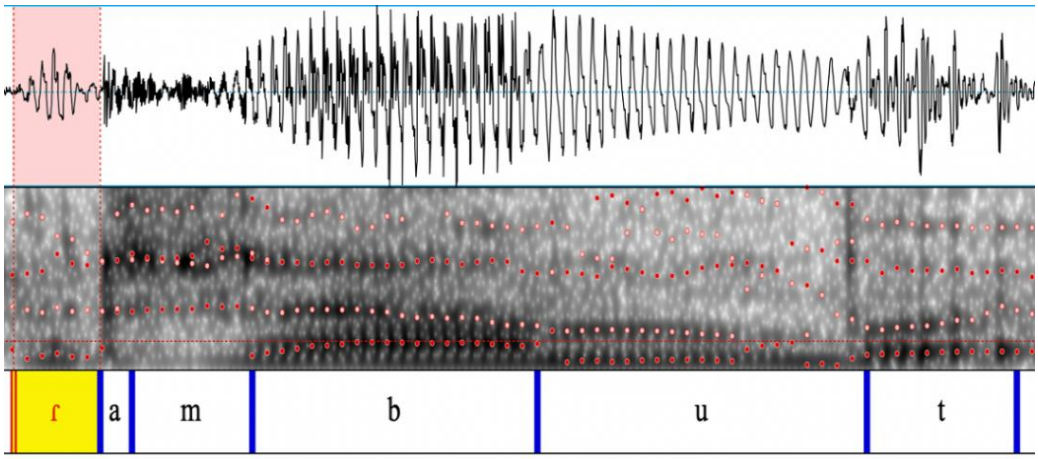

Figure 1a. An Example of a Tap [ɾ] in Word-Initial Position in the Malay Word *Rambut* ('Hair')

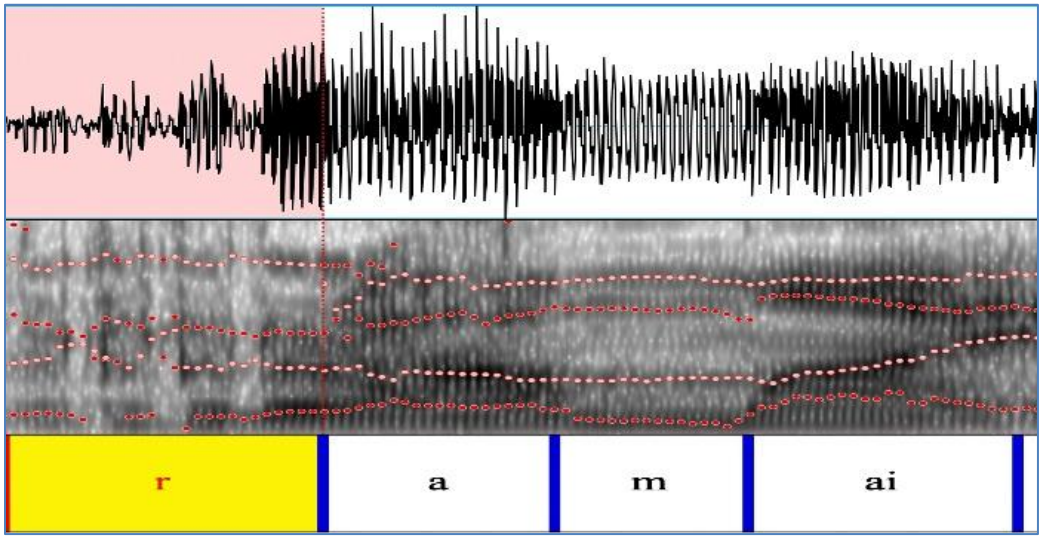

Figure 1b. An Example of a Trill [r] in Word-Initial Position in the Malay Word *Ramai* ('Many')

2.2.3 *Phonological Realisations of /r/ in Malay*

The phonological realisation of the /r/ phoneme in Malay demonstrates considerable variation, shaped by regional dialects, sociolinguistic factors, and individual speaker habits. Although Standard Malay is the official language of Malaysia and is widely used in education, media, and

government, its pronunciation norms are not uniformly realised across speakers. As previously mentioned, the /r/ phoneme in particular, exhibits a wide range of surface forms, including the tap [ɾ], the trill [r], and, in some dialects, approximants [ɹ] or fricatives such as [ʁ], [ɣ] or [ʕ] (Clynes & Deterding, 2011; Shahidi & Rahim, 2010). These variants are typically not phonemically contrastive within Malay but emerge as allophonic or contextually conditioned realisations.

One of the most consistent observations in previous studies is the omission or weakening of /r/ in word-final position, particularly among southern Malay dialects (Shahidi & Rahim, 2010). In such cases, words like *besar* ('big') may be realised as [bəsa], with the final /r/ entirely absent. This phenomenon is often attributed to dialectal simplification and a general tendency to reduce articulatory effort in coda positions. Conversely, in word-initial and intervocalic positions, speakers are more likely to produce an audible /r/, with some favouring a trill and others a tap depending on speech rate, emphasis, or formality.

The choice between a tap and a trill in connected speech appears to be governed more by phonotactic and prosodic context than by lexical distinction. Schmittauer (2024, p. 101) for example, found four variants of /r/ associated with different environments in her study of Klang Valley Malay: tap /ɾ/ was commonly produced in word-initial and intervocalic contexts as well as in consonant clusters; the approximant /ɹ/ was commonly found in preconsonantal contexts; zero realisation tended to occur word-finally; and the alveolar trill /r/, the least common variant in her data, tended to occur in word-initial or word-final contexts especially in read speech. It is posited that trills are more frequently observed in word-initial position or when the /r/ is adjacent to consonant clusters, possibly due to the increased articulatory force required to initiate the segment (Dhananjaya et al., 2012). Taps, on the other hand, commonly occur between vowels, where a more subtle articulatory gesture suffices.

Additionally, the influence of second-language acquisition may also contribute to /r/ variability. Many non-Malay Malaysians are multilingual, with different first languages (e.g., Cantonese, Hokkien, Mandarin, Tamil, Iban, Kadazan) and apart from English, may also speak other languages like Arabic. These languages may have their own *rhotic* systems, which may influence speakers realisation of /r/ in Malay. For instance, the influence of English may encourage an approximant-like [ɹ] and mirror zero-realisation in word-final position (Pillai, 2015) while Arabic may introduce uvular or pharyngeal rhotics (Clynes & Deterding, 2011; Idris Aman, 2000).

Despite this diversity, there remains a lack of systematic acoustic data detailing how these variants are distributed and realised in natural speech, particularly in Standard Malay as spoken by educated urban speakers. The present study seeks to address this gap by providing an empirical characterisation of the /r/ phoneme as produced in connected speech by a sample of native Malay speakers, with specific focus on their realisations of tap [ɾ] and trill [r] in Standard Malay speech.

The following section outlines the methodology employed to analyse the Malay speech data, which contains recorded instances of both tap [ɾ] and trill [r] realisations of the /r/ phoneme in connected speech.

3. Methodology

As noted earlier, initial auditory observations confirmed that the dataset includes both tap [ɾ] and trill [r] realisations produced by Standard Malay speakers. This section outlines the preparation of the speech materials, the criteria used for participant selection, and the procedures employed in the treatment and analysis of the data.

3.1 Speech Materials

For this study, a set of 18 Malay words containing the phoneme /r/ was selected and categorised based on the position of /r/ within each word: initial, medial, and final. Each category consisted of six common Malay words (see Appendix). To ensure consistency in the research context and minimise contextual complexity, all target segments are adjacent to the vowel /a/. This approach helps to facilitate accurate examination of the spectrogram and ensures the validity and reliability of the findings (Shahidi et al., 2012).

3.2 The Speakers

As this was a preliminary study, a small group of participants was selected. Four native speakers of Standard Malay comprising two males and two females aged between 18 and 25 years were recruited. Informal interviews conducted prior to recording confirmed that all participants were from the south-west coast of Peninsular Malaysia, specifically from Kuala Lumpur and Selangor. The selection was intentionally limited and controlled to minimise potential variation in speech due to regional or dialectal influences, thereby allowing a better observation of the patterns of /r/

realisations. Each participant signed a consent form, granting permission for their speech to be recorded and used for the study in accordance with ethical guidelines.

The study was conducted at the International Islamic University Malaysia, Kuantan Campus. All the speakers used Standard Malay. In this study, speakers of the same age group and ethnicity (Malay), and geographical origin (Kuala Lumpur and Selangor) were selected to minimise variations in pronunciation.

3.3 The Recording, Transcription and Annotation

All recordings were conducted in a quiet, low-reverberation room to minimise background noise and ensure clear audio quality. Participants were seated comfortably approximately 20 to 30 cm from a high-quality unidirectional condenser microphone, which was connected to a laptop via an audio interface. Recordings were made using Praat [version 5.3.61] (Boersma & Weenink, 2013). with a sampling rate of 44.1 kHz and a 16-bit resolution.

Participants were instructed to read aloud the word list (see Appendix), which were presented in a randomised order via a computer screen. To ensure natural and fluent speech production, and to embed the target word in a consistent prosodic and phonological environment, each participant was asked to read the word list in the context of a carrier phrase:

Saya sebut ___ sekarang
('I say ___ now')

Prior to the actual recording session, participants were given time to familiarise themselves with the word list and ask questions if needed. Multiple takes were allowed for any items that were unclear. All sessions were conducted individually, and participants were reminded to speak at a natural speaking pace and volume.

Each recording session lasted approximately 10 to 15minutes per participant. Following the session, the audio files were reviewed for clarity and completeness. High-quality tokens were then selected for acoustic analysis, and all recordings were saved in WAV format for subsequent processing in Praat.

The recordings of the 72 tokens underwent both orthographic and phonetic transcription based on auditory and simultaneous visual inspection of the spectrogram. The target sounds were

segmented and transcribed phonetically in TextGrids in Praat (see Figures 1a and 1b). The authors, who comprised a linguist and two speech-language therapists, then checked each annotated TextGrid and associated WAV files to ensure that the segmentation and transcription were correct.

3.4 Data Analysis

After the data were annotated and transcribed, the analysis was conducted in two phases: acoustic and statistical. This section outlines the procedures used in each phase to examine the phonetic realisations of the /r/ phoneme.

3.4.1 *Acoustic Analysis*

The acoustic analysis was conducted to examine the realisations of the phoneme /r/, specifically the tap [ɾ] and trill [r], through their measurable acoustic properties. For each token, both the first formant (F1) frequency and segmental duration were measured to characterise the articulatory and spectral distinctions between the two variants.

The F1 frequency was measured at the temporal midpoint of the /r/ segment to minimise transitional effects from surrounding vowels. Given that F1 correlates with tongue height and is influenced by glottal aperture and breathiness (Dankovičová, 1999; Howson et al., 2014), it served as a useful indicator of articulatory variation between the tap and trill. The measurements were taken with settings adjusted to reflect the vocal tract characteristics of adult speakers (i.e., maximum formant set to 5000 Hz for female speakers and 4500 Hz for males). This approach aligns with prior studies highlighting that the F1 of /r/ variants can reveal articulatory configuration and degree of constriction (Dhananjaya et al., 2012).

In addition to formant frequency, segmental duration served as a key acoustic parameter in distinguishing the tap [ɾ] and trill [r] realisations of the phoneme /r/. Duration was measured in milliseconds (ms) from the onset to the offset of the /r/. The tap [ɾ] typically exhibits a short, single-contact articulation, resulting in a brief duration, often less than 50 ms (Ladefoged & Maddieson, 1996). In contrast, the trill [r] involves multiple rapid tongue-tip vibrations, leading to a relatively longer segmental duration. This temporal difference is not merely articulatory but also reflects the biomechanical requirements for sustained tongue oscillation in trills, as suggested by Dhananjaya et al. (2012), who argued that trills tend to have longer durations due to reduced excitation coupling between the vocal folds and the tongue tip.

The acoustic data were subsequently compiled and statistically analysed to identify patterns in the distribution of tap and trill variants across phonological contexts, as well as their spectral and temporal characteristics.

3.4.2 *Statistical Analysis*

To answer the first research question, which is to acoustically describe the tap and trill realisations, descriptive statistics were first calculated for both the values of the F1 and segmental durations**.** These included the minimum**,** maximum**,** and mean for each variant across word positions (initial, medial, and final). This step was essential to identify overall trends and variation patterns in the data.

Meanwhile, for the second research question, to analyse the differences in the acoustic characteristics of the Malay /r/ sounds across speakers and word positions, the Kruskal–Wallis H test was utilised. This non-parametric statistical test was chosen as an alternative to one-way ANOVA due to the small sample size and the non-normal distribution of the data. Separate Kruskal–Wallis tests were conducted for the F1 values and the durations of /r/ realisations. The test compared median F1 values and durations across four participants, as well as across two-word positions (initial and medial), as in word-final position, the data were omitted due to consonant elision. This approach allowed for the evaluation of potential differences in the production of trill and tap variants of /r/ without assuming normality, thus ensuring the robustness of the statistical interpretation (Field, 2018; Pirinen et al., 2024).

4. Results

The analysis focused on two primary acoustic parameters: F1 frequency (Hz) and segmental duration (ms), assessed across different word positions (initial, medial, and final), in line with the research questions concerning phonetic realisation and contextual distribution. Preliminary findings indicated that the speakers produced /r/ as either a tap or trill in free variation irrespective of the phoneme's position within the word while they tended not to be realised in word-final position (zero realisation).

4.1 The Acoustic Characteristics of Tap [ɾ] and Trill [r]

The comparison of the two variants revealed systematic acoustic distinctions (see Table 2). In terms of duration, trills consistently exhibited longer durations across word positions compared to taps. For instance, in the word *ramai*, the trill had a mean duration of 127.7 ms, whereas in *rasa*, the tap variant had a mean duration of 84.0 ms. In medial position, similar patterns emerged: the trill in *draf* had a mean duration of 161.0 ms, while the tap in *dram* was relatively shorter (M = 103.7 ms). This is reflective of the way in which a tap is typically produced, i.e., with a single, brief tongue contact, resulting in shorter and more uniform durations. The shorter duration of a tap can be related to its phonetic efficiency, especially in rapid or connected speech, which may explain its wider distribution in medial contexts.

Differences in F1 frequency between taps and trills were noted although they were less consistent. F1 values for trills tended to show greater variability, possibly reflecting the complex tongue movements and more open glottal configuration associated with trilling (Howson et al., 2014). For example, in *ramai*, the F1 for trills ranged from 553.4 Hz to 882.9 Hz (M = 677.6 Hz), while in *sarapan*, it ranged from 616.8 Hz to 859.2 Hz. The tap, such as those in *rasa* and *sarapan*, generally exhibited narrower F1 ranges and slightly higher mean values although some overlaps with the values of the trill was evident. These F1 distinctions may be influenced by the degree of tongue bunching and glottal aperture, both of which affect the resonance of the vocal tract during the production of the trill. Trills, requiring sustained articulatory tension and airflow, might have led to more variable resonance patterns, and hence, the broader F1 spread.

Table 2. F1 Values (Hz), Duration (ms), and the Phonetic Realisation of /r/ Across Word Positions

		Phonetic realisation													
		[ɾ]							**[r]**						
Position of /r/	**Word**	**n**	**Hz**			**ms**			**n**	**Hz**			**ms**		
			Min	**Max**	**Mean**	**Min**	**Max**	**Mean**		**Min**	**Max**	**Mean**	**Min**	**Max**	**Mean**
Initial	*ramai* (‘many’) /ramaɪ/	1	-	-	632.6	-	-	148.0	3	553.4	882.9	677.6	84.0	162.0	127.7
	rambut (‘hair’) /ram.bʊt/	0	-	-	-	-	-	-	4	668.6	734.3	716.2	89.0	162.0	134.0
	rasa (‘taste’) /ra.sa/	3	634.3	759.5	696.9	80.0	88.0	84.0	2	634.0	634.8	634.4	120.0	128.0	84.0
	raya (‘celebration’) /ra.ja/	1	-	-	576.3	-	-	121.0	3	586.7	663.1	614.3	124.0	175.0	155.3
	sarapan (‘breakfast’) /sa.ra.pan/	2	652.0	837.7	744.8	87.0	133.0	110.0	2	616.8	859.2	738.0	86.0	104.0	95.0
	gembira (‘happy’) /gəm.bi.ra/	1	-	-	532.4	-	-	151.0	3	439.3	685.2	539.4	97.0	170.0	144.7
Medial	*kraf* (‘craft’) /kraf/	3	609.4	646.7	622.4	61.0	87.0	76.0	1	-	-	671.4	-	-	105.0
	prasekolah (‘preschool’) /pra.sə.kɔ.lah/	2	590.4	655.1	622.7	99.0	105.0	102.0	2	650.7	841.7	746.2	95.0	110.0	102.5
	skrap (‘scrape’) /skrap/	3	580.8	856.6	689.4	71.0	98.0	86.3	1	-	-	638.1	-	-	101.0
	strategi (‘strategy’) /stra.tə.gi/	3	560.7	598.1	577.7	68.0	89.0	81.0	1	-	-	419.7	-	-	107.0
	draf (‘draft’) /draf/	2	624.5	635.5	630.0	84.0	128.0	106.0	2	587.6	660.4	624.0	154.0	168.0	161.0
	dram (‘drum’) dram/	3	601.9	700.6	662.7	87.0	137.0	103.7	1	-	-	704.4	-	-	270.0
Final	*lapar* (‘hungry’) /la.par/				-			-				-			-
	hantar (‘send’) /han.tar/				-			-				-			-
	besar (‘big’) /bə.sar/				-			-				-			-
	biar (‘let something be’) /bi.jar/				-			-				-			-
	kejar (‘chase’) /kə.d͡ʒar/				-			-				-			-
	luar (‘outside’) /lu.ar/				-			-				-			-

4.2 The Phonological Contexts of Tap [ɾ] and Trill [r]

To answer the second research question, the number of occurrences for both [ɾ] and [r] was calculated (Table 3). The distribution of the /r/ phoneme revealed a positional preference between the tap [ɾ] and the trill [r]. In word-initial position, the trill was more frequent, occurring 17 times compared to seven occurrences of the tap. Conversely, in the medial position, the tap dominated with 16 instances, while the trill appeared only eight times. This suggests a tendency for Standard Malay speakers in this study to produce [r] more often at the beginning of words and [ɾ] more frequently in word-medial position.

Table 3. Distribution of Tap [ɾ] and Trill [r] Realisations by Word Position in Standard Malay

Position	[ɾ]	[r]
Initial	7	17
Medial	16	8

To determine whether the observed acoustic differences between the tap and trill were statistically significant, A Kruskal-Wallis–Wallis H test was applied to compare the distributions of F1 and duration across the two variants. The analysis revealed no statistically significant differences in F1 values across the four participants ($\chi^2 = 3$, $p = 0.08$) nor in the duration of /r/ production ($\chi^2 = 3$, $p = 0.06$). Similarly, comparisons of F1 values and durations between initial and medial word positions also indicated no significant differences ($p = 0.08$). Although no statistical significance was found, descriptive observations suggested that /r/ sounds produced at the word-initial position tended to exhibit higher F1 values and longer durations compared to those in word-medial position.

5. Discussion

The results of this study highlight several key insights into the acoustic realisation of the phoneme /r/ among Standard Malay speakers, particularly regarding its contrastive manifestations. Notably, the findings suggest that Standard Malay aligns with a class of world languages that exhibit trilled /r/ sounds. More specifically, this study's main finding is that speakers alternated between producing the trill [r] and the tap [ɾ]. The latter was more frequent in medial position, while trills appeared more commonly in word-initial position. The finding on taps differs from Schmittauer

(2024) who reported that they were commonly produced in word-initial position. However, although not common in her study, trills were also found in word-initial environments. In addition to taps and trills, a consistent deletion of /r/ in word-final position was found in this study, which is similar to Schmittauer (2024). The different realisations of the tap and trill in this study may be related to prosodic and articulatory factors. This is because initial positions often allow more time for complex articulations like the trill, while medial environments favour shorter, more efficient sounds such as the tap. This observation aligns with cross-linguistic trends, as discussed by Ladefoged and Maddieson (1996), where the distribution of /r/ variants often depends on syllable structure and speech rhythm.

In relation to word positions, our findings found that word-initial /r/ consistently yielded the highest F1 and longest duration, reinforcing its association with the articulation of a trill. In contrast, medial /r/ tended to have shorter durations and lower F1 values, consistent with the realisation of a tap and its reduced articulatory demands (Ahmad, 2006). The differences in F1 and duration also served to acoustically differentiate between the tap [ɾ] and trill [r]. Thus, from an acoustic perspective, F1 values and duration were key indicators in distinguishing between the realisation of a tap and trill in this study (Howson et al., 2014). Although the female speakers generally produced higher average F1 and longer segmental duration than the male speaker, potentially due to their typically clearer and sharper voice qualities (Shinde et al., 2017), the overall acoustic characteristics across all speakers were not significantly different.

In fact, the statistical results support the auditory observation that Standard Malay speakers produce both tap and trill variants of /r/, with distinct acoustic signatures. These findings align with previous research such as Dhananjaya et al. (2012) and Howson et al. (2014). As mentioned previously in this paper, the former noted that trills typically exhibit longer durations due to the physiological requirements for multiple tongue-tip vibrations, particularly in word-initial position where the onset of vocal tract movement demands greater aerodynamic coordination. Further, Howson et al. (2014) associated variations in F1 with articulatory breathiness and glottal width. Taps, produced with a single, brief tongue contact, typically exhibit shorter durations and slightly lower F1 values, reflecting a quicker, less complex articulatory gesture. These findings reinforce the notion of phonetic and phonological variation of the realisations of /r/ and contribute to understanding how these variants pattern across speech contexts.

In addition, understanding such acoustic features (e.g., F1 and duration) can help to objectively assess pronunciation variations. Putri, Subagia and Pratama (2021), for example, indicated that difficulties associated with the production of [r] can be detected by observing the differences in the acoustic features of this sound produced by people with dysarthria and those with typical speech. Moreover, given that final /r/ omission is a common feature in Standard Malay, intervention can be shifted on appropriate contextual use rather than enforcing a rigid [r] articulation in all word positions.

To reiterate, trills were associated with higher F1 values and longer durations compared to taps. Additionally, the fluctuations in F1 and duration were observed across word positions, implying that other phonetic or prosodic factors influence /r/ production beyond positional placement alone. This distinction in acoustic parameters also supports the interpretation of the two possible variants as distinct allophonic realisations within the phoneme system of Standard Malay. Thus, the variations observed in this study may have phonetic and clinical relevance for understanding patterns of trill and tap production in connected speech. Additionally, in natural speech, only using one realisation of /r/, for example the trill, may render speech pragmatically odd or unnatural within the Malay-speaking context.

6. Conclusion

This study demonstrated that /r/ may be realised as two main variants in Standard Malay: the trill [r] and the tap [ɾ], which occur interchangeably depending on phonetic context. The application of acoustic analysis in this study has provided a detailed account of the acoustic properties of /r/, facilitating a more precise understanding of its phonetic realisation in Standard Malay. The ability to examine these acoustic patterns contributes to broader discussions on phonetic variability and language standardisation. Understanding the phonetic realisation of /r/ in Standard Malay has implications for the teaching and learning of Malay, speech therapy, and second language acquisition. Better understanding of the phonetic variants of /r/ can contribute to improved pronunciation instruction, (Shahidi et al., 2012). This is particularly important in a multilingual context like Malaysia, where learners may be exposed to various phonetic influences from different languages and dialects.

The findings can be useful for assessment and intervention strategies used in clinical settings by speech-language therapists in managing speech sound disorders. This variability can

be particularly pronounced, especially in Malaysia, where multiple languages and Malay dialects coexist. Understanding these variations may assist them in identifying and differentiating typical and atypical pronunciations of Malay /r/. These include the realisation of a trill or tap and even the omission of final /r/ in connected Standard Malay speech. Such awareness can help prevent unnecessary intervention where variations may be perceived as disordered speech rather than as a natural phonetic diversity (Fabiano-Smith, 2019).

In relation to contributing to clinical applications, the findings can contribute to standardised assessment tools specific to the Malay language for screening and assessing individuals with speech disorders. The need for language-specified assessment tools has been discussed and highlighted in past studies where it has been shown that available assessment tools may lack validity for differential diagnosis in bilingual or multilingual children (Fabiano-Smith, 2019, Goldstein & Gildersleeve-Neumann, 2015).

The clinical implications of acoustically characterising the Malay phoneme /r/ can, therefore impact assessment, intervention, and the overall understanding of speech sound disorders in diverse linguistic contexts. By enhancing diagnostic accuracy, informing culturally relevant assessments and guiding targeted interventions, this study can significantly support more precise clinical assessment and intervention in the future.

6.1 Limitations of the Study

Although the limited number of speakers constrains the study's generalisability, its significance lies in serving as an initial step towards a more comprehensive analysis of the production of /r/ in Malay. Future studies should build on these findings by expanding the sample size and incorporating a wider range of phonetic contexts to deepen our understanding of Malay phonology. Inter-speaker differences, such as age, regional background, and sociolinguistic influences, are crucial factors that require broader representation in order to draw more comprehensive conclusions. Future research should therefore incorporate a larger and more diverse speaker pool to validate these preliminary findings.

6.2 Recommendations for Future Research

However, despite this limitation, the study provides valuable insights into the acoustic realisation of /r/ in Standard Malay. As a preliminary investigation, it establishes an empirical foundation for

future research by identifying key trends in /r/ production, such as the alternation between trill [r] and tap [ɾ] and the omission of /r/ in word-final. These initial findings contribute to our understanding of Malay phonetics and highlight areas for further exploration, particularly in relation to dialectal variations and speech processing.

References

Abu Bakar, M. (2019). Sebutan Johor-Riau dan sebutan baku dalam konteks identiti masyarakat Melayu Singapura [The Johor-Riau and standard pronunciation in the context of the identity of Singapore Malay society]. *Issues in Language Studies*, *8*(2), 61–78. https://doi.org/10.33736/ils.1521.2019

Ahmad, Z. (2006). *The phonology-morphology interface in Malay: An optimality theoretic account*. Pacific Linguistics.

Austin, P. K., & Pillai, S. (2020). Language description, documentation and revitalisation of languages in Malaysia. *Journal of Modern Languages*, *30*(1), 1–5. https://doi.org/10.22452/jml.vol30no1.0

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* [Computer programme]. Retrieved from www.praat.org

Clynes, A., & Deterding, D. (2011). Standard Malay (Brunei). *Journal of the International Phonetic Association*, *41*(2), 259–268. https://doi.org/10.1017/s002510031100017x

Dankovičová, J. (1999). *The domain of articulation rate variation in Czech*. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, & A. C. Bailey (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 429–432). University of California, Berkeley.

Dhananjaya, N., Yegnanarayana, B., & Bhaskararao, P. (2012). Acoustic analysis of trill sounds. *The Journal of the Acoustical Society of America*, *131*(4), 3141–3152. https://doi.org/10.1121/1.3688470

Fabiano-Smith, L. (2019). Standardized tests and the diagnosis of speech sound disorders. *Perspectives of the ASHA Special Interest Groups, 4*(1), 58–66. https://doi.org/10.1044/2018_pers-sig1-2018-0018

Goldstein, B. & Gildersleeve-Neumann, C. (2015). Bilingualism and speech sound disorders. *Current Developmental Disorders Reports*, *2*(3), 237–244. https://doi.org/10.1007/s40474-015-0049-3

Gurevich, N. (2021). *Acoustic correlates of rhotic allophony: A cross-linguistic perspective*. In R. C. Mora & M. J. Solé (Eds.), *Phonological typology* (pp. 115–139). Oxford University Press.

Howson, P., Komova, E. & Gick, B. (2014). *Czech trills revisited: An ultrasound EGG and acoustic study. Journal of the International Phonetic Association*, *44*(2), 115–132. https://doi.org/10.1017/s0025100313000339

Idris Aman. (2000). *Bahasa dan identity* [Language and identity]. Dewan Bahasa dan Pustaka.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.

Pirinen, F. A. (2024). Keeping Up Appearances—or how to get all Uralic languages included into bleeding edge research and software: Generate, convert, and LLM your way into multilingual datasets. In *Proceedings of the 10th International Workshop on Computational Linguistics of Uralic Languages (IWCLUL 2024*) (pp. 142–153). Association for Computational Linguistics. Retrieved from https://aclanthology.org/2024.iwclul-1.16/

Pillai, S. (2015). Rhoticity in Malaysian English: The emergence of a new norm? In U. Gut, E.-M. Wunder and R. Fuchs (Eds.), *Universal or diverse paths to English phonology* (pp. 23-40). Mouton de Gruyter.

Putri, D. A. D. P., Subagia, I. K., & Pratama, A. D. Y. (2021). Acoustic features of phonemes as therapy reference for dysarthria disorder: the case of /r/ alveolar trill. *Linguistics and Culture Review*, *6*, 129–143. https://doi.org/10.21744/lingcure.v6ns2.1980

Schmittauer, K. (2024). A study of the variants of /r/ in Klang Valley Malay [Unpublished master's dissertation]. Universiti Malaya.

Shahidi, A. H., & Rahim, R. (2010). Kajian akustik realisasi kontras penyuaraan bunyi plosif Bahasa Melayu [Acoustic study of the realisation of the plosives voicing contrast in Malay]. *GEMA Online Journal of Language Studies, 12*(2), 297–317. Retrieved from https://journalarticle.ukm.my/4941/1/pp%2520745_760.pdf

Shahidi, A. H., Latiff, A. A., Syabri, M. S., Fauzi, M. F., & Makhtar, R. (2020). Penelitian akustik dialek Kedah [Acoustic research on the Kedah dialect] . *Jurnal Melayu*, Special Issue, 577–594. Retrieved from https://ejournal.ukm.my/jmelayu/article/view/44411/11414

Shinde, D. K., More, Y. N., Sonawane, A. S., Pawade, V. S., & Kulkarni, K. V. (2017). Comparison of male and female vowel formants by native speakers of Marathi language. *International Journal of Engineering Research and General Science*, *5*(1), 173–178. Retrieved from https://oaji.net/articles/2017/786-1488211081.pdf

Zharkova, N. (2016). Ultrasound and acoustic analysis of sibilant fricatives in preadolescents and adults. *Journal of the Acoustical Society of America*, *139*(5), 2342–2351. https://doi.org/10.1121/1.4947046

Appendix

Eighteen (18) target words with /r/ in different word positions

Position of /r/	**Word**
Initial	*ramai* (many) /ramaɪ/
	rambut (hair) /ram.bʊt/
	rasa (taste) /ra.sa/
	raya (celebration) /ra.ja/
	sarapan (breakfast) /sa.ra.pan/
	gembira (happy) /ɡəm.bi.ra/
Medial	*kraf* (craft) /kraf/
	prasekolah (preschool) /pra.sə.ko.lah/
	skrap (scrap) /skrap/
	strategi (strategy) /stra.tə.gi/
	draf (draft) /draf/
	dram (drum) /dram/
Final	*lapar* (hungry) /la.par/
	hantar (send) /han.tar/
	besar (big) /bə.sar/
	biar (let something be) /bi.jar/
	kejar (chase) /kə.dʒar/
	luar (outside) /lu.ar/