

A Study of Sociolinguistic Variation in Malaysian Mandarin Tones

Xin Ren
renxin0221@jnu.edu.cn
College of Chinese Language and Culture, Jinan University
&
Faculty of Languages & Linguistics, Universiti Malaya
50603 Kuala Lumpur, Malaysia
<https://orcid.org/0009-0000-3172-3015>

Received: 1 July 2025

Accepted: 14 August 2025

Abstract

In multilingual and multiethnic Malaysia, Malaysian Mandarin (MalM) is a heterogeneous variety. Due to the linguistic diversity in Malaysia, MalM has drawn scholarly attention to the phonetic variation of this variety. This study investigates the citation tone production in MalM across language and geographical groups. Seventy-one fluent Mandarin speakers were recruited, stratified by their dominant home languages and geographical regions. Findings from the acoustic and statistical analyses suggest a close resemblance of citation tones in MalM and Putonghua, the standard Mandarin used in mainland China, except for T3's tone contour. Additionally, there are some notable differences correlated with their primary language and their geographical regions in the tonal realizations in T3 and T4. This suggests that the dominant language and geographical background of the speakers possibly influence citation tone production in MalM. Overall, this study identifies that MalM is not a homogeneous variety in the heterogeneous Malaysia and supports the perspective that MalM is a unique variety within the Global Chinese paradigm.

Keywords: Global Chinese, Malaysian Mandarin, Multilingual Speakers, Sociophonetic Study, Tonal variation

1. Introduction

The national and cultural boundaries of Mandarin have been expanded, indicating the inclusive and hybridized nature of this language (Wang, 2010). Currently, Mandarin is considered a newly emerging global language. Global Chinese has been proposed to conceptualize the international spread of Mandarin. Mandarin not only serves as a common language in China, but also as a lingua franca or common language among the overseas Chinese communities in multilingual societies (Goh, 2017). Thus, understanding the features of different Mandarin varieties is vital for enhancing intercultural communication. Malaysian Mandarin (MalM) is a typical overseas Mandarin variety. Malaysia is a multiethnic society with three predominant ethnic groups: Malay, Chinese, and Indian, along with a minority of indigenous groups. Correspondingly, Malaysia is multilingual, with Malay as the national language (Pillai et al., 2010), English as the second language (Wang, 2016), and Tamil and Mandarin used in the local Indian and Chinese communities, respectively (Wang, 2016). Khoo (2017) indicated that the Chinese community in Malaysia remains in a unique dialectal grouping in many townships, as major southern Chinese dialects have retained their strongholds as regional lingua franca, such as Cantonese in Kuala Lumpur and Ipoh, and Hokkien in Penang.

In contrast, more recent research suggested that the Chinese community in Malaysia is becoming increasingly Mandarin dominant, as Mandarin has replaced the position of the southern Chinese dialects in most functional domains, including the home domain (Wang, 2024). Thus, the home dominant language might be heterogenous across the local Chinese community in Malaysia. Flege et al. (2003) suggested that bilinguals differ in their patterns of language use, which contribute to the differences in their speech production. Thus, the differences in the home language of Chinese Malaysians might result in the differences in their speech production. Additionally, this phonetic variation is interwoven with geographic distribution. Based on the personal observation, Guo (2017) suggested that Chinese Malaysians' accents vary across different states in Malaysia. Specifically, the central and southern parts of Peninsular Malaysia represent the so-called "federal accent", while the northern and eastern parts of Peninsular Malaysia also exhibit different accents (Guo, 2017). However, thus far, few studies have investigated the roles of the dominant home language and geographical distribution in the phonetic realizations of MalM. Hence, Malaysia offers a fascinating case to investigate the effects of sociolinguistic factors in phonetic variation and change in Mandarin. Furthermore, Mandarin is a tonal language, where the lexical tones are used to contrast semantic meanings. A systematic review by Xu and Zhang (2024) suggests that citation tones,

the lexical tones in isolation, serve as the entry into larger linguistic units. Against these backdrops, this study reports sociophonetic variation in the citation tones in MalM based on the acoustic experimental data. Two research questions were addressed in this study:

- (1) What are the acoustic features of MalM tones?
- (2) To what extent do language and geographical backgrounds influence the tone production of Chinese Malaysians?

2. Literature Review

The lexical tones of Putonghua, the standard Mandarin in mainland China, have been extensively explored for decades (e.g., Bei, 2021; Chao, 1930; Deng et al., 2006; Shi, 2020; Shi & Ran, 2011). Four lexical tones have been well-established in Putonghua: Tone 1 (T1: high-level), Tone 2 (T2: mid-rising), Tone 3 (T3: low-dipping), and Tone 4 (T4: high-falling) (Chao, 1930). Chao's tone letters have been employed to describe how the four lexical tones are transcribed: T1 as 1, T2 as 1, T3 as ㄨ and T4 as ㄨ. Each tone letter represents a simplified time-pitch contour drawn to the left of a vertical reference line that serves as the tone stave (Chao, 1930). Chao's numerals are often adopted as an alternative way to transcribe the four lexical tones in Mandarin. Specifically, based on Chao's numerals (1930), a speaker's normal pitch range is divided into five-level tone numerals from 1 to 5. Hence, the four tones (T1 to T4) in Putonghua can be described as [55], [35], [214], and [51] according to Chao's numerals. Among the lexical four tones, T3 varies based on the preceding tones. A citation T3 (214) becomes a mid-rising tone when it precedes another T3 (Yip, 2002). This phonological alternation is known as T3 sandhi. Additionally, the citation T3 becomes a low-falling tone (21) when it precedes any tones other than T3, and this phonological alternation is called half-T3 sandhi (Yip, 2002). The dipping T3 mainly occurs in isolation or before pauses (Fon & Hsu, 2007). Chao's tone representation system is a convenient abstraction used to understand tones. However, the practical productions of speakers do not always follow this abstract system, and therefore some variations in tone value might occur in Mandarin varieties. Based on previous acoustic and perceptual results, Shi (2020) suggested that in Putonghua, T2 as [35], [24], or [12], and T3 as [214], [212], or [101] are acceptable, as these slight differences do not compromise intelligibility. This indicates the dynamic nature of tones. Deng et al. (2006) suggested that the four tones are acoustically realized as [55], [35], [212], and [51]. In Bei's (2021) acoustic research, the tonal realizations of T1, T2, and T4 align with those in the study by Deng et al.'s (2006). However, T3 in Bei's (2021) research is realized as [213].

A handful of studies have investigated tonal variation in MalM (e.g., Chen, 2007; Huang, 2016; Ren & Chiew, 2024). According to previous studies, MalM tones show substantial phonetic variation from Putonghua. Based on the impressionist approach, Chen (2007) depicted the four tones in MalM as [33]/[44], [24/35], [21], and [31/53]. Studies also focused on tones in specific MalM varieties. For example, Huang (2016) investigated the Mandarin tone production of Chinese Malaysians in Penang, a state located on the northwest coast of Peninsular Malaysia. The speakers in this study used Hokkien as their first language and based on acoustic analysis of the tones, the four tones were [33], [23], [21], and [53]. Another acoustic study (Khoo, 2018) examined the MalM tones produced by speakers in Kluang, Johor, located in the southern part of Peninsular Malaysia. Khoo (2018) distinguished four tones, represented as [55], [21], [21] and [32/21/51] based on the reading of disyllabic words and interviews. A more recent study by Ren and Chiew (2024) investigated Chinese Malaysians whose home dominant language is at least one southern Chinese dialect. Their study showed that the main tonal realizations of MalM tones are phonetical [44], [223], [31] and [52].

In these descriptive studies focusing on the MalM tones, the focus tends to be on speakers from the same language and geographical background. Despite a larger geographical coverage of speakers, Ren and Chiew (2024) did not discuss the role of geographical factors in their speakers' productions. Additionally, the speakers who use Mandarin as the dominant home language have been rarely investigated despite the increasing role of Mandarin in the Chinese community (Pillai & Zainuddin, 2025). Thus, the effects of sociolinguistic factors in shaping Chinese Malaysians' productions have received less attention, although sociophonetic studies have progressed rapidly. Sociophonetics is emerging as a subfield of its own, which is the interface of sociolinguistics and phonetics (Baranowski, 2013). Language contact is an important theoretical space for Chinese sociophonetic studies (Zhang et al., 2024). Previous studies have discussed the factors that drive language contact, such as language dominance and geographical distribution. Importantly, bilingual speakers seldom possess an equal command of their languages (Grosjean, 1998). The dominant language is referred to the language that a bilingual more accesses to and activates in a daily basis (Harris et al., 2006). However, multiple factors have been reported as crucial in deciding the dominant language, such as regular use of the language (Tsui et al., 2019) and language proficiency (Birdsong, 2006). Considering the cases of Chinese Malaysians in the current study, speakers differ in the language use pattern at the home domain. Thus, the dominant home language becomes a critical factor when comparing speech production within the Chinese community in Malaysia. In addition, as mentioned earlier in this paper, Malaysia is heterogeneous as different regional lingua franca

are used. This indicates that speakers may be varied in the quantity of input of languages in social contexts. Accordingly, geographical distribution seems to be another crucial factor in shaping Chinese Malaysians' speech production. This study addresses the gap by investigating tone production among Chinese Malaysians from different language and geographical backgrounds.

3. Methodology

3.1 Speakers

A total of 71 speakers participated in this research, aged from 22 to 25 years old (mean = 22, SD = 1.09). As Pillai et al. (2010) suggested, the recordings from females and males are better analyzed separately due to pitch differences between gender. Importantly, it is well documented that adult males appear to have lower mean F0 and narrower F0 range, while females appear to have higher mean F0 and wider F0 range (Lai & Kuang, 2020). Additionally, Ren and Wang (2025) suggested that female and male are in different stages in tonal changes in Macanese Mandarin, with female leading this change. Hence, all speakers were female in this study to avoid the gender differences. The younger generation was selected because they have typically grown up in a multilingual setting and have a relatively consistent education background. A total of 36 speakers were in the dialect group because they used one southern Chinese dialect as the home-dominant language, i.e., Hokkien, Cantonese, and Hakka, the three major southern Chinese dialects used in the local Chinese community. As these Chinese dialects all belong to the southern Chinese dialect (Choo, 2015), these speakers were treated as a single group. A total of 35 speakers were in the Mandarin group, as their dominant home language was Mandarin. Table 1 shows the geographic distribution of the speakers in the two language groups. Guo (2017) suggests that the Mandarin accent varies across central, northern and eastern regions in Peninsular Malaysia. However, the study does not provide details on how these geographical divisions were defined. Thus far, no standardized geographical division has been applied in linguistic research, particularly in Mandarin in Malaysia. Accordingly, to avoid any potential bias in the linguistic research, the present study, in exploratory nature, adopted the geographical distribution based on the administrative classifications of Peninsular Malaysia used for government (Ministry of Housing and local government Malaysia, 2005). This was also because this geographical division aligns with Guo's (2017) study and was likely to address the current research gaps. Given the lack of standardized divisions in linguistic research, it is acknowledged that alternative geographical classifications may be used in other

studies. As displayed in Table 1, this study categorized these speakers based on the following regions: central, northern and eastern areas of Peninsular Malaysia, as well as East Malaysia (EM). East Malaysia was included as a separate geographic group due to its distinct historical and geographical context. All information pertaining to demographic and language backgrounds were based on speakers' answers in a questionnaire. No speaker reported any speaking or hearing disorders.

Table 1. Geographic Distributions of the Speakers

Language group	Regions	State/city	Number of speakers
Dialect	Central and southern	Kuala Lumpur, Selangor, Negeri Sembilan, Johor	16
	Northern	Penang, Perak	12
Mandarin	Eastern	Pahang, Kelantan	3
	East Malaysia	Sarawak, Sabah	5
	Central and southern	Selangor, Negeri Sembilan, Johor	10
	Northern	Penang, Perak, Kedah	11
	Eastern	Pahang	2
	East Malaysia	Sarawak, Sabah	11

3.2 Speech Materials and Data Collection

To remove using Putonghua as the baseline, the speech material was designed based on Chinese dialectology. In Chinese dialectology, the Chinese language is considered to have developed from Middle Chinese, which features eight tones (Middle Chinese Tone 1 to Middle Chinese Tone 8) (Chen, 2000). Hence, the monosyllabic words were compiled according to the eight tones, as shown in the Appendix. The syllable structures were also balanced according to Middle Chinese phonology, including open syllables (CV), syllables with a nasal ending (CVN), and checked syllables (syllable ending with a stop coda, or a glottal stop, CVP) (Chen, 2000). Additionally, the vowel height was balanced by including both high and non-high vowels, as well as front and non-front vowels. All the words were common and frequently used by Chinese Malaysians. A word list by Ren and Chiew (2024) was adopted to elicit the speech data, as this word list fulfilled all the criteria.

The speech data were recorded remotely using the smartphones and earphones due to the constraints caused by the COVID-19 pandemic at the time of data collection. A handful of studies have validated the reliability of smartphone recordings in investigating language variation and change (e.g., Guan & Li, 2021; Hilton & Leemaan, 2021). Stimuli were presented on Google slides one-by-one in Simplified Chinese character. To avoid any effects of order,

the elicitation order was randomized, and two fillers were added at the beginning and end of the slides. Speakers were required to read the wordlist in a natural way twice with a break between each set.

3.3 Segmentation and Data Analyses

Speech data were transcribed manually via Praat (Boersma & Weenink, 2021). The tonal domain was measured based on the onsets and offsets of the second formant of the vowel. The boundaries of tonal domain were adjusted according to the spectrogram via ProsodyPro (Xu, 2013).

The F0 values of the tonal domain were extracted at 10 equidistant points. However, the first and last 10% points were removed to avoid syllable-edge effects, resulting in eight equidistant points (Peng, 2006). To eliminate the inter-speaker variation, the raw F0 data in Hz was transformed to T-value via the formula displayed in (1). The normalized F0 values were scaled to the range of [0, 5] to align with Chao's (1930) five-scale tone notation, with each interval corresponding to one numeral in Chao's (1930) numerals.

$$T_i = [(\log P_i - \log P_{\min}) / (\log P_{\max} - \log P_{\min})] \times 5 \quad (1)$$

The F0 contours were visualized via line charts. F0 onset, F0 offset and F0 range were used as the acoustic correlates in the statistical validation. F0 range was calculated based on Zhang and Gu (2023), using the formula (2):

$$F0_{\text{range}} = F0_{\text{offset}} - F0_{\text{onset}} \quad (2)$$

The F0 value differences between offsets and onsets were mainly used for straight F0 contours (Zhang & Gu, 2023). However, an initial falling portion might be present in the rising tones in MalM (Ren & Chiew, 2024), this study used F0 value differences between offset and minimum F0 to calculate the rising tones. Additionally, for falling tones, the absolute values were used to show F0 range because it captures the exact magnitude of the pitch movement regardless of the direction.

All statistical analyses were performed via R (R Development Core Team, 2023). Linear Mixed-effects models were performed via the *lme4* R package (Bates et al., 2015). F0 onset, F0 offset, and F0 range were used as the dependent variables, while language background and geographic background were modeled as the fixed variables. Speakers and tokens were involved as the random intercepts. Separate models were fitted for each tone category (T1, T2, T3 and T4) to simplify the interpretation. This study approximated the degrees of freedom using the Kenward-Roger method with the *lmerTest* package (Kuznetsova et al., 2017) and the

pbkrtest package (Halekoh & Højsgaard, 2014) in R. The Tukey method was used to adjust the *p*-values in post-hoc analyses. An alpha level of *p* was set at < 0.05 for all statistical tests in this research.

In the model fitting, this study followed the ‘keep it maximal’ rule by Barr *et al.* (2013) by including both by-speakers and by-token random slopes, and their intercepts for all the relevant fixed effects. Following Wu *et al.* (2023), data with positive skewness were logarithmically transformed. Subsequently, this study performed model criticism by excluding data points whose standardized residuals were greater than 2.5 standard deviations (Wu *et al.*, 2023). The diagnostic plots suggested that the residuals were nearly normally distributed and did not violate the assumption of linearity or homoscedasticity. Therefore, the models were valid for all statistical tests. In the present research, both acoustic and statistical analyses were performed on the main F0 trajectories, the trajectories with the highest frequency of occurrences. This was done because the present study aimed to examine the role of sociolinguistic factors in influencing the main tonal realization in MalM, but not to provide a detailed description of the speakers’ tone productions.

4. Findings

4.1 Acoustic Realizations of Tone Contours

Figure 1 offers an overview of the comparison of the main F0 trajectories across the four tones based on each language group in each of the four regions. In each plot, T-values are presented on the Y-axis while the normalized time is shown on the X-axis with eight points. The shaded area of each F0 contour shows 95% confidence intervals.

Figure 1 shows that the two language groups were generally similar in their major tonal realizations. T1 was realized as a high-level tone, T2 was produced as a mid-rising tone, T3 was uttered as a mid-falling tone, and T4 was a high-falling tone. T1 was generally parallel in each region in each language group, despite a slightly higher F0 height in the Eastern region, particularly in the dialect group. In terms of T2, all groups showed an initial fall followed by a flat stretch before a final rise. Interestingly, in terms of T2, the speakers in the dialect group of the Eastern region started and ended higher than those in the other regions in each language group. In T3, F0 onset was generally in mid to high tone height. For example, F0 onsets of T3 were around 3 in Central-southern, Northern, and EM regions in both language groups. Nevertheless, T3 started relatively higher in the Eastern region across the language group. F0 offsets of T3 were generally around 2; however, in the dialect group, T3’s offsets were

evidently higher in the Eastern and EM regions than those in the other regions in both language groups. Regarding T4, its F0 onsets were similar across each region in both language groups. However, T4 showed a much higher F0 offset in the Eastern region compared to other regions across both language groups. Overall, it appears that speakers in the Eastern region had higher F0 onsets and offsets, particularly in the dialect group.

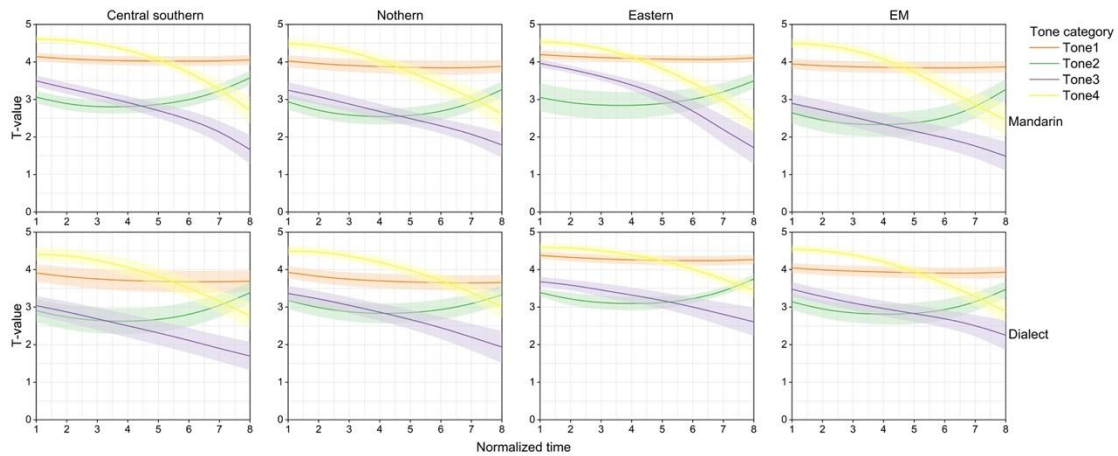


Figure 1. Main F0 Trajectories of Four Tones in MalM in Relation to Language and Geographical Groups

4.2 Statistical Analyses of Tone Contours

Statistical analyses were performed to validate the acoustic observations and to examine the subtle variation that the acoustic analyses do not capture. Regarding T1, the results of the modelling reveal no main effect of geographic language, language background, and their interactions on the F0 onsets, F0 offset, and F0 range of T1 (all $p > 0.05$). This indicates that the main F0 trajectories of T1 were consistent across geographical and language backgrounds.

Similar to T1, in T2, the results of the mixed-effects modelling suggest no main effect of geographic language, language background, and their interactions on the F0 onset, F0 offset, and F0 range (all $p > 0.05$). Thus, the main F0 trajectories of T2 were similar across the language and geographical groups.

In relation to T3, geographic background shows a main effect on the F0 onset ($\chi^2(3) = 8.25$, $p = 0.04$). More importantly, the interaction between geographic and language backgrounds emerged as a significant predictor for the T3's F0 onsets ($\chi^2(3) = 12.04$, $p = 0.007$). Post-hoc analyses identify a marginal main effect between the Eastern region and the EM region ($\beta = 0.21$, $SE = 0.08$, $t = 2.52$, $p = 0.066$). Additionally, post-hoc analyses suggest that the dialect group in the Eastern region had higher F0 onsets than the Mandarin group in the

EM region ($\beta = 0.27, SE = 0.09, t = 3.13, p = 0.0498$). The post-hoc results also show that while the dialect group in the Northern region had higher F0 onsets than the Mandarin group in the EM region ($\beta = 0.19, SE = 0.06, t = 3.32, p = 0.03$), the Mandarin group in the Central-southern group had a significant higher F0 onset than the Mandarin group in the EM region. In terms of the F0 offset, language background has a significant effect ($\chi^2(1) = 9.98, p = 0.002$). This suggested that language background played a role in the realizations of F0 offsets. However, this significance is qualified by the geographic background. In particular, a significant effect is observed on the interaction between language background and geographical background ($\chi^2(3) = 8.02, p = 0.046$). Post-hoc analyses suggest that the dialect group in the Eastern region produced higher F0 offsets than the Mandarin group in the Central-southern region ($\beta = 1.06, SE = 0.34, t = 3.15, p = 0.047$) and in the EM region ($\beta = 1.36, SE = 0.33, t = 4.08, p = 0.003$). Meanwhile, post-hoc analyses indicate that in the EM region, the dialect group showed a higher offset than the Mandarin group ($\beta = 0.96, SE = 0.29, t = 3.36, p = 0.03$). Similar to F0 onsets and offsets, language background reaches a main effect on F0 range ($\chi^2(1) = 20.67, p < 0.001$), indicating that the dialect group showed a narrower F0 range than the Mandarin group. More importantly, the interaction between language background and geographical background has a significant effect on F0 range ($\chi^2(3) = 15.31, p = 0.002$). Post-hoc analyses suggest that the dialect group in the Eastern region had a narrower F0 range than the Mandarin group in the Central-southern region ($\beta = -0.65, SE = 0.16, t = -4.07, p = 0.003$) and in the Eastern region ($\beta = -0.88, SE = 0.27, t = -3.02, p = 0.004$). Furthermore, post-hoc analyses indicate that the dialect group in the EM region exhibited a narrower F0 range than the Mandarin group in the central-southern region ($\beta = -0.45, SE = 0.14, t = -3.16, p = 0.0046$).

In T4, a significant interaction was observed in F0 onset between geographical and language background ($\chi^2(3) = 8.26, p = 0.04$). Post-hoc analyses suggest that in the Central-southern region, the dialect group showed a lower F0 onset than the Mandarin group ($\beta = -0.05, SE = 0.02, t = 3.24, p = 0.03$). A main effect of language background on the T4's offset was identified ($\chi^2(1) = 11.93, p < 0.001$), suggesting the dialect group had a higher F0 offset than the Mandarin group. Since the interaction between language and geographical backgrounds achieved a marginal significance ($\chi^2(3) = 6.21, p = 0.1$), the post-hoc analyses were conducted to avoid any overlook. The post-hoc analyses suggest that the dialect group in the Northern region had a higher F0 offset than the Mandarin group in the EM region ($\beta = 0.57, SE = 0.18, t = 3.24, p = 0.03$). Similarly, language background showed a significant effect on the F0 range ($\chi^2(1) = 20.64, p < 0.001$), indicating a narrower F0 range in the dialect group. Due to the marginal statistical significance in the interaction between language and geographical

backgrounds ($\chi^2(3) = 5.16, p = 0.16$), post-hoc analyses were performed. The results suggested that the dialect group in the Eastern region had a narrower F0 range than the Mandarin group across regions (Central-southern: $\beta = -0.53, SE = 0.14, t = -3.82, p = 0.003$; Eastern: $\beta = -0.64, SE = 0.20, t = 3.3, p = 0.02$; EM: $\beta = -0.61, SE = 0.14, t = 3.58, p = 0.0003$; Northern: $\beta = -0.50, SE = 0.14, t = 3.58, p = 0.008$).

For ease of interpretations, Table 2 summarizes the statistical findings in relation to language and geographical groups. It is important to note that the feature in the Difference column refers to the feature of the first group listed compared to the second group in the Group / Region column. For example, “Dialect (Eastern) vs. Mandarin (EM)” with “higher” means the Dialect group in the Eastern region has a higher F0 onset than the Mandarin group in the EM region.

Table 2. Tones in Relation to Language and Geographical Groups

Tone	Parameter	Group / Region	Difference
T3	F0 Onset	Dialect (Eastern, Northern) vs. Mandarin (EM)	higher
		Mandarin (Central-Southern) vs. Mandarin (EM)	higher
	F0 Offset	Dialect (Eastern) vs. Mandarin (Central-Southern, EM)	higher
		Dialect (EM) vs. Mandarin (EM)	higher
F0 Range	Dialect (Eastern) vs. Mandarin (Central-Southern, Eastern)	narrower	
	Dialect (EM) vs. Mandarin (Central-Southern)	Narrower	
T4	F0 Onset	Dialect (Central-Southern) vs. Mandarin (Central-Southern)	lower
	F0 Offset	Dialect (Northern) vs. Mandarin (EM)	higher
	F0 Range	Dialect (Eastern) vs. Mandarin (all regions)	narrower

Based on the acoustic findings, the dialect group in the Eastern region generally had a higher F0 onset and offset, as well as a narrower F0 range in T3 and T4. As Table 2 displays, these differences were significant between the dialect group in the Eastern region and the Mandarin group in certain regions. Additionally, the statistical studies echoed the acoustic results that the dialect in the Eastern region had a higher F0 offset in T3 than Mandarin group in the Central-Southern and EM regions, and a narrower F0 range in T3 than the Mandarin group in the Central-Southern and Eastern regions, and a narrower F0 range than the Mandarin group in the Central-Southern region. Likewise, the dialect group in the EM region showed a lower F0 offset than the Mandarin group in the EM region. Additionally, the dialect group in

the Eastern region also exhibited a narrower F0 range in T4 compared to the Mandarin group across all regions. In T4, the dialect group in the Central-Southern region showed higher F0 onset than the Mandarin group in the Central-Southern region, while the dialect group in the northern region displayed a lower F0 offset than the Mandarin group in the EM region.

5. Discussion

5.1 Tonal Features of MalM

The first research question proposed in the present study: What are the acoustic features of MalM tones? From the findings, the main F0 trajectories were similar across the language and geographical backgrounds. In general, T1 was a high-level, T2 was a mid-rising, T3 was a mid-falling, and T4 was a high-falling tone. This tone inventory largely aligned with that of Putonghua, where the four tones were respectively high-level, mid-rising, low-dipping, and high-falling. Their similarities may be due to the prevailing norm in Chinese education in Malaysia, because all speakers in this study had finished six years of Chinese education. However, unlike Putonghua, T2 had a long stretch before the final rise in Malaysia. This finding aligned with prior literature on MalM (Ren & Chiew, 2024) as well as other Mandarin varieties, such as Taiwanese Mandarin (Xiong & Li, 2018), Macanese Mandarin (Ren & Wang, 2025), and Singaporean Mandarin (Lee, 2010). Xiong and Li (2008) argued that the presence of this stretch is due to the influence of Hokkien in Taiwan. In contrast, Ren and Wang (2025) suggested that this stretch might be attributed to the principle of least effort. This principle indicates that speakers tend to convey their meanings with the minimal articulation effort (Martinet, 1952, 1955). This study supported Ren and Wang's (2025) findings that this stretch might be conditioned by the principle of least effort instead of the influence of the local southern Chinese dialects. On the one hand, the speakers in the present study varied in their primary home language from Hokkien, Cantonese, Hakka to Mandarin; on the other hand, this stretch may imply tonal reduction in Mandarin (Huang, 2017, Ren & Chiew, 2024), and the presence of the stretch seems to be a widespread feature in Mandarin varieties. Thus, this feature was more likely to be influenced by a universal principle rather than linguistic dominance.

Beyond T2, a salient variation that featured the tone inventory of MalM was the mid-falling T3. In Putonghua, the dipping T3 simply occurs in the citation form and before pauses, while the falling T3 is commonly found in natural speech. This study used the word list to elicit data, a method to investigate the lexical tones in the citation form. However, the mid-falling

T3 was the main F0 trajectory, while the dipping T3 occurred less frequently, which is in line with Ren and Chiew's (2024) findings. They attributed this feature to the influence of larger exposure to the falling T3, and thus, the falling contour is considered the default tone contour. The present research echoes the results of Ren and Chiew's (2024) study. Shi (2020) indicated that the essence of T3 in Mandarin is the falling contour rather than the dipping contour because the falling T3 is closer to the natural speech, occurring more frequently in natural speech than the dipping T3. Speakers are therefore more exposed to the falling variant rather than the dipping variant. Due to the potential larger quantity of input of the falling T3, the speakers are more likely to adopt this variant rather than the dipping one in their speech production.

5.2 Dominant Home Language and Geographical Factors

As previously mentioned, the main F0 trajectories were generally consistent across speakers' groups, particularly for T1 and T2. The statistical results also confirmed that no significant difference was observed in F0 onsets, F0 offsets, and F0 range of T1 and T2. Hence, T1 and T2 were comparatively stable across different regions and language groups. Conventionally, phonemic tones were classified into dynamic or contour tones, such as falling and rising tones, and static or level tones, such as high-level tones, mid-level tones (Abramson, 1976). In MalM, T1 was a level tone, a typically static tone. As Abramson (1976) suggested, rapid F0 movement is required for dynamic tones. This long stretch in MalM T2 indicates the reduced rapid F0 movement in the rising tone compared to the steep falling T3 and T4 in MalM. This confirms that T2 is more static than T3 and T4 in MalM. In addition, a dynamic tone is more likely to become a static one, suggesting possibly reduced contours in dynamic tones (Pittayaporn, 2007). Specifically, a change from high-falling tones to high-level tones is more possible than a change in the opposite direction (Pittayaporn, 2007). Notably, the reduced variant is characterized by a flatter contour with a smaller amount of F0 excursion and a lower F0 offset value (Pittayaporn, 2007). In Putonghua, T2 is a mid-rising tone without a preceding stretch; however, T2 in MalM was a rising tone with a long stretch. This long stretch in MalM T2 indicates a smaller amount of F0 excursion, suggesting that T2 is undergoing tonal reduction. Correspondingly, MalM T2 may be undergoing sound change from a dynamic tone to a static tone and therefore is more stable than T3 and T4 in MalM. Accordingly, T1 and T2 were more stable across language and geographical backgrounds.

As anticipated, the acoustic and statistical results confirmed that T3 and T4 exhibited larger instability across language and geographical groups. The dialect group in the Eastern region tended to be different from other regions and the Putonghua counterparts, while the

Mandarin group in the Central-southern and EM regions appeared to show more resemblance to Putonghua. The phonetic variation could stem from their dominant home language and the heterogeneous linguistic environment across Malaysia. As these differences consistently occurred in the dialect group, this variation might be due to the influence of their primary language (the southern Chinese dialects) on their tone production. This echoed Huang's (2016) findings that Hokkien, the dominant Chinese dialect in Penang, may contribute to the narrower pitch range in MalM. In addition to the effect of their primary home language, the geographical background was also found to influence the tonal articulation. A typical example was that the speakers in the dialect group of the Eastern region not only showed a narrower F0 range in T3 but also in T4. In this study, the Eastern group included speakers from Kelantan and Pahang. In Kelantan, the Malays account for 95.5% of the total population, while the Chinese only constitute 2.5% (Department of Statistics Malaysia, 2020). Likewise, in Pahang, the Malays account for 81% of the state's total population, while the Chinese make up 14.7% (Department of Statistics, 2020). As the majority ethnic group in Kelantan and Pahang is Malay, the Malay language may be more significantly dominant in these two states. Therefore, Chinese Malaysians in these two states are more likely to be exposed to the Malay language in their daily lives. Hence, this study supports Huang's (2016) suggestion that the more compressed tone space in MalM may be attributed to the influence of Malay, as Malay is a non-tonal language, which only contrasts high and low. Thus, the findings provided further evidence in support of this assertion.

Unlike the dialect group in the Eastern region, the Mandarin group—particularly from the Central-Southern and EM regions—showed greater resemblance in their articulation toward Putonghua, whether in F0 onset, F0 offset, or F0 range in terms of T3 and T4. The Central-Southern region included speakers from Kuala Lumpur, Selangor, Negeri Sembilan, and Johor. Based on a survey of 5,433 Chinese Malaysians from Kuala Lumpur, Selangor, and Johor, Wang (2024) found that Mandarin was the most frequently used language among Chinese Malaysians in these regions, prevailing over the use of the southern Chinese dialects in their daily communication. This suggests that Mandarin plays a vital role in both the home and other functional domains in these areas. Thus, tonal articulation in these areas appears to be less influenced by the southern Chinese dialects. Additionally, population distribution further supports this trend. The Chinese account for 41.6%, 27.3%, 23.9% and 32.8% of the total population in Kuala Lumpur, Selangor, Negeri Sembilan, and Johor, respectively (Department of Statistics Malaysia, 2020). Unlike the Eastern region, where the Malay population is dominant, the Central-Southern region features a more linguistically diverse

context, where Mandarin also exerts a substantial influence on the Chinese community. This may reduce their exposure to Malay and the Chinese dialects. Similarly, among the 11 speakers from the Mandarin group in the EM region in this study, ten were from Sarawak and one was from Sabah. Pua and Ting (2016) found that Mandarin ranked first in the frequency of language use across most functional domains, such as transaction, religion, media and friendship in their survey on Sarawak. Meanwhile, the Chinese account for 23.8% of the state's total population. Collectively, the sociolinguistic contexts in both the Central-Southern and EM regions tend to support greater maintenance of Putonghua-like tonal features, due to stronger Mandarin use and reduced interference from local Chinese dialects or Malay.

Overall, this study confirms that the speakers in different regions may have different tonal variation, or different varying degrees of tonal variation in T3 and T4. This study provides some evidence to support Guo's (2017) assertion that Chinese Malaysian's accent differed based on their geographical distribution. In Malaysia, Chinese Malaysians are exposed to different Chinese dialects, alongside other languages such as Malay and English, representing this country's multilingual diversity. The speakers in this study either use Mandarin or the southern Chinese dialects as their dominant home language since birth, and learn Mandarin, Malay and English at a young age. The linguistic repertoire for these Chinese Malaysians are bound to be more complex than the Putonghua speakers. In Malaysia, different states have different population distribution and different regional lingua franca in the Chinese community, such as Cantonese in Kuala Lumpur, and Hokkien in Penang. These distinctions across the states contribute varying degrees of influence on their tonal articulation. In a multilingual society like Malaysia, it is challenging to classify or separate the tonal realizations of Chinese Malaysians, as well as the potential factors into distinctive categories. However, these tonal distinctions and the complex influencing factors confirm that MalM is a heterogeneous variety in Global Chinese.

6. Conclusion

The present study investigated the citation tone production of Chinese Malaysians and examines the effects of primary language and geographical regions on tonal variation. Based on these results, we can find that MalM contains internal variation and may be reductive if we consider it as a single, homogenous variety. This study, exploratory in nature, serves as a springboard and provides nuanced insights to the effects of dominant home language and geographical distribution on MalM.

However, this study is not without limitations. Ideally, in studies of this nature, the number of speakers should be better balanced for a more valid comparison. However, due to constraints caused by the COVID-19 pandemic, it was challenging to achieve such a balance. Additionally, the classification of speakers into different geographical backgrounds was based on Guo's (2017) study, with reference to a government's document issued by the Ministry of Housing and local government Malaysia (2005). There is an evident need to establish a more precise division of the geographical regions of the speakers to minimize potential bias. Furthermore, as this study focused primarily on the main F0 trajectories, the tone variants and the role of duration as an acoustic cue warrant further investigation. Finally, the role of tonal variation in shaping different accents of Mandarin within the Chinese community in Malaysia may be a direction for future perceptual research. Despite these limitations, this study serves as a valuable basis for more comprehensive cross-regional and cross-linguistic comparisons.

References

- Abramson, S. (1976). Static and dynamic acoustic cues in distinctive tones. *Journal of the Acoustical Society of America*, 59(1), S42. <https://doi.org/10.1121/1.2002690>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Baranowski, M. (2013). Sociophonetics. In R. Bayley, R. Cameron & C. Lucas (Eds.), *The Oxford handbook of sociolinguistics* (pp. 403–424). Oxford University Press.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bei, X. (2021). *Sui, Gang, Ao sandi Putonghua yuyin xide yanjiu [A study on Mandarin phonetic acquisition in Guangzhou, Hong Kong and Macau]*. Shehui Kexue Chubanshe (China Social Sciences Press).
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56(Suppl.1), 9–49. <https://doi.org/10.1111/j.1467-9922.2006.00353.x>
- Boersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer* (version 6.2.04) [Computer program]. <https://www.fon.hum.uva.nl/praat/>
- Chao, Y. R. (1930). A system of tone letters. *Le Maître Phonétique*, 30, 24–27.

- Chen, J. Y. (2007). Malaixiya Huayu diaozhi de yanbian jian yu Putonghua bijiao [Tone value evolution in Malaysian Mandarin and its comparisons with Putonghua]. *Xiandai Yuwen [Modern Chinese]*, 5, 34–35.
- Chen, M. Y. (2000). *Tone sandhi patterns across Chinese dialects*. Cambridge University Press.
- Choo, Y. (2015). *Variant and invariant: a preliminary approach to the exploration of the emergence of Singapore Mandarin*. [Unpublished doctoral dissertation]. Nanyang Technological University.
- Department of Statistics Malaysia. (2020). *Current population estimates, Malaysia, 2020*. <https://open.dosm.gov.my/dashboard/kawasanku> (accessed July 31, 2025)
- Deng, D., Shi, F., & Lv, S. (2006). Putonghua yu Taiwan Guoyu shengdiao de duibi fenxi [The contrast on tone between Putonghua and Taiwan Mandarin]. *Shengxue Xuebao [Acta Acoustica]*, 31(6), 536–541.
- Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40, 467–491. [https://doi.org/10.1016/S0167-6393\(02\)00128-0](https://doi.org/10.1016/S0167-6393(02)00128-0)
- Fon, J., & Hsu, H. J. (2007). Positional and phonotactic effects on the realization of dipping tones in Taiwan Mandarin. In C. Gussenhoven & T. Riad (Eds.), *Tones and tunes: Experimental studies in word and sentence prosody* (pp. 239–269). Mouton de Gruyter.
- Goh, Y. S. (2017). *Teaching Chinese as an international language*. Cambridge University Press.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131–149. <https://doi.org/10.1017/S136672899800025X>
- Guan, Y., & Li, B. (2021). Usability and practicality of speech recording by mobile phones for phonetic analysis. *Proceedings of the 12th International Symposium on Chinese Spoken Language Processing*, Hong Kong, 24 – 27 January 2021. 1–5. <https://www.proceedings.com/content/058/058078webtoc.pdf>
- Guo, X. (2017). Malaixiya huayu gaishuo [An outline of Malaysia Huayu]. *Global Chinese*, 3(1), 61–83.
- Halekoh, U., & Højsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models – The R package pbrtest. *Journal of Statistical Software*, 59(9), 1–30. <https://doi.org/10.18637/jss.v059.i09>
- Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression and representation* (pp. 257–283).

Multilingual Matters.

- Hilton, N. H., & Leemann, A. (2021). Editorial: Using smartphones to collect linguistic data. *Linguistics Vanguard*, 7(s1), 1–32. <https://doi.org/10.1515/lingvan-2020-0132>
- Huang, T. (2016). *Topic in Malaysian Mandarin Phonetics and Phonology* [Unpublished doctoral dissertation]. National Tsing Hua University. <https://hdl.handle.net/11296/393u9b>.
- Huang, K. (2017). From pitch contour variation to tone change an analysis of phonological representations of the tones in Taiwan Mandarin. *International Journal of Chinese Linguistics*, 4(2), 273–307. <https://doi.org/10.1075/ijchl.16016.hua>
- Khoo, K. U. (2017). Malaysian Mandarin variation with regard to Mandarin globalization trend: issues on language standardization. *International Journal of the Sociology of Language*, 2017(244), 65–86. <https://doi.org/10.1515/ijsl-2016-0057>
- Khoo, K. U. (2018). Malaixiya “Juluan huayu” diaocha yanjiu: Yige teshu de shehui yuyan bianyi ge’an fenxi [Research on Juluan Mandarin in Malaysia: A special case study of social language variation]. In K. U. Khoo (Ed.), *Malaixiya huayu yanjiu lunji* [A collected papers of Malaysia Mandarin] (pp. 183–207). Huashe yanjiu zhongxin [Center for Malaysian Chinese Studies].
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lai, W., & Kuang, J. (2020). The effect of speaker gender on Cantonese tone perception. *The Journal of the Acoustical Society of America*, 147(6), 41119–4132. <https://doi.org/10.1121/10.0001411>
- Lee, L. (2010). The tonal system of Singapore Mandarin. In L. E. Clements & C. M. Liu (Eds.), *The 22nd North American Conference on Chinese Linguistics (NACCL-22) & the 18th International Conference on Chinese Linguistics (IACL-18)* (pp. 345–362). Harvard University.
- Martinet, A. (1952). Function, structure, and sound change. *Word*, 8, 1–32.
- Martinet, A. (1955). *Economie des changements phonétiques (Economy of phonetic changes)*. A. Francke Verlag.
- Ministry of Housing and local government Malaysia. (2005). <https://jpspn.kpkt.gov.my/jpspn/resources/Images%20JPSPN/Sumber%20Rujukan/Pelan%20Strategik%20Kebangsaan/ExecSum-Final-Report.pdf>.
- Peng, G. (2006). Temporal and tonal aspects of Chinese syllables: A corpus-based comparative

- study of Mandarin and Cantonese. *Journal of Chinese Linguistics*, 34(1), 134–154.
<https://www.jstor.org/stable/23754151>
- Pillai, S., Mohd. Don, Z., Knowles, G., & Tang, J. (2010). Malaysian English: An instrumental analysis of vowel contrasts. *World Englishes*, 29(2), 159–172.
<https://doi.org/10.1111/j.1467-971X.2010.01636>
- Pillai, S., & Zainuddin, S. Z. (2025). Top-down policies and the language profiles of Malaysians in a multilingual language ecology . In P. Siemund, G. Stein & M, Vida-Mannl (Eds.), *World Englishes in their local multilingual ecologies* (pp. 116-142). John Benjamins.
- Pittayaporn, P. (2007). Directionality of tone change. *Proceedings of the 16th International Congress of Phonetic Sciences ICPhS XVI, 6-10 August 2007, Saarbrücken, Germany*.
<http://www.icphs2007.de/conference/Papers/1516/1516.pdf>
- Puah, Y. Y., & Ting, S. H. (2016). Dominant language–Ethnicity notions of the Malaysian Chinese in Sarawak: Paternity or Patrimony. *Asian Studies Review*, 41(1), 1–20.
<http://dx.doi.org/10.1080/10357823.2016.1249340>
- R Development Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.<https://www.r-project.org/>
- Ren, X., & Chiew, P. S. (2024). An acoustic study on the citation tones in Malaysian Mandarin. *Pertanika Journal of Social Sciences and Humanities*, 32(4), 1367–1389.
<https://doi.org/10.47836/pjssh.32.4.07>
- Ren, X., & Wang, H. (2025). A sociophonetic study of the lexical tones in Macau Mandarin. *Sage Open*, 15(1), 1–15. <https://doi.org/10.1177/21582440251323630>
- Shi, F., & Ran, Q. (2011). Putonghua shangsheng de benzhi shi dipingdiao: Dui Hanyu pingdiao de shengdiao ganzhi yanjiu de zai fenxi [The nature of the Mandarin rising tone: A reanalysis of “A study on tone perception of Mandarin level tones”]. *Zhongguo Yuwen [Chinese Language]*, 6, 550–555 + 576.
- Shi, F. (2020). On the essential tones in Standard Chinese: Concurrent discussion on the nature of five-degree method. *International Journal of Chinese Linguistics*, 7(1), 140–157.
<https://doi.org/10.1075/ijchl.00008.shi>
- Tsui, R. K.-Y., Tong, X., & Chan, C. S. K. (2019). Impact of language dominance on phonetic transfer in Cantonese–English bilingual language switching. *Applied Psycholinguistics*, 40(1), 29–58. <https://doi.org/10.1017/S0142716418000449>

- Wang, N. (2010). Global English(es) and Global Chinese(s): Toward rewriting a new literary history in Chinese. *Journal of Contemporary China*, 19(63), 159–174. <https://doi.org/10.1080/10670560903335884>
- Wang, X. (2016). The Chinese language in the Asian diaspora: A Malaysian experience. In G. Leitner, H. Azirah & H. G. Wolf (Eds), *Communicating with Asia: The future of English as a global language* (pp. 205–215). Cambridge University Press.
- Wang, X. (2024). Malaixiya de zhongwen chuancheng yu chuanbo [The inheritance and spread of Chinese in Malaysia]. *Guoji Zhongwen Jiaoyu Xuebao [International Journal of Chinese Language Education]*, (16), 69–84.
- Wu, S., Liu, D., & Huang, S. (2023). The effects of over- and under-specified linguistic input on L2 online processing of referring expressions. *Journal of Psycholinguistics*, 52(1), 283–305. <https://doi.org/10.1007/s10936-022-09879-3>
- Xiong, Z., & Li., A. (2008). Taiwan Guoyu yu Putonghua de Danzidiao Duibi Yanjiu [A study of the citation tone between Taiwan Guoyu and Putonghua]. In J. Zhou & K. Dong (Eds.), *Haixia Liangan Yuyan yu Yuyan Shenghuo Yanjiu [Research on language and language situation across the straits]* (pp. 315–332). The Commercial Press.
- Xu, C., & Zhang, C. (2024). A cross-linguistic review of citation tone production studies: Methodology and recommendations. *Journal of the Acoustical Society of America*, 156, 2538–2565. <https://doi.org/10.1121/10.0032356>
- Xu, Y. (2013). ProsodyPro-A tool for large-scale systematic prosody analysis. *Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)* (pp. 7–10). Aix-en-Provence, France.
- Yip, M. (2002). *Tone*. Cambridge University Press.
- Zhang, J., Tan, W., & Strelluf, C. (2024). Sociophonetics and Chinese. In C. Strelluf (Ed.), *The Routledge handbook of sociophonetics* (pp. 615–633). Routledge.
- Zhang, W., & Gu, W. (2023). F0 range instead of F0 slope is the primary cue for the falling tone of Mandarin. *Journal of the Acoustical Society of America*, 153(6), 3439. <https://doi.org/10.1121/10.0019712>

Appendix

Tone category	Chinese character	Pinyin	English Glossary	Tone category	Chinese character	Pinyin	English Glossary
Middle Chinese Tone 1	他	tā	him	Middle Chinese Tone 2	婆	pó	elderly woman
	天	tiān	sky		题	tí	topic
	波	bō	wave		拿	ná	take
	芳	fāng	fragrance		娘	niáng	mother
	哥	gē	elder brother		雷	léi	thunder
	偏	piān	biased		肥	fēi	fat
	多	duō	many		房	fáng	house
	飞	fēi	fly		田	tián	field
	加	jiā	add		财	cái	wealth
Middle Chinese Tone 3	比	bǐ	compare	Middle Chinese Tone 4	怒	nù	anger
	体	tǐ	body		路	lù	road
	古	gǔ	ancient		贺	hè	congratulate
	粉	fěn	powder		饭	fàn	meal
	展	zhǎn	display		念	niàn	read
	坦	tǎn	level		共	gòng	together
	彩	cǎi	color		耐	nài	endure
	口	kǒu	mouth		漏	lòu	leak
	走	zǒu	walk		队	duì	team
Middle Chinese Tone 4	妇	fù	woman	Middle Chinese Tone 6	失	shī	lose
	弟	dì	younger brother		吃	chī	eat
	米	mǐ	rice		脚	jiǎo	foot
	夏	xià	summer		刻	kè	carve
	染	rǎn	dye		拆	chāi	dismantle
	满	mǎn	full		割	gē	cut
	老	lǎo	old		吉	jí	lucky
	道	dào	path		则	zé	rule
	在	zài	exist		尺	chǐ	ruler
Middle Chinese Tone 7	课	kè	lesson	Middle Chinese Tone 8	罚	fá	punish
	爸	bà	father		十	shí	ten
	替	tì	replace		昨	zuó	yesterday
	变	biàn	change		学	xué	study
	痛	tòng	pain		直	zhí	straight
	欠	qiàn	owe		入	rù	enter
	对	duì	correct		六	liù	six
	嫁	jià	marry		白	bái	white
	带	dài	belt		绝	jué	cut off

(Ren & Chiew, 2024, p. 1389)