Pertanika J. Soc. Sci. & Hum. 32 (4): 1367 - 1389 (2024)

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

An Acoustic Study on the Citation Tones in Malaysian Mandarin

Xin Ren and Poh Shin Chiew*

Department of Malaysian Languages and Applied Linguistics, Faculty of Languages and Linguistics, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

ABSTRACT

In the Global Chinese (GC) framework, Malaysia plays a crucial role in the spread and legacy of Mandarin. Although Malaysian Mandarin (MalM) has been extensively studied, its phonetic aspects have been neglected. Thus, this paper examines the four citation tones in MalM, in addition to the "checked quality"—termed the "fifth tone" ('T5') of MalM—by using Chinese dialectology supported by smartphone-based recordings. The findings suggest that, in MalM, the four tones contrast in tone height, tone value, and tone duration, with the main realizations being the high-level, low-level-rising, mid-falling, and high-falling contours, respectively. Other than the main tonal realizations, a number of variants are depicted. Furthermore, while MalM and *Putonghua* tones have evolved rather similarly, the checked quality—which is absent in *Putonghua* but was part of Middle Chinese, the shared ancestor of *Putonghua* and MalM—remains in MalM. However, there is a lack of distinctive phonemic contrast between Tone 4 and 'T5,' and there is insufficient evidence to consider 'T5' as a new tone category in this study. Unlike *Putonghua*, MalM is undergoing tonal reduction, like Singaporean and Taiwanese Mandarin. These findings indicate that, although Malaysian Mandarin (MalM) is undergoing substantial variations compared to *Putonghua*, it shares similarities with other Mandarin varieties, as there is some overlap in their tonal features. This study makes an empirical contribution to the body of research on tonal variations in the GC framework and echoes GC's call for a pluricentric approach to Mandarin varieties.

Keywords: Acoustic analysis, citation tones, Global Chinese, Malaysian Mandarin, tonal realizations

ARTICLE INFO

Article history:
Received: 08 February 2024
Accepted: 25 July 2024
Published: 19 November 2024

DOI: https://doi.org/10.47836/pjssh.32.4.07

E-mail addresses:
tva180025@siswa.um.edu.my (Xin Ren)
chiewpohshin@um.edu.my (Poh Shin Chiew)
*Corresponding author

INTRODUCTION

Today, Mandarin is spoken by a regional and global community and is, therefore, a language with global ownership.

ISSN: 0128-7702
e-ISSN: 2231-8534

Accordingly, the Global Chinese (GC) concept has proposed challenging the traditional monolingual ideology, indicating a conceptual shift and a pluricentric approach to investigating Mandarin varieties. The increased worldwide usage of Mandarin has also led to innovations in its use, as it is employed by people from different language backgrounds and assumes diverse functions and forms in distinct settings. It is no longer relevant to associate Mandarin exclusively with *Putonghua*, the standard Chinese used in mainland China. Thus, GC advocates for a new research trend that uses *Putonghua* as a reference point rather than as a target for Mandarin varieties spoken outside of mainland China. The concept of GC has been influential in raising awareness about the diversity of Mandarin varieties and has driven scholarly attempts to study them.

Given the establishment, persistence, and spread of the Chinese language and culture in Malaysia, this Southeast Asian country is a crucial component within the broader notion of GC (Diao, 2018; D. M. Xu & Wang, 2009). As a multiracial society, Malaysia has Malay, Chinese, and Indian populations, as well as other ethnic groups. Malay is the official language, and English serves as the second language. Mandarin and Tamil are widely spoken in the Chinese and Indian communities, respectively. Aside from Mandarin, prominent Chinese dialects, such as Hokkien, Cantonese, and Hakka, are also used in Malaysia's Chinese community. As a multiethnic and multilingual country, Malaysia is a rich site for studying linguistic variation and change; its linguistic diversity has aroused interest in Malaysian Mandarin (MalM; Guo, 2022; Y. Li, 2010, 2016; X. M. Wang, 2020). These studies mainly focus on the lexical and syntactic particularities of MalM, but acoustic analysis of the phonetic characteristics, particularly tonal characteristics, has been largely neglected.

In Mandarin, a syllable in isolation carries a citation tone and contrasts word meanings. *Putonghua* has four citation tones: Tone 1 (T1) through Tone 4 (T4). The four citation tones are represented in Chao's (1968) five-scale system with five distinctive levels, with 1 indicating the pitch floor and five denoting the pitch ceiling. Following Chao's system, the four citation tones in *Putonghua* contrast as in [55, 35, 214, 51]. When a citation tone is isolated, the tone contour is fairly stable and well-defined. However, when produced in context, the tone contour undergoes certain variations depending on the adjacent tones (Y. Xu, 1997). This phonological process is referred to as tone sandhi. T3 sandhi and half T3 sandhi are probably the best-known phonological processes in *Putonghua* (Duanmu, 2007). For example, T3 becomes T2 if another T3 follows it, and it becomes half T3 when it precedes T1, T2, or T4. These processes generate two allophonic variants of T3, regarded as the sandhi tone.

Thus far, existing research has sparked debate on the tonal properties of citation tones and the number of tone categories in MalM and Singaporean Mandarin (SigM)—either four categories (Tone 1, Tone 2, Tone 3, Tone 4) or five categories (the former tones and the "fifth tone"). T. Huang (2016a, 2016b) and Ng and Chiew (2012) identified the short-falling tone, the 'fifth tone' (T5), as a new tone category. In contrast, C. Y. Chen (1983) and Lee (2010) emphasized that 'T5' is not always distinguishable from T4 and preferred not to classify it as a contrastive tone category. Although previous studies have considered checked tones in the Middle Chinese tone system as a factor in the emergence of 'T5,' there is a lack of systematic analysis. Hence, it is necessary to examine the citation tones of MalM from a diachronic lens using Chinese dialectology methods. Our adoption of the term "fifth tone" in this paper does not indicate our full agreement with this concept; as it is commonly used, we have used it in this paper, but we have done so hesitantly.

Against these backdrops, this study aims to acoustically investigate tonal variations in MalM and compare them to those of neighboring varieties in the GC framework. The following questions guided our study:

1. How are the citation tones realized in MalM?
2. Can the "fifth tone" be considered a new tone category?

Literature Review

A Diachronic Lens: The Evolution of the Citation Tone System in Mandarin

Modern Chinese is believed to have developed from Middle Chinese (MC; circa 200–900 AD) in Chinese dialectology. Before exploring the tone system in MC, understanding the structure of the Chinese syllable is necessary. As M. Y. Chen (2000) illustrated, the Chinese syllable is divided into an optional initial consonant and a final. The initial consonant is the onset; the final includes an optional medial (onglide) consonant and a rhyme. The rhyme consists of a nucleus and an optional coda, the latter of which can either be an offglide or a consonantal ending.

According to *Qieyun* (AD 601)—the Chinese rhyme dictionary of early MC—MC has four tone categories: *ping*, the level tone (平); *shang*, the rising tone (上); *qu*, the falling tone (去); and *ru*, the entering or checked tone (入). Based on their initials, each of these tones is split into one of two registers—the *yin* (upper) register for voiceless initials (阴) or the *yang* (lower) register for voiced initials (阳)—thereby generating eight tones: Tone 1 (MCT1) through Tone 8 (MCT8). In terms of syllable structure, MC tones are divided into two categories—the smooth tone and the checked tone—based on the rhyme type. The smooth tone contains either an open syllable (CV) or a syllable ending with a nasal stop (CVN); the checked tone contains a syllable that is closed by a stop coda (CVP, -p, -t, -k, often reduced to a glottal stop -q; M. Y. Chen, 2000). While *ping*, *shang*, and *qu* are smooth tones, *ru* is a checked tone. Because of the difference in syllable structures, a checked tone is perceived as auditorily shorter than a smooth tone. These eight MC tones underwent various splits and mergers and developed into modern Chinese. Figure 1 displays the evolution of citation tones in *Putonghua*, drawn up based on M. Y. Chen's (2000) and Ma's (2002) research. Two important diachronic changes have occurred in *Putonghua*: all voiced obstruents have become voiceless, and all checked syllables have lost their stop codas (M. Y. Chen, 2000). The MC tone system provides a common frame of reference in investigating the Chinese language (M. Y. Chen, 2000) and in the study of GC (X. M. Wang, 2019).

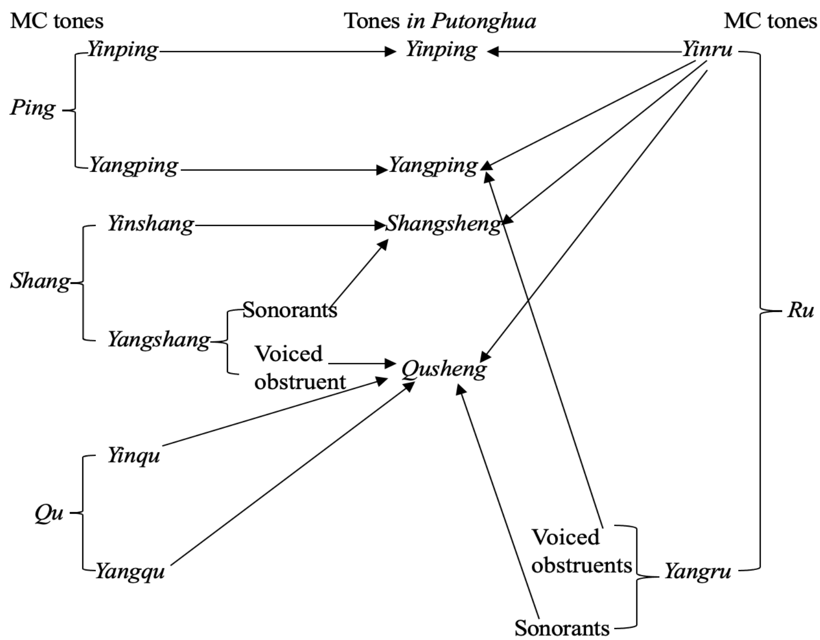

Figure 1. Correspondence between the tone systems in MC and *Putonghua*

Source: M. Y. Chen (2000); Ma (2002)

Previous Studies on the Citation Tones of Mandarin Varieties

The existing studies show that other Mandarin varieties have four citation tones, similar to *Putonghua* (T1, T2, T3, and T4), but with their own tonal properties. *Putonghua* and Taiwan Mandarin (TwM) generally have four-tone categories. In Taiwan Mandarin (TwM), scholars have not yet reached a consensus on whether T2 is dipping (Fon & Chiang, 1999; A. Li et al., 2006) or rising (Deng et al., 2006; Fon et al., 2004; Kuo, 2018; Shi & Deng, 2006), and whether T3 is low-falling (Chiung, 1999; Fon et al., 2004; Kuo, 2018; A. Li et al., 2006; Shi & Deng, 2006) or dipping (Fon & Chiang, 1999). SigM has also drawn attention in recent years. While Chua's (2003) tonal transcription of SigM largely aligns with that of C. Y. Chen (1983) researchers disagree on whether T3 is represented by [11] (C. Y. Chen, 1983) or [211] (Chua, 2003). Another strand of research has looked at the four citation tones in MalM. Based on the impressionist approach, J. Y. Chen (2007) described the four tones as [33/44, 24/35, 21, 31/53]. Following the acoustic approach, T. Huang (2016b) investigated the citation tones of MalM in the Malaysian state of Penang, revealing that MalM has four citation tones: [33, 23, 21, 53]. Despite the inconsistent tone values of the citation tones reported in previous research, their tone contours are still similar: T1 is produced as a level tone, T2 as a rising tone, T3 as a low-falling tone, and T4 as a high-falling tone. Nevertheless, Khoo (2013, 2017) observed that T4 is frequently realized as a high-level tone apart from a high-falling tone. Regarding the acoustic approach, Yeoh (2019) analyzed the variations of T4 in connected speech and confirmed Khoo's (2013, 2017) findings. Based on the tonal properties of the four citation tones, researchers suggested that tonal reduction happens in TwM and MalM (K. Huang, 2017; T. Huang, 2016b).

Not only that, as mentioned in the introduction above, but there is also a question about the tone categories of Mandarin varieties, such as SigM and MalM. Apart from the four citation tones, researchers have proposed an additional tone, the "fifth tone" ('T5'), and sought to identify its tonal properties. The 'T5' terminology implies an independent status in the tone system in MalM and SigM and indicates that these Mandarin varieties may have developed a new tone system compared to *Putonghua*. Interestingly, however, 'T5' is not found in TwM owing to the heavy influence of northern dialects and the intensive teaching and learning of national phonetic symbols (C. Y. Chen, 1983). According to previous studies, 'T5' may be derived from the checked tone in MC (C. Y. Chen, 1983; Choo, 2013). Studies further indicated that the occurrence of 'T5' is constrained by the final types in SigM (Choo, 2013; Goh & Xu, 2003; Ng & Chiew, 2012; X. Y. Xu & Goh, 2001). Similar results have been reported for MalM, which is unsurprising given Singapore's geographic and linguistic proximity to Malaysia. For instance, the "fifth tone" has been mentioned in MalM, and the features of 'T5' in MalM are similar to descriptions by C. Y. Chen (1983) regarding SigM (T. Huang, 2016a, 2016b).

On the other hand, controversy also exists regarding 'T5''s tone value (C. Y. Chen, 1981, 1983; J. Y. Chen, 2007; T. Huang, 2016a, 2016b). C. Y. Chen (1981, 1983) was among the first to report on 'T5'. Using the auditory approach, C. Y. Chen (1983) suggested that, in SigM, 'T5' is a short falling tone that sometimes ends with a glottal stop, but the glottal stop varies in degree of prominence, and 'T5' is sometimes identical to T4. It was either realized as [51] or [53] or varying between [41] and [42] in SigM. In MalM, while J. Y. Chen (2007) suggested that the tone value of 'T5' is either [53] or [42], T. Huang (2016b) indicated that T5 is uttered as [53]. T. Huang (2016b) also indicated that 'T5' is accompanied by a glottal stop or a large number of creaks with a systematic acoustic study. Despite burgeoning research on tonal properties and governing rules on 'T5,' whether 'T5' is a contrastive tone, an allotone, or a variant tone remains a conundrum. Additionally, the reason some non-checked syllables can surface similarly to 'T5' in both SigM and MalM remains in dispute (Choo, 2013; T. Huang, 2016a, 2016b; Ng & Chiew, 2012). In detail, while Huang (2016a) suggested that this phenomenon is probably due to overgeneralization, C. Y. Chen (1981) found that the non-checked syllable, "钢" (iron), a high-level tone in *Putonghua*, was realized as the high-falling tone in SigM. C. Y. Chen (1981) further attributed this to the influence of *Qusheng* in the Teochew dialect. It is further confirmed by Choo (2013). Hence, to date, the tonal properties of MalM are still unclear, and the tone categorization of 'T5' is unsettled.

Past studies on the citation tones in MalM shed light on the citation tones in tonal combinations or connected speech (see T. Huang 2016b; Yeoh, 2019). Such methods may result in tonal variations owing to the adjacent tones or connected speech. Besides, prior studies mainly focused on the production of the old generation (T. Huang, 2016b). Accordingly, by monosyllabic reading from a diachronic lens, the present study reports on the number of tone categories and their tonal properties by involving young MalM speakers in the GC framework.

METHODS

Participants

We recruited participants based on a self-reported questionnaire that collected demographic data and information about language use in different contexts (e.g., at home, at school, and in dealings with the government). Participants consented to be recorded for research purposes. None of them reported any speech impairments. As Chinese Malaysians constitute the primary group contributing to MalM, and as undergraduates represent the young generation who have grown up in a predominantly multilingual context and were exposed to standardization of the Chinese language in Malaysia, known as *Kurikulum Bersepadu Sekolah Rendah*, we recruited 36 Chinese Malaysians aged 20–25 years (with an average age of 22) who had completed at least six years of Chinese-language education and were fluent in Mandarin. All speakers were female to maintain consistency regarding the gender variable. It was done to obviate the influence of specific speaker effects (Jacobi, 2009). It is also attributed to the pitch differences between male and female voices, and it was considered better to avoid analyzing male and female recordings together (Pillai et al., 2010).

As southern Chinese dialects play a crucial role in MalM variations (X. M. Wang, 2019), we selected these participants because they all considered at least a southern Chinese dialect to be their dominant home language. They occasionally used Malay and English besides Mandarin and southern Chinese dialects. Although the participants were from different dialect groups, including Hokkien, Cantonese and Hakka, we will consider them a cohesive group because these three dialects all belong to southern Chinese dialects, and they preserve the checked tone in their phonology. Additionally, as shown in Figure 2, they came from different states in Malaysia.

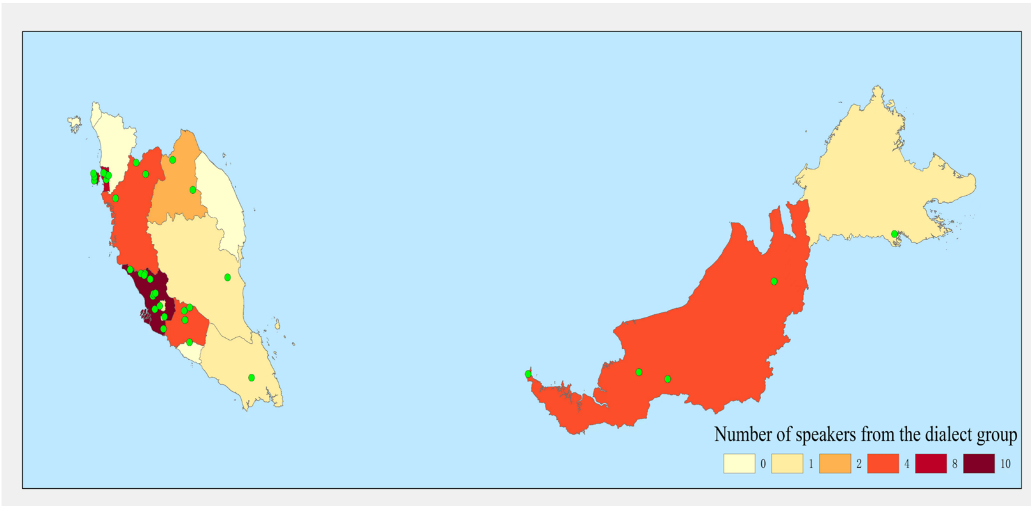

Figure 2. Malaysian map of the participants from their dialect group, mapped according to the locations of the participants' homes

Source: Authors' work

Reading Material

The reading material was designed using the methodological framework of Chinese dialectology, in which the standard method used to investigate the phonological system entails the use of *Fangyan Diaocha Zibiao* (the *Questionnaire of Characters for a Dialect Survey*, hereafter *Zibiao*) (Institute of Linguistics, CASS, 2018) to elicit data. *Zibiao* covers more than 3,700 monosyllabic words arranged based on MC phonology, where the development of a certain dialect with reference to MC can be captured. Based on *Zibiao*, the reading material was selected and compiled to represent MC tone categories, specifically MCT1 through MCT8. While syllables in MCT1 to MCT6 were smooth syllables, those in MCT7 and MCT8 were checked syllables. A total of 72 monosyllabic syllables (eight tone categories x nine monosyllabic syllables) are listed in the Appendix. We used the following criteria to select monosyllabic syllables: (1) *All monosyllabic syllables are real words and are frequently used in MalM* (Nonsense and uncommon syllables were excluded), (2) *Further division of the initial consonant was considered, including both voiceless and voiced initials* (The development of MC tones is sensitive to the initial consonants), (3) *Simplify segmentation* (Syllables with zero-onsets were avoided), and (4) *The finals of monosyllabic syllables in the wordlist included all the following* *final types:* monophthongal, diphthongal, triphthongal, and nasal finals (in Mandarin, diphthongal and triphthongal finals are termed compound finals).

Recording Procedures

Smartphone recordings have been implemented in many acoustic sociophonetic studies, and valuable results have been reported (Kim et al., 2019; Kuo, 2018; Leemann et al., 2016; B. Li et al., 2020; Stanford, 2019). As Stanford (2019) suggested, most smartphones today offer better fidelity and audio range than outdated devices; many sociolinguistic datasets gathered in the field have varying quality in their recordings, and some sociolinguistic studies even compare modern recordings with legacy ones. Guan and Li (2021) and C. Zhang et al. (2021) suggested that smartphones are reliable alternative recorders for general purposes, at least in prosodic aspects. Therefore, smartphone technology was utilized to gather data for the present study. This approach was also chosen for its efficiency in collecting recordings, which could increase the data size and expand the geographic coverage. Additionally, data collection was conducted in 2020, when the COVID-19 pandemic curtailed face-to-face recordings.

As the recordings were made remotely, we prepared an instruction manual for the participants to ensure good sound quality. The participants made the recordings in a quiet room in their homes, which yielded a natural and familiar speaking environment. The stimuli were fully randomized and presented in Google Slides, and a different order was applied for each participant to control the intensity and avoid priming effects (Alzaidi et al., 2019; Clark, 2018). Two monosyllabic syllables were added as fillers to the beginning and end of the Google Slides to avoid order effects (Zhu, 2018). The participants were asked to read the word list once and then take a short break before re-recording the same list. They were instructed to read naturally at a normal speaking rate and upload the recordings to Google Drive.

Data Analysis

Filtering and downsampling were implemented to remove other information in the spectrum and avoid environmental sounds in the recordings. Segmentation, labeling, and annotation were carried out in Praat (Boersma & Weenink, 2021).

Many precautions were taken to ensure good, valid data for analysis, such as providing a clear instruction manual for the participants and manually checking the sample quality. Three participants' recordings were excluded because of the background noise. Therefore, data from 36 participants were analyzed, and the total number of tokens was 5,184 (72 monosyllabic syllables x 2 repetitions x 36 participants = 5,184 monosyllabic syllables). These tokens were also visually checked, and 3.9% (200 tokens) with noise or non-modal voice were removed to avoid inaccurate pitch tracking. In the end, 4,984 tokens were examined in the subsequent analysis.

Bao and Lin (2018) noted that tones in Mandarin are carried by the stable part of the final, which is the vowel segment in this study. The onset and offset of the vowel's second formant (F2) were used to identify and determine the boundaries of each vowel segment (Zheng, 2010). Fundamental frequency (F0) is the basic acoustic parameter of perceived pitch, which functions as a dimension in suprasegmental linguistic systems of tone, intonation, and stress (Lehiste, 1970). Apart from F0, the linguistic pitch is also mediated by the interaction of F0 with other major acoustic parameters, such as duration (Rose, 1989). Thus, F0 and duration were the main acoustic correlates and were measured to define the tonal features in this study. With the help of ProsodyPro (Y. Xu, 2013), a script running under Praat, we hand-inspected vocal pulse marking for errors and corrected them, then extracted acoustic parameters. We extracted F0 from 10 equidistant points and excluded 10% of the leftmost and rightmost F0 values to reduce tone-irrelevant variations.

T-value is superior because of its higher normalization index (Liu, 2007), its accordance with Chao's five-scale system (Liu, 2007; J. W. Zhang, 2014), and its wide usage in the Chinese literature on tones (J. W. Zhang, 2014). Thus, we transformed the raw F0 data into T-value using the formula outlined below (Liu, 2007; Shi & Wang, 2006) to eliminate inter-speaker variation. T-value is divided into five intervals to align with Chao's five-scale system. However, as actual production and perception do not always follow this scale system, Liu (2007) proposed a dynamic strategy where the five intervals are not strictly divided into 0–1, 1–2, 2–3, 3–4, and 4–5, and the T-value has a dynamic range. The five intervals are, therefore, 0–1.1, 0.9–2.1, 1.9–3.1, 2.9–4.1, and 3.9–5.

$$T - value = 5 \times \frac{(logx - logb)}{(loga - logb)} \quad [1]$$

where *x* = observed F0 value*, b* = lowest F0 value*,* and *a* = highest F0 value.

We also extracted duration in milliseconds (ms). Corresponding to the exclusion criteria of F0 points, we only measured 80% of the extracted duration of each syllable. Furthermore, to minimize the individual factors, we performed duration normalization using the following normalization formula (Liu, 2007):

$$Normalized\ duration\ (Dur) = \frac{measured\ duration}{average\ measured\ duration\ of\ a\ speaker} \quad [2]$$

After normalization, we combined qualitative (auditory perception and acoustic identification) and quantitative (statistical analysis) methods to answer the research questions. In the qualitative descriptions, the acoustic realization of the syllables is presented in tone height and tone contour with reference to the spectrogram, auditory perception, and the *T-value*. While tone height is categorized into high, mid, and low based on the five-scale system (Chao, 1967), tone contour is grouped into level, level-rising, dipping, and falling. In quantitative statistics, first, to check how the citation tones in MalM deviate from each other, we performed an extensively applied technique by deciding on the tone contour and tone height; more specifically, the F0 slope and the average F0 (T-value; L. Wang et al., 2018). Accordingly, we obtained the mean F0 by averaging eight points. Considering the definition by L. Wang et al. (2018)—the "coefficient of the linear regression line for 21 points of each tone, with the corresponding real timestamp as the independent variable" (p. 7) —we calculated the F0 slope based on the eight points. Second, we used SPSS (IBM Corp., 2021) to check whether the data were normally distributed and complied with homogeneity of variance. As both assumptions were violated, we used non-parametric tests. After the Bonferroni correction, we set the alpha level to 0.05 for the Kruskal-Wallis test and to 0.008 for the Mann–Whitney test. Given the large sample size of this study, effect size estimates were also reported. As Table 1 displays, following Cohen's (1988) guidance, we set three cut-off points for the non-parametric tests. While the large effect size indicates that tones are distinct in terms of F0 slope, average F0, or tonal duration, the small effect size suggests that tones tend to overlap in these parameters. The moderate effect size falls in between.

Table 1
Effect sizes used in this research

Effect size	Kruskal–Wallis tests by η^2	Mann–Whitney tests by r
Large	0.14 and above	0.5 and above
Moderate	0.06–0.139	0.3–0.49
Small	0.01–0.59	0.1–0.29

Source: based on Cohen (1988); Yamaguchi and Chiew (2019)

RESULTS

Tonal Realizations of the Citation Tones in MalM

Table 2 displays different realizations of the four tones with different frequencies. The main realization of T1 is a high-level tone, followed by high-falling and low-level-rising contours. T2 has more varied realizations, with the level-rising tone being the primary one. Furthermore, the tonal realizations of T2 include high-level, high-falling, mid-falling, and dipping tones. Regarding T3, the main realization has a mid-falling contour; 11.1% (*n*=106) and 8.9% (*n*=85) of the occurrences have dipping and low-level-rising contours, respectively, and 4% (*n*=38) have a high-level contour. Compared with other citation tones, T4 seems to have less variation. The chief production of T4 has a high-falling tone, comprising 98.3% (*n*=1801) of the total. Only 0.3% (*n*=6) of the tokens were uttered with a high-level contour, and 1.4% (*n*=26) of the T4 tokens were produced with a high-level-rising contour.

Hence, the main realizations for T1 through T4 were a high-level tone, a low-level-rising tone, a mid-falling tone, and a high-falling tone. Moreover, the four tones showed different variants of the tone contour. However, while the other variants of T1, T2, and T4 were marginalized, T3 frequently displayed two variants: level-rising and dipping (besides the main realization). Furthermore, 1.6% (*n*=15) and 0.7% (*n*=9) of the T1 and T2 tokens, respectively, were realized as a high-falling tone; hence, they may be 'T5' tokens, as discussed previously.

Table 2
Distributions of tonal realizations of citation tones

Tone contour	**Citation tones**			
	Tone 1	**Tone 2**	**Tone 3**	**Tone 4**
High-level-rising	/	/	/	1.4 (26)
Low-level-rising	0.2 (2)	96.8 (1231)	8.9 (85)	/
Mid-falling	/	0.2 (3)	76 (724)	/
High-falling	1.6 (15)	0.7 (9)	/	98.3 (1801)
High-level	98.2 (909)	/	/	0.3 (6)
Mid-level	/	2.1 (27)	4.0 (38)	/
Mid-dipping	/	0.2 (2)	11.1 (106)	/
Total	100 (926)	100 (1272)	100 (953)	100 (1833)

Note. The value preceding the parentheses represents the occurrence percentage, while the numeral enclosed within the parenthesis denotes the count of the tokens

Source: Authors' work

The tone contour and value of the four tones are shown in panels (a)–(d) in Figure 3. The normalized tone contours in each panel are plotted against eight equidistant points. According to Figure 3, the tone values of T1–T4's main realizations are [44], [223], [31], and [52], respectively. Additionally, as mentioned, T3 contains two frequently occurring variants—level-rising and dipping contours—which could be expressed as [223] and [313]. Furthermore, the high-falling variants in T1 and T2 could be denoted as [52], which aligns with the tone value of the primary production of T4.

Informed by descriptive statistics as discussed above, the realizations of citation tones in MalM appear quite variable. However, it remains unknown how the four tones differ from each other. Following L. Wang et al. (2018), the average F0 was plotted on the overall F0 slope to obtain the variability of each tone, as depicted in Figure 4.

Figure 4 portrays a scatterplot where each plotted symbol represents a token, with the horizontal axis marking the F0 slope and the vertical axis denoting the F0 height. As seen in Figure 4, while realizations of T1 and T4 are higher in F0 height, as they were mostly produced as high tones, the realizations of T2 and T3 are somewhat lower, as they were primarily uttered as middle or low tones. We ran a Kruskal-Wallis test to compare F0 height and

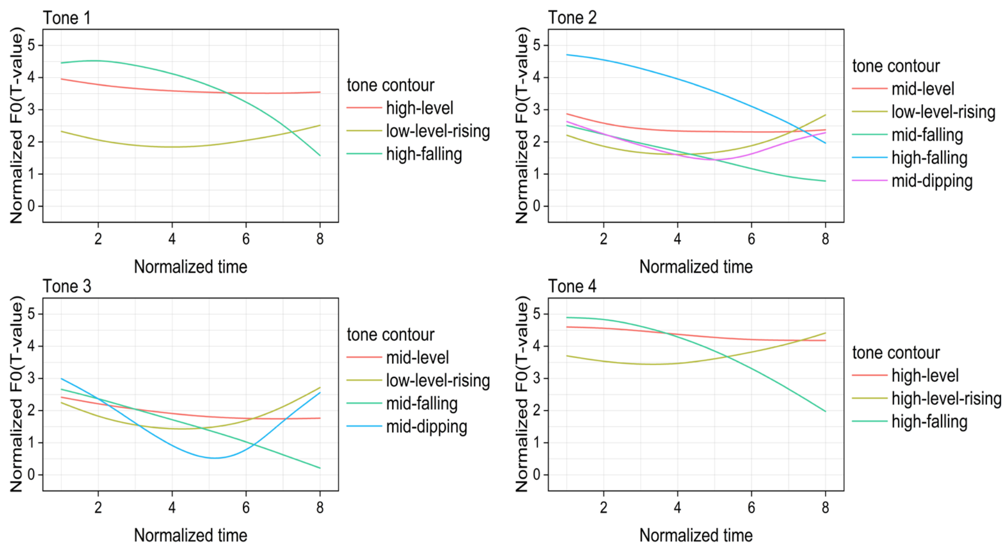

Figure 3. The tone contours of the citation tones in MalM
Source: Authors' work

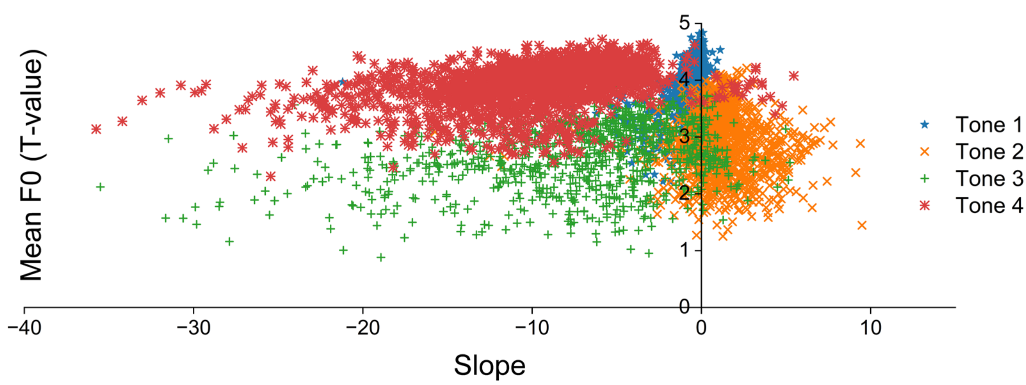

Figure 4. Scatterplot of the citation tones in MalM
Source: Authors' work

identified a significant effect of tone category on F0 height ($\chi^2(3) = 2829.90$, $p < 0.001$, $\eta^2 = 0.57$ [large effect]). Accordingly, F0 height of the tones decreased significantly across the four tones: T4 > T1 > T2 > T3. As Table 3 outlines, the Mann–Whitney test results suggest that F0 height significantly differs in all dimensions. However, the effect sizes reach a large effect between each pair except between T1 and T4 and T2 and T3. The large effect size indicates that differences in F0 height are distinct in all tonal pairs except between T1 and T4 and between T2 and T3. While T1 and T4 belong to a high register, T2 and T3 belong to a fairly low register.

In Figure 4, we can see differences in F0 slope. T1 is mostly located around the vertical axis owing to its fairly flat slopes.

Table 3
Mann–Whitney test results for F0 height and effect sizes. P-*values smaller than 0.008 are in bold*

Tone	*Z*	*P-value*	Effect size
T1 versus T2	Z = -31.05	$\boldsymbol{p}$ **< 0.001**	r = 0.66 [large effect]
T1 versus T3	Z = -33.69	$\boldsymbol{p}$ **< 0.001**	r = 0.78 [large effect]
T1 versus T4	Z = -5.18	$\boldsymbol{p}$ **< 0.001**	r = 0.01 [small effect]
T2 versus T3	Z = -11.25	$\boldsymbol{p}$ **< 0.001**	r = 0.24 [small effect]
T2 versus T4	Z = -40.00	$\boldsymbol{p}$ **< 0.001**	r = 0.72 [large effect]
T3 versus T4	Z = -40.62	$\boldsymbol{p}$ **< 0.001**	r = 0.77 [large effect]

Source: Authors' work

As the main realization of T2 is rising, it is primarily located to the right of the vertical axis. T3 and T4, mostly in a falling contour, are located to the left of the vertical axis. Likewise, a Kruskal-Wallis test identified a significant effect of tone category on the F0 slope ($\chi^2(3) = 3368.52$, $p < 0.001$, $\eta^2 = 0.68$ [large effect]). The F0 slope of tones declined significantly across the four tones: T2 > T1> T3 > T4. As Table 4 displays, the Mann-Whitney test revealed significant differences in the F0 slope between each pair of tones. Nevertheless, the results of effect sizes confirm that differences in F0 slope are distinct between each pair except between T3 and T4. The small effect size in the T3 and T4 pair demonstrates that T3 and T4 are similar in tone contour. In the high register, T1 and T4 differ significantly regarding contour: T4 mostly has a falling target, and T1 chiefly has a level target. In the low register, T2 has a predominantly rising target, and T3 mostly has a falling target.

Table 4
Mann–Whitney test results for F0 slope and effect sizes. P-*values smaller than 0.008 are bold*

Tone	*Z*	***P*-value**	Effect size
T1 versus T2	Z = -29.22	$\boldsymbol{p}$ **< 0.001**	r = 0.62 [large effect]
T1 versus T3	Z = -26.75	$\boldsymbol{p}$ **< 0.001**	r = 0.62 [large effect]
T1 versus T4	Z = -40.42	$\boldsymbol{p}$ **< 0.001**	r = 0.77 [large effect]
T2 versus T3	Z = -35.71	$\boldsymbol{p}$ **< 0.001**	r = 0.76 [large effect]
T2 versus T4	Z = -45.91	$\boldsymbol{p}$ **< 0.001**	r = 0.82 [large effect]
T3 versus T4	Z = -15.13	$\boldsymbol{p}$ **< 0.001**	r = 0.29 [small effect]

Source: Authors' work

Figure 5 presents the total duration of the four tones in MalM. The tone duration decreased across the four tones: T1 > T2 > T3 > T4. A Kruskal-Wallis test was run, and a significant effect of tone category was identified on the tonal duration ($\chi^2(3) = 2005.77$, $p < 0.001$, $\eta^2 = 0.43$ [large effect]). It indicates that T1 has the longest duration,

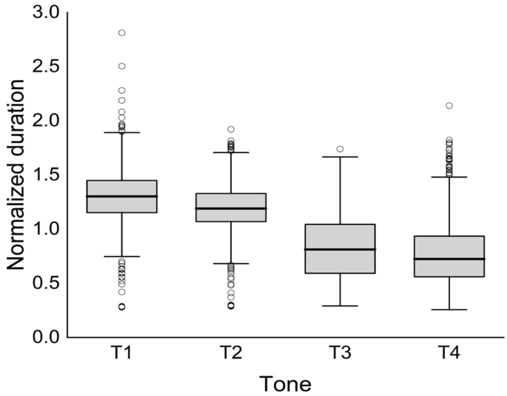

Figure 5. Boxplot for the normalized duration of the citation tones
Source: Authors' work

followed by T2, T3, and T4. Additionally, as Table 5 shows, the Mann–Whitney test revealed significant differences in the tonal duration between each pair of tones. Despite statistical significance, a small effect size was reached in the T1 and T2 and T3 and T4 pairs. The small effect size suggests that the tonal duration differences are not prominent in these two pairs.

Acoustic Realizations of 'T5'

As suggested in prior literature, one of the salient MalM features is 'T5'. Despite its common usage and our conformity with terminology, some controversy may surround 'T5''s source and acoustic productions. Therefore, we have attempted to shed light on this important topic.

Given the debate on 'T5', the first step involved defining 'T5' and establishing inclusion criteria. First, we only included tokens as 'T5' if they were derived from the checked tone in MC. Although some non-checked syllables were realized similarly to 'T5,' the governing rules of their occurrences differed, as discussed in the literature review. Accordingly, one token of 时 "time" was excluded from the analysis. It is also because this individual instance may be due to chance. Second, we only examined and compared the 'T5' tokens derived from T1, T2, and T3 in Mandarin, as previous research suggests that T4 and 'T5' cannot be consistently distinguished. Accordingly, of all 1,249 checked tone tokens, 23 tokens from nine participants were transcribed as 'T5'. To identify its acoustic properties, we compared 'T5' with T4 in terms of tone contour, tone value, duration, and voice quality.

As presented in Figure 6, while 'T5' was produced as a falling pitch that started

Table 5
Mann–Whitney test results for the tonal duration. P*-values smaller than 0.008 are bold*

Tone	*Z*	*P*-value	Effect size
T1 versus T2	$Z = -10.50$	$\boldsymbol{p < 0.001}$	$r = 0.22$ [small effect]
T1 versus T3	$Z = -27.59$	$\boldsymbol{p < 0.001}$	$r = 0.64$ [large effect]
T1 versus T4	$Z = -33.40$	$\boldsymbol{p < 0.001}$	$r = 0.64$ [large effect]
T2 versus T3	$Z = -26.44$	$\boldsymbol{p < 0.001}$	$r = 0.56$ [large effect]
T2 versus T4	$Z = -34.53$	$\boldsymbol{p < 0.001}$	$r = 0.62$ [large effect]
T3 versus T4	$Z = -5.18$	$\boldsymbol{p < 0.001}$	$r = 0.10$ [small effect]

Source: Authors' work

slightly lower and fell slightly earlier than T4, they are similar in tone contour and tone value. Both ‘T5’ and T4 can be denoted as [52], making it challenging to consistently distinguish them based on tone contour and tone value. Additionally, as displayed in Figure 7, the normalized duration of ‘T5’ is comparatively shorter than that of T4. Despite differences in duration, this minor difference may be difficult for listeners to perceive. Furthermore, Figure 8 shows the productions of voice quality of ‘T5’ and T4. Figure (8a), ‘T5’, the first case on the left, ends with a glottal stop (denoted by an arrow). Apart from this, ‘T5’ in this study was primarily realized with creaks (indicated by an arrow), as in the middle of Figure (8a). Like the case on the right, ‘T5’ was sometimes produced in normal voicing. As displayed in Figure (8b), T4 on the left was realized in creaks (shown by an arrow), and on the right, it was produced in normal voicing. In most cases, T4 was accompanied by creaks. In *Putonghua*, creaks often occur in the middle of T3 and at the end of T4, as creaks are a tonal feature of “lowest-ness.”[1] in Mandarin (Kuang, 2017). Therefore, when comparing ‘T5’ and T4 in terms of voice quality, they are both largely identical and consistently indistinguishable acoustically.

As noted above, earlier work on ‘T5’ also considered the role of finals, which indicates that the source of ‘T5’ is also related to finals (Choo, 2013; Ng & Chiew, 2012). Therefore, we investigated the final types of ‘T5’. While 91.3% (*n*=21) of ‘T5’ consisted of monophthongal finals, only 8.7% (*n*=2) comprised diphthongal finals. This finding is consistent with the results of Choo (2013). Additionally, C. Y. Chen (1983) and T. Huang (2016a) demonstrated that ‘T5’ emerged mostly from T1 in both SigM and MalM, followed by T2 and T3. We revisited this issue and found that 65.2% (*n*=15) of ‘T5’ emerged from T1, 34.8%

1 Creaks mostly occur in the lowest T3 and can be seen in the low-pitch target of T4.

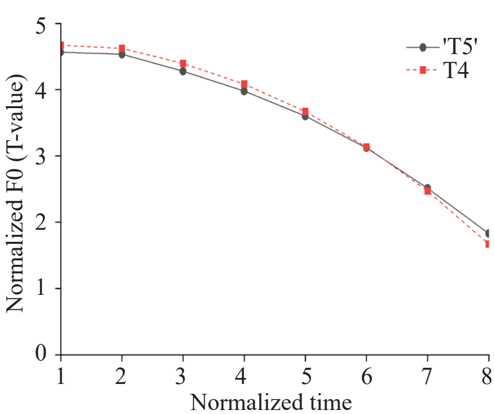

Figure 6. The tone contours of ‘T5’ and T4
Source: Authors’ work

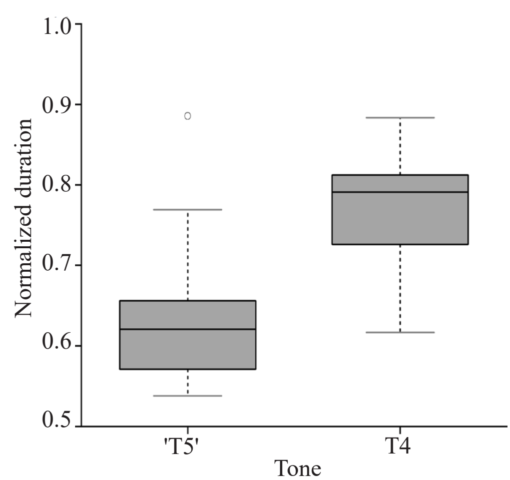

Figure 7. Boxplot of the normalized durations of ‘T5’ and T4
Source: Authors’ work

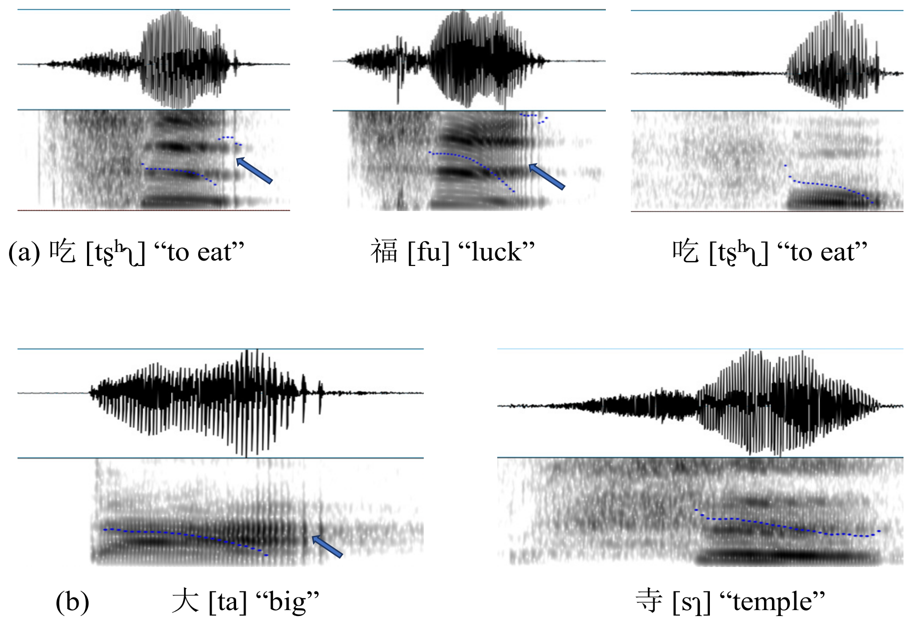

Figure 8. Examples of the spectrograms of ‘T5’ (a) and T4 (b)
Source: Author’s work

(*n*=8) of ‘T5’ emerged from T2, and none of ‘T5’ emerged from T3. Accordingly, our results align with those of C. Y. Chen (1983) and T. Huang (2016a). More tokens of Mandarin T1 and T2 were subject to a change to ‘T5’.

DISCUSSION

Our first concern is understanding the tonal properties of the citation tones in MalM. We found that the tone contour of the four citation tones varied within and across speakers. The acoustic and quantitative analyses suggest that the four tones in MalM contrast with each other in terms of F0 height, F0 slope, and tone duration and that the main tonal realizations are phonetic, as in [44, 223, 31, 52], and opposed to [55, 35, 214, 51] in *Putonghua*. Among the four citation tones in MalM, T3 is the most difficult to characterize. Unlike the other three tones, the central realization of T3 was inconsistent across speakers. In this study, two speakers differed from the rest of the participants, as T3 was predominantly produced as a dipping contour by these two speakers. Furthermore, this study is the first to report that a dipping tone contour is still the basic contour in MalM. It is expected because Mandarin teaching in Malaysia adheres to *Putonghua* standards, and the dipping T3 is merely observed in deliberate, emphatic, or yes-no questions (Fon & Hsu, 2007). Aside from a dipping tone, T3 was also produced as mid-falling and low-level-rising contours, with a mid-falling contour as the main acoustic realization in MalM. As demonstrated earlier, in *Putonghua,* sandhi

T3 contains the falling and rising contours. Thus, the T3 sandhi rule might explain why T3 has these two variants in MalM. It also closely reflects Sanders's (2008) description of the tone contour of T3 in TwM: T3 occurs more frequently as a falling contour in natural speech than as the citation allotone. Based on the findings above, the falling contour may become the T3 basic contour in TwM as the new generation mainly perceives a falling contour for T3 (Sanders, 2008).

K. Huang (2017) suggested that "another way to explain…tonal variations is to treat them as a kind of sound change" (p. 301). Compared to *Putonghua*, T2 and T3 are considered contour reductions in MalM, as they are produced as level-rising and falling contours. In contrast, they are expressed as rising and dipping contours in *Putonghua*. Pittayaporn (2007) indicated that a dynamic tone is more prone to evolving into a static one than moving in the opposite direction. Therefore, this variation implies that the F0 excursion of the contour tones tends to decrease, resulting in flattened contours in MalM. Correspondingly, T. Huang (2016b) observed contour leveling in T2 in MalM both acoustic and perceptually, suggesting that T2 may eventually be neutralized to a level tone. Likewise, T2 and T3 were sometimes produced as a level contour, which may predict the directionality of tonal change in MalM. Tonal reduction is not confined to MalM; it is also commonly found in SigM and TwM (K. Huang, 2017; Lee, 2010). However, our data and analysis are insufficient to draw a firm conclusion about this prevalence, but it would be worth examining in future research.

Additionally, based on the findings, we can answer the second research question. First, with reference to sources of 'T5', findings from previous studies and the current study cannot provide a conclusive answer, as final types and tone types may also play an active role. Second, besides sources, 'T5' is largely identical to T4 in tonal properties, indicating the lack of distinctive phonemic contrast between these two tones. Despite the significant difference in duration between them, this small quality of difference may not be perceived by listeners. Therefore, this may cause difficulty in distinguishing 'T5' syllables from T4 syllables. Third, 'T5' varies across individuals and may not occur consistently in a single person's pronunciation. For instance, as a salient feature of 'T5', the glottal stop is unstable and thus an optional element. Accordingly, this evidence seems to support the argument that 'T5' is not systematic. Hence, based on the current study's findings, there is inadequate evidence to claim that 'T5' is a new tone category in MalM. Instead, 'T5' is a variant tone. In line with this view, C. Y. Chen (1983) stated that "it is sometimes referred to as 'the short-falling pitch' in preference to 'the fifth tone' as the latter seems to suggest more strongly a completely detached identity" (p. 97). It is also supported by Lee (2010), who stated that his study on SigM focused purely on the four tones given the declining use of 'T5' and the unreliable discrimination between

T4 and 'T5', even for speakers who exhibit this contrast. Our findings indicate a need for more research on 'T5', especially through a perceptual lens.

CONCLUSION

This study examined the tonal properties of the four citation tones and 'T5'. Similar to previous research, four citation tones, T1 to T4, are observed in MalM, and the four tones contrast with each other. However, different variant tones occur in each of the four tones, among which T3 is the most difficult to characterize. Like the neighboring varieties of SigM and TwM, T2 and T3 are undergoing tonal reduction in MalM. Additionally, some similarities and differences in 'T5' exist between MalM and SigM. The tone contour, tone value, and voice quality of 'T5' in MalM differ from those in SigM. However, similar to SigM, our research confirmed that the occurrence of 'T5' in MalM is not only related to the MC-checked tone but also associated with final types and tone categories. More importantly, the acoustic realizations and findings of 'T5' answer our second question that 'T5' is insufficient to be deemed a new tone category in MalM.

This article has clear implications for the GC paradigm. On the one hand, this research pushes forward the empirical results on Mandarin varieties from the perspective of GC, increasing the empirical rigor of this theoretical paradigm; on the other hand, the focus, as well as the outcome, has, as a result, shifted from a conventional emphasis on accuracy in language usage to a more positive exploration of the features of MalM in a multilingual environment. This echoes GC's call for a conceptual shift and a pluricentric way to explore Mandarin varieties. In addition, the findings complement the previous research by focusing on the productions of the young generation and provide insights into the debates of Mandarin 'T5' categorization in prior studies from a diachronic perspective.

The findings of this study can inform Mandarin pedagogy in Malaysia to emphasize awareness of and exposure to the diversity of Mandarin varieties rather than the "correctness" of pronunciation. For example, the primary tone contour of T3 in MalM is mid-falling rather than mid-dipping. The mid-falling T3, however, does not comprise intelligibility, as the mid-dipping contour merely occurs in deliberate and emphatic speech. Hence, it is not necessary to produce a T3 dipping tone in the same way as a speaker of *Putonghua* would. Consequently, it is important to consider teaching Mandarin pronunciation in Malaysia more carefully rather than rigidly adhering to a native model.

There are several potential improvements for future studies, such as including more participants from different age groups and genders to enhance the generalizability and representation of the research results. While smartphone recording offers convenience, it falls short of preserving the original voice quality and may compromise the reliability of the acoustic analysis of segmental features. Therefore, Linear PCM, an uncompressed audio recording format, is essential for accurate acoustic studies of voice quality. In addition, given the use of single-word lists as stimuli and the limited number of 'T5' tokens, results may not reflect the full picture of 'T5'. Future studies could also ascertain the existence of 'T5' by using minimal pairs as stimuli and investigating the tonal phenomenon in polysyllabic words and connected speech.

ACKNOWLEDGMENTS

This study was supported by the Universiti Malaya Faculty of Languages and Linguistics Research Grant (UMFLLRG), project number UMG013N-2023. The authors are indebted to the participants who consented to participate in the data collection and to the reviewers for their valuable comments.

REFERENCES

Alzaidi, M. S., Xu, Y., & Xu, A. Q. (2019). Prosodic encoding of focus in Hijazi Arabic. *Speech Communication, 106*, 127-149. https://doi.org/10.1016/j.specom.2018.12.006

Bao, H. Q., & Lin, M. C. (2018). *Shiyan yuyinxue gaiyao* 实验语音学概要 [The outline of experimental phonetics]. Beijing University Press.

Boersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer* (Version 6.2.04) [Computer program]. http://www.praat.org/

Chao, Y. R. (1967). *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press.

Chao, Y. R. (1968). *A grammar of spoken Chinese.* University of California Press.

Chen, C. Y. (1981). Bendi Huayu li shengdiao de wenti 本地华语里声调的问题 [Tones in Singapore Mandarin]. *Yuwen*, *7*, 38-42.

Chen, C. Y. (1983). A fifth tone in the Mandarin spoken in Singapore. *Journal of Chinese Linguistics, 11*(1), 92-119.

Chen, J. Y. (2007). Malaixiya Huayu diaozhi de yanbian jian yu Putonghua bijiao 马来西亚华语调值的演变兼与普通话比较 [The evolution of the tonal value of Malaysia Mandarin and its comparison to Putonghua]. *Xiandai Yuwen, 5*, 34-35.

Chen, M. Y. (2000). *Tone sandhi: Patterns across Chinese dialects*. Cambridge University Press. https://doi.org/10.1017/CBO9780511486364

Chiung, W. V. (1999, October 28-31). *Taiwanese and Taiwan Mandarin tones: Tonal drift?* [Paper presentation]. 32nd International Conference on Sino-Tibetan Language & Linguistics, Illinois, United States.

Choo, Y. (2013, June 21-23). Fangyan ganrao yu yuyan jiechu de juxian-yi xinjiapo huayu diwusheng wei ge'an方言干扰与语言接触的局限—以新加坡华语第五声为个案 [Constraints of dialect interference in language contact-a case study of the fifth tone in Singapore Mandarin] [Paper presentation]. *The* 25th North American Conference on Chinese Linguistics (NACCL-25). Michigan, United States.

Chua, C. L. (2003). *The emergence of Singapore Mandarin: A case study of language contact* (Publication No. 3113654) [Doctoral dissertation, The University of Wisconsin]. The University of Wisconsin - Madison ProQuest Dissertations & Theses.

Clark, L. (2018). Priming as a motivating factor in sociophonetic variation and change. *Topics in Cognitive Science, 10*(4), 729-744. https://doi.org/10.1111/tops.12338

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Deng, D., Feng, S, & Lv, S. (2006). Putonghua yu Taiwan guoyu shengdiao de duibi fenxi 普通话与台湾国语的对比分析 [The contrast on tone between Putonghua and Taiwan Mandarin]. *Shengxue Xuebao* [*ACTA Acustica*], *31*(6), 536-541. https://doi.org/10.15949/j.cnki.0371-0025.2006.06.011

Diao, Y. B. (2018). Yuyan anquan shijiao xia de quanqiu huayu jiqi yanjiu 语言安全视角下的全球华语及其研究 [Global Chinese and its research in the perspective of language security]. *Yunnan Shifan Daxue Xuebao* (*Zhexue Shehui Kexue Ban*), *50*(3), 25-34.

Duanmu, S. (2007). *The phonology of standard Chines*e. Oxford University Press. https://doi.org/10.1093/oso/9780199215782.001.0001

Fon, J., & Chiang, W. Y. (1999). What does Chao have to say about tones? A case study of Taiwan Mandarin. *Journal of Chinese Linguistics, 27*(1), 13-37.

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. https://doi.org/10.1515/9783110207576.2.239

Fon, J., Chiang, W. Y., & Cheung, H. (2004). Production and perception of the two dipping tones (tone 2 and tone 3) in Taiwan Mandarin. *Journal of Chinese Linguistics, 32*(2), 249-281.

Goh, Y. S., & Xu, X. Y. (2003). Yuyan yiwei guocheng zhong de fangyan yijie bianyi 语言易位过程中的方言移借变异 [The variation of dialectal borrowing in language shift]. In J. A. Zou & R. J. You (Eds.), *Yuyan jiechu lunji* (pp. 395-412). Shanghai Educational Publishing House.

Guan, Y., & Li, B. (2021). Usability and practicality of speech recording by mobile phones for phonetic analysis. *12th International Symposium on Chinese Spoken Language Processing*, 24-26. https://doi.org/10.1109/ISCSLP49672.2021.9362082

Guo, X. (2022). *Quanqiu huayu yufa: Malaixiya juan* 全球华语语法: 马来西亚卷 [Global variation of Chinese grammar: Malaysia volume]. The Commercial Press.

Huang, K. (2017). From pitch contour variation to tone change: An analysis of the 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

Huang, T. (2016a). *On the "fifth tone" of Malaysian Mandarin.* https://www.academia.edu/36168817/On_the_Fifth_tone_in_Malaysian_Mandarin

Huang, T. (2016b). *Topics in Malaysian Mandarin phonetics and phonology* [Doctoral dissertation, National Tsing Hua University]. National Digital Library of Theses and Dissertations in Taiwan. https://hdl.handle.net/11296/393u9b

IBM Corp. (2017). *IBM SPSS Statistics for Macintosh* (Version 25.0). Armonk, NY.

Institute of Linguistics, CASS. (2018). *Fangyan diaocha zibiao* 方言调查字表 [Questionnaire of Chinese characters for dialect survey] (2nd ed.). The Commercial Press.

Jacobi, I. (2009). On variation and change in diphthongs and long vowels of spoken Dutch. [Doctoral dissertation, University of Amsterdam]. UvA-DARE (Digital Academic Repository). https://hdl.handle.net/11245/1.316951

Khoo, K. U. (2013). Lun "huayu" yu malaixiya huayu yanjiu 论"华语"与马来西亚华语研究 [On "huayu" and its research]. In K. U. Khoo (Ed.), *Malaixiya huayu yanjiu lunji* (pp. 240-257). Huashe Yanjiu Zhongxin.

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

Kim, C., Reddy, S., Stanford, J. N., Wyschogrod, E., & Grieve, J. (2019). Bring on the crowd! Using online audio crowd-sourcing for large-scale new England dialectology and acoustic sociophonetics. *American Speech, 94*(2), 151-194. https://doi.org/10.1215/00031283-7251252

Kuang, J.-J. (2017). Creaky voice as a function of tone categories and prosodic boundaries. *Interspeech*, 3216-3220. https://doi.org/10.21437/Interspeech.2017-1578

Kuo, J. (2018). A large-scale smartphone-based sociophonetic study of Taiwan Mandarin. *Asia-Pacific Language Variation, 4*(2), 197-230. https://doi.org/10.1075/aplv.18005.kuo

Lee, L. (2010). The tone system of Singapore Mandarin. In L. E. Clements & C. M. Liu (Eds.), *Proceedings of the 22nd North American Conference on Chinese Linguistics* (*NACCL-22*) *& the 18th International Conference on Chinese Linguistics* (*IACL-18*; pp. 345-362). Harvard University.

Leemann, A., Kolly, M.-J., Purves, R., Britain, D., & Glaser, E. (2016). Crowdsourcing language change with smartphone applications. *PLOS ONE, 11*(1), Article e0143060. https://doi.org/10.1371/journal.pone.0143060

Lehiste, I. (1970). *Suprasegmentals*. MIT Press.

Li, A., Xiong, Z., & Wang, X. (2006). Contrastive study on tonal patterns between accented and standard Chinese. In Q. Huo, B. Ma, E. S. Chng, & H. Li (Eds.), *Proceedings of the 5th International Symposium on Chinese Spoken Language Processing* (pp. 157-168). Springer.

Li, B., Guan, Y., & Chen, S. (2020). Carryover effects on tones in Hong Kong Cantonese. *10th International Conference on Speech Prosody 2020,* 489-493. https://doi.org/10.21437/SpeechProsody.2020-100

Li, Y. (2010). *Quanqiu huayu cidian* 全球华语词典 [*A dictionary of Global Huayu*]. The Commercial Press.

Li, Y. (2016). *Quanqiu huayu dacidian* 全球华语大词典 [*A comprehensive dictionary of Global Huayu*]. The Commercial Press.

Liu, L. (2007). *Jianghuai fangyan shengdiao shiyan yanjiu he zhedu fenxi* 江淮方言声调实验研究和折度分析 [*An experimental study and deflection analysis of tone systems in Jianghuai dialects*]. Bashu Publishing House.

Ma, J. L. (2002). *Hanyu tonglun* 汉语通论 [*The general theory in Chinese*]. Jiangsu Classics Publishing House.

Ng, Y. Y., & Chiew, P. P. (2012). The origin of the fifth tone in Singapore Mandarin. In *Proceedings of the URECA@NTU 2011-2012*. Nanyang Technology University. http://hdl.handle.net/10356/67854

Pillai, S., Don, Z. M., Knowles, G., & Tang, J. (2010). Malaysian English: An instrumental analysis of vowel contrast. *World Englishes, 29*(2), 159-172. https://doi.org/10.1111/j.1467-971X.2010.01636.x

Pittayaporn, P. (2007). Directionality of tone change. *16th International Congress of Phonetic Sciences*. http://www.icphs2007.de/conference/Papers/1516/1516.pdf

Rose, P. J. (1989). On the non-equivalence of fundamental frequency and pitch in tonal description. In D. Bradley, E. J. A. Henderson, & M. Mazaudon (Eds.), *Prosodic analysis and Asian linguistics*: *To honour R. K. Sprigg* (pp. 55-82). Pacific Linguistics.

Sanders, R. (2008). Tonetic sound change in Taiwan Mandarin: The case of Tone 2 and Tone 3 citation contours. In M. K. M. Chan & H. Kang (Eds.), *20th North American Conference on Chinese Linguistics (NACCL-20;* pp. 87-107). The Ohio State University.

Shi, F., & Deng, D. (2006). Putonghua yu Taiwan Guoyu de yuyin duibi 普通话与台湾国语的语音对比 [A phonetic contrast of Putonghua and Taiwan Mandarin]. In D.-A. Ho, S. Cheung, W. Y. Pan, & F. X. Wu (Eds.), *Shanggao shuichang: Ding Bangxin xiansheng qizhi shouqing lunwenji* (pp. 371-393). Institute of Linguistics, Academia Sinica.

Shi, F., & Wang, P. (2006). Beijinghua danziyin shengdiao de fenzu tongji fenxi 北京话单字音声调的分组统计分析 [*A statistical analysis of tone groups in Beijing Mandarin*]. *Dangdai Yuyanxue*, *8*(4), 324-333; 379-380. https://doi.org/10.3969/j.issn.1007-8274.2006.04.003

Stanford, J. N. (2019). *New England English large-scale acoustic sociophonectis and dialectology*. Oxford University Press. https://doi.org/10.1093/oso/9780190625658.001.0001

Wang, L., Gussenhoven, C., & Liang, J. (2018). How to pronounce a low tone: A lesson from Kaifeng Mandarin. *Journal of the International Phonetic Association, 50*(2), 199-219. https://doi.org/10.1017/S0025100318000270

Wang, X. M. (2019). "Gu, Fang, Pu, Wai"—lun quanqiu huayu yanjiu de sige shijiao "古、方、普、外"—论全球华语研究的四个视角 [On the four perspectives of Global Chinese studies: classical Chinese, dialects, Putonghua and foreign languages]. *Global Chinese, 5*(1), 45-57. https://doi.org/10.1515/glochi-2019-0003

Wang, X. M. (2020). Cong Malaixiya huayu kan quanqiu huayu de kuajing guihua 从马来西亚华语看全球华语的跨境规划 [On language planning of Global Chinese based on Malaysian Mandarin]. *Global Chinese, 6*(1), 165-177. https://doi.org/10.1515/glochi-2020-0008

Xu, D. M, & Wang, X. M. (2009). Quanqiu huayu shequ shuolue 全球华语社区说略 [Research on Global Chinese Community]. *Jilin Daxue Shehui Kexue Xuebao, 49*(2), 132-137.

Xu, X. Y., & Goh, Y.-S. (2001). Xinjiapo huayu "diwusheng" yishi guochengzhong de "cihui kuosan" xianxiang 新加坡华语"第五声"遗失过程中的"词汇扩散"现象 [The lexical diffusion in the loss of the "fifth tone" in Singapore Mandarin]. *Proceedings of the 5th Phonetic Conference of China*. Tsinghua University Press.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics, 25*, 61-83. https://doi.org/10.1006/jpho.1996.0034

Xu, Y. (2013). ProsodyPro: A tool for large-scale systematic prosody analysis. *Tools and Resources for the Analysis of Speech Prosody* (*TRASP 2013*). Aix-en-Provence, France.

Yamaguchi, T., & Chiew, P. S. (2019). Is there conflation? An acoustic analysis of vowels in Japanese English. *Asian Englishes, 22*(1), 35-51. https://doi.org/10.1080/13488678.2018.1563665

Yeoh, Y. Y. (2019, October 21). *Malaixiya huayu qusheng yanjiu* 马来西亚华语去声研究 [On Qusheng in Malaysian Mandarin]. Malaixiya Huayu yanjiu luntan [Forum on Malaysian Mandarin Studies], Sepang, Malaysia.

Zhang, C., Jepson, K., Lohfink, G., & Arvaniti, A. (2021). Comparing acoustic analyses of speech data collected remotely. *The Journal of the Acoustical Society of America, 149*, 3910-2916. https://doi.org/10.1121/10.0005132

Zhang, J. W. (2014). *A sociophonetic study on tonal variation of the Wúxī and Shànghǎi dialects*. Netherlands Graduate School of Linguistics / Landelijke (LOT).

Zheng, Y. L. (2010, May 28-30). *Shilun yuyin shengxue canshu* 试论语音声学参数 [On acoustic parameters in Phonetics]. Dijiujie Zhongguo yuyinxue xueshu huiyi, Tianjin, China.

Zhu, X. N. (2018). *Yuyin xue* 语音学 [*Phonetics*] (3rd ed.). The Commercial Press.

APPENDIX

Target words in the word list

MC tones	Mandarin word	Phonetic form	Glossary	MC tones	Mandarin word	Phonetic form	Glossary
	低	ti	low		扶	fu	help
	天	tʰian	sky		时	ʂʅ	time
	猪	tʂu	pig		茶	tʂʰa	tea
	车	tʂʰɤ	vehicle		逃	tʰau	escape
MCT1	高	kau	tall	MCT2	鞋	ɕiɛ	shoes
	花	xua	flower		桥	tɕʰiau	bridge
	伤	ʂaŋ	injury		房	faŋ	house
	浇	tɕiau	water		田	tʰian	field
	灯	tɤŋ	lamp		陈	tʂʰən	surname
	比	pi	compare		父	fu	father
	打	ta	hit		社	ʂɤ	bureau
	古	ku	ancient		米	mi	rice
	粉	fən	pink		件	tɕian	piece
MCT3	展	tʂan	exhibit	MCT4	动	tuŋ	move
	等	tɤŋ	wait		满	man	full
	口	kʰou	mouth		老	lau	old
	火	xuo	fire		坐	tsuo	to sit
	水	ʂuei	water		罪	tsuei	crime
	注	tʂu	infuse		大	ta	big
	怕	pʰa	afraid		树	ʂu	tree
	替	tʰi	replace		寺	sɿ	temple
	帐	tʂaŋ	curtain		阵	tʂən	position
	唱	tʂʰaŋ	sing		病	piŋ	disease
MCT5	汉	xan	a dynasty of China	MCT6	共	kuŋ	together
	盖	kai	cover		害	xai	harmful
	化	xua	change		夏	ɕia	summer
	跳	tʰiau	jump		队	tuei	team
	失	ʂʅ	lose		罚	fa	punish
	吃	tʂʰʅ	eat		十	ʂʅ	ten
	福	fu	blessing		笛	ti	whistle
	刻	kʰɤ	to carve		舌	ʂɤ	tongue
MCT7	笔	pʰi	pen	MCT8	俗	su	custom
	割	kʰɤ	cut		达	ta	arrive
	黑	xei	black		六	liou	six
	百	pai	hundred		白	pai	white
	铁	tiɛ	iron		贼	tsei	thief