Journal of Foreign Language Teaching and Learning
Volume 8, No. 1, January 2023
Available online at: https://journal.umy.ac.id/index.php/FTL/issue/view/964
e-ISSN: 2580-2070, p-ISSN: 2527-7650

Article History:
Submitted:
16 October 2022
Reviewed:
30 November 2022
Edited:
30 January 2023
Article Accepted:
3 March 2023

Check for updates

Production of Mandarin Chinese Tones by Sri Lankan CFL Learners: An Acoustic Analysis

Noel Dassanayake

Department of Languages, Faculty of Social Sciences and Languages, Sabaragamuwa University of Sri Lanka, Sri Lanka

Author email: noel@ssl.sab.ac.lk

DOI: https://doi.org/10.18196/ftl.v8i1.16536

Abstract

Tonal variation in Mandarin forms a relationship with the meanings of words at a lexical level. Tones and tone sandhi are considered distinctive features of Mandarin Chinese phonetics; thus, acquiring accurate Mandarin pronunciation is challenging for speakers of other languages. The present study examined the production of Mandarin tones by Sri Lankan learners through acoustic analysis of f0 using Praat. The study participants were seven (n=7) undergraduate students at intermediate-level Chinese. Each participant recorded a sample of 20 elements, totaling 140 at three tiers: isolation, disyllabic form, and synthesis. The recorded acoustic data were analyzed using Praat 5.4.04 and SPSS Statistics 21. The results indicated that the mean f0 values of the realization of the four tones by the informants are heterogenous at all three tiers. The realization of T3 demonstrated a low accuracy, while T2 was realized with the highest accuracy at all three tiers. As revealed by the tone sandhi analysis, participants have resorted to the underlying representation of the tone sandhi than the surface realization.

Keywords: Sri Lankan Mandarin Learners; Tone Production; Tone Sandhi; Acoustic Analysis

Introduction

Researchers have approached teaching Mandarin pronunciation from different perspectives over time (Třísková, 2017; Halliday, 2014; Biduri, 2017). Mandarin Chinese syllables usually comprise initial, final, and tone. The initial usually is a consonant, and the final comprises vowel(s) or vowel(s) followed by nasal. The tone is the pitch variation in sounds, and semantic encoding in Mandarin Chinese depends on lexical tones (Chen et al., 2013). Second and foreign language researchers have taken different approaches to solve issues of teaching Mandarin Pronunciation. Machine-aided modeling, perception, and analysis of Mandarin pronunciation have become popular in recent studies (Zheng et al., 2007; Beasley & Wang, 2021).

There is little argument for the fact that Mandarin Chinese consists of one of the most intricate phonetic systems among other international languages, largely due to its tonal variation. Mandarin tonal variation is one of the most distinctive supra-segmental features of Chinese phonology, which is also often perceived as a relatively challenging area in the pedagogy of the Chinese language. Accuracy in perception and production of tones is paramount to Mandarin Chinese learners for communicating successfully in the target language (Li & Lee, 2021). Jiang and Cohen (2012) categorized tonal and pitch variation as a key feature that makes it distinctively different from other languages where minor alterations may also result in a change of meanings.

Although there have been considerable developments in teaching the Chinese language in Sri Lanka, a systematic approach to the production and perception of Mandarin tones has hardly been taken in recent research. The present study examines the accuracy of tone production by Sri Lankan Mandarin Chinese learners with acoustic analysis tools. Two key attributes assure the originality of the study of it. From a methodological perspective, acoustic analysis tool and their attributes have hardly been used in the limited number of studies conducted on the production of Mandarin Chinese tones by Sri Lankan learners. From a theoretical perspective, recent studies on the production of Mandarin Chinese tones have not distinguished between the production of tones in isolation and synthesis, most of which have overlooked the latter. The present study examines the acoustic attributes of tone productions by Sri Lankan Mandarin Chinese learners. The key problems addressed by the study include how Sri Lankan learners produce Mandarin Chinese tones in isolation and synthesis and how the pitch contours of each tone differ under these two different circumstances. It is expected that the results of this study will be supportive for recognizing the scope and level of erroneous tone production by Sri Lankan Chinese Language learners who will be instrumental in designing teaching material and approaches.

Literature Review

Pitch is significant in all languages, but its role differs from one language to another. (Zhang, 2018, p. 26) Although Mandarin Chinese and other non-tonal languages use prosodic variations for various purposes, Mandarin Chinese uses tonal variations phonologically contrastively at a lexical level (Hallé et al., 2004, p. 5). However, the classification of languages into tonal and non-tonal languages and the approaches to such classifications have been controversial. According to Duanmu (2004, p. 893), the typology of tone and non-tone languages can be analyzed from four perspectives: lack of distinction between tone and non-tone languages, similarities among tone and non-tone languages, differences within tone languages, and lack of guidelines for research. It could also be observed that the levels and extents of tonal influence vary from one language to another. Contour tone languages use temporally changing pitch trajectories, while register tone languages use only level pitches. While some languages apply tone values to limited syllables, others apply them to all syllables (Best, 2019, p. 1).

As emphasized in the introductory section, tonal variation is a significant feature of Mandarin Chinese, and researchers have attempted to analyze them from different perspectives over time. Standard Chinese (hereafter referred to as SC) consists of four tones, usually T1 through T4. According to Duanmu (2002), primary acoustic correlates of tones are two-dimensional. The first and the most dominant is the fundamental frequency (F0), and the second is ‘murmur’ or ‘breathiness,’ of which the latter often correlates with a broader formant width and flatter spectral envelope. Duanmu further suggested that factors like consonant voicing, aspiration, vowel height, the position of the syllable, or stress factors can affect f0. Zhiming claimed that tone consists of two aspects, register and contour, and the register node and contour node play different conceptual roles (Zhiming, 1999, p. 3).

The four tones in SC are generally identified according to pitch fluctuation in speech. The first tone, which is generally called the píngshēng, has a flat high-level tone which is usually identified as a 55 pitch with the symbol ̄ on top of pinyin, which is also called *yinping*. The second tone is a rising tone often identified with the diacritic á. The third tone is realized as a low-falling tone which also can occur as a low-dipping tone in isolation or at the end of a phrase. The fourth tone is falling (Liu & Xiao, 2021, p. 144). The four tones in Mandarin Chinese are often identified as T1, T2, T3, and T4.

There are substantial dialectical differences in tonal variation and pitch contour in Chinese. According to Liu et al. (2020), while some dialects differ from SC in segmental and tonal levels, other dialects within the Mandarin family largely overlap with segmental information with SC. Liu et al. claimed that tones are systematically mapped between Xi'an Mandarin and Standard Chinese. Cantonese is one of the most widely spoken dialects in the Chinese language and has distinctive prosodic features in contrast to SC. Matthews and Yip (2013) claim that one of the key prosodic features of Cantonese, in contrast to SC, is the absence of tone sandhi. There are nine tones in Cantonese with two variants of the high tone, which differs from Hongkong Cantonese, which only uses one high tone (So, 1997).

SC tones may undergo alterations depending on the preceding tone of the previous syllable, usually referred to as tone sandhi. Tone sandhi occurs when two adjacent tones have 'conflicting collocations,' which do not invite easy natural transition but require manipulation of the larynx. It does not occur in compatible collocations such as the high tone followed by a falling tone which allows a smooth transition (Wee, 2019). According to Shih (2013) Mandarin third-tone sandhi provides strong evidence for a hierarchically organized prosodic structure and allows researchers to investigate some important phonological concepts. ss

Mandarin Chinese tones do not function in isolation but occur in connected speech. Thus in authentic pronunciation, they may undergo co-articulatory and tone sandhi alterations according to context. Coarticulation refers to the minor alterations of tones at a phonetic level, and tone sandhi refers to categorical changes at the phonological level (Zhang, 2018, p. 8). According to Yang (2015), tone sandhi refers to the rule-governed tonal changes in some specific contexts, and Mandarin Chinese has three tone sandhi rules, namely the neutral tone sandhi, T3 tone sandhi, and "yi-bu-qi-ba" rule (p. 9). Xu (1997) claimed that some tonal transition has to occur when two lexical tones are produced in sequence, especially when there is uninterrupted voicing through the two syllables that carry them (p. 62).

Perception, Production, and Teaching of Mandarin Chinese Tones

The influence of tonal variation at the lexical level demands that Chinese language learners acquire a systematic knowledge of the MC sound system. Xing (2006, p. 99) proposed that since the knowledge of the Chinese sound system is useful for Chinese language students, they should become familiar with it along with Chinese characters. Kan and Ito (2020) investigated MC tones and proposed that auditory cognitive ability in the second language is an essential factor in the perception and production of pronunciation (p. 12). They also emphasized the importance of polysyllable perceptual training and tone comparison.

Researchers have analyzed the perception and production of Mandarin Chinese tones and their relationship from different perspectives over the years. After a methodical analysis of the perception and production of Mandarin Chinese tones, Elliott concluded that while learners find it difficult to perceive and produce some tones, other tones only pose either perception or production issues to the learner (Elliott, 1991, p. 196). According to Moore and Jongman (1997), the perception of tones is a "talker-contingent" process and that tones are not dependent on concrete acoustic values but rather contrast with other tones in utterances as well as of f0 range to gain a relative identity (p. 1875).

According to Zhang (2018), factors such as native language background, music background, Chinese language level, tone acquisition methods, and stimulus materials influence the perception of Mandarin Chinese tones (p. 602). Native language phonology is a key factor influencing MC tone production by speakers of other languages. Zhang and Chen (2018) have examined the influence of L1 Kazakh and L2 Russian on the tone production of L3 Mandarin and claim that the effects of L1 on L3 Mandarin production are relatively higher than the influence of L2 (p. 240). Recent studies signify that the speakers' L1 and L2 segmental and supra-segmental features can influence the acquisition of MC tonal variation. The general assumption is that L2 lexical tone discrimination poses more threats to learners whose L1 does not rely on tonal variation at the lexical level (Wang, 2013, p. 144).

Studies have revealed that orthography could also influence the production of Mandarin Chinese tones by L2 learners. Tonal variation is not Chinese's only specialty, making it stand out from other global languages. Chinese characters account for one of the most effective writing

Available online at: https://journal.umy.ac.id/index.php/FTL/issue/view/964
e-ISSN: 2580-2070, p-ISSN: 2527-7650

systems and have brought new paradigm shifts to the word concept. As claimed by Pellatt et al. (2014), it has been observed that the 'visual' nature of Chinese characters makes their meaning directly accessible visually. Therefore, it is a question of how learners perceive tones' segmental and supra-segmental features when morphemes are presented as Chinese characters and pinyin. An empirical study by Li et al. (2014) has discovered that speakers with low proficiency in MC tend to be more prone to errors in tone production when words are presented in pinyin compared to characters.

Although the four tones in Mandarin Chinese are generally identified with distinctive pitch and contour features, their realization in actual speech varies from one context to another. Yang et al. (2017) claimed that the duration of tones might vary depending on different speech contexts. In isolated monosyllabic occurrences, the four tones have demonstrated significant variations in length (p. 1410). The T4 has the lowest length, while the T3 has the highest length. Han et al. (2018 p. 101) claimed that although the realization of T3 and T4 are quite distinct when produced in isolation, they can hardly be discriminated from each other in connected speech. Chen and Peng (2016) examined the perception of T1 and T2 by native Chinese speakers and claimed that tones are influenced in a contrastive manner only within normal speech context (p. 260).

Recent studies show that segmental features could influence the realization of Mandarin tones. Xu and Xu (2003) revealed that consonant aspiration remarkably affects the following onset F0 (p. 19). Studies have revealed that speech tempo also could affect the realization of Mandarin tones. Based on an acoustic analysis, Tang and Li (2020) claimed that Mandarin tones exhibit flatter and higher pitch contours in fast speech compared to speech of normal tempo. In addition, it has also been found that the tone area in fast speech is relatively smaller than in normal speech (p. 1940).

According to Van de Weijer and Sloos (2014), the acquisition of Mandarin Chinese in the progressive sequence is T1>T4>T2 and T3. Sun et al. (2018) claimed that pronunciation of T2 and T3 is the most challenging for speakers of non-tone languages owing to the similar F0 contours of Tone 2 and 3 (p. 2549). Recent studies on the perception of MC tones show that both foreigners' and native Chinese speakers' perceptions and production of tones differ from one tone to another. As claimed by Rhee et al. (2020), children's production of tones achieves and exceeds the adult level of tonal distinction at the age of 7 to 8, and the allophonic spectral cues also exhibit a sharp growth spurt between ages 4 to 5 and 7 to 8 (p. xx).

Different approaches have been taken over time in teaching and training non-native speakers on acquiring MC tones. Teaching Mandarin tones and tonal variation to speakers of other languages has been perceived as one of the most challenging tasks. It has received the attention of many researchers over the world. Learners of different language backgrounds encounter different issues in perceiving and producing the four tones in Chinese. According to a study by Guo and Tao (2008), the sample American students demonstrated higher performances in T1 and T4 than in T2 and T3. According to Dong et al. (2013), some Japanese learners find the initial syllables more difficult to perceive than the final syllables, which is not the tendency in production. Elliot (1991) stressed the need to make students consciously aware of their tone errors to solve their tone perception and production issues (p. 196). According to Han et al. (2018), exaggerated pronunciation produced in teaching styles does not influence the average pitch and the pitch range of MC tones (p. 732).

F0 adjustments in the production of tones and intonation are the two key challenges to Mandarin Chinese tone perception, and beginners are the ones who struggle most with the identification of tones (Yang & Chan, 2010, p. 26). Mother tongue or L2 influence is another key challenge in the perception and production of MC tones by non-native speakers. According to Rungruang and Mu (2017), who have taken CFL students in Thailand as a sample, L1 has considerably influenced the acquisition of tones by Thai speakers (p. 112). They further claimed that the realization of the MC T4 is much closer to the falling tone in Thai. Li (2015) found that Thai speakers outperformed native-English speakers in the accuracy of pronouncing tonal variation of MC and further proposed that L1 of non-native speakers can significantly influence the realization of MC tones (p. 129).

Methodology

Participants

Journal of Foreign Language Teaching and Learning
Volume 8, No. 1, January 2023
Available online at: https://journal.umy.ac.id/index.php/FTL/issue/view/964
e-ISSN: 2580-2070, p-ISSN: 2527-7650

Key participants of the present study were ten undergraduate students from a Sri Lankan higher education institute who were studying for a BA Honors Degree Programme in the Chinese Language, which constituted a convenience non-probability sampling. All informants had at least three years of Chinese language learning experience and had learned MC from native-Chinese and local Sri Lankan teachers. All respondents were between the ages of 20-25 and were all Sri Lankans by nationality.

Instruments and Recording

The audio recording platform Yubao Sheluji was used to record audio data; all audio files were recorded as .wav files. Yubao Sheluji (语保摄录机) is an audio recording software specially designed for Natural Language Processing which allows recording a transcribed list of sounds in the form of a spreadsheet. After recording each file, the software automatically sorts stores and renames the list of output audio files according to the spreadsheet. The respondents were provided with a spreadsheet that included four monosyllabic words with four tones, nine disyllabic words with tone sandhi combinations, and seven sentences, each containing a monosyllabic character in the first set. The participants were only provided a very short time to familiarize themselves with the words before recording to ensure their tone production's genuineness.

Table 1 – Sample Word/Phrase List

In isolation	Disyllabic	*Sandhi*	In synthesis
妈	妈妈	你好	他妈妈去商店了 tā māmā qù shāngdiànle
鱼	钓鱼	不对	我们常去钓鱼 wǒmen cháng qù diàoyú
椅	椅子	一次	这是爸爸的椅子 zhè shì bàba de yǐzi
怕	很怕	一般	我很怕这种够 wǒ hěn pà zhè zhǒng gòu
		一毛	我去过北京一次 wǒ qùguò běijīng yīcì
			斯里兰卡人一般喝红茶 sīlǐlánkǎ rén yībān hē hóngchá
			请给我一毛钱 qǐng gěi wǒ yī máo qián

Methods

Praat 5.4.04 was used as the acoustic analysis tool that allows a graphical and statistical analysis of audio files in Natural Language Processing. Praat recorded the f0 of tones in specific intervals within the total length from 10% to 90% and automatically calculated statistical data such as the Standard Deviation and Mean frequency of the sound. Praat also offered a variety of graphical representations for a sound, including waveforms, spectrograms, and text grid analysis. The f0 of each of the 120 audio files was analyzed from multiple perspectives using the pitch function of Praat. Each syllable with the respective tones was extracted using Praat in disyllabic and synthesized productions, and the f0 was extracted periodically from 10% to 90%. The realization of each tone was analyzed in three dimensions: isolation, disyllabic form, and synthesized speech. The pitch contour of each of the three Tone sandhi rules was separately analyzed. A Qualitative Comparative Analysis (QCA) was conducted by comparing the circumstances and causal contributions of the production of four tones that resulted in accurate or inaccurate production of the same.

Findings

Results of the Acoustic Analysis

The acoustic analysis demonstrated that the informants' realization of Mandarin Chinese tones is diverse in isolated production and synthesis. Informants demonstrated a significantly low accuracy in the production of T3 and a relatively higher accuracy in the production of T2, which could be related to the influence of their L1 and L2 intonation pitch patterns. The onset of the informants is different within the production of the same tone, which is especially evident in T2 and T3. There is an abstract relationship between the production of tones by the informants in isolation, disyllabic form, and synthesis. Realizing the T4 sandhi with Bù demonstrates the lowest accuracy out of the three sandhi rules.

Isolated Pronunciation

Table 1 to 4 demonstrates the systematic analysis of pitch levels of the four tones produced by the informants. Production in isolation depicts the speaker's perception of the four tones. Table 1 shows that the standard deviation of pitch height in T1 is less than 10 Hz in four speakers (57.14%), and except for S3, all other speakers' standard deviation remains below 20 Hz. The high average pitch of S3 is the exceptionally high pitch level at the offset.

Table 2 - f0 of the production of T1(妈) in isolation

f0	S1 Hz	S2 Hz	S3 Hz	S4 Hz	S5 Hz	S6 Hz	S7 Hz
90%	295.59	346.37	529.82	299.14	329.18	303.72	277.07
84%	295.97	346.00	474.68	296.29	326.03	302.69	276.41
50%	299.85	343.18	311.19	279.88	318.67	298.70	258.48
16%	304.56	340.09	305.45	274.91	314.00	292.04	242.92
10%	308.68	338.72	304.54	274.10	313.42	291.06	239.93
Average	300.93	341.68	364.28	281.94	319.83	298.35	259.77
Std. dev.	5.201	6.732	92.58	10.44	5.598	4.489	14.14

Table 2 demonstrates the periodic pitch analysis of T2, and the standard deviations of S1 to S6 are close to each other except for S7. The average pitch ranges from 214 Hz to 355 Hz, and the average pitch at 10% is 235.47 Hz. The average pitch at 90% is 333.80 Hz.

Table 3 - f0 of the production of T2(鱼)in isolation

f0	S1 Hz	S2 Hz	S3 Hz	S4 Hz	S5 Hz	S6 Hz	S7 Hz
90%	366.35	262.69	363.56	317.56	352.09	416.09	258.31
84%	336.73	260.04	346.52	293.65	348.16	412.34	256.23
50%	234.51	225.86	214.29	240.84	339.29	361.90	201.50
16%	226.87	199.46	206.46	223.58	333.40	316.89	189.16
10%	225.65	199.11	204.99	217.58	328.99	283.15	188.85
Average	264.82	230.02	251.32	261.10	349.54	355.94	214.41
Std. dev.	52.23	57.37	64.22	60.3	44.52	51.94	27.26

Periodic pitch analysis of T3 is depicted in Table 3, and the average pitch ranges from 206.60 Hz to 269.72 Hz. The standard deviation shows a much higher diversity than T1 and T2. The average pitch at 10% is 205.81 Hz which at 90% rises to 247.16 Hz. However, the difference between the pitch heights of informants at each periodic level has a high diversity than T1 and T2.

Table 4 - f0 of the production of T3(椅)in isolation

f0	S1 Hz	S2 Hz	S3 Hz	S4 Hz	S5 Hz	S6 Hz	S7 Hz
90%	90.05	262.38	251.77	251.77	293.39	324.35	256.41
84%	332.27	258.83	250.12	250.12	290.25	264.01	253.47
50%	241.89	220.10	238.02	238.02	259.03	211.20	197.94
16%	228.81	192.69	224.27	224.27	226.91	177.19	183.99
10%	227.98	190.33	220.30	220.30	224.14	174.70	182.94
Average	269.72	224.08	237.62	237.62	263.95	221.85	206.60
Std. dev.	63.66	26.87	12.38	12.38	42.94	51.12	31.29

Table 4 includes the periodic pitch analysis of T4 produced in isolation. It indicates that the standard deviation has the highest diversity in the realization of T4 by the informants, ranging from 5.67 Hz to 135.2 Hz.

Table 4 - f0 of the Production of T4 (怕) in isolation

f0	S1 Hz	S2 Hz	S3 Hz	S4 Hz	S5 Hz	S6 Hz	S7 Hz
90%	102.95	317.49	327.36	180.86	242.81	362.80	264.31
84%	111.85	318.58	335.86	187.90s	244.57	375.60	270.36
50%	256.93	325.76	357.34	288.53	268.90	457.83	287.19
16%	296.93	330.79	371.76	308.71	562.63	531.60	304.03
10%	300.37	332.54	373.12	311.51	570.60	534.61	304.79
Average	235.18	325.11	354.51	263.93	356.44	454.23	286.96
Std. dev.	71.51	5.672	17.36	53.77	135.2	76.6	14.73

Figure 1 demonstrates the onset of each informant in pronouncing the four tones in isolation. The onset of T1 has an average pitch of 294.8 Hz with a standard deviation of 29.77.

T2 has been realized with an average onset pitch of 254.41 Hz, and the standard deviation is 54.25. The onset of T3 has an average pitch height of 243.98 Hz, and the standard deviation is 32.46. The average pitch height of T4 at onset is 340.8 Hz, and the standard deviation is 37.76, while the average onset of T4 is the highest out of the four tones, and T3 has recorded the lowest averages.

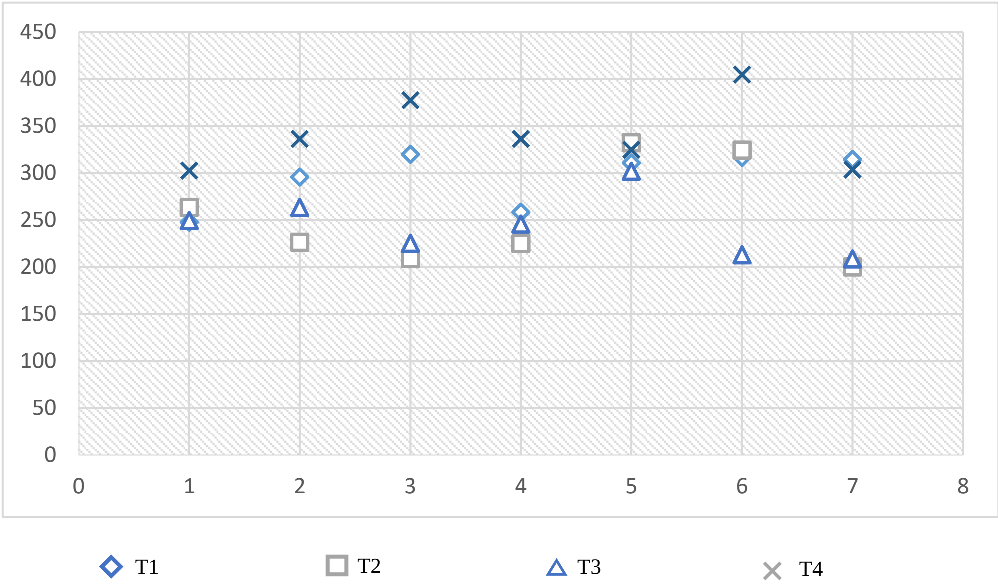
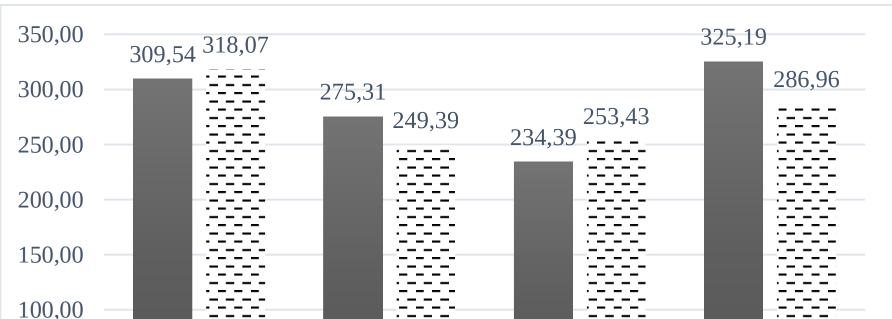


Figure 1 Average Onset of Each Tone

Isolated Monosyllables vs. Isolated Disyllabic Production

Figure 2 shows that the average pitch height is higher when T1 and T3 are produced in disyllabic forms, while in T2 and T4, the average pitch height is higher when pronounced in isolation. The analysis of the disyllabic word 妈妈 māma demonstrates that speakers 6 and 7 have attempted to realize both tones as a T1 despite the normal realization of the 2nd T1 as a neutral tone. Only speakers 1 and 3 have realized the second T2 as a neutral tone depicted by the abrupt pitch slope at offset in contrast with the onset.



Journal of Foreign Language Teaching and Learning
Volume 8, No. 1, January 2023
Available online at: https://journal.umy.ac.id/index.php/FTL/issue/view/964
e-ISSN: 2580-2070, p-ISSN: 2527-7650

One of the significant findings from the analysis of T2 is that in disyllabic production, several speakers have intuitively corrected erroneous productions they made in isolation. For example, speaker 5 has realized that T2 much similar to a T1 with a standard deviation of 44.52 Hz, and the same speaker has produced T2 in disyllabic 钓鱼 diàoyú closer to a T2 with a standard deviation of 60.92 Hz. T3 records the highest number of erroneous realizations. 85% of the informants have produced the T3 in isolation with a higher offset pitch value than at the offset. In the disyllabic form, speakers 1, 2, and 4 do not demonstrate a rising pitch but a linear falling or rising similar to either T2 or T4.

Only speakers 5,6, and 7 show a falling rising pitch curve which is also not as distinct as a Standard Chinese T3 pitch curve. There is no significant difference between the accuracy of T3 production by the speakers in isolation and disyllabic counterparts. Speakers 2 and 5 have produced T4 with less accuracy in isolation, while the same speakers have produced T4 with higher accuracy in the disyllabic forms. For example, in the case of speaker 2, who has produced T4 in isolation with a standard deviation of 5.672 Hz, has produced T2 in the disyllabic 很怕 hěnpà with a standard deviation of 66.03 Hz.

Pronunciation in Synthesis

Figure 3 shows the realization of T4 in isolation and synthesis by the seven speakers. Pitch analysis shows a significant difference between the average pitch height at onset in isolation and synthesis. The average pitch height in isolation (342.77 Hz) is higher than in T1than at synthesis (330.8 Hz). Similarly, the average pitch is higher in isolated (269.67 Hz) production of T2 than in synthesis (251.48 Hz). T3 has recorded a higher pitch height in synthesis (278.85 Hz) than in isolation (225.91 Hz). Average pitch height of T4 in synthesis (313.84 Hz) than in isolated production (295.07 Hz).

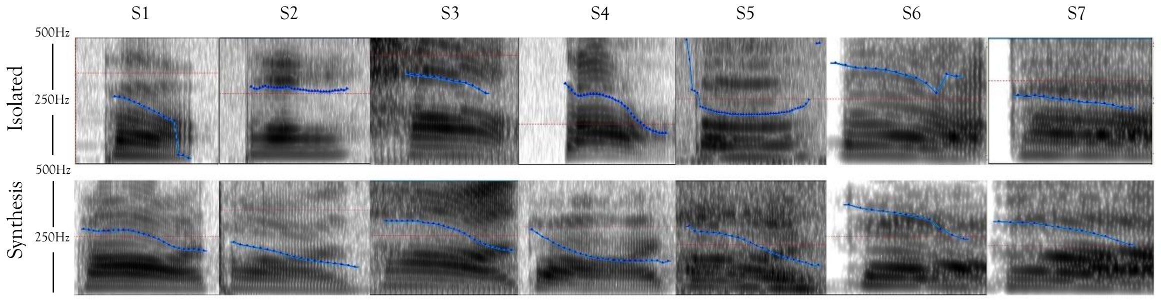


Figure 3 - Realization of T3 in Isolation and Synthesis

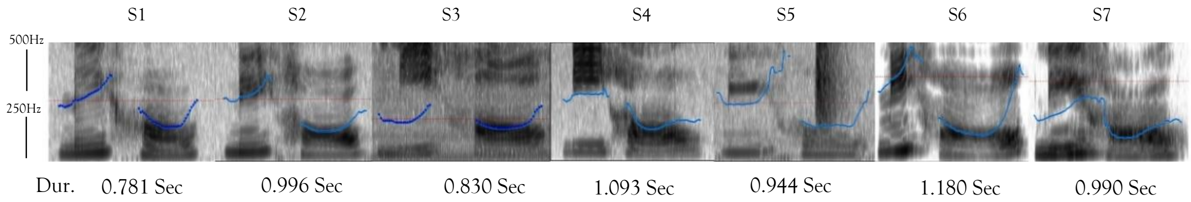


Figure 4 - Realization of T3 *Sandhi*

Third-tone sandhi was analyzed using the phrase 你好 with a 3+3 combination which is often realized as 2+3. Acoustic data demonstrate that the first T3 has been realized as a T2 in 5 cases, while the remaining two cases slightly differ. Speaker 3 has produced the first T3 much similar to a T3, while Speaker 4 has realized the first T3 is much similar to a T1 with a pitch variation of 90% -10% = 25.09 Hz. The total duration of the sandhi varies from 0.781 sec. to 1.180 sec. The average pitch of the second T3 ranges from 187.72Hz to 221.37Hz, with an average of 207.93Hz. From Figure 4, except for speaker 4, all other speakers' realization of the second T3 is more inclined towards a T2 with a higher pitch level at the offset. The average onset of T3 sandhi is 282.414 Hz, and the average offset is 286.928 Hz.

T4 Sandhi with Bù

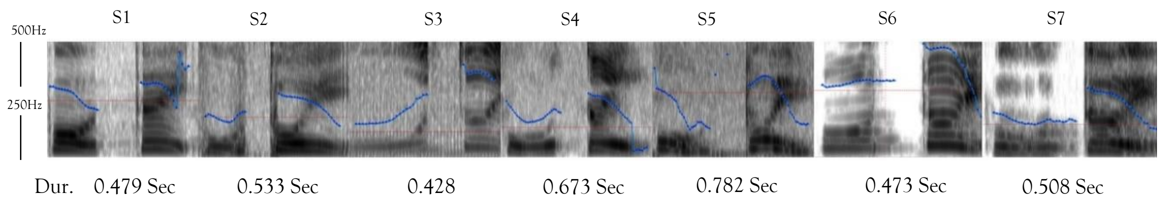


Figure 5 - Realization of T4 *Sandhi*

Figure 5 reveals that tone sandhi with Bù has been realized with much diversity compared to T3 sandhi. Speakers 1 and 5 have realized the 4+4 sandhi where the first T4 has not changed. Speaker 4 has realized the first T4 as a rising tone similar to a T3. All speakers have produced the second T4 as a T4 itself, and thus it could be inferred that the mismatching of the tone sandhi has only occurred in the first T4, unlike in T3 sandhi where the mismatch could occur at both T3s. In normal speech, unlike in the case of speaker 1, who is faster than the rest of the sample, the length of 4+4 realizations is longer than 2+4, created by the disruptive pause generated by two consecutive 4 tones. The average pitch at onset is 280.1 Hz, and at an offset, the average is 242.37 Hz.

T1 Sandhi with Yī

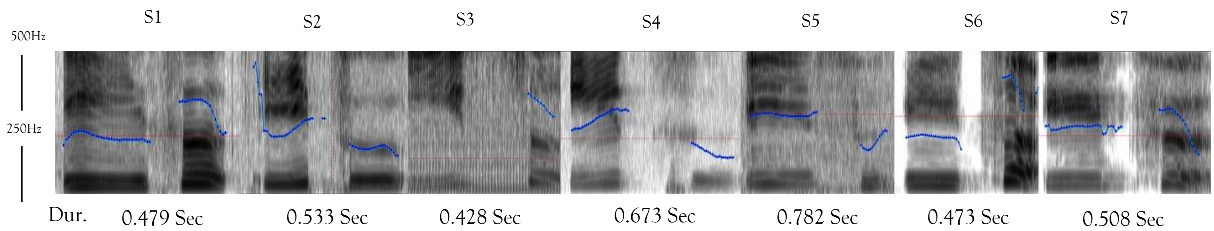


Figure 6 Realization of T1 *Sandhi*

The realization of tone sandhi with 一 Yī is much as diverse as the T4 tone sandhi with 不 Bù. Speakers 1, 2, 5, 6, and 7 have attempted to realize a pure T1 which has resulted in the mispronunciation of 次 (T4), while speakers 3 and 4 have produced the first T4 as a T2, which is the accepted tone sandhi. The contrast between the length of the two syllables is very high. The first syllable has an average length of 0.308 sec. At the same time, the average length of the second syllable is 0.171 sec. The average pitch at onset is 299.81 Hz, and the average pitch at offset is 246.771 Hz.

Discussion

The above acoustic analysis provides several significant implications for Sri Lankan learners' production of Mandarin Chinese tones. The periodic analysis of f0 in isolated production shows that the onset of different speakers is at different levels. It is one of the key reasons behind erroneous production, especially in the case of T2 and T3. For example, speaker 5 has produced T2 with a relatively high onset (327.5 Hz), leading to less diversity with an offset similar to the realization of a T1. Regarding accuracy, T2 shows the highest accuracy in both isolation and synthesis, while T3 shows the lowest accuracy. T1 demonstrates a relatively higher accuracy compared to T4.

Relative analysis of tone production in isolation, disyllabic form, and synthesis shows that the relationship between tone production in these three circumstances is abstract on most occasions. In other words, while some tones have been realized with high accuracy in isolation, some speakers have performed better in synthesis than in isolation. In the case of T4, the accuracy is significantly higher in synthesis than in isolation, and acoustic data demonstrate a significant difference between pitch height in isolation and synthesis. T2 demonstrates a relatively higher accuracy in both isolation and synthesis. Realization of T3 depicts the lowest accuracy in synthesis; except for one speaker, all other speakers have produced T3 inaccurately. Of the four tones in Mandarin, T3 has the greatest number of variants, and owing to its low register, T3 plays a significant role in the fluctuating pitch level in sentences Zhang (2018, p. 9).

The accuracy of Mandarin Chinese tones in synthesized speech is highly dependent on the awareness of tone sandhi rules of non-native speakers. In the tone sandhi f0 analysis of the present study, it is evident that the informants are less aware of the tone sandhi rules in Mandarin Chinese, which is especially depicted in the realization of T3 sandhi. The disparity between the surface and the underlying tone representations in T3 tone sandhi results in the realization of T3+T3 as a T2+T3 (Gao et al., 2021, p. 2). Despite this phenomenon, 42.8% of the speakers (n=3) have realized the underlying tone representation against the surface-level realization of the first T3 as a T2.

Realizing the T4 sandhi with Bù demonstrates the lowest accuracy of the three sandhi rules evaluated in the present study. Despite the underlying tonal combination of 不对 bùduì being T4+T4, at surface level, it is realized as a T2+T4. In Standard Mandarin, bù will change from a falling tone [51] into a rising tone [35] whenever another falling tone follows it (Chen, 2018). ss57% of the informants of the present study have resorted to the underlying representation of bù with a falling tone.

1st tone sandhi rule is generally referred to as the "yi-bu-qi-ba" sandhi rule owing to its application to four frequently used words, such as 一 yī (one), 不 bù (not), 七 qī (seven), and 八 bā (eight) (Yang, 2015, p. 14). Realization of the 1st tone sandhi with 一 yī also shows a relatively low accuracy. Like the two other sandhi realizations, most informants have attempted to produce T1 in its underlying representation. Only speakers 3 and 4 have resorted to the intuitive production of T1 as a T2 in sandhi with T4. Although all the informants are undergraduates who have been exposed to specific instructions and training on tonal variation and tone sandhi rules in the initial Mandarin lessons, they are not applied in the realization of tone sandhi in synthesized speech.

Production of Mandarin Chinese tones has been perceived as challenging for speakers of non-tonal languages. Wu and Lin (2008) claimed that the native language background is a key influential factor in the realization of Mandarin tones (p. 185). The acoustic analysis depicts that T2 has shown a relatively higher level of accuracy both in isolation and synthesis compared to the accuracy of the other three tones. According to a study conducted on the acquisition of intonation by Vietnamese L2 English speakers, the transfer from L1 is evident in the realization of intonation at the offset of Wh questions (Nguyễn & Đào, 2018, p. 10). The intonation patterns available in the participants' L1 (Sinhala) and L2 (English) have significantly influenced the higher accuracy in their production of Mandarin T2. The higher accuracy of the production of T2 could be associated with the intonation patterns of Sinhala, in which the inquisitive intonation ends with a higher f0, similar to the T2 in Mandarin Chinese. Other substantial differences, such as vowel length and the place and manner of articulation of sounds in the two languages, could also be associated with the erroneous production. The collection of vowels in Sinhala consists of long and short vowels, whereas Mandarin Chinese does not have such differences in vowel length.

The present study's findings are also consistent with Tang and Li (2020), who claimed that Mandarin Chinese pitch contours at fast speech are much flatter and tone spaces are narrower than normal speech (p. 1940). Out of all the seven informants of the present study, speaker 2 records the fastest speech. The pitch contours of speaker 2 are much flatter than the other speakers, especially in T1 (Std. dev = 6.732) and T4 (Std. dev = 5.672). However, the average speed of speakers remains at a low level. It could be hypothesized that lack of analytic-linguistic instruction and phonological knowledge are key influencers behind the erroneous tones production by Sri Lankan Mandarin learners.

Conclusion and Implication

The present study examined the production of Mandarin Chinese tones by Sri Lankan speakers in acoustic analysis. The study's findings reveal the diversity and inconsistency in the production of Mandarin Chinese tones by Sri Lankan learners. The mean f0 values of the realization of the four tones by the informants are heterogenous in isolation, disyllabic form, and synthesized speech. T3 has been realized with much lower accuracy than other tones, and T2 demonstrates the highest accuracy. As evidenced by the tone sandhi analysis, speakers have attempted to resort to the underlying representation of the tone sandhi than the surface realization. It could be inferred from the acoustic analysis and the existing literature that the lack of analytic linguistic instruction on the accurate production of Mandarin Chinese tones is a strong reason behind the erroneous production of tones.

The present study's findings would be instrumental for designing, scrutinizing, and rediscovering teaching learning material, methods, and curricula for teaching the Chinese language in Sri Lanka. Considering the dearth of acoustic analysis conducted on the production of Mandarin Chinese tones in Sri Lanka, the methodological and technical implications of the study would offer future studies a foundation for conducting further studies in this area. However, as with many other studies, the present study has limitations. The study has only analyzed the production of Mandarin Chinese tones, and further studies could be conducted on the perception of tones by Sri Lankan Chinese language learners. In addition, since this study has only analyzed the production of Mandarin Chinese tones by Sri Lankan learners from an acoustic perspective, there is much room for further studies on this from a pedagogical perspective.

References

Beasley, W., & Wang, L.-C. C. (2021). Using an iPhone or iPad for Tonal Activities in Teaching Mandarin Chinese. *International Journal of Education*, *13*(2). https://doi.org/10.5296/ije.v13i2.18675

Best, C. T. (2019). The diversity of tone languages and the roles of pitch variation in non-tone languages: Considerations for tone perception research. In *Frontiers in Psychology* (Vol. 10, Issue FEB). https://doi.org/10.3389/fpsyg.2019.00364

Biduri, F. N. (2017). Teaching Chinese Language Pronunciation for Early Childhood Education in Jakarta. *OALib*, *04*(12). https://doi.org/10.4236/oalib.1104215

Chen H. (2018). *A Comparative Study of Tone 4 Sandhi in Standard Mandarin and Luoyang Dialect*.

Chen, F., & Peng, G. (2016). Context Effect in the Categorical Perception of Mandarin Tones. *Journal of Signal Processing Systems*, *82*(2). https://doi.org/10.1007/s11265-015-1008-2

Chen, N. F., Shivakumar, V., Harikumar, M., Ma, B., & Li, H. (2013). Large-scale characterization of mandarin pronunciation errors made by native speakers of European languages. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. https://doi.org/10.21437/interspeech.2013-553

Dong, Y., Tsubota, Y., & Dantsuji, M. (2013). Difficulties in perception and pronunciation of mandarin chinese disyllabic word tone acquisition: A study of some Japanese university students. *27th Pacific Asia Conference on Language, Information, and Computation, PACLIC 27*.

Duanmu S. (2002). *The phonology of Standard Chinese*. Oxford University Press.

Duanmu, S. (2004). Tone and Non-Tone Languages: An Alternative to Language Typology and Parameters. *Language and Linguistics*, *5*(4).

Elliott, C. E. (1991). The relationship between the perception and production of Mandarin tones: An exploratory study. *University of Hawai'i Working Papers in English as a Second Language*, *10*(2).

Gao, F., Lyu, S., & Lin, C. J. C. (2021). Processing Mandarin Tone 3 Sandhi at the Morphosyntactic Interface: Reduplication and Lexical Compounds. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.713665

Guo, L., & Tao, L. (2008). Tone production in Mandarin Chinese by American students: A case study. *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, *1*.

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, *32*(3). https://doi.org/10.1016/S0095-4470(03)00016-0

Halliday, M. A. K. (2014). Notes on teaching Chinese to foreign learners. *Journal of World Languages*, *1*(1). https://doi.org/10.1080/21698252.2014.893675

Han, Y., Goudbeek, M., Mos, M., & Swerts, M. (2018). Audio-visual analyses of differences between natural and teaching styles for mandarin tone production. *Proceedings of the International Conference on Speech Prosody*, *2018-June*. https://doi.org/10.21437/SpeechProsody.2018-148

Jiang, X., & Cohen, A. D. (2012). A critical review of research on strategies in learning Chinese as both a second and foreign language. *Studies in Second Language Learning and Teaching*, *2*(1). https://doi.org/10.14746/ssllt.2012.2.1.2

Kan, M. S., & Ito, A. (2020). Language cognition and pronunciation training using applications. *Future Internet*, *12*(3). https://doi.org/10.3390/fi12030042

Li, J., Xu, R. B., & Mok, P. P. K. (2014). Effect of Orthography on L2 Production of Mandarin Tones. *4th International Symposium on Tonal Aspects of Languages*.

Li, Y. (2015). English and Thai Speakers' Perception of Mandarin Tones. *English Language Teaching*, *9*(1). https://doi.org/10.5539/elt.v9n1p122

Li, Y., & Lee, G. (2021). *The Effect of Perceptual Training on Teaching Mandarin Chinese Tones*. https://doi.org/10.1007/978-981-15-3809-4_5

Lin, Q. (2017). The Diachrony of Tone Sandhi: Evidence from Southern Min Chinese. In *PQDT - Global*.

Liu, J., & Xiao, C. (2021). *Tone Category Learning Should Serve Tone Word Learning: An Experiment of Integrating Pronunciation Teaching in the L2 Chinese Curriculum*. https://doi.org/10.1007/978-981-15-3809-4_6

Liu, M., Chen, Y., & Schiller, N. O. (2020). Tonal mapping of Xi'an Mandarin and Standard Chinese. *The Journal of the Acoustical Society of America*, *147*(4). https://doi.org/10.1121/10.0000993

Matthews, S., & Yip, V. (2013). Cantonese: A Comprehensive Grammar. In *Cantonese: A Comprehensive Grammar*. https://doi.org/10.4324/9780203835012

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America*, *102*(3). https://doi.org/10.1121/1.420092

Nguyễn, A. T. T., & Đào, Đ. M. (2018). The acquisition of question intonation by Vietnamese learners of English. *Asian-Pacific Journal of Second and Foreign Language Education*, *3*(1). https://doi.org/10.1186/s40862-018-0044-4

Pellatt, V., Liu, E. T., & Chen, Y. Y. Y. (2014). Translating Chinese Culture: The process of Chinese-english translation. In *Translating Chinese Culture: The Process of Chinese-English Translation*. https://doi.org/10.4324/9781315857022

Rhee, N., Chen, A., & Kuang, J. (2021). Going beyond F0: The acquisition of Mandarin tones. *Journal of Child Language*, *48*(2). https://doi.org/10.1017/S0305000920000239

Rungruang, A., & Mu, Y. (2017). Mandarin Chinese Tonal Acquisition by Thai Speakers. *Asian Social Science*, *13*(5). https://doi.org/10.5539/ass.v13n5p107

Shih, C. (2013). Mandarin third-tone sandhi and prosodic structure. In *Studies in Chinese Phonology*. https://doi.org/10.1515/9783110822014.81

So, L. K. H. (1997). Tonal Changes in Hongkong Cantonese. In *One Country, Two Systems, Three Languages: A Survey of Changing Language Use in Hong Kong Book*. Multilingual Matters.

Sun, Y., Kyaw, W. T., Zhang, J., & Sagisaka, Y. (2018). Analysis of L2 learners' progress of distinguishing Mandarin tone 2 and tone 3. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2018-September*. https://doi.org/10.21437/Interspeech.2018-1983

Tang, P., & Li, S. (2020). The acoustic realization of Mandarin tones in fast speech. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2020-October*. https://doi.org/10.21437/Interspeech.2020-1274

Třísková, H. (2017). Acquiring and Teaching Chinese Pronunciation. In *Educational Linguistics* (Vol. 31). https://doi.org/10.1007/978-3-319-54027-6_1

Van de Weijer, J., & Sloos, M. (2014). The four tones of Mandarin Chinese. Linguistics in the Netherlands, 31, 180-191. https://doi.org/10.1075/avt.31.13wei

Wang, X. (2013). Perception of mandarin tones: The effect of L1 background and training. *Modern Language Journal*, *97*(1). https://doi.org/10.1111/j.1540-4781.2013.01386.x

Wee, L-H. (2019). *Phonological tone*. Cambridge University Press.

Wu, X., & Lin, H. (2008). Perception of Mandarin Tones by Mandarin and English Listeners. *Journal of Chinese Language and Computing*, *18*(4).s

Xing, J. Z. (2006). Teaching and learning Chinese as a foreign language: A pedagogical grammar. In *Teaching and Learning Chinese as a Foreign Language: A Pedagogical Grammar*.

Xu, C. X., & Xu, Y. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association*, *33*(2). https://doi.org/10.1017/S0025100303001270

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*(1). https://doi.org/10.1006/jpho.1996.0034

Yang, B. (2015). Perception and production of Mandarin tones by native speakers and L2 learners. Springer.

Yang, C., & Chan, M. (2010). The Perception of Mandarin Chinese Tones and Intonation by American Learners. Journal of the Chinese Language Teachers Association, 45(1).

Yang, J., Zhang, Y., Li, A., & Xu, L. (2017). On the duration of Mandarin tones. Interspeech 2017. https://doi.org/10.21437/interspeech.2017-29

Zhang, T. (2018). Foreigners' perception of Mandarin tones: A review. Modern Linguistics, 06(04), 599-606. https://doi.org/10.12677/ml.2018.64070

Zhang, Y., & Chen, Y. (2018). Kazakh learners' production of Mandarin tones in colloquial contexts. 6th International Symposium on Tonal Aspects of sLanguages (TAL 2018). https://doi.org/10.21437/tal.2018-48

Zheng, J., Huang, C., Chu, M., Soong, F. K., Ye, W.-ping. (2007). Generalized segment posterior probability for automatic Mandarin pronunciation evaluation. 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07. https://doi.org/10.1109/icassp.2007.367198

Zhiming, B. (1999). *The Structure of Tone*. Oxford University Press.