



Analysis of Monosyllabic Tones in Mandarin Chinese Produced by Polish Students

Agnieszka Jóźwik¹, Feng Shi²

¹Beijing Language and Culture University, China

²Nankai University, China

¹ jagna.joz@outlook.com ² shifeng@nankai.edu.cn

Abstract

This study provides a phonetic analysis of monosyllabic tones in Mandarin Chinese produced by native Polish speakers with a background in Mandarin as a foreign language. Tonal patterns of Polish students were compared and contrasted with corresponding native-produced tones. Standard Chinese is a tonal language with 4 tones described as high, rising, low (falling-rising or dipping) and falling. Previous studies suggest that native speakers of non-tonal languages find difficulty in tone perception and production. It was suggested that T2 and T3 are particularly problematic for language learners. The results of this study support to a certain extent previous findings concerning T2 and T3 production in Polish speakers, as well as raise points regarding T1 and T4.

Index Terms: Mandarin Chinese, tone pattern, Polish students, language learning

1. Introduction

Standard Chinese (Standard Modern Mandarin) is a tonal language with 4 tones, described as high (T1), rising (T2), low or falling-rising (T3) and falling (T4). Chinese tones are commonly portrayed on a 5-level chart and represented by their initial, final and in case of the third tone, mid-syllable value on the chart. A fifth tone, also called neutral tone or zero tone, can be observed in weak syllables and is considered to be an extension of the preceding tone, therefore it has no defined contour and is usually not mapped on the chart. Following the aforementioned 5-level notation, first proposed by Yuan Ren Chao in 1968, Standard Mandarin T1 is described as a 55 tone, T2 is described as 35, T3 is described as 214 and T4 is described as 51. [1]

Previous research suggests that learners of Mandarin as a foreign language experience difficulties mainly in perceiving and producing T2 and T3 [2], [3], [4]. Yanping Wu's 2015 case study of Polish learners' tone pronunciation errors presents examples of some T2 and T3 errors [5]. According to the study, main difficulty lies in the low starting point of T2. Learners, who may be uncomfortable with the low pitch, begin T2 on a higher note, gradually lowering the pitch until reaching the lowest point at 1/3rd of the syllable duration, then gradually rising to the highest point located by the end of the syllable [5]. This results in producing a dipping contour, which is considered characteristic to T3; and can result in the two tones becoming indistinguishable. However, it is worth noting that Wu's study included only 2 participants with a brief, 3-month history of studying Mandarin and no previous exposure to Mandarin speaking environment.

This study aims to provide an analysis of monosyllabic tones produced by Polish native speakers with upper-

intermediate to advanced levels of Mandarin; living and studying in Beijing. This study of Polish learners' patterns in Mandarin tones can be used for educational purposes thanks to providing better understanding of common mistakes in tone production. It can also potentially become a preliminary study and a starting point for future research in prosodic patterns in Chinese as L2 and prosody-based Native Language Identification.

2. Materials and Methods

2.1. Participants

Six Polish native speakers (3 female and 3 male) were recruited for this study. Mean age of participants was 25.5 years. At the time of recording all participants were students of Beijing Language and Culture University in Beijing, China and had passed Chinese Proficiency Test (*Hanyu Shuiping Kaoshi* or HSK) level 5, which is compared to CEFR levels of B2/C1. Participants have lived in Beijing for 2 to 4 years prior to the recording. None of the participants had reported history of hearing or speaking problems. Participants were not paid for their attendance in the study.

2.2. Materials

Material for this study consists of monosyllabic words representing four Mandarin tones (high tone, rising tone, falling-rising tone and falling tone). Neutral tone or zero tone was not taken into account in this research, since it does not occur in monosyllabic words. In line with previous research on the topic, only syllables with a voiceless initial were chosen. Syllables with initial nasal consonant, lateral consonant, "zero initial" (absence of initial consonant), as well as characters with multiple readings were avoided in order to maintain consistency within pitch patterns [6]. To ensure that participants would be able to recognize the sinograms and provide a correct reading, suitable monosyllabic words were chosen from word lists recommended for HSK levels 1 to 5. For each of four Mandarin tones ten monosyllabic words were chosen, 40 syllables in total. 8 sample monosyllabic words used in this study are presented in Table 1.

2.3. Procedure

Each participant was given 5 minutes to prepare and review the sinograms before the recording in order to limit pronunciation errors stemming from misreading or not recognizing the sinogram. Each participant was provided with the same material and read the material 3 times, providing the total of 120 isolated syllable realizations.

Recordings were edited and processed in Praat, tone boundaries were manually annotated based on pitch contour.

A Praat script was used to automatically extract pitch value in Hz for 10 points on each tone curve. Each speaker's average pitch value for each of 10 points was subsequently transformed to a value in range from 0 to 5 according to the following T-value formula proposed by Feng Shi [7]:

$$T = \frac{\lg x - \lg \min}{\lg \max - \lg \min} \times 5$$

(1)

where x is speaker's average pitch value for a given point, \min is speaker's minimal average pitch and \max is speaker's maximal average pitch.

Values obtained using this formula were used to form a 5-level tone chart displaying the relative pitch height and tone contour. This allows visual assessment and comparison of tones between speakers regardless of their absolute pitch, age, gender and other personal factors.

In order to describe the tones by their values as discussed in 1.1, T-values are rounded up to the nearest threshold bigger than the given value, thus a tone with starting point value equal to 1.5, mid-syllable value of 0 and final value of 2.5 would be described as 213.

Table 1: Sample target monosyllabic words. The columns contain: target sinogram, pinyin and tone information.

他	tā	T1
多	dūo	T1
牌	pái	T2
毒	dú	T2
普	pǔ	T3
敢	gǎn	T3
库	kù	T4
去	qù	T4

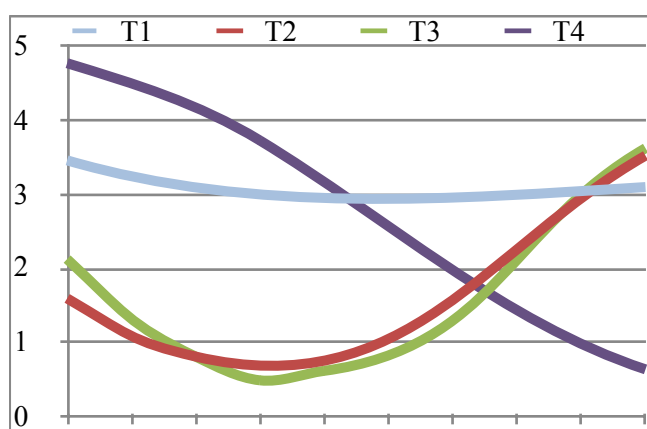


Figure 1: Tone contours of T1, T2, T3 and T4 produced by Polish speakers

3.

Results

3.1. Single tone analysis

Average tone contour of all tones for all speakers is presented in Figure 1. As seen on the 5-level chart, T1 average value is 44, T2 average value is 214, T3 average value is 314 and T4 average value is 51.

Average T-value for starting point of T1 equals 3.46 with standard deviation of 0.73. Initially, tone contour slowly declines and reaches the lowest point (2.95) around 60% into the syllable. Average ending point of T1 is 3.1. T1 produced by Polish learners is characterized by a mostly flat contour with an optional falling element at the beginning, however there are visible discrepancies in tone height. For instance, while T1 pattern produced by Male 1 can be described as a high-level 55 tone, Female 3 produced a much lower 43 tone. Average pattern of T1 is 44.

Average T-value for starting point of T2 equals 1.59, from that point tone contour dips and reaches the lowest point of 0.68 around 40% into the syllable. T2 reaches the highest point, 3.53 at the very end of the syllable. Tone contour for T2 is not as uniform as in T1, but for all speakers T2 starts with a short falling element, low flat element and final rising element that constitutes over 50% of syllable length. There are visible discrepancies in tone values, for instance 2 male speakers produced a low rising 13 tone while one female speaker produced a high rising 35 tone. Average pattern of T2 is 214.

T3 average initial value equals 2.13. T3 reaches its lowest point (0.48) around 40% into syllable duration. From that point T3 contour gradually rises and reaches its peak value of 3.62 at the end of the tone. For all speakers T3 retains the characteristic dipping contour with a relatively high rise towards the end. The lowest tone value observed was 1.13, the highest one was 3.15. Main disparity in height was observed at the initial point, where values ranged from 0.57 to 3.47. Average tone pattern of T3 is 314.

T4 initial average T-value is 4.77, final value is 0.63. The falling contour characteristic to tone 4 is visibly produced by all participants. Aside from abnormally low starting point by Female 1, all speakers initial T-value was equal to 5. Observed T4 patterns include 51, 52 and 41, average pattern is 51.

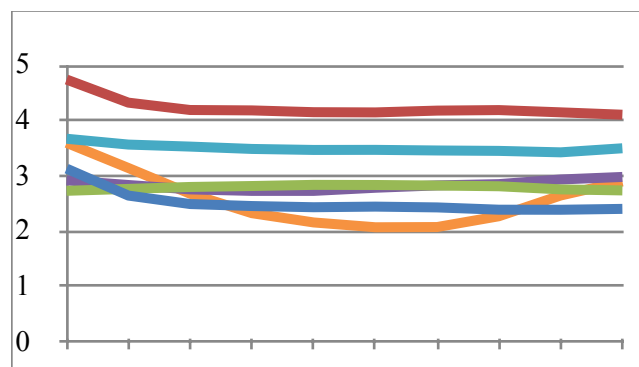


Figure 2: T1 patterns of Polish speakers

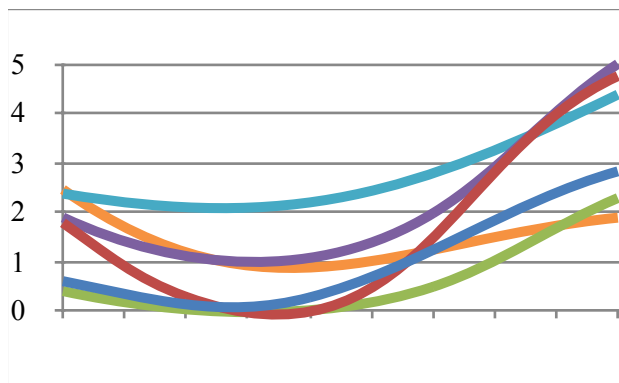


Figure 3: T2 patterns of Polish speakers

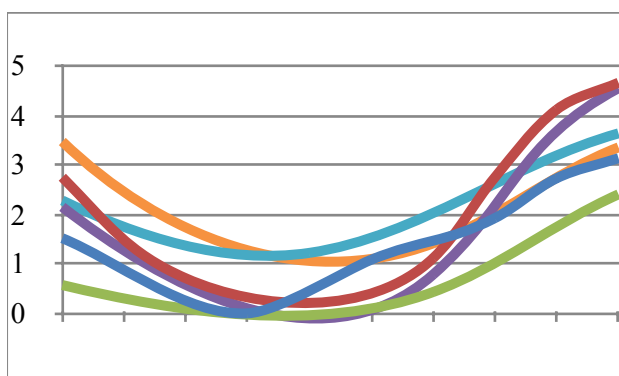


Figure 4: T3 patterns of Polish speakers

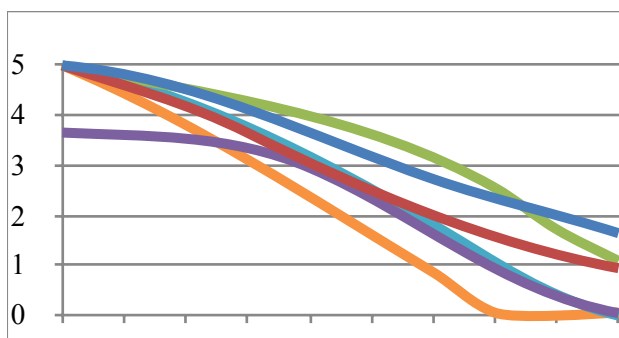


Figure 5: T4 patterns of Polish speakers

Tone patterns of all participants are presented in Figure 2 to Figure 5.

3.2. Polish-Chinese comparison

Polish tone values obtained in this study were compared and contrasted with Chinese native speakers monosyllabic tones. Data used for comparison was collected in 2017 by Zixia Fan. For her study she recruited 4 Chinese native speakers studying in Beijing, aged 22-26 years old; participants were asked to read 40 monosyllabic words, 10 for each tone. [8]

T1 produced by Chinese native speakers has a mainly flat contour, initial T-value of 4.51, mid-syllable value of 4.32 and final value of 4.24, thus the pattern is described as 55. Polish realization of T1 distinctly lower and initiates with an angled declining element that is not present in Chinese pattern.

T2 produced by Chinese natives is a consistently rising tone with initial T-value of 2.07, mid-syllable value of 2.67 and final value of 4.21, tone pattern is described as 35. Polish realization of T2 is generally lower and involves a dipping shape, which is not present in native speech.

Mandarin T3 is a dipping tone, it is described as a 214 tone, however previous research consider other patterns like 213 or 113 also common and correct [6, pp. 102]. In Fan's study T3 initial T-value is 2.29, mid-syllable value and the lowest point 0.08 and final value 2.18, resulting in a 313 pattern. Initial value and around 30% of the tone produced by Polish students is consistent with native Chinese results from Fan's study, but the initial value is higher than allowed by Shi Feng's findings. The mid-syllable and final tone values are consistent with native Chinese T3, with a very pronounced dipping contour.

T4 produced by Chinese native speakers is a consistently falling tone with the initial T-value close to 5 and final value being 0.49. T-values of Polish students' T4 are in agreement with these results, while the contour itself is less curved than it is in case of native speakers.

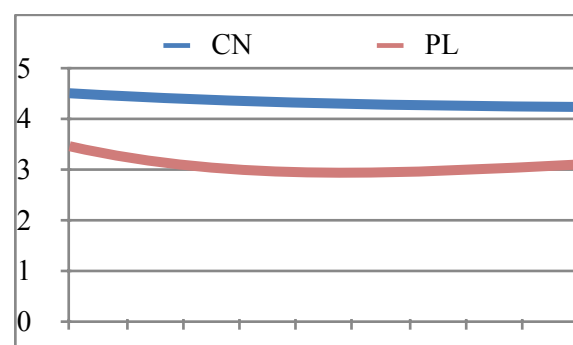


Figure 6: T1 patterns of CN natives and PL students

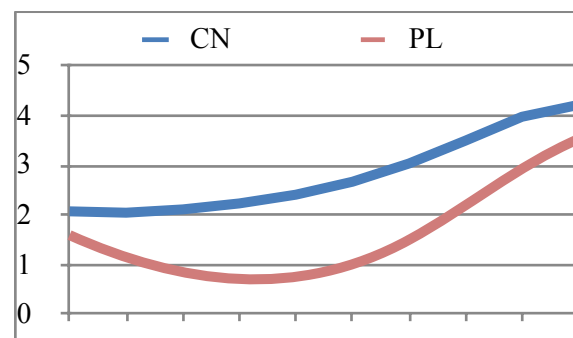


Figure 7: T2 patterns of CN natives and PL students

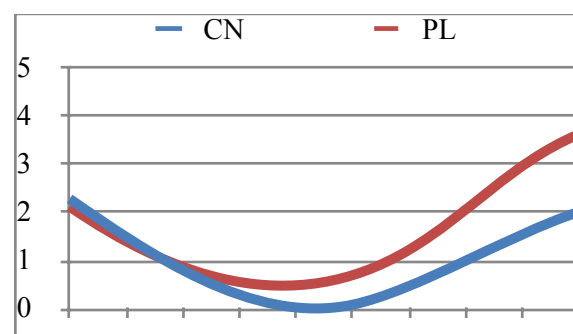


Figure 8: T3 patterns of CN natives and PL students

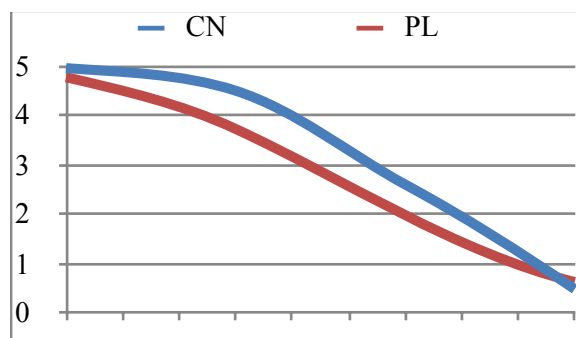


Figure 9: T4 patterns of CN natives and PL students

Comparisons of tonal patterns of Mandarin native speakers and Polish students are presented in Figure 6 to Figure 9.

4. Discussion and conclusions

The present study aimed to examine tone patterns in Standard Mandarin produced by Polish students. Tonal patterns emerging from this study are: 44 for Tone 1, 214 for Tone 2, 314 for Tone 3 and 51 for Tone 4.

In comparison with Chinese native speakers observed T1 was lower and included a falling element at the beginning. This result is consistent with Wu's previous case study. T2 measured in the present study was evidently lower and more dipped, with its contour resembling T3. Contrary to Wu's research, Polish T3 initial value turned out to be higher than native Chinese T3 initial value. Mid-syllable and end-syllable T3 pattern turned out to be consistent with native Chinese T3 pattern. This suggests that in fact T3 is not problematic in isolated monosyllabic reading. Both initial and final values of T4 are in line with native Chinese values.

This study supports previous claims that one of the main difficulties for students of Chinese as a foreign language lies in distinguishing T2 and T3. The two tones were pronounced with a nearly identical contour, having indistinguishable mid-syllable and final values (Figure 10). However, results suggest that for Polish students this issue manifests mostly in T2, as opposed to for instance Thai students, who tend to mispronounce T3. More data and further analysis is required to determine the nature of this T2-T3 merger. Based on those results it could be hypothesized that for Polish learners isolated tone acquisition is easier for T4 and T3 than for T1 and T2.

It is worth noting that the material consisted of isolated syllables and was read by the participants with a chance to prepare and review the sinograms beforehand. This decision might have resulted in better reliability, limited mispronunciations stemming from unrecognized characters and diminished stress or pressure perceived by participants. On the other hand, this caused the material to be unnatural and students' behavior to be artificial and exceptionally careful. Therefore, results obtained in this study cannot be extrapolated to describe natural spontaneous speech.

It should be emphasized that this is a small-scale preliminary study, aiming to investigate the basic features of tonal patterns produced by Polish speakers. In order to assess the significance of these findings, a larger study as well as static and perceptual analyses will be conducted in the future.

Mandarin tones, especially T3 and T4 are subject to sandhi rules when connected with other tones. It has been shown that tone production when sandhi rules are applied differs from

native speakers [9]. Therefore in order to successfully research tone patterns of Polish students, bisyllabic tone combinations, tones in sentences and tones in spontaneous speech must be examined.

5. References

1. Y. R. Chao, *A grammar of spoken Chinese*, Berkeley and Los Angeles: University of California Press, 1968.
2. C. Y. Lee, L. Tao, and Z. S. Bond, "Identification of acoustically modified Mandarin tones by native listeners," *Journal of Phonetics*, vol. 36, no. 4, pp. 537-563, Oct. 2008.
3. Y. Wang, A. Jongman, and J. Sereno, "Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training," *The Journal of the Acoustical Society of America*, vol. 113, no. 2, pp. 1033-1043, Jan. 2003.
4. J. H. Jin, "The crucial reason and solution of problems in Chinese production by Japanese learners", *Journal of Shandong Teachers' University*, 5: 100-102, 1996.
5. Y. P. Wu, "BoLan LiuXueSheng HanYu DanZiDiao FaYin PianWu YanJiu [Study on pronunciation errors in monosyllabic tone in Chinese by Polish learners]", *XianDaiYuWen [Modern Chinese]*, vol. 12 no. 1, pp. 107-110, 2015.
6. F. Shi, *YuYinGeJu [Patterns in Phonetics]*, Beijing: Shangwu, 2008.
7. F. Shi, "TianJin FangYan ShuangZiZu ShengDiao FenXi [Analysis of bi-syllabic tones in Tianjin dialect]", *YuYanYanJiu [Studies in Language and Linguistics]*, no. 1, pp. 77-90, 1986.
8. Z. Fan, "The Experimental Analysis of Monosyllabic Tone in Mandarin of Thai Students", Unpublished Manuscript, Beijing Language and Culture University, 2017.
9. S. Chen, Y. He, C.W. Yuen, B. Li and Y. Yang, "Mechanisms of Tone Sandhi Rule Application by Non-native Speakers" in *INTERSPEECH 2017 – 18th Annual Conference of the International Speech Communication Association*, September 8–12, Stockholm, Sweden, Proceedings, 2017, pp. 1760–1764.