INTERMINGLING TONE SYSTEMS: THE RELATIONSHIP OF NANNING MANDARIN TO NANNING CANTONESE AND STANDARD MANDARIN

Xinzi Hou[1], Liang Zhao[1], Eleanor Chodroff[2]

[1)]Department of Language and Linguistic Science, University of York, York, United Kingdom
[2)]Department of Computational Linguistics, University of Zürich, Zürich, Switzerland
xinzi.hou@york.ac.uk, liang.zhao@york.ac.uk, eleanor.chodroff@uzh.ch

ABSTRACT

Nanning Mandarin, the lingua franca of a multilingual city in China, presents a special case of lexical tone development arising from language contact between Nanning Cantonese and Standard Mandarin. Previous research has suggested that the realisation of Nanning Mandarin tones derives from the transfer of Nanning Cantonese to Standard Mandarin. However, this research was limited to auditory methods, differed in reports of the tone system, and lacked a direct acoustic comparison of all three language varieties. In the present study, we conducted an acoustic-phonetic analysis of Nanning Mandarin tones from 30 talkers, and directly compared the f0 contours of Nanning Mandarin with those in Nanning Cantonese (48 talkers) and Standard Mandarin (21 talkers). Though the four-tone Nanning Mandarin tone inventory shows a resemblance to Standard Mandarin, the phonetic realisation of Nanning Mandarin tones exhibits considerable similarity to that of the primary source language, Nanning Cantonese in f0 height and contours.

Keywords: lexical tone, dialects, intermingling communities, Mandarin, Cantonese

1. INTRODUCTION

Language contact creates a unique environment for tone systems to develop. As a multilingual city, Nanning, China presents one such environment for investigating the relationships between lexical tone systems. Nanning is the capital of the Guangxi Zhuang Autonomous Region in southern China. Traditionally, the dominant language of the city has been Nanning Cantonese, a Sinitic language, but the city has also had a relatively high number of Pinghua (Sinitic) and Zhuang languages (Tai) speakers. Since the 1950s, Standard Mandarin (or Putonghua, also a Sinitic language) has been officially promoted as the lingua franca throughout China and is the required language in education, news media and government affairs. Moreover, Nanning has received an influx of immigrants with different language backgrounds from within and beyond Guangxi, providing further reason to use the lingua franca. In the 1990s, the strong nationwide promotion resulted in the widespread prevalence of a distinct Mandarin dialect, namely Nanning Mandarin, which has become the current lingua franca in the region.

Nanning Mandarin (NM) is the local version of Standard Mandarin (SM) produced by native speakers of the traditional regional languages (mainly Nanning Cantonese, NC) speakers. It is usually categorised as a type of Local Mandarin (or Local Putonghua) [4] or an interlanguage produced by L1 Nanning locals in approximating Standard Mandarin as an L2 language [1]. While Nanning Cantonese was the primary L1 for older generations, Nanning Mandarin is now frequently acquired as the native language. Though Nanning Mandarin retained the four tone categories from its target language (Standard Mandarin), the phonetic realisation of those tones has been noted to be unique.

Few studies have directly examined the lexical tone systems of Nanning Mandarin, and they mostly rely on auditory transcription. One study reported the tone inventory of the interlanguage between Nanning Cantonese and Standard Mandarin, which was presumably Nanning Mandarin; however, the methodology was not made explicit [12]. In addition, Fan [21] conducted an auditory analysis of the tone inventory in Nanning Mandarin from one older and one younger speaker, and found that the differences between generations were minor. Both studies identified a level tone for NM Tone 1, a rising tone for NM Tone 2, and a low falling tone for NM Tone 3; however, they differed in the description of NM Tone 4, as well as the exact Chao tone numbers assigned to many of these tones. Two additional studies have indirectly examined Nanning Mandarin through an auditory comparison of Nanning Cantonese and Standard Mandarin [9, 12].

Overall, previous studies have found that Nanning Mandarin has a narrower pitch range than Standard Mandarin and the absence of a dipping tone [9, 10, 12, 18]. A similar observation was reported in an intonational study of Nanning Mandarin [10]. These studies conjectured that the identified patterns can be explained as a partial transfer of the phonetic tone realisation from Nanning Cantonese to Mandarin. All four studies indicated intermediate and unstable

phonetic realisations between Nanning Cantonese and Standard Mandarin [23]. Nevertheless, discrepancies still remain regarding the precise characterisations of the NM tone system.

In the present study, we extend previous research on the NM lexical tone system by directly examining the relationship of the NM tone system to the NC and SM tone systems. In contrast to previous studies, we assess these relationships using acoustic-phonetic analysis from a large number of NM, NC, and SM speakers. The presented research aims to demonstrate the acoustic representations of tone in Nanning Mandarin and conduct a comparison of these with the historically dominant source language, Nanning Cantonese, and the target language, Standard Mandarin. This study sheds new light on the research on tone in language contact environments by expanding the documentation of tonal languages with systematic descriptions of two understudied Sinitic varieties, Nanning Mandarin and Nanning Cantonese. The findings further our understanding of tone in complex language contact, as well as source-language effects on interlanguages.

2. METHODS

2.1. Materials

For Nanning Mandarin and Nanning Cantonese, a wordlist containing 72 monosyllabic words and 72 disyllabic words was designed based on the diachronic correspondence between Mandarin and Cantonese tone categories. Nanning Mandarin and Standard Mandarin have four citation tones, while Nanning Cantonese has six citation tones that have historical correspondences with Mandarin tones [11, 19]. The present study focuses on the monosyllabic words alone, so only these will be described in detail. The 72 monosyllabic words cover all four tone categories in Nanning Mandarin with 12 to 24 words for each NM tone and all six tone categories in Nanning Cantonese with 12 words for each NC tone.

All words are shared between Nanning Mandarin and Nanning Cantonese and recognised in spoken language, though a few sound old-fashioned. The words in both languages have similar phonetic contexts with a V, CV or CVN structure. Examples in Nanning Mandarin and Nanning Cantonese include 姑 '*aunt*' /ku^{55}/–/ku^{55}/; 有 '*have*' /jɐu^{24}/–/jou^{214}/. For clearer segmentation and to minimise reading errors, syllables with multiple readings were disambiguated with a disyllabic context and instructions for the participant to only read the target word.

For Standard Mandarin, ten unique syllables were used to create a wordlist of 40 monosyllabic words, with each syllable realised in one of the four SM lexical tone categories (see [24] for more detail).

2.2. Participants

For Nanning Mandarin and Nanning Cantonese, 52 native speakers were recruited; 48 were bilingual speakers and four were Nanning Mandarin monolingual speakers. With a few exceptions, speakers were born and raised in Nanning; started speaking Nanning Mandarin or Nanning Cantonese before the age of 12; had not stayed outside of Nanning for more than one year within the past ten years; did not have any explicit training or test in Standard Mandarin; and did not have close family members or friends who spoke Northern Mandarin dialects. Participants who did not meet the above criteria went through additional validation of their NM accent for further exclusion (see Section 2.4).

Participants were roughly balanced for gender (27 females and 25 males) and age (12 each for age 24–39, 40–49 and 50–59, 16 for age 60–69). Speakers aged 30 to 60 formed the majority because they were less likely to have settled elsewhere in recent years and have had less exposure to Standard Mandarin at school. Therefore, they were less affected by other languages or dialects than the younger speakers.

For Standard Mandarin, 21 fluent speakers of Standard Mandarin (13 females and eight males aged 23–41) were recruited from six Mandarin-dominant cities of China [24]. Though these cities have their own regional dialects, the participants reported an average score of 4.14 on a 1–5 scale (5 as native-like fluency) on their fluency in Standard Mandarin. Each audio used for analysis was also manually checked for its representativeness of Standard Mandarin.

2.3. Procedure

For Nanning Mandarin and Nanning Cantonese, recordings were conducted in person in a quiet room or booth in a location convenient for the participant in Nanning. The instructions and wordlist were presented in Simplified Chinese on a Macbook Pro via PsychoPy v.2021.2.3 [20]. The production was recorded using Awesome Voice Recorder X v2.0.4 on an iPhone 7 smartphone. The audio was recorded in WAV format with a sampling rate of 22,050 Hz and a bit rate of 320 kbps. Before the study, participants completed a demographic questionnaire, and received the instructions and overview of the wordlist. This was then followed by five practice words.

The recording involved two repetitions of four blocks: monosyllabic and disyllabic words in each of Nanning Mandarin and Nanning Cantonese. Language order was counterbalanced across bilingual participants, and word order was randomised within

each block. Target words were presented one by one on the screen; the participant controlled the progression by pressing any button on the keyboard. Participants were allowed to repeat a token if needed for correction. In total, 7486 monosyllabic NM and 6905 monosyllabic NC words were collected.

For Standard Mandarin, recordings were collected remotely via participant-controlled smartphones. Participants were instructed to install a specific recording app and record themselves while reading the words presented on the Gorilla Experiment Builder [19]. Words were presented trial by trial and fully randomised (see [24] for more detail).

2.4. Data processing

Recordings were automatically segmented and aligned using Praat [8] and the Montreal Forced Aligner (MFA) [18] at the syllable and phone levels. Given the formal register of reading, it was possible that some speakers produced tokens with a strong SM accent. Validation of the naturalness of the NM-accent was performed for 26 speakers based on their demographic background and an initial auditory evaluation was performed by the first author. Six native NM auditors listened to three to five random tokens from each of these speakers and rated their accents on a scale from 1 to 5. NM speakers rated as 4 or above were retained in the acoustic analysis and 20 NM speakers were thus excluded.

For all three language varieties, f0 values were automatically extracted from 11 equidistant points in the sonorant portion of each syllable. The first point was excluded as its value was highly affected by the preceding consonant and was frequently undefined. After averaging by speaker and tone, the f0 values were standardised by speaker-specific f0 ranges for each language, and then transformed into the tone values in Chao tone numbers ranging from 0 to 5, using the Log T-value formula [2]:

$$LT = \frac{\log_{10} mean - \log_{10} min}{\log_{10} max - \log_{10} min} \times 5 \quad (1)$$

Reference to Chao tone numbers in the paper thus reflects the LT value. Tokens with at least six usable f0 values and speakers with at least 72 (50%) usable tokens were retained for acoustic analysis (two speakers were excluded). After a combination of validation tasks, there were 6365 tokens from 48 NC speakers (24 female, 24 male) and 3732 tokens from 30 NM speakers (14 female, 16 male).

3. RESULTS

Figure 1 displays the f0 contours in Chao tone numbers for each tone and language variety; Table 1 shows the Chao tone numbers for each tone and language variety. Both varieties of Mandarin have four phonological tones, whereas Cantonese has six phonological tones, with a likely tone merger between Tone 2 and Tone 6.

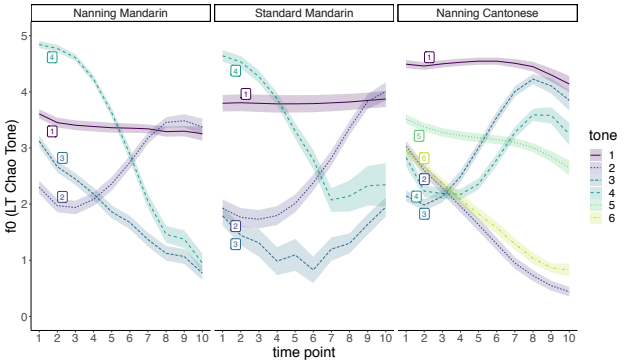

Figure 1: f0 contours in Chao tone numbers of Nanning Mandarin, Standard Mandarin, and Nanning Cantonese for each phonological tone. Each ribbon represents 1 standard error from the mean.

Tone	Nanning Mandarin	Standard Mandarin	Tone	Nanning Cantonese
1	44	44	**1**	55
2	34	24	**2**	31
3	41	212	**3**	35
			4	34
4	52	53	**5**	43
			6	31

Table 1: Chao tone numbers for each phonological tone in Nanning Mandarin, Standard Mandarin, and Nanning Cantonese. Tones are aligned according to their historical correspondence.

At a glance, the NM tone–contour mappings reflect those of Standard Mandarin, but with a noticeable lack of dipping in NM Tone 3 and a considerably higher onset in NM Tone 4. When we zoom out to the use of the f0 space, however, Nanning Mandarin appears to contrast with Standard Mandarin and shares greater resemblance to Nanning Cantonese. Unlike Standard Mandarin where a High–Low contrast is utilised, Nanning Mandarin displays a High–Mid–Low contrast, most saliently at the onset, resembling Nanning Cantonese. Specifically, NM Tone 4 begins in the high range, NM Tone 1 and Tone 3 begin in the middle range, and NM Tone 2 begins in the low range.

To investigate whether individual NM tones share greater similarity with Standard Mandarin or Nanning Cantonese, we examined the overall similarity of each NM f0 contour to similarly-shaped NC and SM f0 contours using the root-mean-squared

deviation (RMSD) between the averaged f0 contours from all ten timepoints in Chao tone numbers. Lower RMSD values correspond to a higher similarity between contours. As shown in Table 2, all four NM f0 contours were most similar to NC f0 contours, although some were closely followed by SM tones.

To further investigate differences in the exact contour independent of f0 height and any potentially remaining individual variation, the RMSD values were also calculated after *z*-scoring the raw f0 in hertz by speaker and tone. The results still indicated that all four NM tones shared greatest similarity to NC tones, though the falling SM Tone 4 became substantially more similar to NM Tone 3 and Tone 4.

NM Tone	Tone Comparisons			
NM1 (Level)	**NC1**	**NC5**	**SM1**	
	1.12	0.26	0.50	
	0.56	0.27	0.70	
NM2 (Rising)	**NC3**	**NC4**	**SM2**	
	0.59	0.20	0.38	
	0.23	0.19	0.39	
NM3 (Mid-fall)	**NC2**	**NC6**	**SM3**	**SM4**
	0.30	0.26	1.11	1.40
	0.12	0.11	1.05	0.32
NM4 (High-fall)	**NC2**	**NC6**	**SM3**	**SM4**
	1.77	1.72	2.43	4.00
	0.25	0.23	1.13	0.38

Table 2: RMSD values on the Chao tone numbers listed first in each cell, and on the *z*-scored values, second. Each row presents a comparison of an NM tone to tones with a similar contour in Nanning Cantonese and Standard Mandarin.

4. DISCUSSION

Overall, the lexical tone system of Nanning Mandarin resembled Nanning Cantonese in implementing a three-level tone system with High–Mid–Low contrasts in the phonetic space, in line with transcriptions in previous studies [9, 10, 12, 18, 20]. Moreover, individual contours were most similar to those of Nanning Cantonese in both Chao tone numbers and *z*-scored hertz, but closely followed by some SM tones. This suggests a clear influence of Nanning Cantonese on Nanning Mandarin.

Nevertheless, individual NM tones still approximated those of Standard Mandarin in f0 contours. Though NM Tone 3 did not dip like SM Tone 3, it did exhibit a similar initial fall. In Standard Mandarin, the realisation of Tone 3 is highly contextual, and critically only dips in phrase-final position [8, 13]. In Nanning Mandarin, it may be that only the core portion of Tone 3 was realised without the boundary rise. The absence of the Tone 3 prepausal rise resulted in two falling tones (NM Tone 3 and Tone 4). The dispersed f0 onsets and different slopes of the two falling tones could reflect a need for tonal contrast. The dispersed f0 onsets also correspond to the mid and high levels of the three-level tone system that is common to Nanning Cantonese, but distinct from Standard Mandarin.

With respect to Nanning Cantonese, we also observed a merger of Tone 6 (traditionally low-level [3–6, 11, 13, 15]) with Tone 2 (mid-falling), consistent with findings in previous auditory [3, 5, 12] and acoustic studies [15, 17]. This tone merger resulted in a lack of a low-level tone in the inventory, potentially accounting for why NM Tone 3 was not realised even as the non-phrase-final SM Tone 3 (21 contour). The falling NC Tone 6 could have potentially served as the initial template for NM Tone 3, giving rise to a mid- instead of low-falling tone (41 contour).

The findings of both Nanning Mandarin and Nanning Cantonese show similarity to the data in previous literature, but with some minor differences in tone values and contours. Results of the three language varieties show that Nanning Mandarin, as a variety of Standard Mandarin, has presented a hybrid tone system with a clear and large influence from Nanning Cantonese in its approximation to Standard Mandarin. Whether these similarities and differences are significant and affect pitch perception, however, remains to be tested.

5. CONCLUSION

The present study investigated the lexical tone system of Nanning Mandarin in comparison to Nanning Cantonese and Standard Mandarin. The results showed that the NM tone–contour mappings were clearly inherited from Standard Mandarin, but displayed considerable similarity to and influence from Nanning Cantonese. Overall, this study provided acoustic data of two understudied Sinitic language varieties in Nanning. The findings set the basis for further investigation of the relation between tone and intonation in Nanning Mandarin, as well as variation in tone realisation arising from complex language contact, from Nanning to a broader range of world languages.

6. ACKNOWLEDGEMENTS

The authors would like to thank Chenzi Xu, Dom Watt, and the anonymous reviewer for helpful feedback and suggestions. We also thank the NM auditors for their help with data annotation. The final author was supported by SNSF grant 208460.

7. REFERENCES

[1] L. Selinker, 'Interlanguage', *International Review of Applied Linguistics in Language Teaching*, vol. 10, no. 1–4, pp. 209–232, Jan. 1972.

[2] F. Shi, '天津方言双字组声调分析 [Analysis of bi-syllabic tones in Tianjin dialect]', *语言研究 [Studies in Language and Linguistics]*, no. 01, pp. 77–90, 1986.

[3] J. Hong, '南宁白话的语音特点 [Phonetic Features of Nanning Cantonese]', *广西师范大学学报(哲学社会科学版) [Journal of Guangxi Normal University (Philosophy and Social Science Edition)]*, no. 01, pp. 47–54, 1989.

[4] Y. Yao, '应该开展对"地方普通话"的研究 [Research Should be Conducted on "Local Standard Mandarin"]', *语文建设 [Language Planning]*, no. 03, pp. 20–23, 1989.

[5] J. Xie, '南宁白话同音字汇 [Nanning Cantonese homophone glossary]', *方言 [Fangyan]*, no. 04, pp. 286–303, 1994.

[6] H. Yang, *南宁话音档 [Sound archive of Nanning languages]*. Shanghai: Shanghai Educational Publishing House, 1997.

[7] 广西壮族自治区地方志编纂委员会 [Guangxi Zhuang Autonomous Region Local Documentation Compilation Committee], Ed., *广西通志·汉语方言志 [Comprehensive documentation of Guangxi: Sinitic dialects]*. Nanning: Guangxi People's Press, 1998.

[8] P. Boersma, 'Praat, a system for doing phonetics by computer', *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2002.

[9] M. Yip, *Tone*. Cambridge University Press, 2002.

[10] H. Yang, '南宁普通话的语调、语气词及其声学表现 [Intonation and Modal Particles Used in Nanning Putonghua and Their Manifestations in Acoustics]', *语言研究 [Study in Language and Linguistics]*, vol. 23, no. 04, pp. 104–109, 2003.

[11] Y. Liang, '"南宁普通话"韵母研究 [A Study of the Finals in Nanning People's Oral Mandarin]', Master, Guangxi University, 2005.

[12] J. Yu, Ed., *广西推广普通话的理论与实践研究 [A Study of the theory and practice of promoting Putonghua in Guangxi]*. Shanghai: Shanghai Educational Publishing House, 2008.

[13] Y. Lin and F. Qin, *广西南宁白话研究 [A study of Cantonese in Nanning, Guangxi]*. Guilin: Guangxi Normal University Press, 2008.

[14] F. Shi and Q. Ran, '普通话上声的本质是低平调——对《汉语平调的声调感知研究》的再分析 [A Comment on "Mandarin Tone Perception: A Report on a Low Level Tone"]', *中国语文 [Studies of the Chinese Language]*, no. 6, pp. 550-555+576, 2011.

[15] T. Li, '浅谈南宁白话对南宁方言普通话的声调影响 [A brief discussion of the impact of Nanning Cantonese on the tones of Nanning dialect Mandarin]', *剑南文学(经典阅读) [Jiannan Literature (Classics Reading)]*, no. 2, pp. 97–98, 2012.

[16] T. Luo, '南宁白话单字调实验语音研究 [An Experimental Study on the Mono-Syllable of Nanning Cantonese]', Master, Guangxi University, 2015.

[17] K. Wang, '南宁白话(新派)单字调声学实验研究 [An acoustic experimental study on the monosyllabic tones in Nanning Cantonese from the younger generation]', *语文学刊 [Journal of Language and Literature Studies]*, no. 04, pp. 37–39, 2015.

[18] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, 'Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi', in *Proceedings of the 18th Conference of the International Speech Communication Association*, 2017, pp. 498–502.

[19] A. Anwyl-Irvine, J. Massonnié, A. Flitton, N. Kirkham, and J. Evershed, 'Gorillas in our Midst: Gorilla.sc, a new web-based Experiment Builder'. bioRxiv, p. 438242, 22-Oct-2018.

[20] J. Peirce *et al.*, 'PsychoPy2: Experiments in behavior made easy', *Behav Res*, vol. 51, no. 1, pp. 195–203, Feb. 2019.

[21] F. Fan, '南宁普通话语音研究 [Nanning Mandarin Phonetic Research]', Master, Guangxi University, 2020.

[22] X. Li and M. Xiang, *汉语方言学基础教程 [Basic tutorials in Chinese dialectology]*, 2nd ed. Beijing: Peking University Press, 2020.

[23] J. Lyu and J. Yu, '多语接触与通语变异——广西"南普"形成及变异研究之考察 [Multi-language Contact and Standard Language Variation: A Study on the Form and Variation of "Nanpu" in Guangxi]', *西南交通大学学报(社会科学版) [Journal of Southwest Jiaotong University (Social Science)]*, vol. 22, no. 01, pp. 47–55, 2021.

[24] L. Zhao and E. Chodroff, 'The ManDi Corpus: A Spoken Corpus of Mandarin Regional Dialects', in *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 2022, pp. 1985–1990.