Individual Differences in Tone Perception and Production in the Process of Dialect Emerging: A Case study of Elementary School Children in Changsha

Yi Zhang[1] Lishan Li[1] Xiaoying Xu[1,2*]

[1]The School of Chinese Language and Literature, Beijing Normal University, Beijing, China;

[2]Research Center for Collationand Standardization of Chiense Characters, Beijing, China.

xuxiaoying2000@bnu.edu.cn

Introduction

Over time, with the official promotion of Standard Mandarin, Changsha's language use has changed, characterized by the displacement of the Changsha Dialect and the emergence of Changsha Plastic Mandarin, a non-standard Mandarin accent that features the speech of young, urban residents and has crystallized over the past few decades. Changsha Plastic Mandarin is a product of Standard Mandarin and local dialect contact and is used in informal settings. Its segments are similar to Standard Mandarin but with different tones; the tonal features are more akin to those in Changsha Dialect than in Standard Mandarin. Numerous studies have examined the rise of Changsha Plastic Mandarin and described its acoustic features(Jin and Niu, 2010).

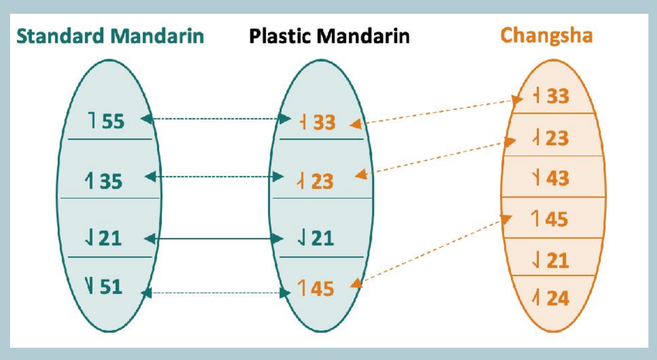

FIG.1. The relationship between the tonal values of Standard Mandarin, Plastic Mandarin, and Changsha Dialect (Xu, 2022).

This study explores the perceptual discrimination of T1 and T4 in Changsha Plastic Mandarin among primary school children and the natural production of single-word tones. The main issues to investigate are:

(1)Whether there are significant individual differences in identification accuracy of T1 and T4 and whether age is a factor;

(2)The correlation between identification accuracy and utilization of perceptual cues (F0 height, F0 movement, and voice quality);

(3)Whether the natural production of single-word tones by Changsha elementary school children is closer to Changsha Plastic Mandarin or Standard Mandarin and its correlation with recognition accuracy.

Participants

All experiments were conducted at Lushan International Experimental Primary School. Forty-four elementary students from grades 1 to 6 participated. All were from Changsha, raised speaking SM and what they considered as CPM, but not speaking the CD. The group included 22 girls and 22 boys, aged 6-12, with average ages of 8.47 and 8.81 years respectively.

Experiment 1: natural stimuli perception

Experiment 1 was conducted using natural stimuli(("搭" (da_T1), "大" (da_T4), "低" (di_T1), and "地" (di_T4)) to explore the subjects' discrimination accuracy of T1-T4 in Changsha Plastic Mandarin.

Table 1: *Summary of the mixed-effects regression model for accuracy of identifying natural stimuli (syllable reference level = da. *p < 0.05; **p < 0.01; ***p < 0.001)*

	Est	SE	df	t	*p*
(Intercept)	0.50	0.10	42.03	4.96	0.00***
syllable(di)	0.04	0.00	3451.00	15.24	0.00***
age	0.02	0.01	42.02	1.95	0.06

Results in Table 1 indicate no age-related differences in accuracy; however, accuracy for the /di/ syllable significantly exceeds that for the /da/ syllable. To further demonstrate the individual differences among participants and provide a basis for grouping in subsequent research, Participants were divided into three groups based on their recognition rates for different syllable stimuli: Group A (accuracy ≥ 15/16), Group B (accuracy ≥ 9/16 and < 15/16), and Group C (accuracy < 9/16). Each Participant was categorized twice based on identification accuracy for /da/ and /di/. The distribution of participants per group and syllable is listed in Table 2.

Table 2: *Grouping: Participants Categorized Twice Based on Identification Accuracy for /da/ and /di/.*

	da	di
A	2	10
B	34	27
C	8	7

Experiment 2: the utilization of perceptual cues

Experiment 2 examines how three participant groups use perceptual cues from T1 and T4's acoustic features in CPM, focusing on F0 height, F0 movement, and voice quality. We artificially synthesized a total of 80 stimuli(2 syllables * 2 voice qualities * 4 F0 slopes * 5 F0 heights.), and the F0 range of the stimuli is shown in the FIG.2.

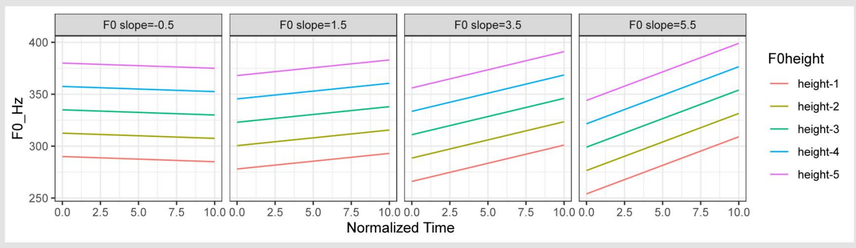

FIG. 2: F0 range diagram. X-axis: Normalized Time. Y-axis: F0 in Hz. Each subplot shows stimuli with the same F0 slope. Grayscale differentiates F0 heights, from height-1 to height-5, increasing sequentially.

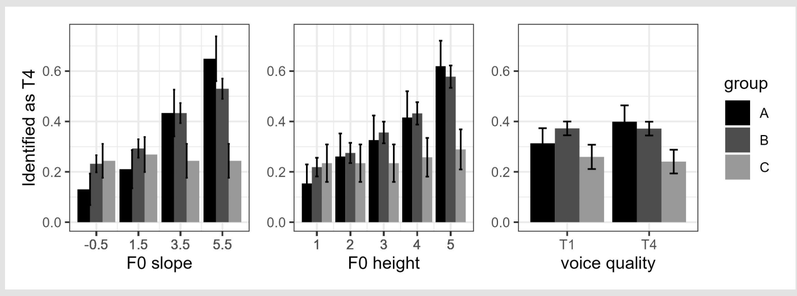

FIG. 3:T*he identification frequency of T4 stimuli, grouping in grayscale. Features such as F0 slope, F0 height, and voice quality are displayed from left to right, with error bars indicating a 95% confidence interval.*

FIG.3 shows differences in perceptual cue utilization among groups: F0 height and slope impact Groups A and B's perception($p < 0.001$), and voice quality affects only Group A($p=0.03$), with minimal effects on Group C from all three factors($p>0.05$).

Experiment 3: production

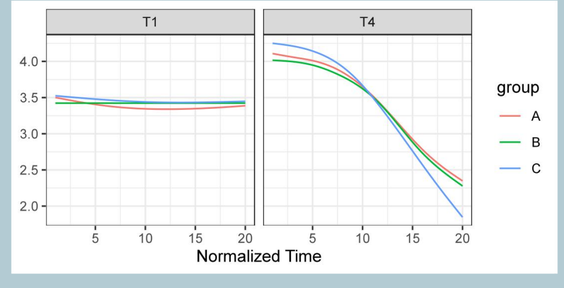

FIG. 4 :*Normalized F0 of T1 and T4 syllables by groups A, B, and C.Y-axis: normalized F0, X-axis: standardized time. Different lines indicate different groups, with the left chart for T1 and the right for T4.*

The normalized curves of the F0 outputs for each group in T1 and T4 are shown in figure 4. From FIG. 4, these curves align more closely with SM than CPM.

Meanwhile, through statical analysis, there is no significant differences in T1 tone characteristics between groups. For T4 tones, A and B show no differences; A and C differ in quadratic and slope, not intercept; B and C differ only in slope. This suggests that A and B's ability to distinguish T1 and T4 tones in CPM leads to significantly gentle slopes in T4 tone output compared to C.

We also studied T1 and T4 voice quality characteristics across groups. We described voice quality using harmonics differences (H1-H2, H2-H4) after formant corrections. No significant differences were found in T1 and T4 outputs using /da/ or /di/ as syllables ($p>0.05$), nor between groups ($p>0.05$), suggesting that the outputs are closer to SM in voice quality, which means on the difference in voice quality between T1 and T4, without group differences.

Discussion

- **Individual difference**

According to Experiment 1, not all primary school students in Changsha who do not speak CD possess the ability to differentiate T1 and T4 in CPM. The participants in our study all come from the same primary school with a similar language environment. In our survey of their family language backgrounds, we did not identify any factors related to this difference in accuracy. This indicates that for this emerging dialect, attitudes, production, and perception vary between individuals, suggesting that the status of this dialect is in a dynamic state and less stable than local dialects or standard languages.

- **The contradiction between production and perception**

There appears to be a contradiction between perception and production: Participants in Groups A and B can perceptually distinguish tones T1-T4, yet their outputs resemble those of Group C, who can't differentiate these tones, leaning more towards SM than CPM. Despite perceptual distinction, they fail to produce these tones, presenting an apparent contradiction.

This discrepancy can be explained by the fundamental differences in experimental approaches: Standard production tasks allow free cue production, whereas perception experiments involve passive reception and recognition of artificially synthesized stimuli(Schertz and Clare, 2020).

- **The correlations among identification accuracy, the utilization of perceptual cues, and production**

Integrating the findings from Experiment 1, Experiment 2, and Experiment 3, our study revealed significant correlations among differentiation accuracy, the utilization of perceptual cues, and production traits.

References

[1] Kuang, J., & Cui, A. (2018). Relative cue weighting in production and perception of an ongoing sound change in Southern Yi. Journal of Phonetics, 71, 194–214.

[2] Jin, S., & Niu, F. (2010). A Study on the Solidification of Changsha Local Mandarin: A Case Study of the Solidification of Local Mandarin, Journal of Language and Writing Application 04, 41-49. 劲松 & 牛芳.(2010).长沙地方普通话固化研究——地方普通话固化的个案调查.语言文字应用(04),41-49.

[3] Xu, C. (2022). Investigating the tonal system of Plastic Mandarin: A cross-varietal comparison. (Doctoral dissertation). University of Oxford.

[4] Schertz, J., & Clare, E. (2020). Phonetic cue weighting in perception and production (2nd ed.). John Wiley & Sons, Ltd.

Acknowledgements

We would like to express our gratitude to the teachers and students from Lushan International Experimental Primary School for their assistance in data collection.