Acoustic Attributes of Citation Tones in Standard Chinese Produced by Prelingually Deaf Adults

Jie Hou* Yu Chen* Yutong Xing* and Jianwu Dang†
* Tianjin University of Technology, Tianjin, China
E-mail: chenyu@tjut.edu.cn Tel: +86-22-60216328
† Tianjin University, Tianjin, China

***Abstract*—Based on perception judgment and acoustic analysis with the data of ten prelingually deaf adults (PDAs) and ten normal hearing adults (NHAs), the present paper investigated the performance of the four citation tones in Standard Chinese produced by prelingually deaf adults. Overall, the error rate of perception judgment for PDAs was 12.95%; however, the error rates of Tone 2 and Tone 3 were 30.70% and 19.85% respectively, which were much higher than that of Tone 4 and Tone1. As the results of acoustic analysis, although the general performance of the deaf females was significantly different from NHAs, they approached or reached the level of the control group on some parameters, while deaf males faced greater challenge than deaf females. Therefore, this study confirmed that PDAs still have some impairments on tone production of Standard Chinese.**

I. Introduction

It is known that, normal hearing children are sensitive to tone perception as early as 5 months [1]; and they could acquire all four lexical tones in Standard Chinese in 28 months or before the age of three [2], [3]. However, because of the congenital deafness or hearing impairment before four years old, the acquisition of Mandarin tonal system is an arduous task for the prelingually deaf children (PDC). Therefore, as a hotspot in the area of Chinese language rehabilitation and its related disciplines, tone acquisition of PDC had attracted a lot of attention in the past few decades, both on tone perception [4], [5], [6], tone production [7], [8], [9] and the reciprocal relation between tone perception and production [10], [11], [12].

As for tone production, it has been proved that PDC have poor abilities to generate tones, compared with normal hearing children. In a study of evaluating four PDC's tone production (age range: 4.00 to 8.75 years), Xu et al. indicated that the mean scores of intelligibility of deaf children ranged from 0.25 to 8.5, which are obviously lower than the standard of normal hearing children; and the tone curves produced by deaf children tended to be flat in all cases [7]. In another study that compared the speech intelligibility of 26 congenitally deaf children (mean age of 5.9 years, implanted age at 3.5 years) with 26 normal hearing children (mean age of 5.84 years), Huang et al. found that the average correct rates of tones for deaf group and control group were 54.76% and 75.78% respectively; the difference between the two groups was statistically significant [8]. Interestingly, though PDC have poor capabilities on tone production, many studies still reported that some deaf participants could perform as well as the normal hearing children [13]. This may suggest that, despite trapping in hearing difficulties, PDC could acquire the lexical tones well, with the help of devices and practice.

This study is sponsored by The National Undergraduate Training Program for Innovation and Entrepreneurship (201710060001) and partly by The National Social Science Fund of China 'A study on the speech behavior and its neural mechanism of Prelingually deaf adults' (17BYY166).

Furthermore, previous studies on PDC had demonstrated that the four citation tones in Standard Chinese have different hierarchies of difficulty. By analyzing the results of tone confusion matrix, Han et al. found that, Tone 2 was the most severely impaired, followed by Tone 3 and Tone 4; while Tone 1 was the easiest one to produce for cochlear implant children [9]. He also found that PDC produced Tone 1 frequently even the target tone was not it. From these studies, one could conclude that most studies of tone production were using qualitative listener judgments as analysis method, rather than acoustic analysis.

As for PDAs, few studies were concerned about their tone production. In an earlier study about Mandarin citation tone patterns, Chen et al. studied four prelingually deaf females aged from 19 to 22 and suggested that: Tone 2 was the most difficult tone for PDAs, while Tone 3 was easier than Tone 2 but harder than the other two tones [14]. It could be inferred roughly that PDAs' performance is similar to that of PDC on tone production. However, limited by the samples, their results may not be robust. Therefore, it is necessary and urgent to consider the production conditions of PDAs: with many years of hearing reconstruction and long experience of pronunciation practice, could they correctly produce the citation tones in Standard Chinese? And what are the difficulties of tone production for them?

Keeping the aforementioned thoughts in mind, this paper will compare the acoustical attributes of the citation tones between prelingually deaf adults and normal hearing adults. By doing this, we will examine the performance of tone production of prelingually deaf adults, and find their impairments in generating citation tones. The remaining parts of this paper are organized as follows: In Section 2, we will provide an introduction about the participants, materials and recording, and data processing. Section 3 will present the results of experiment and Section 4 will discuss our findings. Then in Section 5, we will conclude this paper.

II. Experiment

A. Participants

Ten participants (5 males and 5 females), aged from 19 to 22, were recruited from Technical College for the deaf, Tianjin University of Technology. They are all prelingually deaf people who lost their hearings in born (n=4, 2 males) or before four years old (n=6, 3 males). All of them had hearing aids or received cochlear implantations at the time of data recording. And all of them speak Standard Chinese as their first oral language.

As for the control group, 10 age-matched normal hearing adults (5 males and 5 females) were recruited from Tianjin University of Technology. They are speaking Standard Chinese as their oral language. Besides, all of them have no historical hearing diseases or speech disorders.

B. Materials and Recording

The materials for recording were selected from the Mandarin Speech Test Materials (MSTMs) which contained 350 mono-syllabic words, 400 bi-syllabic words, 150 sentences and one long passage [15]. The current study used the mono-syllabic word list and had balanced the number of words with the tones. So the wordlist employed in this study consisted of 281 tokens: they were 71 tokens for Tone 1, 70 tokens for Tone 2, 72 tokens for Tone 3 and 68 tokens for Tone 4.

A recording system was developed by Matlab 12.0b for presenting wordlist and recoding data. Participants were guided to access this recording system in a speech lab to record their speech production. After the data recording, each item was stored as an individual audio file in .wav format (mono soundtrack, 16 KHz, 32-bit resolution).

C. Data Processing

Several steps were adopted to process the collected speech recording:

First, the recording of deaf participants was perceptually judged by two well-trained Chinese linguistics students separately for the accuracy of tone production. They were required to judge whether the tone of each token was 'right' or 'wrong'; for the 'wrong' cases, they were asked to point out the error types as follows: Tone 1 was produced as Tone 2, Tone 3, Tone 4 or Tone X (atypical tone); Tone 2 was produced as Tone 1, Tone 3, Tone 4 or Tone X; Tone 3 was produced as Tone 1, Tone 2, Tone 4 or Tone X; and Tone 4 was produced as Tone 1, Tone 2, Tone 3 or Tone X. The wrong tokens were not be used for further analysis.

Second, the remaining data were imported into Praat for acoustic analysis. The F0 values of each token were extracted from 11 equal time interval points, from $F0_1$ to $F0_{11}$. Then three acoustic parameters were obtained in this study:

* $F0_1$: Fundamental frequency of the first point on the tone contour
* F0-Difference: F0 range, is obtained by calculating the difference of fundamental frequencies between the first point and the last point ($F0_1$ - $F0_{11}$)
* Duration: Time length between $F0_1$ and $F0_{11}$

F0-Difference can show the tendency of F0 change, which is important for tones. Tone 1 is a high-level tone, so the F0-Difference is very small. Tone 2 is a middle-rising tone and its value is negative and large. Tone 4 is a high-falling tone and the value is positive and large. As for Tone 3, although it is labelled as a falling-rising tone, its rising is not such dramatic as Tone 2, and it is not the critical distinctive feature for Tone 3 [16]. Namely, the F0-Difference of Tone 3 may be negative, and its value is relatively small. Together with $F0_1$ and Duration, the three parameters in this study are effective in describing tonal properties and differentiating the four tones in Standard Chinese.

At last, *z-scores* of the three parameters were calculated for normalization and comparison purposes. And all of the parameters (Z-$F0_1$, Z-Difference and Z-Duration) were entered in SPSS for statistical analysis.

III. Results

A. Perception Judgment

The tone confusion matrix is displayed by Table 1. Overall, there are 340 counts in 2626 (12.95%) which are classified as errors. In detail, they are 199 of Tone 2, 130 of Tone 3 and 11 of Tone 4. That is to say, except for the case of Tone 1, the error rates of Tone 2, Tone 3 and Tone 4 are 30.70%, 19.85%, 1.73% separately. Therefore, the order of error rates of the four lexical tones in Standard Chinese is T2 > T3 > T4 > T1 for prelingually deaf adults.

B. Acoustic Analysis

Descriptive Analysis

Table II shows the differences of the acoustical attributes among the four participant groups. Taking Tone 1 as example, the $F0_1$ and F0-Difference of DFs are 10.07 Hz and 7.7 Hz lower than that of NFs respectively, while the $F0_1$ and F0-Difference of DMs are much higher than that of NMs (48.13Hz of $F0_1$; 24.47 Hz of F0-Difference); the average Duration of DFs is 28.74 ms longer than that of NFs, while the average Duration of DMs is similar to that of NMs.

MANOVA Analysis

A series of ANOVAs with the *z-scores* data indicated that the differences between DFs and NFs, DMs and NMs, DFs and DMs were reserved, while the differences between NFs and NMs were neutralized. Therefore, the *z-score* data of normal

TABLE I
Tone confusion matrix of tone production of the prelingually deaf adults (PDAs)

Target	Responses				
	Tone1	Tone2	Tone3	Tone4	ToneX
Tone1	686	0	0	0	0
Tone2	22	449	138	7	32
Tone3	7	56	525	3	64
Tone4	0	2	2	626	7

TABLE II
DESCRIPTIVE STATISTICS ON DIFFERENT TONES AND DIFFERENT PARAMETERS

		DFs		DMs		NFs		NMs	
		MEAN	*SD*	*MEAN*	*SD*	*MEAN*	*SD*	*MEAN*	*SD*
Tone1	$F0_1$	272.68	49.24	200.21	38.29	282.75	31.50	152.08	17.82
	F0-Difference	-9.12	28.90	24.15	22.70	-1.42	13.59	-0.32	7.98
	Duration	344.64	122.96	330.67	87.16	315.90	57.21	334.17	71.09
Tone2	$F0_1$	212.52	28.27	152.26	21.21	212.35	20.19	118.86	15.26
	F0-Difference	-72.94	25.94	-37.86	29.99	-70.63	24.41	-47.58	14.54
	Duration	351.60	120.04	274.56	49.43	321.15	46.89	308.50	75.20
Tone3	$F0_1$	219.09	43.68	164.66	25.04	209.47	22.86	121.24	24.22
	F0-Difference	-9.37	45.45	7.79	21.50	-8.71	21.83	-5.03	20.43
	Duration	471.70	151.97	354.28	67.06	432.42	69.89	409.12	95.67
Tone4	$F0_1$	295.63	43.64	225.79	37.81	322.93	45.67	181.84	25.70
	F0-Difference	102.37	46.86	94.89	41.76	142.49	56.68	79.29	26.70
	Duration	212.28	83.20	222.05	95.20	213.45	70.60	223.31	70.23

*DFs refers to deaf females; DMs refers to deaf males; NFs refers to normal hearing females; NMs refers to normal hearing males; Hz for $F0_1$ and F0-Difference; ms for Duration

hearings are combined as the control group in this study for further analysis.

The MANOVA considered the Participant (three levels: DFs, DMs and NHAs/Control) and Tone (four levels: T1, T2, T3 and T4) as the independent factors with the confidence level of 0.001, and a series of subsequent Kruskal-Wallis tests were performed to confirm the findings of MANOVA test. Results of MANOVA indicated that, although the homogeneity of covariance matrices [Box's M (5605.731), F (66, 1e+007) = 84.642, $p < 0.001$] was significant, all the independent factors reported significant multivariate effects, while Participant contributed the most for the model [Partial η^2: Participant = 0.174; Tone = 0.092; Participant*Tone = 0.046, Power =1.000].

Tests of Between-Subjects Effects indicated that, except for the case of Tone for Z-Duration, significant univariate effects were found on all the remaining situations.

Results of post hoc tests reported that, the values of Z-Difference and Z-Duration were significantly different between each two of the three participant groups (See Fig. 1). For the case of Z-$F0_1$, as shown by Fig. 1, while DMs were significantly different from the other two groups, there was no significant difference between DFs and Control (I-J = 0.100, SD= 0.051, p = 0.140). Thus, except the Z-$F0_1$ of DFs, the performances of PDAs were significantly different from the norms. In particular, comparing with the cases of the DFs, the divergences of Z-Difference and Z-$F0_1$ of DMs were impressively divergent from the norms. In contrast, the divergences of the Z-Duration of DMs seemed a little better than their female peers.

To explore the performance of the three groups on each tone, the data was split by Tone and then reanalyzed with a one-way MANOVA. The results indicated that the main effects of Participant on all tones' data were significant at $p < 0.001$ level [Partial η^2: Tone 1 = 0.251; Tone 2 = 0.224; Tone3 = 0.227; Tone 4 = 0.174, Power = 1.000]. Moreover, significant univariate main effects for Participant were obtained for almost all the data of the four tones, for which the only exception was Participant on T4's Z-Duration [F (2, 1290) = 0.226, p = 0.798, partial $\eta^2 < 0.001$, Power = 0.003].

1) Z-$F0_1$: Since there were no significant differences of Z-$F0_1$ on T1, T2 and T4, and the difference of T3 between DFs

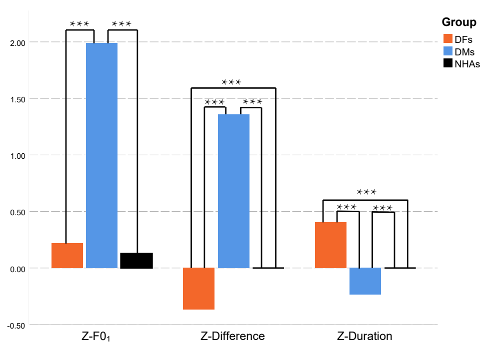

Fig. 1. Mean *z-scores* and the significance of the differences grouped by Participant at p <.001 level

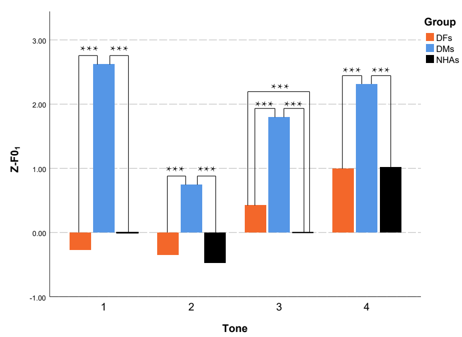

Fig. 2. Z-$F0_1$

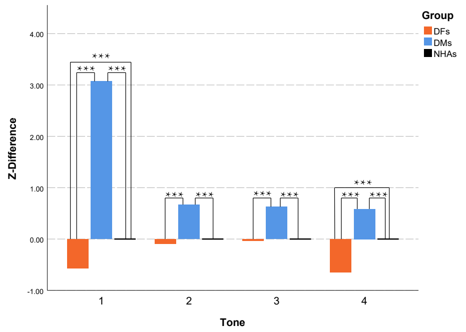

Fig. 3. Z-Difference

and control group was much smaller than that of DMs and Control, it could be safe to infer that DFs performed better than DMs in setting the starting pitch of the tones. Especially in T1, T2 and T4, DFs were approaching the level of their normal hearing counterparts.

2) Z-Difference: Besides the insignificant differences of Z-Difference on T2 and T3 between DFs and control group, Fig. 3 also revealed that the divergence of T1 between DFs and Control was much smaller than that of DMs and Control. Furthermore, for all four tones, the deaf groups demonstrated a unitedly reverse direction of divergences.

3) Z-Duration: For the cases of Z-Duration, as shown by Fig. 4, post hoc tests failed to report any significant differences between DMs and Control on T1, and between all comparing pairs of T4. Except for T4, the three tones left again showed a consistent tendency that the DFs' Z-Duration was positive, while the DMs' direction was negative. Thus, it seemed that the DMs tended to realize a hypo-durational tone, while the DFs liked to perform a tone with a lengthened duration.

Taking together, besides some exceptions such as the Z-Duration of T1, the divergences of DFs' data from the norms were much smaller than the DMs. Therefore, the above analysis revealed that the performance of DFs seemed better than DMs.

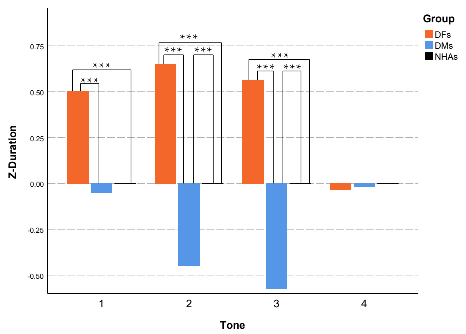

Fig. 4. Z-Duration

IV. Discussion

Results of the present paper indicate that DFs can't keep well control on $F0_1$ of Tone 3, as well as F0-Difference of Tone 1 and Tone 4; except these cases, they can reach the level of participants with normal hearing on these two parameters. In contrast, DMs perform worse than DFs in this respect. Therefore, it could conclude that PDAs, especially for DMs, still have some impairments on tone production with many years of hearing reconstruction and abundant experience of hearing practice.

What is particularly noteworthy is the deaf groups' deficiency in the dimension of tonal duration. Except for Tone 4, DFs tend to produce tones with longer durations while DMs always produce tones with shorter durations for the other three tones. Although their tonal durations of Tone 1 and Tone 4 are standard, the DMs' durations of Tone 2 and Tone 3 are much shorter than that of NHAs. Therefore, one could conclude that DMs still should put extra effort on tonal duration. As for DFs, they have adopted a strategy with the overlong durations to make tones more 'recognizable' or 'full'.

Another point which is worth discussing is that why do PDAs always confuse Tone 2 and Tone 3 in production. In fact, the confusion between Tone 2 and Tone 3 is a common phenomenon in Standard Chinese. Li and Thompson proposed that the confusions between Tone 2 and Tone 3 exist persistently in the lexical tone acquisition of normal children' language development until they could accurately produce all four tones [17]. And even adults frequently misidentify Tone 3 as Tone 2 when presented with monosyllabic words in isolation [18]. Besides, the confusion of Tone 2 and Tone 3 is also found on learners who use Chinese as a second language [19]. Therefore, it is natural that PDAs share the same pattern with the above-mentioned groups in confusing Tone 2 and Tone 3.

Furthermore, Tone 2 and Tone 3 have some features in common: both of them have an initial declining and then a rising; and the two tones have the similar initial F0 values. Thus, just as this study shows, those similarities may make the distinction between Tone 2 and Tone 3 much harder. From the results, one could see that DMs' $F0_1$ and F0-Difference are far from the standard. DFs can master the these two attributes well, yet they can't generate the correct $F0_1$ of Tone 3. Meanwhile, both DFs and DMs can't control the Duration of Tone 2 and Tone 3 well. Therefore, to capture the distinction of Tone 2 and Tone 3 in monosyllabic words is still a tough work for PDAs.

V. Conclusions

As a brief conclusion, this study reconfirmed that, Tone 1 was the easiest one among the four citation tones in Standard Chinese for PDAs; Tone 2 and Tone 3 were the hardest to produce and they were easily confused with each other. By analyzing the acoustic data, this study also found that PDAs still had some impairments on tone production compared with NHAs; deaf females performed better than deaf males.

In the future, we will devote to compare the tone production between participants with hearing aids and participants with cochlear implants; and investigate the tone production of PDAs in polysyllabic words and continuous speech.

REFERENCES

[1] H. H. Yeung, K. H. Chen, and J. F. Werker, "When does native language input affect phonetic perception? The precocious case of lexical tone." *Journal of Memory & Language*, vol. 68, no. 2, pp. 123–139, 2013.

[2] Z. Hua and B. Dodd, "The phonological acquisition of Putonghua (Modern Standard Chinese)." *J Child Lang*, vol. 27, no. 1, pp. 3–42, 2000.

[3] F. Chen, G. Peng, N. Yan, and L. Wang, "The development of categorical perception of Mandarin tones in four- to seven-year-old children," *Journal of Child Language*, vol. 44, no. 6, p. 1, 2016.

[4] T. S. Huang, N. M. Wang, and S. Y. Liu, "Tone perception of Mandarin-speaking postlingually deaf implantees using the Nucleus 22-Channel Cochlear Mini System." *Annals of Otology Rhinology & Laryngology Supplement*, vol. 166, no. 9, p. 294, 1995.

[5] J. L. Wu and H. M. Yang, "Speech perception of Mandarin Chinese speaking young children after cochlear implant use: effect of age at implantation." *International Journal of Pediatric Otorhinolaryngology*, vol. 67, no. 3, pp. 247–253, 2003.

[6] S. Zhu, L. L. Wong, and F. Chen, "Development and validation of a new Mandarin tone identification test." *Int J Pediatr Otorhinolaryngol*, vol. 78, no. 12, pp. 2174–2182, 2014.

[7] L. Xu, Y. Li, J. Hao, X. Chen, S. A. Xue, and D. Han, "Tone production in Mandarin-speaking children with cochlear implants: a preliminary study." *Acta Oto-Laryngologica*, vol. 124, no. 4, pp. 363–367, 2004.

[8] C. Y. Huang, H. M. Yang, Y. J. Sher, Y. H. Lin, and J. L. Wu, "Speech intelligibility of Mandarin-speaking deaf children with cochlear implants." *International Journal of Pediatric Otorhinolaryngology*, vol. 69, no. 4, pp. 505–511, 2005.

[9] D. Han, N. Zhou, Y. Li, X. Chen, X. Zhao, and L. Xu, "Tone production of Mandarin Chinese speaking children with cochlear implants," *International Journal of Pediatric Otorhinolaryngology*, vol. 71, no. 6, pp. 875–880, 2007.

[10] S. C. Peng, J. B. Tomblin, H. Cheung, Y. S. Lin, and L. S. Wang, "Perception and Production of Mandarin Tones in Prelingually Deaf Children with Cochlear Implants," *Ear & Hearing*, vol. 25, no. 3, pp. 251–264, 2004.

[11] L. Xu, X. Chen, H. Lu, N. Zhou, S. Wang, Q. Liu, Y. Li, X. Zhao, and D. Han, "Tone perception and production in pediatric cochlear implants users," *Acta oto-laryngologica*, vol. 131, no. 4, p. 395, 2011.

[12] N. Zhou, J. Huang, X. Chen, and L. Xu, "Relationship between tone perception and production in prelingually deafened children with cochlear implants," *Otology & Neurotology*, vol. 34, no. 3, pp. 499–506, 2013.

[13] J. Tan, R. Dowell, and A. Vogel, "Mandarin Lexical Tone Acquisition in Cochlear Implant Users With Prelingual Deafness: A Review." *American Journal of Audiology*, vol. 25, no. 3, p. 1, 2016.

[14] Y. Chen, Y. Chen, J. Zhang, J. Zhang, H. Lin, J. Wei, J. Dang, "Mandarin citation tone patterns of prelingual Chinese deaf adults," *APSIPA ASC 2016*, pp. 1-4, 2016.

[15] H. Zhang, S. Wang, J. Chen, L. Wang, G. Y. Shao, L. I. Yu-Ling, and W. B. Jiang, "The Mandarin Speech Test Materials (MSTMs):Development and application," *Chinese Scientific Journal of Hearing & Speech Rehabilitation*, 2008.

[16] Y. J. Wang and X. H. Qin, "The identification and discrimination of Tone2 and Tone3 in Mandarin: The Influence of Experimental Design on Categorical Perception of Tones," *Linguistic Sciences*, vol. 4, pp. 337–352, 2015. (In Chinese)

[17] C. N. Li and S. A. Thompson, "The acquisition of tone in mandarin-speaking children," *Journal of Child Language*, vol. 4, no. 2, pp. 185–199, 1977.

[18] C. K. Chuang and S. Hiki, "Acoustical features and perceptual cues of the four tones of standard colloquial chinese," *The Journal of the Acoustical Society of America*, vol. 52, no. 1A, pp. 146–146, 1972.

[19] A. H. Wang, "Discussion on the teaching of chinese tone characteristics," *Language Teaching and Linguistic Studies*, vol. 3, pp. 70–75, 2006. (In Chinese)