Study on the Tones Biases of Mandarin Speaker in Amdo Tibetan Areas Based on Statistics

Gan Zhenye*†, Jiao Yi*, Yang Hongwu*†, Zhao Gaungying*, Song Zhimeng*
*College of Physics and Electronic Engineering, Northwest Normal University
†Intelligent Information Processing Center of Gansu Province
Lanzhou, China
E-mail: ganzy@nwnu.edu.cn 1767013387@qq.com 732637787@qq.com guangying_zhao@sina.com

Abstract—Tone learning is a major difficulty when students in Amdo Tibetan area learn mandarin. This paper uses experimental phonetics, comparative analysis, and theory and methods of biases analysis to investigate and analyze the pronunciation of mandarin in the Amdo Tibetan area. The perception experiment and similarity experiment were used to analyze the tone of mandarin in the Amdo Tibetan area. The experimental results show that in the process of learning mandarin, the students in Amdo Tibetan area learn from the highest to the lowest in order of tone4, tone1, tone3 and tone2. The most prone to sound in the process of speaking mandarin is tone4, and the most common pronunciation in their biased pronunciation is tone4. The similarity detection has a higher diagnostic accuracy rate, and tone1 has the best effect. The tone2 and tone3 are easy to judge the correct pronunciation as the biased pronunciation, and tone4 is easy to judge the biases pronunciation as the correct pronunciation.

Keywords: pronunciation biases, perception, similarity detection, tone.

I. INTRODUCTION

In recent years, Intra-Chinese teaching and teaching Chinese as a foreign language have become a hot topic. The study is mainly focused on teaching Chinese as a foreign language. It takes Botswana, Tamil, Turkish, Chile, Indonesia and other countries to learn Chinese pronunciation and meta consonant pronunciation biases [1-5]. It is found that in tone biases, tone1 and tone4 are easier to grasp than tone3 and tone2. In vowel and consonant biases, high and middle vowels, aspiration and non-aspiration consonants are more likely to make biases. In the teaching of Chinese as a foreign language, the students whose mother tongue is Uygur have great problems in the acquisition of intonation owing to the negative transfer of their mother tongue [6,7]. This problem also exists for Chinese learners in Amdo Tibetan areas. Previous studies have mainly studied the influence of the mother tongue in Chinese in Amdo Tibetan areas on the Chinese language acquisition and the causes of the biases and the characteristics of the tone [8-10]. Knowing the causes of biases and the characteristics of biases contribute to language learning for learning languages, but tone learning is also of utmost importance for mandarin learning. In the acquisition process of mandarin, the main difficulties are the pronunciation of initials, finals and tones. The tones are the most sensitive and influential factors [11], and the tones can better reflect the characteristics of mandarin Chinese and form mandarin "caliber" [12]. The main reason for the phenomenon of foreigners who speak Chinese when they speak Chinese is the tone and the higher level of speech than the tone [13], so the study of tone is very important in the acquisition of mandarin. Due to the particularity of mandarin, the same vowel may express completely different meanings in different tones, such as "qu3" and "qu4", so tone learning is increasingly crucial to speak mandarin. Tibetan belongs to the Sino-Tibetan language family. Tibetan is divided into three major dialects: Weizang dialect, Kangba dialect and Amdo dialect. Only Amdo dialect has no distinctive tone, while mandarin is a tonal language [14]. Tone is a special difficulty for the native speaker who has no tonal language [15]. Therefore, for Tibetans whose mother tongue is Tibetan, the study of tone in mandarin is a major difficulty. When students in Amdo Tibetan area learn mandarin, there exists the phenomenon of Tibetan tune, which is mainly manifested in the pronunciation of tones. Therefore, this paper takes the study of mandarin tonality in Amdo Tibetan area as the research object, and studies the acquisition of mandarin tones in Amdo Tibetan areas.

In this paper, from the perspective of tone, using experimental phonetics, mandarin is regarded as the target language of second language acquisition, and the study of mandarin tone acquisition of Undergraduates in Amdo Tibetan area is carried out. This experiment is divided into two parts: the perception experiment and the similarity test experiment. The main purpose of the perception experiment is to distinguish the biases of Amdo Tibetan students when they speak mandarin, and to make statistics and analysis of the biases. In the similarity test, the intonation similarity of mandarin corpus of Amdo Tibetan students was calculated with standard mandarin as the control group. Finally, the results of the perception experiment and the similarity experiment are analyzed to determine the degree of biases and provided learning basis for mandarin learners in Amdo Tibetan area.

II EXPERIMENTAL DESIGN

This experiment consists of two parts: perception experiment and similarity test. The perception experiment is mainly based on the students who live in the mandarin environment, and the students in the Amdo Tibetan area speak the tone of the ordinary corpus. These students have the first grade in the mandarin Proficiency Test (They can speak standard mandarin). The tone of the auditory results is obtained, and the results of the perception are analyzed to illustrate the tone mastery and tone biases types of the mandarin students in the Amdo Tibetan area. The similarity detection experiment first calculates the similarity of each tone of standard mandarin. Since the fundamental frequency of the tone is different when different people speak mandarin, the same tone will have a slight difference. The corpus of the same tone is "correspondence in turn" for similarity detection, and the average test results are taken as the evaluation criteria. The similarity of each tone of standard mandarin and Amdo Tibetan college students' mandarin is calculated. The experimental results under the evaluation criteria are obtained and analyzed.

A. Perception experiment

Eight subjects (4 males and 4 females) were undergraduates of the Northwest Normal University. They are all from Amdo Tibetan area at the age of 22. Their mother tongue are Amdo Tibetan. They have no language and hearing impairment. Their parents are Amdo Tibetans. Their communication language is Amdo Tibetan. Mandarin is their second language. The experimental corpus was recorded by eight subjects. Tone acquisition is more difficult in continuous language flow, so the text used in corpus recording is 40 long sentences, each long sentence contains four short sentences, the last syllable of each short sentence is tone1, tone2, tone3 and tone4, respectively. After the corpus recording is completed, the last syllable of each voice is segmented by Cool Edit to obtain a monosyllabic voice. A total of 1280 (40*4*8=1280) voices per tone.

The perception experiment was completed by 4 (2 males and 2 females) college students at the Northwest Normal University. The four students scored in the mandarin Proficiency Test. They lived in the mandarin environment. In the process of listening, two students (one male and one female) listened to the voice of the same participant and distinguished their voice with correct speech and biased speech. If there were inconsistent judgments, they were handed over to a third person for judgment, and the results of all voices were obtained. Statistics on the number of biases in mandarin speech in Amdo Tibetan students, analysis of the type of biases, and calculation of the bias rate of tonal biases (bias rate = number of biases in the tone / total number of occurrences of the tone).

Let select an Ando Tibetan student to speak the pronunciation data of mandarin to make the tone curve of mandarin in the Amdo Tibetan area. Using the standard mandarin in the similarity experiment makes the standard tone curve, as shown in Fig. 1 and Fig. 2.

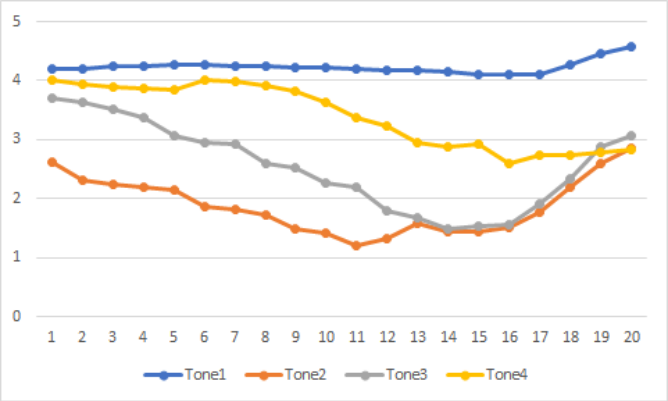

Fig. 1 Amdo Tibetan students speak the tone curve of mandarin

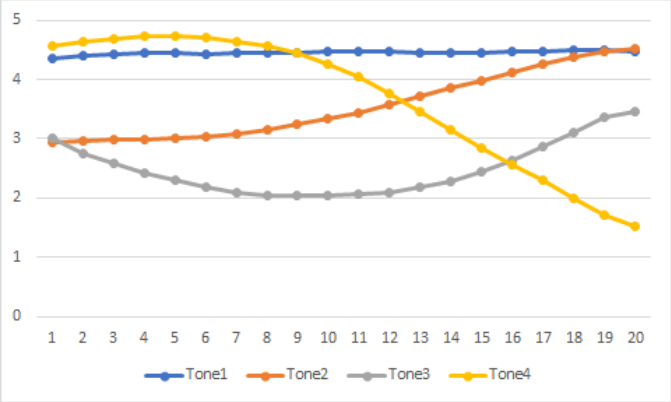

Fig. 2 standard mandarin tone curve

In the Fig. 1 and Fig. 2, the abscissa is the fundamental frequency point, and the ordinate is the fundamental frequency value after normalization. After normalization, the biases caused by the singular fundamental frequency value can be eliminated, and the comparison is easier.

Fig. 1 shows the Amdo Tibetan students speaking the tone curve of the mandarin. Each tone is averaged from 40 samples. Fig. 2 shows the standard mandarin tone curve. It can be seen from Fig. 1 and Fig. 2 that students in Amdo Tibetan area say that the tone curve of mandarin is different from the tone curve of mandarin. The difference between the tone1 curve and the standard curve in Fig. 1 is the smallest, but there is a bias in the tail. The difference between the tone2 curve and the standard curve is the largest, and it can be seen that it is completely biased. When the tone3 curve is compared with the standard curve, it is found that they both fall first and then rise, but the rising inflection point in Fig. 1 is at the tail and Fig. 2 is at the middle front; The tone4 curve shows a downward trend compared with the standard curve, but the decrease in Fig. 1 is much smaller than that in Fig. 2. On the whole, the curve of Fig. 1 will appear as a rise or a straight line at the end, which may be caused by the end of the Amdo Tibetan language.

Judging from the perception results of mandarin Chinese corpus in Amdo Tibetan students, there were 657 biases in 1280 voices, and the bias rate reached 51.33%. Among them, the situation of males and females students' biases and the tone deviations are different. The following table shows the

statistical tables of males and females and the tone deviations.

Tab. 1 students in Amdo Tibetan areas speak the pronunciation biases of mandarin and male students.

Gender	Males	Females
Number of biases	348	309
Bias rate	54.38%	48.28%

It can be seen from Tab. 1 that in the process of learning mandarin, the bias rate of males is greater than that of females. Under the same environment, females are more capable of learning new languages than males. However, as a whole, due to the influence of their native language, even in the language environment of mandarin, they are more serious when they are learning mandarin. The mastery of tone is not quite good. The overall bias rate has also reached more than 50%, so they will have great difficulties in learning.

Tab. 2 students in Amdo Tibetan areas say mandarin.

Tone	Tone1	Tone2
Number of biases	153	251
Bias rate	47.81%	78.44%
Tone	Tone3	Tone4
Number of biases	202	51
Bias rate	63.13%	15.94%

It can be seen from the perception experiment of tone biases that the students of Amdo Tibetan College will appear when speaking mandarin are concentrated in tone2 and tone3. In contrast, tone1 and tone3 are better. The highest bias rate is tone2, followed by tone3, the lowest bias rate is tone4, followed by tone1. When listening to the corpus of Amdo Tibetan students, it was found that the main reason for the deviation of tone1 was that the pronunciation was low during the pronunciation process, and it was easy to make tone1 into tone2 and tone3. The most prone to the biases of tone2 and tone4 is that they are easily confused and cannot be distinguished well. Tone4 is the least biased in listening experiments, and tone4 is the easiest to master. Tibetan students cannot distinguish the characteristics of tones very well in the process of tone acquisition, so there is a common biases in all the biases that it is easy to make all the tones into tone4.

B. similarity experiment

The similarity experiment is to calculate the similarity between the standard mandarin and the Amdo Tibetan students' mandarin tone curve, and to give a measure to the mandarin students in Amdo Tibetan area. The standard mandarin corpus used here was recorded by college students at Northwest Normal University, with a total of 4 students (2 males and 2 females).The text used in the recording of the corpus is the same as the listening experiment, and the monosyllabic speech is processed in the same way as the auditing experiment, with 40 tones and 640 (4*4*40) voices. The corpus of eight subjects was used in the acquisition of mandarin in the Amdo Tibetan area.

The fundamental frequency is extracted for each voice. The specific method is to use the Praat software to mark each voice, and then extract the fundamental frequency, and extract the base frequency value of each voice by 10, 20, 50, so that the obtained result is more convincing.

The fundamental frequency values of each speech extraction are normalized. The purpose of the normalization process is to reduce the random difference and extract constant parameters. The specific method is the Min Max Scaler (as shown in Equation 1), scaling the feature to a given minimum and maximum value, or you can convert the maximum absolute value of each feature to a unit size. This method is a linear transformation of the original data, normalizing the data to [0,1]. This scales the fundamental frequency of each speech to between [0,1], which eliminates the anomaly and makes the calculation more reliable.

$$X_0 = \frac{X_i - L}{H - L} \tag{1}$$

Where X_i is the fundamental frequency value of the sampling point, L is the lower limit frequency value of the subject's tuning domain, and H is the upper limit frequency value of the subject's tuning domain. In the similarity calculation, the similarity is calculated using the correlation coefficient (as shown in Equation 2).

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}} \tag{2}$$

Where n is the number of base frequency points, and X_i and X are the fundamental frequency value and the mean value, respectively. The value of r is between -1 and +1, but in the experiment of this paper, r takes the absolute value.

The data obtained by normalizing the obtained fundamental frequency values were used as inputs for the similarity experiment. Draw the tone curve of each tone, and eliminate the voice with non-standard pronunciation, get 146 voices of tone1, 150 voices of tone2, 153 voices of tone3, and 160 voices of tone4. The similarity calculation is performed for each speech of each tone and the remaining speech, resulting in tone1 21170 (146*145=21170) results, tone2 22350 (150*149=22350) results, tone3 23256 (153*152=23256) Results, tone4 25440 (160 * 159 = 25440) results. The average of each result is used to obtain the criterion for each tone, as shown in Tab. 3.

It can be seen from Tab. 3 that when the number of extracted fundamental frequencies is different, the similarity obtained has a certain difference, but the difference is not significant. In the four tones, the similarity between tone1 and tone3 is small, and tone2 and tone4 are relatively large. The reason for this large difference is that their adjustment values are different, as shown in Fig. 2. Tone1 has a value of 55, and

44 can be considered as unbiased during normal communication. The tone curve of tone 1 is a straight line with a slope close to 0, but there is a small fluctuation when extracting the fundamental frequency. This fluctuation is uncertain, resulting in the fundamental frequency fluctuation of each voice is different, so the similarity of tone 1 is relatively small. On the contrary, although tone2 and tone4 have different fundamental frequency fluctuations, they have the same fluctuation trend and get a greater similarity. Tone 3 is a tortuous tone, so it will also be affected in the process of similarity calculation, and the similarity will be smaller than tone 2 and tone 4. However, in the saliency test, all similarities are significantly correlated. In the later experiment, the number of fundamental frequencies used is 20, so the criterion of similarity is based on the results of extracting 20 fundamental frequencies from standard mandarin, tone1 0.44, tone2 0.87, tone3 0.68, tone4 0.81.

Tab. 3 standard mandarin similarity test table

Tone	Tone1			Tone2		
Number of pitch points	10	20	50	10	20	50
Similarity	0.45	0.44	0.45	0.88	0.87	0.86
Tone	Tone3			Tone4		
Number of pitch points	10	20	50	10	20	50
Similarity	0.71	0.68	0.68	0.82	0.81	0.8

A speech in standard mandarin is randomly selected as a control, and the similarity is detected. When the similarity is greater than the experimental results, it is judged to be correct. The correct number and the correct rate are counted (the correct rate = the correct number of the tone/the total number of occurrences of the tone). The following table is the result of similarity detection.

Tab. 4 Amdo Tibetan students say the mandarin similarity test

Tone	Tone1	Tone2
Correct number	157	62
The correct rate	49.06%	19.38%
Bias rate	50.94%	80.62%
Tone	Tone3	Tone4
Correct number	106	261
The correct rate	33.13%	81.56%
Bias rate	66.89%	18.44%

It can be seen from Tab. 4 that the results of the similarity test show that the bias rate is quite different, the tone4 bias rate is the smallest, and then the tone1, tone3, and tone2 are consistent with the auditory experiment results. However, the bias rate of the similarity detection is high when the magnitude of the bias rate is compared. The reason for this phenomenon may be the biases caused by subjective judgment in the hearing, or the difference in numerical value of the machine detection is easier to distinguish than the human ear, and the fundamental frequency is worthy of a few changes, the human ear cannot hear, but small changes in the machine can be distinguished.

III COMPOSITE EVALUATION INDEX

This paper mainly tests the mandarin in the Amdo Tibetan area from the auditory experiment and the similarity test to determine the correct rate and the bias rate, but there is a certain gap between the correct rate and the bias rate. In order to test the accuracy between the two experiments and the practicability of similarity detection, this paper verifies the experimental results, and the verification results are divided into four types： True Acceptance (TA), the correct pronunciation in the auditory experiment is also detected as correct pronunciation in the similarity detection. True Rejection (TR), the pronunciation of the bias in the listening experiment is also detected as a biased pronunciation in the similarity detection. False Acceptance (FA), the pronunciation of the bias in the perception experiment is detected as the correct pronunciation in the similarity detection. False Rejection (FR), the correct pronunciation in the perception experiment is detected as a false pronunciation in the similarity test. The four results usually use three evaluation indicators: False Acceptance Rate (FAR), False Rejection Rate (FRR), and Diagnostic Accuracy (DA). Their calculation formulas are as follows.

$$FAR=\frac{FA}{FA+TR} \quad (3)$$

$$FRR=\frac{FR}{FR+TA} \quad (4)$$

$$DA=\frac{TA+TR}{TR+TA+FA+FR} \quad (5)$$

FAR represents the percentage of correct pronunciation when perception biases pronunciation is detected by similarity test, FRR represents the percentage of correct pronunciation when perception biases pronunciation is detected by similarity test, and DA represents the percentage that perception biases pronunciation is consistent with the result of similarity test.

Tab. 5 composite evaluation results

Tone	Tone1	Tone2	Tone3	Tone4
FAR	16.34%	8.76%	13.37%	43.13%
FRR	20.95%	40.02%	33.05%	11.15%
DA	81.25%	84.06%	79.38%	83.75%

The results of the composite evaluation showed that each tone had a higher diagnostic accuracy rate when combined with the perception test and the similarity test. However, tone2 and tone3 have a higher error rejection rate, which is

easy to detect the pronunciation as a false pronunciation, and tone4 has a higher error acceptance rate, and it is easy to detect the false pronunciation as the correct pronunciation. Tone1 has a lower error acceptance rate and false rejection rate, and has a higher correct diagnosis rate, which is relatively best. Tone2 is the best in detecting false, and it also achieves the highest correct diagnosis rate, but the error rejection rate is relatively high, and it is easy to judge the correct pronunciation as a false pronunciation. In comparison, tone3 has the worst diagnostic effect and a high false rejection rate, which is the worst effect. Tone4 is the easiest to detect false pronunciations as correct pronunciation,but works best on the detection of correct pronunciation.

IV CONCLUSION

This paper analyzes the mandarin spoken by college students in Amdo Tibetan Area through audition experiment and similarity test. Due to the influence of mother tongue and growing environment, students in Amdo Tibetan Region have increasingly difficult to master the tone when speaking the tonal language of mandarin. During the experiment, it was found that the mastery of tone2 and tone3 was still at a low level, and the tone could not be distinguished well.There is a widespread situation in which tones are made into the most easy to tone4. This paper proposes to use the tone similarity detection to detect the pronunciation biases. This paper proposes to use the tone similarity detection to detect the pronunciation biases. According to the test results obtained by the evaluation index, each tone has a higher correct diagnosis rate, but the tone2 and tone3 are easy to judge the correct pronunciation as a bias pronunciation, and tone4 is easily to detect the biases pronunciation as the correct pronunciation. The experimental results in this paper still have a lot of optimization space. One reason for the high error rate of tone2 and tone4 in detection is that the corpus is too small, resulting in poor results. One can reduce the bias rate by increasing the corpus, and increase the volatility.

REFERENCES

[1] Feifei Z. Experimental study on Botswana Elementary Chinese learners of the Tone Errors[D]. Shanghai Normal University,2018.

[2] Huan C. Experimental Analysis of Chinese Phonetic Errors of Tamil Learners[D]. Lanzhou University,2018.

[3] Difei W.A Study on the Acquisition of Intonation Among Intermediate Level Mandarin Learners of Turkish Languages [D]. Jinan University,2018.

[4] Jinjin Y. The Experimental Research on Mandarin Tone Acquisition of Chilean College Students [D]. Jinan University,2018.

[5] MUHAMMAD RIEZAL ADAM. Error Analysis on Learning Chinese Language Tone by Indonesian Student [D]. Nanchang University,2018.

[6] Yali L. The Relationship between Chinese Tone Perception and Vocal Errors in Xinjiang Nationality Students[C]// The 10th Youth Academic Conference of the Chinese Acoustics Society. 0.

[7] GULNUR Arkin, ASKAR Hamdulla. Tone analysis of non-native Chinese speakers based on rules and statistics[J]. Applied acoustics, 2018, v.37(03):63-70.

[8] Haifeng D. The Influence of Tibetan Ando Dialect Word Stress on the Acquisition of Mandarin Chinese Tone[D]. Minzu University of China, 2012.

[9] Shuping L. An Experimental Analysis of Tibetan (Anduo Dialect) Students Learning Chinese Monosyllabic Tone Errors[D]. Minzu University of China, 2006.

[10] Haifeng.Experimental Study on Tibetan Students' Speaking of Monosyllabic Tone System in Mandarin Chinese[D]. Minzu University of China, 2008.

[11] Gui Qin, LIU Ying, YU Ze. Roles of Lexical Tone and Neural Mechanisms Underlying Processing of Mandarin Speech Melody [J]. Progress in psychological science, 2012, 20(3):338-343.

[12] Rulong L. The effect of tone on the finals of the sound[J]. Language teaching and research, 1990(1):89-95.

[13] Lin. Speech research and teaching Chinese as a foreign language[J]. World Chinese teaching, 1996(3):20-23.

[14] Jianji L. Interlanguage Theory and Analysis of Foreigners' Pronunciation Errors in Learning Chinese[J]. Language teaching and research(03):44-56.

[15] Ru z. Research on Automatic Evaluation Methods of Mandarin Pronunciation[D]. Harbin Institute of Technology,2013.