Perception and Production of Mandarin Monosyllabic Tones by Amdo Tibetan College Students

Zhenye Gan[1,2(✉)], Jiafang Han[1,2], and Hongwu Yang[1,2]

[1] College of Physics and Electronic Engineering, Northwest Normal University, Lanzhou 73000, China
ganzy@nwnu.edu.cn
[2] Engineering Research Center of Gansu Province for Intelligent Information Technology and Application, Lanzhou 73000, China

Abstract. The purpose of the work is to research the error patterns of production and perception of Mandarin monosyllabic tone by college students from Amdo Tibetan agricultural and pastoral areas, and make the analysis of the causes of acoustics in both errors. We do the work through the two experiments of perception and production of tone. We use the methods of combining the speech engineering and experimental phonetics. Results show that the error rate of tone perception is highly correlated [r = 0.92] with that of tone production. The level of Mandarin in Amdo Tibetan agricultural area is higher than that in pastoral area both in terms of tone perception and production. The hierarchy of difficulty for the four grades in agricultural area is as follows: sophomore > freshman > junior > senior, pastoral area is as follows: freshman > sophomore > junior > senior. The hierarchy of difficulty for the four tones both in agricultural and pastoral areas is as follows: Tone 2 > Tone 3 > Tone 1 > Tone 4. Tone 2 and 3 are most likely to be confused. There is no obvious tone shape bias of the four tones, but the tone domain is narrow and the location of the tone domain is lower than standard Mandarin both in agriculture and pastoral areas.

Keywords: Amdo tibetan college students · Agriculture and pastoral areas
Tonal production and perception

1 Introduction

Amdo Tibetan is different from Mandarin. The most significant difference is that Mandarin is a tonal language while Amdo Tibetan is not. Previous studies have found that the background of native language influenced the learning of the tone of the second language [1, 2]. Therefore, students in the Amdo Tibetan area have difficulty in the learning of Mandarin tone. Liu [3] had pointed out that tone was an important factor,

Foundation item: National Natural Science Foundation of China (Grant No. 61262055, 11664036).

M. Zhang et al. (Eds.): NLPCC 2018, LNAI 11108, pp. 16–26, 2018.
https://doi.org/10.1007/978-3-319-99495-6_2

which directly affected the level of Mandarin. Tone should be the focus of research in the process of learning Mandarin.

In recent years, studies of Mandarin tone have focused on the learning situation of international students [4–8]. There were relatively few studies on the learning of Mandarin tone by Tibetans, especially the Amdo Tibetan. Literatures [9–11] conducted related research on the analysis of Mandarin tone of Amdo Tibetans to summarize the characteristics of Mandarin monosyllables tone errors by Amdo Tibetans. Most of the studies focused on the research of tone pronunciation by the experimental phonetics and linguistic methods. And the selection of experimental personnel was limited to a certain area. However, learning a language was as important as hearing and pronunciation. Only byhearing correctly can you have a correct pronunciation [12]. For most of the college students from the Amdo Tibetan area, they used Tibetan as their primary language both in primary and middle school, supplemented by Mandarin. When they came to the University of Han nationality and entered into the whole Mandarin environment suddenly, which makes it difficult for them to communicate and learn in life. Therefore, it is necessary to start research from the university to get a more comprehensive law.□

This paper was based on the needs of the Mandarin learning by college students from Amdo Tibetan area. According to the different level of Mandarin in the internal of Amdo Tibetan area and the sudden change of the Mandarin environment of college students from the Amdo Tibetan area, the college students from Amdo Tibetan agriculture and pastoral areas were taken as research subjects, the research subjects were in the four grades of freshman to senior. Agricultural area refers to areas where crops are the main source of life while pastoral area refers to areas where livestock and cattle are the main sources of life. In this paper, we selected suitable experimental subjects through field investigations to research the tone perception and production of Mandarin by college students from Amdo Tibetan area. We used the methods of combining the speech engineering and experimental phonetics. The main works of the paper were as follows. Firstly, we make analysis of the perception bias of the Mandarin tone by college students from Amdo pastoral and agriculture areas. Secondly, we make analysis of the error patterns of production of Mandarin tone by college students from Amdo pastoral and agriculture areas. Thirdly, we make the analysis of the causes of acoustics in both two errors.

2 Perception of Mandarin Tone

2.1 Experimental Corpus

Literature [13] once pointed out that the tone acquisition of Central Asian students had more difficulties in continuous speech than that in a single word. Therefore, the corpus used in this experiment consists of 50 long sentences in Mandarin. Each long sentence consists of 4 short sentences. Each short sentence has 5 monosyllables. The tone of last syllables of each short sentence is Tone 1, Tone 2, Tone 3, and Tone 4. The experimental corpus is 200 monosyllables and distributed evenly among four tones. The corpus is covering all monosyllable structures basically.

2.2 Experimental Subjects

The experimental corpus of standard Mandarin was produced by a postgraduate, a woman of Han nationality. The subjects were college students from Amdo Tibetan area and none of them had difficulty in hearing and speaking.

2.3 Experimental Process

Step1, Standard Mandarin was recorded by the above-mentioned female, which was recorded with 16 kHz and 16 bit in a sound proof studio. We denoise and segmented the speech signals into perception speech signals.
Step2, we designed a questionnaire according to the perception speech signals.
Step3, the subjects were asked to listen to the speech signals and write out the most probable tone marks of the monosyllabic they had heard. The tone of the monosyllables would be marked as follows: T1, T2, T3 and T4, which represent the four tones.

A total of 185 questionnaires were distributed and we finally obtained134 valid data. We removed some results such as an outgoing result that is not in line with the general rule. The valid data included 60 in pastoral area and 74 in agriculture area. The error rate of each tone was separately counted. All data were tested for normal distribution and then analyzed.

2.4 Experimental Results

As shown in Figs. 1 and 2, it can be seen that the perceptual error rate of college students in pastoral area is higher than that in agriculture area. Tones 2 and 3 have higher error rates in the four tones. Tone 4 is best judged, which is related to the reason that the tone of Amdo Tibetan had only one habit tone.

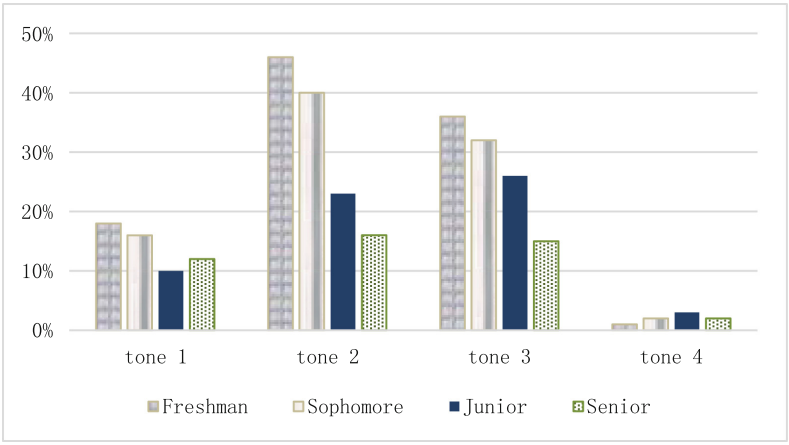

Fig. 1. The result of tone perception in pastoral area

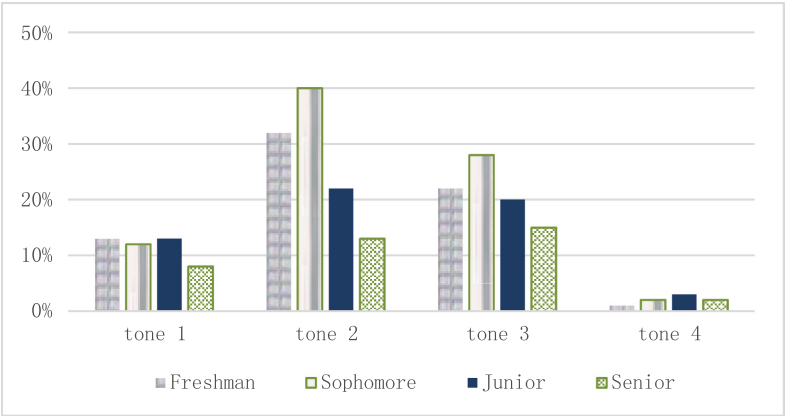

Fig. 2. The result of tone perception in agriculture area

It can be concluded that the hierarchy of difficulty for the four tones both in agriculture and pastoral areas is as follows: Tone 2 > Tone 3 > Tone 1 > Tone 4. The hierarchy of difficulty for the four grades in pastoral area is as follows: freshman > sophomore > junior > senior, agriculture area is as follows: sophomore > freshman > junior > senior. We find that the perception proficiency of freshman is better than sophomore in agriculture area which has stayed in school for one year. This is more relevant to that freshman has more chances to contact Chinese than sophomore in agriculture area. The progress of society is affecting their Mandarin proficiency. It will be further analyzed by using the follow experiment of tone production.

3 Production of Mandarin Tone

The experiment of tone production includes two parts: the tone annotation and the production of tone pronunciation experiment. We will first make tonal annotation experiment to select appropriate speech corpora. And then make further analysis of the error pattern of production in tone pronunciation.

Tone annotation is that several listeners judge the correct tone and the type of tone by subjects, and then decide the label of the tone. The tone of the monosyllables would be labeled as follows: T1, T2, T3 and T4, which represent the four tones.

3.1 Experimental Corpus

To make the result of the comparison between the tone perception and production correctly, the corpus used in this experiment is chosen from the perceptual corpus. It consists of 40 long sentences in Mandarin. There are a total of 160 monosyllables and distributed evenly among the four tones.

3.2 Experimental Subjects

Experimental subjects included 16 speakers of Amdo Tibetan. None had difficulty in hearing and speaking. The detailed information of the 16 speakers of Amdo Tibetan is as showed in the following Table 1. Through a large number of questionnaires, 16 speakers whose native language was Amdo Tibetan were selected. They had the same background in learning Mandarin.

Table 1. Experimental subjects information table

Area	Grade				
	Freshman	Sophomore	Junior	Senior	Total
Agriculture	2	2	2	2	8
Pastoral	2	2	2	2	8
Total	4	4	4	4	16

3.3 Experimental Process

In this experiment, 6 listeners of Han nationality were invited and didn't know the original correct tone. Each subject had 160 speech data. 6 listeners listened to the data of 16 speakers and judged the tone of the last word of each short sentence, and then made a mandatory tone judgment and labeled it as T1, T2, T3 andT4.

3.4 Experimental Results

The following Table 2 shows the result of comparing the results of the annotation in pastoral area with standard Mandarin. Table 3 shows the result of comparing agriculture areas with the standard mandarin. T1, T2, T3 and T4 represent the four tones.

Table 2. Results of annotation of Pastoral area

Correct tone	Tone annotation			
	Compare annotation result with correct tone			
	T 1	T 2	T 3	T 4
T 1	**75.0%**	7.5%	5.0%	5.0%
T 2	10.0%	**55.0%**	22.5%	5.0%
T 3	7.5%	27.5%	**65.0%**	5.0%
T 4	7.5%	10.0%	7.5%	**85.0%**

Comparing the results of the two groups, the accuracy of the Tone 4 is the highest both in the agriculture and pastoral areas and the lowest accuracy is Tone 2. We can also see the obvious differences from the comparison between the agriculture area and the pastoral area. The pronunciation accuracy of the four tones of the agriculture area is higher than that of the pastoral area. The overall accuracy rate is higher than 50%. From

Table 3. Results of annotation of agriculture area

Correct tone	Tone annotation			
	Compare annotation result with correct tone			
	T 1	T 2	T 3	T 4
T 1	**80.0%**	7.5%	2.5%	5.0%
T 2	10.0%	**67.5%**	17.5%	2.5%
T 3	7.5%	20.0%	**70.0%**	2.5%
T 4	2.5%	5.0%	10.0%	**90.0%**

Tables 2 and 3, we can also find that Tone 2 and 3 are most easily confused. It is probably because of the similar phonetic properties in Tone 2 and Tone 3, which both have a rising portion.

Later we will make further error analysis of the correct pronunciation signals in the production of Mandarin tone experiment.

3.5 Correlation Between Perception and Production

A comparison of perception and production is shown in Table 4. The error rate of tone perception is highly correlated [r = 0.92] with that of tone production. The formula of r is as follows (1):

$$r(x, y) = \frac{Cov(xy)}{\sqrt{D(X)D(y)}} \tag{1}$$

The formula of r is the cross-correlation function of the two sequences. Where x and y are the correct rates of the four tones in pastoral and agricultural areas (x: 75%, 55%, 65%, 85%; y: 80%, 67.5%, 70%, 90%).

Table 4. Overall error rate of perception and production

Experimental type	Tone			
	Error rate			
	T 1	T 2	T 3	T 4
Production	22.5%	28.75%	32.5%	12.5%
Perception	13%	29%	24%	2%

The tonal production error rate ranges of 12% to 33%, the perception error rates ranges of 2% to 29%. The level of tone perception is greater than that of production.

4 Production of Mandarin Tone Pronunciation

We will do an analysis of the error pattern production of tonal pronunciation with the speech signal of the correct annotation result from the preceding annotation experiment. As shown in Fig. 3, the experiment of production of Mandarin tone pronunciation is roughly divided into the following sections.

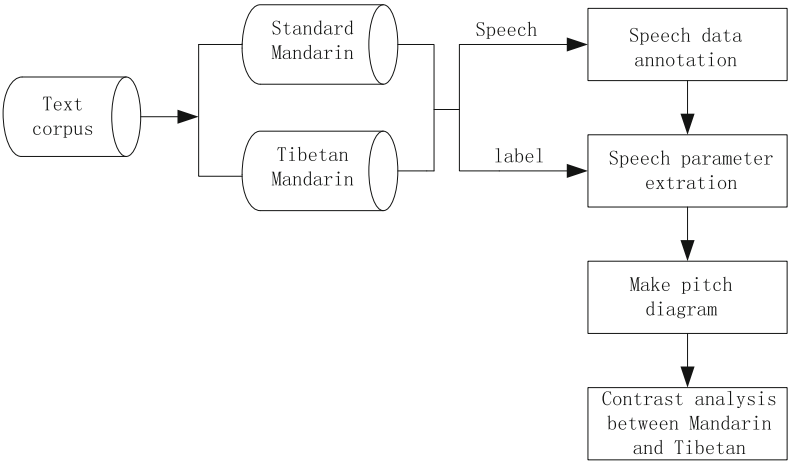

Fig. 3. The process of production of Mandarin Tone pronunciation

4.1 Experimental Corpus

A total of 2350 monosyllables were correctly pronounced according to the results of Tables 2 and 3. We classified the 2350 monosyllabic counterparts into 2 groups according the area of the subjects. According to the grade, speech data of 2 groups were divided into four groups separately. There were a total of 8 groups of speech data recorded by Amdo Tibetan, and with another group of standard Mandarin speech data recorded by native Mandarin speakers. There are a total of 9 groups of speech data.

4.2 Experimental Process

Step 1, the data of each group was divided into Tone 1, Tone 2, Tone 3, and Tone 4, which was used as the experimental speech signal of tone production.
Step2, we labeled the experimental speech signals by Visual Speech software, and then normalized the duration of syllables by straight algorithm to extract and modify the pitch [14].
Step 3, a program was written to realize the extraction of the maximum, minimum, and T value of the pitch data of each syllable obtained. The detailed approach was to divide the pitch data of each syllable into 10 equal parts by the duration. Each of which takes their average value to obtain the pitch data of the 10 average points that we need.

We used the formula proposed by [15] to normalize to convert the pitch data to T values, and then made pitch contours with T value. The formula of normalized processing was as follows (2):

$$T = 5 * \frac{\log_x - \log_L}{\log_H - \log_L} \quad (2)$$

Where H and L are the highest and lowest f0 of a group of all people's regulatory domains, and X is any given point of a pitch contour. The output of T is a value between 0 to 5.

The data obtained in the preceding step were statistically analyzed to obtain tone curves of 8 groups Mandarin monosyllables by Amdo Tibetan students. We made comparison between the tone curves of the speech data of 8 groups and compared the speech data of 8 groups with the standard Mandarin.

4.3 Experimental Results

As showed in Figs. 4 and 5, this is comparison graph of the monosyllable tone curves between standard Mandarin monosyllables and Mandarin by students from Amdo pastoral and agriculture areas. We make compares from the following aspects. Firstly, we make analyze between the four grades that freshman to senior. It is clear that compared with the standard Mandarin from the two figures, the monosyllable tone shape of the four grades both in Amdo pastoral and agriculture areas has obvious differences both in tone shape and tone domain. Secondly, we make analyze between the four tones that from Tone 1 to Tone 4. Tone 1, Tone 2, and Tone 3 of Mandarin monosyllable by students both in Amdo pastoral and agriculture areas have obvious differences in the tone shape and tone domain. It is still different from Mandarin in tone domain while the tone shape of Tone 4 is similar to standard Mandarin.

The result of tone production by college students from Amdo pastoral area. We compare the difference of the tone shape curves in Fig. 4 between Mandarin monosyllabic by college students from Amdo pastoral area and standard Mandarin.

Tone 1: The tone shape of freshman students from Amdo pastoral area shows a downward trend. The tone value is (23). The tone shape of junior shows a trend of decreasing first and then increasing, and the tone value is (34). The tone shape of sophomore and senior are similar to standard Mandarin. The location of the tone domain of the four grades is lower than standard Mandarin. In general, the Mandarin proficiency of senior and junior are better than sophomore and freshman.

Tone 2: In terms of tone shape, the shape tail of the tone by freshman, sophomore and senior are not rising enough and the middle of the tone shape is decreasing too much. The tone shape of junior is closest to standard Mandarin among the four grades. In terms of tone value, the tone value of sophomore and senior are closer to standard Mandarin than junior and freshman.

Tone 3: In terms of tone shape, the tone shape of the freshman, junior and senior are similar to standard Mandarin. The shape tail of the tone by freshman and senior are not rising enough. The height of the tone shape of junior at the beginning is not sufficient. The tone shape of sophomore is very different from standard Mandarin. The shape of

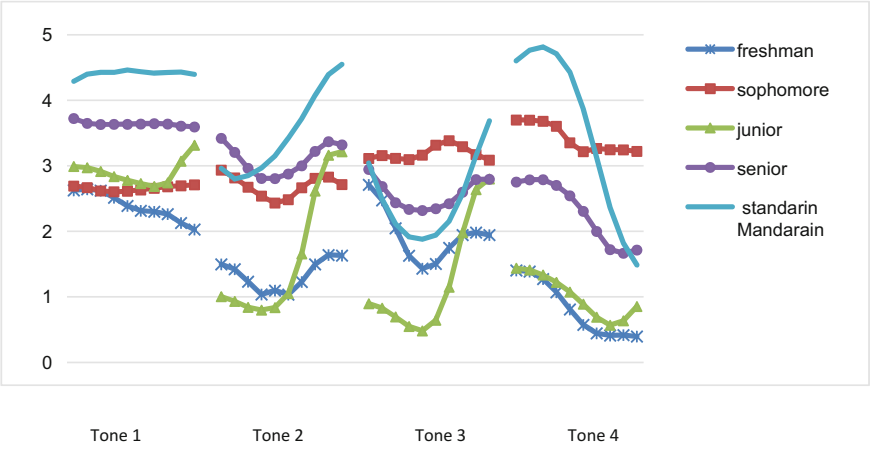

Fig. 4. Mandarin monosyllabic tone shape curves of students from Amdo pastoral area (Note: The ordinate is the normalized T value)

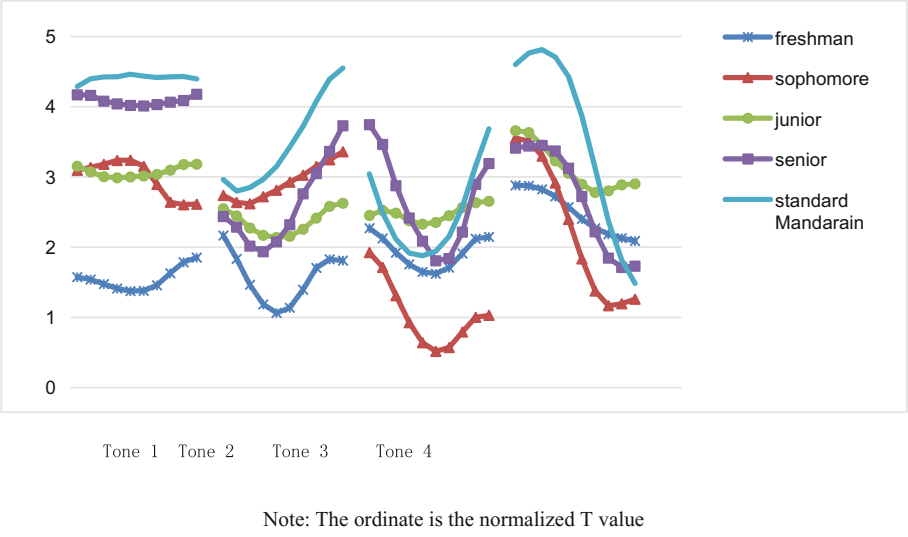

Fig. 5. Mandarin monosyllabic tone shape curves of students from Amdo agriculture area (Note: The ordinate is the normalized T value)

Tone 3 emphasizes the rising and falling lifting too much and ignores the need to reduce tone in the middle of the Tone 3, which is leading to a very strange shape.

Tone 4: The tone shape of the four grades from pastoral area is similar to standard Mandarin. But the tone domain is narrower than standard Mandarin. In general, the Mandarin proficiency of senior and sophomore are better than junior and freshman.

As above analyzed, the hierarchy of difficulty for the four grades from pastoral area is as follows: freshman > sophomore > junior > senior.

The result of tone production by college students from Amdo agriculture area. We compare the difference of the tone shape curves in Fig. 5 between Mandarin monosyllabic by college students from Amdo agriculture area and standard Mandarin.

Tone 1: The tone shape of freshmen shows a trend of decreasing first and then increasing, the tone value is (12), which is much lower than the standard Mandarin. The tone shape of sophomore shows a trend of decrease, the tone value is (34).The tone shape of junior and senior are similar to standard Mandarin. In general, the Mandarin proficiency of senior and Junior are better than sophomore and freshman.

Tone 2: The tone shape and tone domain of junior and senior are both similar to standard Mandarin. Senior and junior students have a higher level of Mandarin than freshman and sophomore. In general, the tone shape of Tone 2 is similar to the shape of Tone 3. It is probably because of the similar phonetic properties in Tone 2 and Tone 3, which both have a rising portion.

Tone 3: In terms of tone shape, the middle of the tone shape of junior is not decreasing enough. The tail of the tonal shape by freshman, sophomore is not rising enough. The tone shape of senior students is very close to standard Mandarin. In terms of tone value, the difference between sophomore and standard Mandarin is relatively big.

Tone 4: The tone shape of the four grades is both similar to standard Mandarin. The tone domain is a little narrower than the standard Mandarin. In general, the level of Tone 4 in the agriculture area is higher than the pastoral area.

In summary, it is evident that the hierarchy of difficulty for the four grades in pastoral area is as follows: freshman > sophomore > junior > senior, agriculture area is as follows: sophomore > freshman > junior > senior. We find that the Mandarin proficiency of students in pastoral area was increasing as the increasing of grade level, the poorest pronunciation of Mandarin tones by sophomore in agriculture area is due to the fact that the society is continuously improving. The freshman in agriculture area compared to pastoral area has more chances to contact Mandarin.

We compared Fig. 4 with Fig. 5, the pronunciation of college students in agriculture area is better than that in the pastoral area. It can be concluded that the hierarchy of difficulty for the four tones both in agriculture and pastoral areas is as follows: Tone 2 > Tone 3 > Tone 1 > Tone 4. In terms of tone domain, the agriculture area is obviously better than the pastoral area. The tone domain of the pastoral and agriculture area is narrow, and the location of the tone domain is much lower than standard Mandarin.

5 Conclusion

In this paper, the college students from the Amdo rural and pastoral areas were selected for experiment subjects. Two experiments were conducted on the error patterns of Mandarin tone perception and production by college students from Amdo Tibetan area.

From the above researches, it can be found that tone is a major difficulty for Amdo Tibetan students. And the confusion of Tone 2 and Tone 3 is a common error for

students from Amdo Tibetan area. It is probably because of the similar phonetic properties in Tone 2 and Tone 3, which both have a rising portion. The error rate of tone perception is highly correlated [r = 0.92] with that of tone production. Most of them can imitate the tone shape of the four tones, but the tone domain is difficult to imitate. It is related to that Amdo Tibetan is a non-tonal language. The Amdo Tibetan has only one habit tone, which makes it difficult for them to grasp the tone domain of the four tones correctly. In addition, there is a big difference between college students from the Amdo agriculture and pastoral areas on the Mandarin proficiency. The reason of the difference is that the economic in agriculture area is higher than that in pastoral area, and there are fewer people speak Mandarin in pastoral area than that in agriculture area. In addition to the negative transfer of their native language, the different economic development in the internal of the Amdo Tibetan area has a great influence on Mandarin learning.

References

1. Ding, H.: Perception and production of Mandarin disyllabic tones by German learners. Speech Prosody **184**, 378–381 (2012)
2. Weltens, B.: Language attrition in progress. Language **66**(3), 37–49 (1986)
3. Yali, L.: Perception and production of Mandarin tones of pupils from Xinjiang ethnic areas. J. Tsinghua Univ. **53**(6), 823–827 (2013)
4. Gao, M.: Perception and production of mandarin tones by swedish students. Iran. J. Radiol. **29**(4), 533–538 (2010)
5. Hoffmann, C.: Disyllabic Mandarin lexical tone perception by native Dutch speakers: a case of adult perceptual asymmetry. J. Acoust. Soc. Am. **134**(5), 4230 (2013)
6. Zhang, J.: Influences of vowels on the perception of nasal codas in Mandarin for Japanese and Chinese natives. J. Tsinghua Univ. **57**(2), 164–169 (2017)
7. Lijuan, G.: Tone production in Mandarin Chinese by American students. In:Chan, K.M., Kang, H. (eds.) Conference 2008, NACCL, Ohio, vol. 9999, pp. 123–138 (2008)
8. Ding, H.: An investigation of tone perception and production in German learners of Mandarin. Arch. Acoust. **36**(3), 509–518 (2011)
9. Haifeng, D.: A Study of single Chinese characters of Tibetan learners of Amdo Tibetan. J. Lang. Lit. Stud. **14**, 17–18 (2011)
10. Yan, L.: Acoustic research on the monosyllabic pitch pattern of Amdo-Mandarin Inter language and Tibetan Amdo Dialect. Hanzangyu Xuebao **1**, 51–63 (2015)
11. Wu Yong, F.: A study of International Error on Putonghua Inter language of Tibetan language Anduo Dialect Region. J. Qinghai Nationalities Univ. (Soc. Sci.) **38**(02), 135–140 (2012)
12. Yali, L.: The train methods of perception of Mandarin tones of pupils in Xinjiang ethnic areas. Tech. Acoust. **33**(s2), 328–330 (2014)
13. Wenbo, C.: Analysis of Chinese phonetic teaching for students in Central Asia. Lang. Transl. **3**, 73–76 (2009)
14. Irino, T.: Speech segregation using an auditory vocoder with event- synchronous enhancements. IEEE Trans. Speech Audio Process. **14**(6), 2212–2221 (2006)
15. Feng, S.: Exploration of Experimental Phonology. Beijing Peking University Press, Beijing (2009)