Research on the Lip Shape Characteristics of the Vowels in Tibetan Xiahe Dialect

Shengyin Zhu[1,a]

[1] College of Chinese Language and Literature of Northwest Normal University Lanzhou Gansu China
[a] Email: 87961192@qq.com

ABSTRACT
This paper studies the characteristics of the dynamic lip shape of vowels in the Tibetan Xiahe dialect, and illustrates some relationships between vowels through experiments. The experimental results show that the value of the lip width is large when flattened vowels are pronounced, while it is smaller when rounded vowels are pronounced. The lip height is an important indicator to distinguish each vowel. The lip protrusion is prominent in distinguishing rounded vowels. When flattened vowels are pronounced, the protrusion of both lips is smaller than the natural state while the protrusion of the lips is larger than the natural state when rounded vowels are pronounced. The value of the mouth corner stretch and the protrusion value of the lower lip have a certain similarity. If the protrusion of the lower lip cannot be obtained, the value of the mouth corner stretch can be used to judge the changes of the lower lip protrusion. In addition, the experimental results show that it is feasible to collect the dynamic three-dimensional coordinate data of the lip shape of the speakers with the help of high-precision equipment to study the dynamic characteristics of the lip shape in linguistics.

Keywords: *Tibetan Xiahe dialect, vowels, lip shape features*

1. INTRODUCTION

The change of tongue position and lip shape is an important indicator to distinguish different vowels, and it is also a basic problem in phonetics and phonology research. Many scholars have given certain standards for the division of vowels from different angles. A.M. Bell proposed a set of criteria for dividing vowels in Visible Speech (1867), using the height, the front and the rear of the tongue position to determine vowels. According to the height, it is divided into high, medium and low, and according to the front and the rear, it is divided into front, mixed and rear[1]. In addition, Bell added two other factors that describe the quality of vowels, namely rounded, flattened and pharyngeal lip width. Linker used the factor analysis program (PARAF AC) in 1982 to analyze the lip shape characteristics of English, Chinese (Cantonese), French and so on, and pointed out that the horizontal opening of the lips is the lip flattening, and the ratio of the protrusion and the vertical opening is the lip roundness. Bao Huaiqiao and Lin Maocan collected the lip shape data of five speakers in Summary of Experimental Phonetics, and analyzed it by comparing the average value of different speakers when they pronounce a certain vowel.[2]

Although scholars have given many conclusions through different experimental methods, due to the limitation of experimental conditions, the dynamic changes of lip shape during the pronunciation process cannot be extracted, and the accuracy of the parameters of the collected lip shape is also affected by experimental equipment and algorithms[3]. With the rapid development of science and technology today, we can use digital technology to extract more precise lip shape parameters, so this paper will use the world's most advanced facial motion capture system to collect the feature changes of the dynamic lip shape of the speaker during the pronunciation process. The purpose is to further explain the characteristics of the lip shape on the basis of the existing parameters, and to compare and analyze the different viseme indexes of each vowel. The research results can provide reference for scholars in the field of Tibetan synthesis, and also help to solve the pronunciation problem of Tibetan learners.

2. EXPERIMENTAL PROCEDURE

2.1. Experimental materials

Tibetan Xiahe dialect has a total of 8 vowels, of which 6 are commonly used, namely /i/, /e/, /a/, /ə/, /o/, /u/, and

/ɛ/ and /ɔ/ has a low frequency in Xiahe dialect[4][5], so this paper only compares and analyzes the characteristics of the dynamic lip shape of the six commonly used vowels.

Table 1 is the experimental vocabulary of Tibetan Xiahe vowels. Vowel /i/ corresponds to Tibetan གཉིས (two), pronounced /hɲi/; vowel /e/ corresponds to Tibetan རས (cloth), pronounced /re/; vowel /a/ corresponds to Tibetan ཨ (ah!), pronounced /_a/; vowel /ə/ corresponds to Tibetan ཤུ, pronounced /_ə/; vowel /o/ corresponds to Tibetan ཨོ (oh!), pronounced /_o/; vowel /u/ corresponds to Tibetan དངུལ (silver), pronounced /hɲu/.

Table 1. The experimental vocabulary of Tibetan Xiahe vowels

No	1	2	3	4	5	6
Vowels	/i/	/e/	/a/	/ə/	/o/	/u/
Tibetan	གཉིས	རས	ཨ	ཤུ	ཨོ	དངུལ
IPA	/hɲi/	/re/	/ʔa/	/_ə/	/_o/	/hɲu/
Lip shape	Flattened lips	Flattened lips	Flattened lips	Flattened lips	Rounded lips	Rounded lips
Voice Duration	2.55s	2.15s	2.2s	1.8s	2.2s	2.08s

2.2. The extraction of the lip parameter

The experimental equipment is from the Key Laboratory for Ethnic Minority Language Intelligent Processing, which has a professional motion capture field, top-level equipment at home and abroad and an excellent management team. All data collection and processing are completed in this laboratory.

In the study, seven indicators are mainly explored, including the width of the inner and outer lips, the height of the inner and outer lips, the protrusion of the upper and lower lips, and the stretch of the corners of the mouth. From the collected experimental data, it can be found that the width of the inner and outer lips is similar, so is the height of the inner and outer lips, therefore the width of the inner lip and the height of the inner lip are only discussed as reference indicators. This article will focus on analyzing the values of outer lip width, outer lip height, upper and lower lip protrusion, and mouth corner stretch.

Figure 1 is about the annotation of the feature points of collected facial dynamic data. This annotation method is annotated according to the MPEG-4 (Moving Picture Experts Group) international standard. The outer lip width and outer lip height can be directly extracted according to the annotated points. For the upper lip protrusion and mouth corner stretch, a position needs defining to obtain the relative value. This method is similar to the one with which Wang Anhong (Wang Anhong 2000) obtained the lip protrusion value when dealing with Mandarin speech viseme[6]. However, using this method must ensure that the feature coordinate point of the upper lip protrusion and mouth corner stretch is on the same rigid body as the defined position point. The protrusion of the lower lip is difficult to obtain because it is related to the mandible, and will be rotated and translated by the movement of the mandible during pronunciation. The parameter extraction of the lower lip protrusion in this paper is consistent with the method involved in the study of the three-dimensional viseme features of Tibetan Lhasa dialect by He Xiangzhen (He Xiangzhen 2016) [7] [8].

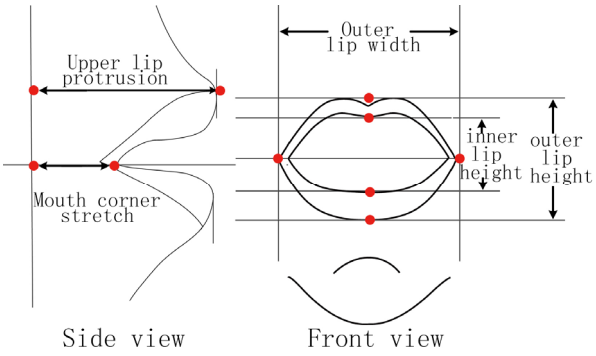

Figure 1 The parameter calibration

Table 2 is the target viseme values of three speakers in different states. As for extracting the outer lip width and outer lip height, the distance of the two points can be calculated directly from the three-dimensional coordinate data. The relative value of the upper lip protrusion and mouth corner stretch is collected according to the defined point. The relative value of the lower lip protrusion is also obtained by calculating the rotation and translation of the mandible. [9] [10] [11]

.

Table 2. Target viseme values of three speakers in different states(unit:cm)

Speaker	Lip shape	Outer lip width	Outer lip height	Upper lip protrusion	Lower lip protrusion	Mouth corner stretch
1#	Natural state	5.7	2.51	3.68	3.7	2.22
	Half open	5.56	3.45	3.6	3.07	1.65
	Full open	5.55	3.83	3.98	3.81	2.52
2#	Natural state	5.9	1.89	3.68	3.84	2.25
	Half open	5.54	3.38	3.59	3.5	1.98
	Full open	5.61	3.48	4.26	4.11	2.62
3#	Natural state	5.68	2.1	3.65	3.67	2.18
	Half open	5.62	2.91	3.61	3.12	1.73
	Full open	5.51	3.79	3.93	3.78	2.49

2.3. The extraction of the lip parameter

Due to the different physiological characteristics of different speakers, there are certain differences in the parameters obtained from the speakers, and the movements and habits of pronunciation are also different. Therefore, the index parameters have individual physiological characteristics when different vowels are pronounced. So the lip shape parameters of different speakers cannot be compared with each other. However, after normalization, similarity analysis can be done.

This paper conducts statistical analysis on the three-dimensional coordinate samples of the three speakers respectively, makes normalization processing of the obtained lip shape data, and then compares the dynamic data of the three speakers after the normalization. By doing so, the commonality of pronunciation habits can be discovered. Combined with the indicators and parameters of the three speakers and the laws of the pronunciation stages, the dynamic features of the lip shape suitable for describing different vowels are proposed.

3. EXPERIMENTAL RESULTS AND ANALYSIS

3.1. Comparison of the lip width

Figure 2 is about the comparison of the outer lip width of vowels in Tibetan Xiahe dialect. According to the experimental results, the average width of the vowel /i/ is larger than the natural state, and the average width of the rest of the vowels is smaller than the natural state, indicating that the lip opening is smaller when the vowel /i/ is pronounced. The average lip width of the vowel /i/ is 5.72 cm, which is 0.02 cm wider than the natural state value. The average outer lip widths of vowels /e/, /a/ and /ə/ are 5.64cm, 5.62cm and 5.63cm respectively, which are all smaller than the natural state value. The average lip widths of the rounded vowels /o/ and /u/ are 5.45cm and 5.49cm, both smaller than the natural state by more than 0.2cm. After comparing the data, it is found that the width of the outer lip changes little when flattened vowels are pronounced, and the width of the outer lip becomes narrower when rounded vowels are pronounced. In his research on rounded vowels in Mandarin, Pan Xiaosheng believed that the physiological movements of the lips in Mandarin pronunciation have certain randomness, but basically the same speaker can always maintain the characteristic that the lip width of rounded vowels is narrower than that of non-rounded vowels. Through analyzing the data from this experiment, it is found that the Tibetan Xiahe dialect also has the same characteristic.

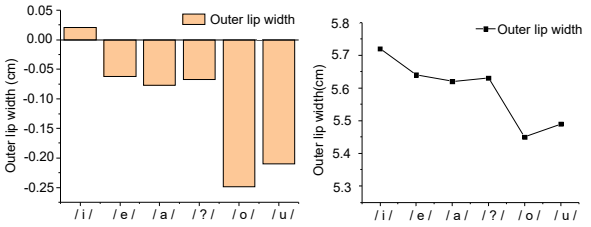

Figure 2 Comparison of the outer lip width of vowels in Tibetan Xiahe dialect

3.2. Comparison of the lip height

Figure 3 is about the comparison of inner and outer lip heights of vowels in Tibetan Xiahe dialect. The maximum value of the vowel /i/ is 3.87cm, and the average value is 3.13cm. The maximum value of the vowel /e/ is 3.69cm and the average is 3.08cm. The maximum value of the vowel /a/ is 3.83cm and the average is 2.96cm. The maximum value of the vowel /ə/ is 4.29cm and the average is 3.3cm. In the vowel experiment, although the vowel /ə/ has a larger opening of the upper and lower lips, the opening of the upper and lower teeth is smaller. The maximum value of the vowel /o/ is 3.27cm and the average is 2.93cm. The maximum value of the vowel /u/ is 3.38cm, and the average value is 2.96cm. According to the data, the vowel /ə/ has a larger opening at the target position, while the vowel /o/ has a

smaller opening at the target position. The maximum and average values of rounded vowels /o/ and /u/ are the closest, indicating that the change of the lip height is relatively stable during pronunciation. In terms of the average, the outer lip height of the vowel /ə/ is the largest, and that of the vowel /o/ is the smallest. Through the comparison of the data of inner and outer lip height, it is found that the inner and outer lip height is an important indicator to distinguish each vowel.

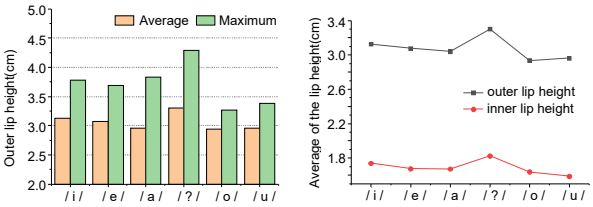

Figure 3 Comparison of the inner and outer lip heights of vowels in Tibetan Xiahe dialect

3.3. Comparison of the lip protrusion

Figure 4 is about the comparison of the protrusion of the lips. The average value of the lip protrusion of the vowels /i/ and /ə/ are smaller than the natural state, and the protrusion value of the upper and lower lips of the vowel /i/ are 3.66cm and 3.69cm. The protrusion values of the upper and lower lips of the vowel /ə/ are 3.66cm and 3.53cm. According to the data, the protrusion value of /ə/ is the smallest among all vowels. The protrusion value of the upper lip of vowels /e/ and /a/ is larger than that of the natural state, and the protrusion value of the lower lip of vowels /e/ and /a/ is smaller than that of the natural state. The protrusion values of the upper and lower lips of the vowel /e/ are 3.72 cm and 3.69 cm respectively. The protrusion values of the upper and lower lip of the vowel /a/ are 3.74 cm and 3.63 cm respectively. The rounded vowels /o/ and /u/ both have a larger lip protrusion than the natural state. The protrusion values of the upper and lower lip of the vowel /o/ are 3.83cm and 3.73cm respectively. The protrusion values of the upper and lower lip of the vowel /u/ are 3.83cm and 3.75cm, respectively. From the comparison of the lip protrusion parameters, when rounded vowels are pronounced, the protrusion values of the upper and lower lip are larger than the natural state, and the upper lip protrusion varies more than the lower lip protrusion.

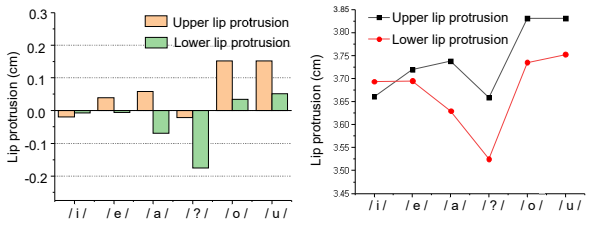

Figure 4 Comparison of the lip protrusion

3.4. Comparison of the stretch of the corners of the mouth

Figure 5 shows the stretch of the corners of the mouth. The changing process of the stretch of the mouth corners is complex, which is not only related to the changes of the lips, but also affected by the movement of the mandible. On the whole, the stretch changes of flattened vowels are smaller than the natural state value, while the stretch changes of rounded vowels are larger than the natural state value. The average stretch value of the vowel /i/ is 2.14 cm, with a minimum value of 2.05 cm. The average stretch value of the vowel /e/ is 2.1 cm, with a minimum value of 2.01 cm. The average stretch value of the vowel /a/ is 2.17 cm, with a minimum value of 1.96 cm. The vowel /ə/ has an average stretch of 1.99cm and a minimum of 1.65cm, and the vowel /ə/ has the smallest stretch of all vowels. The average stretch of the vowel /o/ is 2.37 cm with a maximum of 2.56 cm. The average stretch of the vowel /u/ is 2.43 cm, with a maximum of 2.62 cm, and the vowel /u/ has the largest stretch of all vowels. According to the comparison of the parameters of the mouth corner stretch, when the vowels /o/ and /u/ are pronounced, the value of the forward movement of the mouth corner is larger than the natural state value. On the whole, the stretch degree of rounded vowels is significantly larger than that of flattened vowels, so the stretch of the corners of the mouth is of great significance in distinguishing rounded and flattened vowels. (Zhu Shengyin 2020)Zhu Shengyin, in his research on the characteristics of the lip shape in Tibetan Amdo dialect based on facial motion capture, believes that the stretch of the corners of the mouth is similar to the protrusion of the upper and lower lips. When there is no way to obtain the value of the lip protrusion, the stretch of the corners of the mouth can well reflect the changes of the lip protrusion.[12]

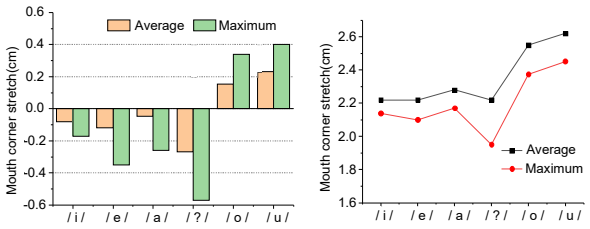

Figure 5 Comparison of the stretch of the corners of the mouth

4. CONCLUSION

In this paper, through the research method of experimental phonetics, the dynamic three-dimensional coordinate parameters of the speaker's lips are collected with the help of high-precision experimental equipment. The experimental results also show that the parameters clearly reflect the differences between different vowels when vowels are pronounced. Especially in the

comparison of the lip protrusion, it is found that the protrusion values of the upper and lower lip are larger than the natural state value when rounded vowels are pronounced, and the upper lip protrusion has greater change than the lower lip protrusion. Therefore, the protrusion of the lips is of great significance in distinguishing rounded and flattened vowels. This experiment compares and analyzes the differences of lip shape characteristics among vowels through various indicators, and achieves ideal results. It shows that it is feasible to use this method to study the dynamic changes of lip shape in linguistics.

This paper is the phased achievement of the Gansu provincial social science planning project " Research on Lip Shape Dynamic Viseme of Vowels in Tibetan Xiahe Dialect and the Database Construction" (2021YB045). It is also supported by the project "Research on Dynamic Viseme of Vowels in Tibetan Xiahe Dialect" for the improvement of young teachers' scientific research ability in Northwest Normal University.

REFERENCES

[1] Bao Huaiqiao, Lin Maocan. Outline of Experimental Phonetics (Updated Edition)[M], Beijing:Peking University Press, 2014.

[2] Pan Xiaosheng, Kong Jiangping. Research on the round lips in Mandarin [J]. Linguistics, 2014.

[3] Pan Xiaosheng. Research on lip co-articulation and visual speech perception in Mandarin Chinese [D]. Peking University, 2011.

[4] Ma Ning. Acoustics of intra-syllable co-articulation in Tibetan Xiahe dialect [D]. Doctoral dissertation of Northwest University for Nationalities, 2013.

[5] Yu Hongzhi, Li Yonghong, etc. Tibetan Dialect Survey [M]. Beijing: Commercial Press, 2011.

[6] Wang Anhong. A Preliminary Study on the Viseme System of Mandarin Chinese [D]. Beijing Language and Culture University, Beijing Language and Culture University, 2000.

[7] He Xiangzhen. Research on the Three Visual Features of Tibetan Lhasa Dialect [D]. Doctoral dissertation of Northwest University for Nationalities 2016.

[8] Wu Zongji, Lin Maocan. Outline of Experimental Phonetics (Updated Edition) (21st Century Chinese Language Professional Planning Textbook) [M]. Peking University, 2014.

[9] Kong Jiangping. Basic Course of Experimental Phonetics [M]. Peking University Press, 2015.

[10] Li Yonghong. Analysis of Acoustic Characteristics of Amdo Tibetan Intonation [D], Master Thesis of Northwest University for Nationalities, 2006.

[11] Wang Shuangcheng, The Complex Vowels and Finals of Amdo Tibetan [J], Tibet Studies 2005 (03).

[12] Zhu Shengyin. A Study on the Lip Shape Characteristics of Tibetan Amdo Dialect Based on Facial Motion Capture [D]. Northwest University for Nationalities Doctoral Dissertation 2020.