A Voice Analysis on the Voiced Consonants of Tibetan Lhasa Based on Mdvp

Mingxiao Gu[a], Dawa Pengcuo[b, *]

Key Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Gansu 730030, China

[a]Gmx_Deity@163.Com, [b]Dwpc@163.Com

*Corresponding Author

Keywords: Mdvp, Tibetan Lhasa, Voiced Consonants, Voice

Abstract: This Paper Uses the Multi-Dimensional Voice Program (Mdvp) and Statistical Methods to Extract Parameters from Eight Voiced Consonants ([m], [n], [ng], [Ny], [l], [W], [r], [J]) in Tibetan Lhasa. These Parameters Are Used to Analyze the Difference between Different Genders. the Data Are from 6 Categories and 33 Parameters, These Parameters Reflect the Voice Characteristics of Voiced Consonants in Tibetan Lhasa. the Study Found That the Mean and Standard Deviation of the Fundamental Frequency, Frequency Jitter, Amplitude Jitter and Other Parameters Are Quite Different between Men and Women. While the Parameters Such as Sample Duration and Pitch Perturbation Are Not Significantly Different in Genders. This Study Has Positive Significance for the Pronunciation of Tibetan Lhasa and Voice Therapy.

1. Introduction

With the Progress of Technology and the Emergence of Advanced Instruments, Experimental Phonetics Has Become a Hotspot in the Study of Phonetics. But Most of the Phonetics Research in Tibetan Language Still Stays in the Traditional Phonetics Method. Due to the Difference between Tibetan and Chinese Pronunciation Systems, Their Phoneme Structure is Also Significantly Different. in Addition, Due to the Particularity of the Tibetan Living Environment, the Voice Quality of Tibetan Are Also Very Different from Chinese. Therefore, the Chinese Voice Model Cannot Be Directly Applied to the Study of Tibetan Pronunciation. Therefore, It is Necessary to Study the Pronunciation of Tibetan Lhasa.

There are 30 consonants in Tibetan Lhasa, ther are [g], [k], [ng], [j], [q], [ny], [d], [t], [n], [b], [p], [m], [z], [c], [w], [x], [s], [y], [r], [l], [h], [hy], [sh], [f], [lh], [a], [gy], [ky], [zh], [ch] (ཀ་ཁ་ག་ང་། ཅ་ཆ་ཇ་ཉ། ཏ་ཐ་ད་ན། པ་ཕ་བ ་མ། ཙ་ཚ་ཛ་ཝ། ཞ་ཟ་འ་ཡ། ར་ལ་ཤ་ས། ཧ་ཨ།)[1]. This classification is according to the Tibetan-Chinese Lhasa Spoken Language Dictionary. If the consonants of Tibetan Lhasa are classified from the turbidity angle, most of the consonants are clear, only eight consonants ([m], [n], [ng], [ny], [l], [w], [r], [j]) are voiced consonants[2]. Since the clear consonant cannot make vocal cords vibrate, it is impossible to obtain functional voice parameters for analysis. But the vibration of the voiced consonants is very clear compared to the pure consonants, and sufficient acoustic voice parameters can be obtained for calculation and analysis[3]. Therefore, the Multi-Dimensional Voice Program (MDVP) is mainly used in this paper to analyze the voiced consonants of Tibetan Lhasa dialect from the perspective of gender[4].

2. Experimental Plan

2.1 Data Collection

The text recorder is 20 college students who are native speakers of Lhasa, 10 men and 10 women, they don't have speech disorder. All signal acquisition work in this experiment was done in a professional voice recording studio. The recording software is Adobe Audition 3.0, with a sampling rate of 44100 Hz and a resolution of 16 bits (according to the condition that MDVP can handle[5]). Voice signal acquisition takes place in a quiet room, the ambient noise is below 45dB.The recorder

  DOI: 10.25236/soshu.2020.009

wears a lavalier microphone, and his (or her) mouth is about 15cm from the microphone. They all take natural comfort, and then smoothly speak the eight voiced consonants ([m], [n], [ng], [ny], [l], [w], [r], [j]). The recorders should stretch the sound to make the consonants last for at least 3 seconds, and read each consonant three times. Finally, we stored all the sounds as voice files in the "*.wav" format.

2.2 Parameter Extraction

The eight voiced consonants of Tibetan Lhasa were extracted by the Kay Company's Multi-Dimensional Voice Program (MDVP)[6]. Multi-Dimensional Voice Program (MDVP) can extract 33 acoustic parameters, which can be divided into 6 categories according to their nature. The six categories of parameters show as following: the first category is "Pitch Basic Parameters", including: (1) Average Fundamental Frequency (F0. Hz), (2) Average Pitch Period (To. Ms), (3) Highest Fundamental Frequency (Fhi. Hz), (4) Lowest Fundamental Frequency (Flo. Hz), (5) Standard Deviation of F0 (STD. Hz), (6) Phonatory F0-Range in semi-tones (PFR); The second category is the "Frequency Jitter Parameter", including: (7) F0-Tremor Frequency (Fftr. Hz), (8) Amplitude Tremor Frequency (Fatr. Hz), (9) Length of Analyzed Sample (Tsam. s), (l0) Absolute Jitter (Jita. Us), (11) Jitter Percent (Jiit. %), (12) Relative Average Perturbation (RAP. %), (13) Pitch Perturbation Quotient (PPQ. %), (14) Smoothed Pitch Perturbation Quotient (sPPQ. %), (15) Fundamental Frequency Variation (vF0. %); The third category is the "Amplitude Jitter Parameter", which includes: (16) Shimmer in dB (ShdB. dB), (17) Shimmer Percent (Shim. dB), (18) Amplitude Perturbation Quotient (APQ. %), (19) Smoothed Ampl. Perturbation Quotient (sAPQ. %), (20) Peak-to-Peak Amplitude Variation (vAm. %); The fourth category is the "Sound Index", which includes: (21) Noise to Harmonic Ratio (NHR), (22) Voice Turbulence Index (VTI), (23) Soft Phonation Index (SPI), (24) F0-Tremor Intensity Index (FTRI. %), (25) Amplitude Tremor Intensity Index (ATRI%); The fifth category is "Voice Clearing Parameters", including: (26) Degree of Voice Breaks (DVB), (27) Degree of Sub-harmonics (DSH), (28) Degree of Voiceless (DUV), (29) Number of Voice Breaks (NVB), (30) Number of Sub-harmonic Segments (NSH), (31) Number of Unvoiced Segments (NUV); The sixth category is the "basic parameters", including: (32) Number of Segments Computed (SEG), (33) Total Number Detected Pitch Periods (PER)[7].

3. Experimental Analysis

Using the Multi-Dimensional Voice Program (MDVP), 33 parameters of the 8 voiced consonants in Tibetan Lhasa dialect were extracted, and the average and standard deviation of the men and women were calculated using these data[8]. As shown in Table 1 - Table 6:

3.1 Analysis of the Fundamental Parameters

Table 1 Pitch Basic Parameter Analysis Results ($\bar{X}\pm s$).

Gender	Number	F0(Hz)	T0(ms)	Fhi(Hz)	Flo(Hz)	STD(Hz)	PFR
Female	10	221.83±8.91	4.53±0.19	263.77±30.37	179.8±16.15	13.55±1.59	7.4±1.36
Male	10	148.8±11.31	6.86±0.58	201.85±26.83	118.09±14.16	15.77±5.7	9.67±0.47

From Table 1, we can see that:

(1) The average value of the female's F0 in the Lhasa Tibetan language is 221.83 Hz, the male is 148.8 Hz, the female is much higher than the male. The standard deviation is 8.9 Hz for female and 11.31 Hz for male.

(2) In the average data of Highest Fundamental Frequency (Fhi) and Lowest Fundamental Frequency (Flo), the female is 263.77 Hz and 179.80 Hz, respectively, the male is 201.85 Hz and 118.09 Hz. From these data, we can see that the female is all higher than the male. In terms of standard deviation, female students are 30.36Hz and 16.15Hz respectively, and male students are 26.83Hz and 14.16Hz respectively, there are little difference between male and female.

(3) The average value of Phonatory F0-Range in semi-tones (PFR) is 7.4 for female and 9.67 for male, the difference is not very large. While the standard deviation is 1.35 for female, and 0.47 for male. Female is more than 3 times higher than male.

On the whole, on the average, the biggest difference between the data of male and female is Average Fundamental Frequency (F0), Highest Fundamental Frequency (Fhi), and Lowest Fundamental Frequency (Flo). On the standard deviation, the largest difference between the genders is Average Fundamental Frequency (F0), Standard Deviation of F0 (STD), Phonatory F0-Range in semi-tones (PFR).

3.2 Analysis of Jitter Parameters

Table 2 Frequency Jitter Parameter Analysis Results ($\bar{X}\pm s$).

Gender	Number	Fftr (Hz)	Fatr (ms)	Tsam (Hz)	Jita (Hz)	Jitt (Hz)	RAP (%)	PPQ (%)	sPPQ (%)	vF0 (%)
Female	10	0	2.88± 2.53	1.59± 0.3	85.55± 17.98	1.88± 0.34	1.07± 0.21	1.11± 0.24	0.8± 0.11	6.09± 0.55
Male	10	0	3.22± 2.68	1.51± 0.22	145.85± 33.63	2.15± 0.56	1.19± 0.41	1.1± 0.23	2.42± 1.16	10.93± 4.7

Table 3 Amplitude Jitter Parameter Analysis Results ($\bar{X}\pm s$).

Gender	Number	ShdB(Hz)	Shim(ms)	APQ(Hz)	sAPQ(Hz)	vAm(%)
Female	10	0.51±0.07	5.19±0.88	4.12±0.63	6.91±0.86 26.4±5.1	
Male 10	0.71±0.13	7.51±1.63	7±2.2	19.74±17.06	35.64±15.27	

The jitter parameters consist of two broad categories, one is the frequency jitter parameter in Table 2, and the other is the amplitude jitter parameter in Table 3. From the above two tables we can see:

(1) The parameter data of Absolute Jitter (Jita), Smoothed Pitch Perturbation Quotient (sPPQ), Fundamental Frequency Variation (vF0), Smoothed Ampl. Perturbation Quotient (sAPQ), Amplitude Perturbation Quotient (APQ), and Peak-to-Peak Amplitude Variation (vAm) have positive related to Total Number Detected Pitch Periods (PER). The average values of Jita, sPPQ, vF0, sAPQ, Shim, APQ and vAm in female are 85.55 Hz, 0.80%, 6.09%, 6.91 Hz, 5.19 ms, 4.12 Hz, 26.40%, and the male are 145.85 Hz, respectively. 2.42%, 10.93%, 19.74Hz, 7.51ms, 7.00Hz, 35.64%. All the values of male are higher than female. As for the standard deviation, the male is also higher than female. The standard deviation data for female is 17.98Hz, 0.11%, 0.55%, 0.86 Hz, 0.88ms, 0.62Hz, 5.09%, and female are 33.63Hz, 1.16%, 4.70%, 17.06Hz, 1.63ms, 2.20Hz, 15.27%.

(2) The parameter data of Length of Analyzed Sample (Tsam) and Pitch Perturbation Quotient (PPQ) were analyzed. There is no significant difference in the parameter data between the mean and standard deviation. The average value of female is 1.59 Hz and 1.11%, male are 1.51Hz and 1.10%. And the standard deviation values of female are 0.30Hz and 0.24%, male are 0.22Hz and 0.23%.

On the whole, on the average, the biggest changes in data for both male and female are Absolute Jitter (Jita), Smoothed Pitch Perturbation Quotient (sPPQ), Fundamental Frequency Variation (vF0), Smoothed Ampl. Perturbation Quotient (sAPQ), but Length of Analyzed Sample (Tsam), Relative Average Perturbation (RAP), Pitch Perturbation Quotient (PPQ) are not significantly different in genders. On the standard deviation, the values of Absolute Jitter (Jita), Smoothed Pitch Perturbation Quotient (sPPQ), Fundamental Frequency Variation (vF0), Shimmer Percent (Shim), Amplitude Perturbation Quotient (APQ), Smoothed Ampl. Perturbation Quotient (sAPQ), and Peak-to-Peak Amplitude Variation (vAm) are much different between genders. But the data of Amplitude Tremor Frequency (Fatr) is basically the same between genders.

3.3 Analysis of Other Parameters

Table 4 Voice Index Analysis Results ($\bar{X}\pm s$).

Gender	Number	NHR	VTI	SPI	FTRI(%)	ATRI(%)
Female	10	0.13±0.01	0.05±0.01	15.33±3.57	0	4.99±4.43
Male	10	0.2±0	0.06±0.05	11.94±2.2	0	2.54±2.72

Table 5 Voice Clearing Parameters Results ($\bar{X}\pm s$).

Gender	Number	DVB	DSH	DUV	NUB	NSH	NUV
Female	10	0	0.65±1.29	38.47±10.55	0	0.2±0.4	19.6±3.88
Male	10	0	0	39.93±14.97	0	0	19±4.55

Table 6 Basic Parameter Analysis Results ($\bar{X}\pm s$).

Gender	Number	SEG	PER
Female	10	52.4±9.89	219.8±65.01
Male	10	50±7.12	147±54.85

Tables 1 to Tables 3 are analytical descriptions of the main parameters of the voice. The other three types of voice parameters are: the voice index of Table 4, the voice clearing parameters of Table 5, and the basic parameters of Table 6.

From Table 4 to Table 6, we can find that major differences are:

(1) The number of Degree of Voiceless (DUV) is not much different between the average of male and female, but the male is higher than the female in the standard deviation data. The average data is 38.47 for female, 39.93 for male. And the standard deviation is 10.55 for female, 14.97 for male.

(2) The average and standard deviation of Total Number Detected Pitch Periods (PER) are higher for female than for male. The average female voice is 219.8, the male voice is 147. The standard deviation data of female voice is 65.01, and the male voice is 54.85.

On the whole, in the average, the Total Number Detected Pitch Periods (PER) data of both male and female change greatly, and the change of other parameters are not very large. In the standard deviation, the difference between the gender is Degree of Voiceless (DUV) and Total Number Detected Pitch Periods (PER), and the rest of the data are basically the same.

4. Summary

From the above analysis data, on the fundamental frequency parameters of the voices in Lhasa Tibetan, the average value for female is much higher than male, while male are higher than female for standard deviation. Among them, the average data of the Highest Fundamental Frequency (Fhi) and the Lowest Fundamental Frequency (Flo) are higher for female than for male, but the data for male and female on the standard deviation is not much different. The parameter data of Absolute Jitter (Jita), Smoothed Pitch Perturbation Quotient (sPPQ), Fundamental Frequency Variation (vF0), Smoothed Ampl. Perturbation Quotient (sAPQ), Amplitude Perturbation Quotient (APQ), and Peak-to-Peak Amplitude Variation (vAm) have positive related to Total Number Detected Pitch Periods (PER). In addition, in the parameter data of Length of Analyzed Sample (Tsam) and Pitch Perturbation Quotient (PPQ), whether it is the average value or the standard deviation, the parameter data of the male and female are not greatly different.

This study is only a basic study on the voice of Tibetan Lhasa, it may has practical and theoretical value for the Tibetan voice acoustic modeling and voice lesions treatment[9].

Acknowledgement

This work was financially supported by NSFC grant fund (No. 61562076).

References

[1] Gesang Jumian, Gesang Yangjing. (2002). Introduction to Tibetan Dialects, Ethnic Publishing House.

[2] Hu Tan, Suonan Zhuoga, Luo Bingfen. (2001). Lhasa colloquial reader, National Publishing House.

[3] Kong jiangping. (2001). The articulation of language, Central university for nationalities press.

[4] Oğuz H, Kiliç M A, ŞAFAK M A. (2010). Comparison of results in two acoustic analysis programs: Praat and MDVP. Turkish Journal of Medical Sciences, Vol.41, No.5, pp.835-841.

[5] Nicastri M, Chiarella G, Gallo L V, et al. (2004). Multidimensional Voice Program (MDVP) and amplitude variation parameters in euphonic adult subjects. Normative study. Acta Otorhinolaryngol Ital, Vol.24, No.6, pp.337-341.

[6] Kay PENTAX. (2007). MDVP Instruction Manual.

[7] Shi Jing, Zhou Ling, Yu Hongzhi. (2010). Research on Multi-dimensional Voice Analysis Software (MDVP). Journal of Changchun Institute of Technology (Natural Science Edition), Vol.11, No.02, pp.102-105.

[8] Zelcer S, Henri C, Tewfik T L, et al. (2002). Multidimensional voice program analysis (MDVP) and the diagnosis of pediatric vocal cord dysfunction. Annals of Allergy, Asthma & Immunology, Vol.88, No.6, pp.601-608.

[9] Hema N, Mahesh S, Pushpavathi M. (2009). Normative data for Multi-Dimensional Voice Program (MDVP) for adults-A computerized voice analysis system. Journal of All India Institute of Speech and Hearing, Vol.28, No.1, pp.1-7.