Study on Fricative Initials of Tibetan Xiahe Dialect Based on Speech and Airflow

Jing Zhang[1,2] and Yonghong Li[2](✉)

[1] Lanzhou University of Finance and Economics, Lanzhou, China
[2] Key Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Lanzhou, China
lyhweiwei@126.com

Abstract. The production of language is a complex physiological process. After years of accumulation, the research methods and fields of language have been improved and expanded. This paper takes the fricative initial of Xiahe dialect as the research object, adopts the method of phonetic experiment to extract the parameters of speech, voice and air flow, reveals the acoustic characteristics of fricative and the role of air flow in the process of fricative pronunciation, complements the understanding of traditional linguistics on pronunciation, and provides more methodological references for the study of Tibetan speech and speech physiology.

Keywords: fricative · Xiahe dialect · speech · airflow rate · airflow quantity

1 Introduction

Xiahe County is under the jurisdiction of Gannan Tibetan Autonomous Prefecture in Gansu Province. Xiahe dialect is a typical representative of Amdo Tibetan, which is between agricultural dialect and pastoral dialect, and part of it is semi-agricultural and semi-pastoral dialect [1]. The initial consonants of Xiahe dialect include voiceless consonants and voiced consonants, with more compound consonants and no tone.

There are 10 fricative initials in the Xiahe dialect. According to the different points of articulation, they can be divided into alveolar, coronal, retroflex, velar and glottal. According to the different pronunciation method, they can be divided into unaspirated voiceless fricative, aspirated voiceless fricative and voiced fricative.

It is difficult to fully explain the principles or laws of some special pronunciation phenomena by means of listening and recording sounds and conventional acous-tic analysis methods. In recent years, with the introduction of speech physiological instruments and equipment, many linguistic phenomena have been deeply analyzed and studied, which has played a positive role in promoting and supplementating the study of speech acoustics. For example, Lv Shiliang et al. [2] studied Mandarin consonants based on airflow and pressure signals. This paper comprehensively analyzes the initial consonants of fricative sounds in Xiahe dialect by combining speech signals, voice signals and airflow signals to show the intuitive performance of different fricatives on the signal and reveal their internal pronunciation characteristics and rules.

C. F. Peng et al. (Eds.): EIMT 2023, AHSSEH 8, pp. 276–282, 2023.
https://doi.org/10.2991/978-94-6463-192-0_37

2 Experimental Method

Speech signal, voice signal and airflow signal are collected simultaneously in the experiment. Airflow signal was collected by Glottal Enterprises MS 110. The calibrated airflow volume was 1.4 l and flow rate was 0.5 l/s. The data collection is carried out in the studio. In this paper, we mainly extract the sound pressure, duration, airflow rate, airflow volume and other parameters. The experimental analysis refers to the physiological description and signal analysis methods of phonemes such as Summary of Experimental Phonetics [3], The dictionary of Modern Chinese Phonetics [4] and Basic Experimental Phonetics [5].

The speaker selected a 30-year-old female who had been living in Xiahe County. We selected the spoken monosyllabic words containing 10 fricative initials in Xiahe Dialect, all of which were selected from the *Survey of Tibetan Dialects* [6]. In order to facilitate the extraction of parameters and observe the acoustic characteristics and airflow characteristics of fricative initials, we try to select monosyllabic vowels as the syllables of finals in the experiment.

3 Unaspirated Voiceless Fricative

Unaspirated voiceless fricatives include [ɕ][ɬ][s][ʂ][x][h]. This paper takes [sa] as an example to analyze. The initial [s] is the alveolar fricative. When pronouncing, the tip of the tongue is slightly lifted close to the back of the upper tooth and the front of the gum to form a slit. At the same time, the soft palate rises to close the nasal cavity, and the airflow quickly passes through the slit to form a fricative.

In terms of waveform, [s] is a non-periodic waveform, the duration is long, about 140 ms. Because the amplitude decreases before the end of the friction segment, and there is a gap between the vowel and the consonant, it is easy to define the boundary of the initial and the final. From the perspective of the sound spectrogram, [s] is reflected as the clutter in the high frequency region, and the spectral energy is concentrated in the high frequency. The spectrum distribution of [s] is different before different vowels. For non-rounded [s], the lower limit frequency of spectrum energy distribution is very high, and the energy is mainly distributed above 4.7 kHz. The lower limit of rounded [s] is lower than that of non-rounded [s]. Figure 1 is the three-dimensional spectrogram of [sa]. Figure 2 is the two-dimensional spectrum of [sa].

From the airflow signal, the airflow duration of [sa] (the first valley position of the airflow signal and the first peak position of the voice signal) is 170 ms, the average airflow rate is 120 ml/s, the maximum airflow rate is 184 ml/s, and the airflow volume is 20.6 ml. Due to the influence of consonants, the airflow velocity of the first 4 periods of vowels is relatively large, and the latter periods gradually enter a stable stage.

Figure 3 is the air flow signal diagram of [sa].

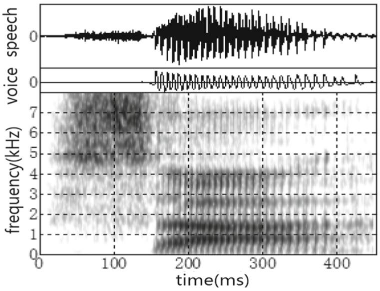

Fig. 1. The spectrogram of [sa]

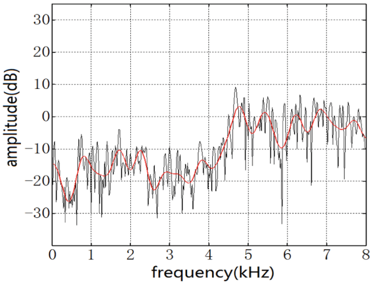

Fig. 2. The spectrum of [sa]

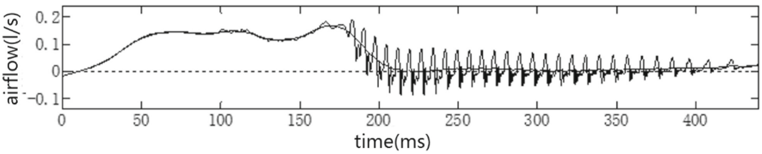

Fig. 3. The air flow signal diagram of [sa]

4 Aspirated Voiceless Fricative

The aspirated voiceless fricative sound in Tibetan Xiahe dialect has only one [s^h], and the pronunciation of [s^h] is divided into two stages: 1) the tip of the tongue is slightly raised close to the back of the upper teeth and the front of the gum to form a slit; at the same time, the soft palate rises to close the nasal cavity, and the air quickly passes through the slit to form a fricative sound; 2) The tongue body lowers, the gap becomes larger, the airflow flows out quickly, and the friction generates turbulence to form an aspiration segment.

In terms of waveform, the friction section of the initial [s^h] is a noise waveform with a duration of about 160 ms, which is divided into two sections: the duration of the friction section is 108 ms, accounting for two-thirds of the total duration of the initial, and the duration of the aspirated section is short, about 52 ms, accounting for one-third of the total duration of the initial. The amplitude of the air supply section is lower than that of the friction section. From the sound spectrogram, [s^h] is reflected as the clutter. The spectrum of the fricative segment gradually increases with the increase of frequency, and the frequency of the energy concentration area is relatively high, which is distributed above 4 kHz. The friction in the aspiration section is small, and the frequency of the energy concentration area is relatively low. The spectral energy distribution in the aspiration section is relatively average, and the energy below 1300 Hz is weak.

Figure 4 is the three-dimensional spectrogram of [s^ha]. Figure 5 is two-dimensional spectrum of friction section. Figure 6 is two-dimensional spectrum of aspirated section.

From the airflow signal, the airflow signal of [s^ha] is obviously divided into two sections: the air flow duration of the fricative section is 124 ms, the average airflow velocity is 124 ml/s, the maximum airflow velocity is 163 ml/s, and the airflow volume is 15.4 ml; the air flow duration of the aspiration section is 94 ms, the average air flow velocity is 270 ml/s, the maximum air flow velocity is 335 ml/s, and the air flow rate is 25.6 ml. Due to the influence of consonants, the airflow rate of the first 6 periods of vowels is relatively large, and the latter periods gradually enter a stable stage. Figure 7 is the air flow signal diagram of [s^ha].

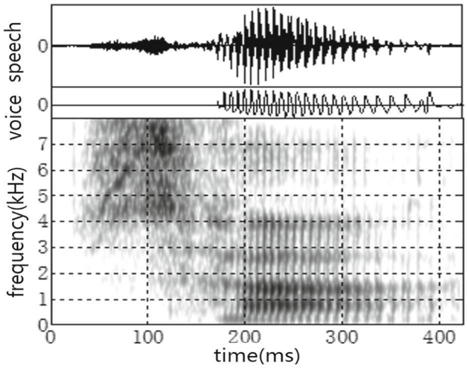

Fig. 4. The spectrogram of [s^ha]

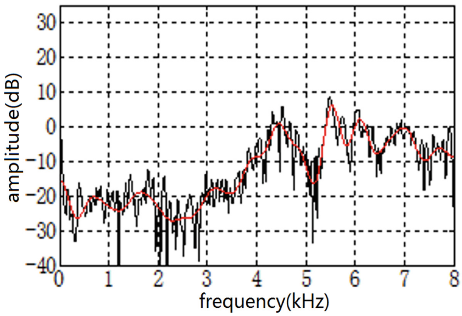

Fig. 5. The fricative segment spectrum

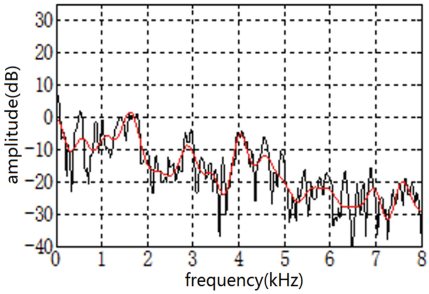

Fig. 6. The aspirated segment spectrum

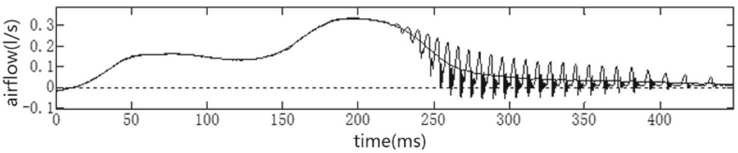

Fig. 7. The air flow signal diagram of [s^ha]

5 Voiced Fricative

The voiced fricatives in Xiahe dialect include[z][ʑ][ʐ]. In this paper, the voiced fricative is analyzed with [ʑ] as an example.

When the initial [ʑ]is pronounced, the tip of the tongue is gently pushed to the back of the lower tooth, and the soft palate rises to close the nasal cavity, and the airflow quickly passes through the slit to form a fricative sound. The vocal cords keep vibrating during the whole pronunciation process.

From the perspective of speech waveform, the duration of [ʑ] is about 106 ms, which is a periodic high-frequency noise signal with the characteristics of voiced and fricative. From the perspective of sound spectrogram, [ʑ] has an acoustic mode similar to the formant of vowels. It has both obvious formant and high-frequency clutter. The F1, F2 and F3 of the stable part are about 0.2 kHz, 1.6 kHz and 2.7 kHz, respectively. The spectrum energy of clutter is mainly concentrated at high frequencies above 4.2 kHz. The F1 of [a] followed by vowel [ʑ] points to low frequency, and the F2 points to high frequency.

Figure 8 is the three-dimensional spectrogram of [ʑa]. Figure 9 is the spectrum of [ʑa].

Figure 10 is the air flow signal diagram of [ʑa]. From the airflow signal, [ʑ] is a periodic arch, the airflow duration is 103 ms, the airflow rate is very small, the average airflow rate is 27 ml/s, the maximum airflow rate is 88 ml/s, and the airflow volume is 2.8 ml. The maximum velocity of the airflow appears in the first few cycles of the vowel, reflecting that the size of the gap or the openmouthed size has a positive relationship with the airflow.

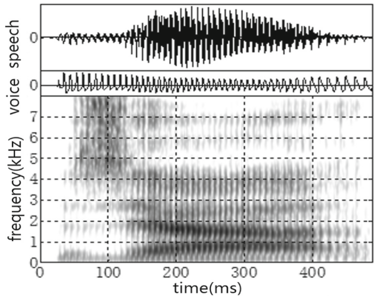

Fig. 8. The spectrogram of [ʑa]

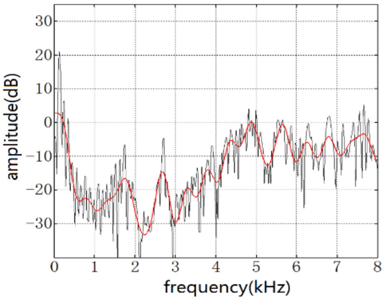

Fig. 9. The spectrum of [ʑa]

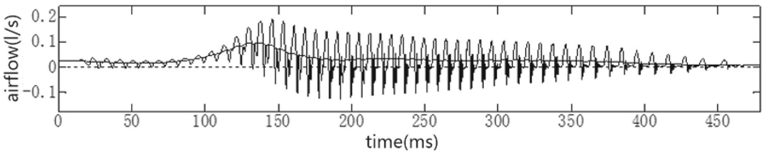

Fig. 10. The air flow signal diagram of [ʑa]

6 Conclusion

Taking Tibetan Xiahe dialect as an example, this paper makes a preliminary analysis of speech signal, voice signal and airflow signal, and focuses on the acoustic and airflow characteristics of fricatives.

From the perspective of acoustic parameters: 1) The duration of fricative initials in Xiahe dialect is longer, between 150–200 ms; 2) According to the different of friction degree of fricative consonants, the amplitude and lower limit of frequency spectrum of fricative consonants are different. The greater the degree of friction is, the higher the lower limit of frequency spectrum is. 3) The voiced fricative initials are frictional sounds while the vocal cords are also vibrating, so they are shown as periodic waveforms on the waveform diagram. In the three-dimensional spectrogram, voiced fricatives have both formants similar to vowels and high frequency clutters.

From the perspective of airflow signal: 1) The airflow of fricatives is gentle from the beginning of sound production to the end of airflow. 2) The average velocity of the air flow in the aspirated fricative is the largest, followed by the unaspirated fricative, and the voiced fricative is the smallest. 3) The air flow [s^h] > [ʂ] > [s] > [ʐ] > [z]. The exhaled air flow rate is related to whether the fricative is aspirated or not, the size of the friction gap and whether the fricative is voiced or not. The airflow volume of the aspirated segment is the largest, and the voiced segment is the smallest. The turbulence generated by the tongue and the gums is larger than the turbulence generated between the teeth. 4) Due to the vibration of the vocal cords, the airflow energy of voiced fricative is very weak, and the airflow also shows a periodic waveform.

Acknowledgement. This work was financially supported by the Fundamental Research Funds for the Central Universities (No. 31920220019) and NSFC grant fund (No. 11964034).

References

1. Wang, Shuangcheng. The phonetic research of Tibetan Anduo dialect, Shanghai: Zhongxi Book Company, 2012.
2. Lv, Shiliang et al. "A study of Mandarin consonants based on airflow pressure signals". In Tianjin: China Academic Conference on Phonetics, 2010:56–61.
3. Bao,Huaiqiao, Maocan Lin, Summary of Experimental Phonetics, Beijing: Peking University Press, 2014.
4. Cai, Lianhong, Jiangping Kong, The dictionary of Modern Chinese Phonetics, Beijing: Tsinghua University Press, 2014.
5. Kong, Jiangping, Basic Course of Experimental Phonetics, Beijing:Peking University Press, 2015.
6. Kong, Jiangping et al. Survey of Tibetan Dialects, Beijing: Commercial Press, 2011.