A Study of Final Vowels in Xiahe Dialect Based on Acoustics and Airflow

Jing Wang and Yonghong Li(✉)

Key Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Lanzhou, China
lyhweiwei@126.com

Abstract. This paper mainly uses the methods of acoustics and speech aerodynamics to study the vowels and vowels of Xiahe dialect, and establishes the vowel pronunciation model of Xiahe dialect. Some praat script programs are used to extract the formants of vowels and draw the vowel ellipses of Xiahe dialect. The airflow data is extracted using the software provided by GLOTTAL. The main conclusions are as follows: (1) The pronunciation duration of vowel [a] is the longest among all vowels; (2) Vowels [i] have the strongest ability to resist synergetic pronunciation both in acoustics and in air flow pronunciation; (3) The similarity between vowels [o] and [u] is the largest, up to 50%; (4) The sound and airflow patterns of vowels [i] and [e] are similar.

Keywords: Acoustics · Airflow · Coarticulation · Xiahe dialect

1 Introduction

Amdo dialect is a Tibetan dialect that retains more features of ancient Tibetan in Tibetan language. Its most special language phenomenon is the existence of compound consonants, and there are also voiceless and voiced consonants in Amdo Tibetan language [1]. This is a completely different language phenomenon from Chinese, which is of great research value. Based on the monosyllabic vowels in Xiahe dialect as the research object, this paper looks at the relationship between the final vowels of Xiahe dialect from the perspective of acoustics and speech aerodynamics. At present, there is relatively little research on the vowels of Xiahe dialect in the academic circle, and Wanmocuo (2021) has sorted out the phoneme distribution of the unit vowels in Anduo Tibetan language, and analyzed the parameters of the unit vowels in Anduo Tibetan language, such as pitch, duration, intensity and formant, using acoustic experimental methods. It is found that in terms of duration, the position of the vowels in the syllable structure, and the difference in the pronunciation mode of the front and rear consonants will affect the duration of the vowels. In addition, the subsequent consonants have more influence on vowels than the preceding consonants [2]. Shawotso (2022) studied the pronunciation mechanism of Tibetan vowels from the perspective of speech aerodynamics, taking the vowels in Anduo dialect of Tibetan as the research object. It was found that the fundamental frequency and sound pressure level of Tibetan vowels were generally affected

X. Yuan et al. (Eds.): ICEKIM 2023, AHCS 13, pp. 1973–1982, 2023.
https://doi.org/10.2991/978-94-6463-172-2_219

by the consonant consonants and the gender of the speaker, while the expiratory duration, expiratory flow, and aerodynamic resistance values were not directly related to the consonant consonants, but closely related to the gender of the speaker [3].

2 Research Method

2.1 Pronunciation Materials

According to the phonology of Xiahe dialect, there are only six vowels in Xiahe dialect, which are list in the Table1:

2.2 Speaker Selection

Because of the instability of air flow pronunciation, this paper chooses six male and female Tibetan student who is in Xiahe dialect with Amdo dialect as its mother tongue. He lived in pastoral area since childhood. The Tibetan pronunciation standard is free of any throat disease, and the vocal cord is in good condition, normal listening ability and normal sound characteristics. Before the experiment, the pronunciation training was carried out for the pronunciation person, and it was required to read three times for each sound according to the pronunciation list.

2.3 Experimental Equipment and Methods

In this experiment, the aerodynamic data acquisition equipment is composed of a circular ventilation breathing velocimeter mask connected to a narrow band pressure sensor (PTL-1) and a separate broadband pressure sensor (PTW-1) (Glotal Enterprises MS 110). The Aeroview system uses an automated factory optimization algorithm to calculate the “average glottal flow resistance” and its reciprocal, the “average glottal flow conductance” during vowel generation. Praat software is used to assist in observing the recording effect, as to ensure the quality of the corpus. The collection of air flow and voice data is carried out in the professional recording room of Northwest Minzu University.

3 Study on the Duration of Monophonic Pronunciation

3.1 Study on the Duration of Vowel Pronunciation

Time length is also a very important parameter for the study of vowel pronunciation. It can not only reflect the pronunciation movement of vowels, but also reflect the degree of synergy between vowels and the front and back sounds. In the actual speech flow, there

Table 1. Xiahe dialect Final vowel table

[i]	[e]	[ə]
[a]	[o]	[u]

Note: The actual pronunciation opening of [e] is small, which should be distinguished from [i]

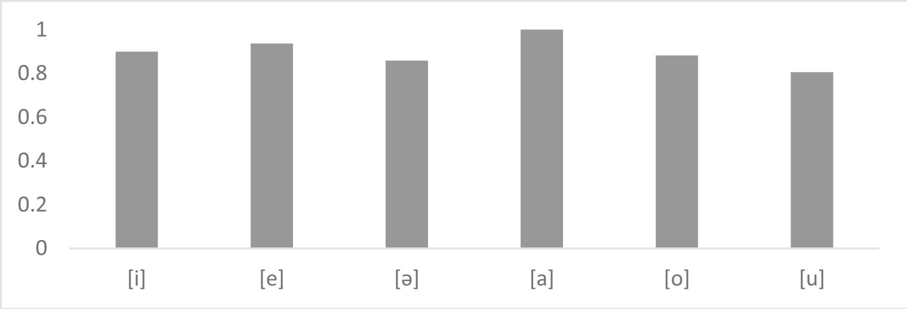

Fig. 1. Vowel duration chart of Xiahe dialect

is a phenomenon that the duration of the air flow is longer than the actual vowel duration. From the beginning to the end of the air flow, there is a section of air flow that does not correspond to the duration of the voice. This section does not produce voice. We call it a paralinguistic phenomenon. Due to the great difference in time length between people or between languages, we choose a relative time length calculation method to analyze the time length of vowels, choose the vowel with the longest absolute time length average value, set its time length as unit 1, and compare the time length of other vowels with the average value of this vowel to obtain the corresponding relative time length. The formula is as 1:

$$\text{Relative duration of a vowel} = \frac{\text{Average absolute duration of the vowel}}{\text{The maximum average of absolute duration of all vowels}} \tag{1}$$

According to the formula, the absolute duration of vowel [a] in Xiahe dialect is the longest, and its value serves as the denominator of the above publicity. After calculation, the following relative duration chart is obtained:

According to the Fig. 1, the duration difference between vowels in Xiahe dialect is not large, with an average difference of about 30-40ms.The order of pronunciation duration is: [a] > [e] > [i] > [o] > [ə] > [u], the duration of [a] is longer and that of [u] is shorter.

3.2 Study on Acoustic Distribution of Monophonic Sound

The concept of voice pattern was first proposed by Professor Shi Feng. He believed that any language and dialect has its own system, and voice pattern is its systematic performance. The analysis of voice pattern is to use the data and charts obtained from voice experiments to investigate the acoustic performance of various phonemes, including the positioning characteristics of phonemes themselves, the internal variation distribution rules, and the overall arrangement relationship [4]. M. Joos (1984) believed that although the frequency of the same vowel formant uttered by different people is different, the relative position of the vowel uttered by each person on the acoustic vowel map is basically stable, and the human brain perceives speech based on this relative position [5]. The experimental method of vowel pattern is based on vowel formant data. The tone quality of vowels is closely related to the formant. Formant is the expression of sound energy accumulation in a specific frequency band. Different vocal cavity shapes will produce different formant patterns. There are three resonance peaks commonly used in speech research: F1, F2, F3.F1 is inversely proportional to the height of tongue position; F2 is proportional to the front and back of tongue position.

This paper uses the praat script program written by Minghui Zhang[1]. After marking the stable section of the vowel formant, use the "_FormantPro[2]". The praat script program extracts the parameters of vowel formants in batches. According to the average value of the extracted formant F1, F2, F3 data, the vowel formant pattern of Xiahe dialect is drawn:

According to the Fig. 2, the mean distribution pattern of formants F1, F2 and F3, vowels can be divided into three groups: the first group is [i] and [e]. The distance between F1 and F2 of these two sounds is relatively far, and the value of F2 is relatively large, which means that the tongue position of these two sounds is front, which is the front vowel. It is found that the F1 of [e] is slightly smaller than the average value of [i], which means that the tongue position of [i] and [e] in Xiahe dialect is similar in pronunciation. The second group is[ə] and [a], their F2 values are smaller than [i] and [e], indicating that their tongue position is lower than [i] and [e], while the first formant of [a] and other vowel ratios are the highest, indicating that [a]'s tongue position is the lowest. The third group has the lowest second formant of [o] and [u], which means that its tongue position is the lowest, which is the rear vowel. The F1 of [o] is higher than [u], which means that its pronunciation tongue position is lower than [u]. The F3 of these two sounds is lower than the first two groups, which is the round lip vowel. The F3 value of [u] is the lowest, and its round lip degree is the largest.Classify the formants of vowels

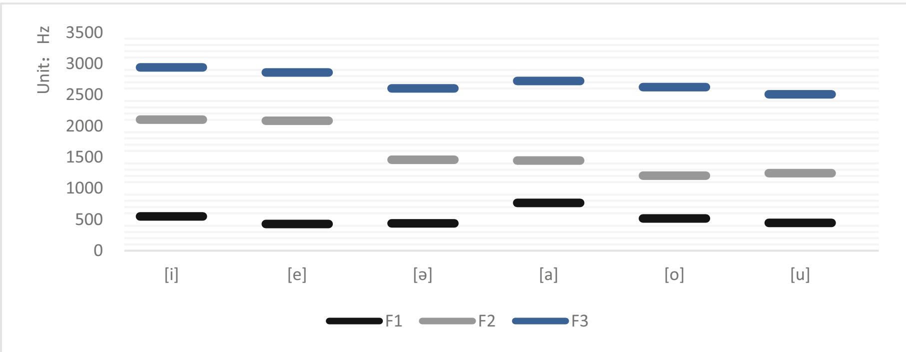

Fig. 2. Monophonic formant pattern

[1] The program of "batch labeling and generating inverse filtering audio" was written by Minghui Zhang (2016 doctoral student of the Simian Institute of Humanities, East China Normal University) on March 30, 2018.

[2] FormantPro "script program written by Professor Yi Xu (Professor of University College London, UK) in 2020, Version 1.4.3.

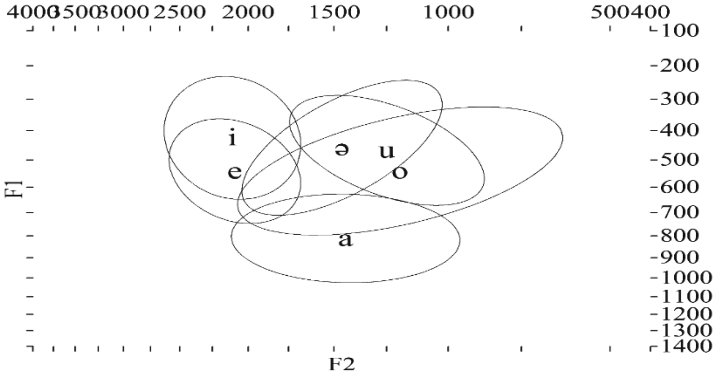

Fig. 3. Monophthong formant ellipse

in each syllable, and use the praat script written by Ling Feng[3] draw vowel ellipses, first generate a "tableofreal" file, and then run the program to generate the following formant vowel ellipses:

The ordinate of the Fig. 3 is F1 and the abscissa is F2. Because the fundamental frequency of male and female pronunciation is quite different, the figure is drawn using male formant parameters. The formant value is relatively low, but it basically reflects the formant distribution pattern of Xiahe vowels. Among them, we can see the distribution pattern of the front, middle and back of the tongue position that is consistent with the grouping of the formant pattern. The two ellipses of the front vowel [i] and [e] have a large overlap. In the sense of hearing, the two vowels in Xiahe dialect are very close to each other. It needs to be carefully identified. The main reason is that the pronunciation opening of the vowel [e] in the actual pronunciation is small, leading to its actual sound value close to [i]. The central vowels mainly include[ə] And [a], [a] have the lowest tongue position, the lowest point is about 924Hz, and the central vowel[ə] A part of the pronunciation tongue position is raised, and the sound value is close to[ɯɿ], However, from the perspective of integrity, they are all marked as[ə].The back vowels are [u] and [o], where the tongue position of [u] is slightly higher than that of [o], and the formant data of [u] and [o] also have some overlap.In order to detect the similarity between two groups of vowels, Mahalanobis Distance is used to calculate the covariance distance of formant data. The formula as 2:

$$D^\wedge M\ \ (x) = \sqrt{\left((x-\mu)^\wedge T \sum {}^\wedge(-1)(x-\mu)\right)} \tag{2}$$

The praat script program used in this part is "using formant data and Markov distance algorithm to quantify the similarity of two kinds of vowels[4]" and the following values are calculated:

[3] Praat script "Draw vowel ellipses based on formant data" written by Professor Ling Feng of Shanghai University.

[4] The script was written by Zhang Minghui (Beijing Huaruibao Technology Co., Ltd.) on February 14, 2020.

Table 2. Final vowel similarity scale table

	[i]	[e]	[ə]	[a]	[o]	[u]
[i]		25%	0%	0%	0%	0%
[e]	23.5%		0%	0%	0%	0%
[ə]	0%	0%		0%	7%	23%
[a]	0%	0%	0%		0%	0%
[o]	0%	0%	0%	0%		50%
[u]	0%	0%	0%	0%	17.6%	

It can be seen from the Table 2 that the sample coincidence rate between the vowels [i] and [e] of Xiahe dialect is 23.5%, and there is a 25% sample coincidence rate between [e] and [i]. Central vowel[ə] There is also a certain degree of similarity between the vowels [o] and [u] [ə]. The sample coincidence rate between [o] and [o] is 7% [ə]. The sample coincidence rate between [u] and [u] is 23%, which also shows that[ə] greater similarity with [u]. The sample coincidence rate between [o] and [u] is 17.6%, and the sample coincidence rate between [u] and [o] is 50%, which is the two vowels with the greatest similarity among all vowels.

3.3 Study on Airflow Distribution of Monophonic Sound

Because the larynx is difficult to access and the glottic pressure is difficult to collect, this paper mainly takes the oral air flow as the main research object, taking the time length as the abscissa and the average air flow speed as the ordinate, to draw the pattern of vowel pronunciation air flow.

According to the above Fig. 4, we can see that the dispersion of the average air velocity distribution of the vowel [e] is very large, which indicates that the vowel is greatly affected by the consonant initials. When the consonant initials are aspirated or fricative, its average air velocity will show a significant increase trend. From the perspective of duration, the duration distribution of vowel [a] has a large dispersion. In addition, vowels [i] [a] [ə] [o]. The density of the air flow pattern of [u] is relatively close, and there is no obvious difference between them. The air flow pattern of [i] is more concentrated and the least affected. Although these sounds are also affected by consonant consonants, the differences are not significant.

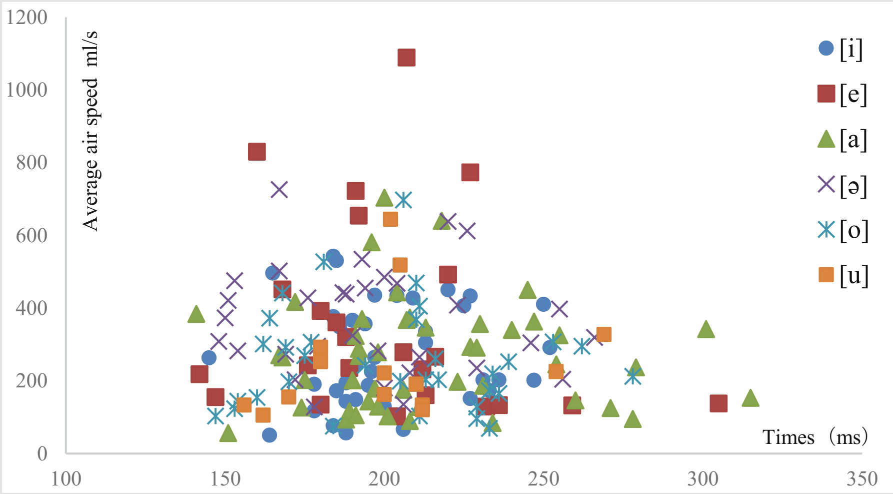

Fig. 4. Vowel airflow pattern

4 Study of Coarticulation in CV/CCV Syllables

Each phoneme in the syllable structure does not exist independently, and the pronunciation between them is bound to be affected by the front and back sounds. This involves the research method of the synergetic pronunciation theory. From a physiological point of view, the various parts of the vocal organs continue to move, and the pronunciation posture of the adjacent segments overlap each other, so that at any point in time, the shape of the vocal tract (especially the tongue) is always affected by the front and back sound segments. Coarticulation describes this phenomenon, because voice realizes its own nature in the change of speech flow [6]. Lindblom (1963) put forward the concept of "acoustic target" in the study of vowel weakening, that is, the formant mode in the ideal state without the influence of context is the target value of vowel formant, but in most cases, this acoustic target cannot be reached. When Fant (1986) mentioned that people use physiological and physical experiments to investigate these discrete speech units, it was found that there is no one-to-one correspondence between discrete speech units and continuous physiological and physical quantities [7]. Zhang Lei (2012) put forward the concept of "Coarticulation coefficient" for the first time in his paper on the study of Chinese Mandarin co-pronunciation [8]. This concept is also a processing method of voice physical parameters based on the mutual loading of voice features. This paper will also use this concept to study the influence of consonants on vowels in CV/CCV syllable structure.

The acoustic coarticulation coefficient of vowels = SD (F2onset)/SD (F2target), where SD is the standard deviation, F2onset is the starting point of the second formant transfer, and F2target is the midpoint of the second formant. The greater the ratio, the greater the influence of vowels on different preceding consonants; vice versa. Based on this coefficient, this paper proposes that the Coarticulation coefficient of the average air velocity of vowels = SD (avg. Airflow onset)/SD (avg. Airflow target). Next, we will

Table 3. Acoustic Coarticulation Coefficient Table of Vowels

	[a]	[i]	[e]	[ə]	[o]	[u]
1	1.30	1.45	0.94	5.83	2.67	2.97
2	3.62	0.82	1.02	1.34	1.29	0.74
3	2.48	1.58	2.69	1.93	2.81	1.36
4	2.92	1.82	2.43	5.09	6.90	1.75
5	1.69	1.04	1.81	0.98	1.49	2.65
6	2.29	2.73	3.22	0.51	3.49	0.41
Mean	2.39	1.63	2.02	2.52	3.10	1.69
SD	0.84	0.69	0.92	2.31	2.03	1.01

study the co-pronunciation of vowels in Xiahe dialect from the perspectives of acoustics and airflow.

4.1 Research on CV/CCV Syllable Coarticulation

Extract the starting point value and midpoint value of the vowel consonant formant F2 of six speakers, and calculate the obtained data according to the acoustic coarticulation coefficient of vowels, which controls the syllable environment of different vowels. The following table lists the acoustic coarticulation coefficients of six vowels uttered by different speakers, and calculates their mean and standard deviation.

According to the Table 3,the mean value of Coarticulation coefficient of six vowels, we can see that the degree of vowels affected by consonants is: [o] > [ə] > [a] > [e] > [u] > [i], it can be seen that [o] is the most affected by consonants, while [i] is the least affected by consonants, which is consistent with Chao.Z's conclusion of studying mandarin vowels.The acoustic coarticulation coefficients of six groups of vowels were statistically analyzed by single-factor repeated measurement. The data were verified by normal distribution before analysis, and all showed normal distribution. The measurement results were $F = 56.097$, $P < 0.01$. There was significant difference between the acoustic coarticulation coefficients of six groups of vowels.

4.2 Research on CV/CCV Syllable Airflow Coarticulation

The same research method was used to study the effect of consonants in CV/CCV structure on the synergetic pronunciation of monophonic pronunciation airflow, extract the average airflow velocity at the phonological junction and the average airflow velocity at the stable segment of vowels, and calculate the standard deviation. The greater the Coarticulation coefficient of the average air velocity of vowels, the greater the influence of consonants on the average air velocity of vowels. The following table shows the Coarticulation coefficient of the average air velocity of six groups of vowels:

According to the Table 4, the average value in the table above, the order of the average air velocity of vowels affected by consonants is [o] > [e] > [a] > [u] = [ə] >

Table 4. Aieflow Coarticulation Coefficient Table of Vowels

	[a]	[i]	[e]	[ə]	[o]	[u]
1	1.93	2.69	1.91	3.43	1.44	0.74
2	0.92	1.35	2.86	3.32	3.35	3.69
3	1.64	2.4	2.38	1.34	1.43	1.65
4	3.52	1.37	1.74	1.88	4.92	1.13
5	2.18	2.65	4.03	1.96	3.69	2.89
6	4.08	1.08	5.15	2.31	5.53	4.19
Mean	2.38	1.92	3.01	2.37	3.39	2.38
SD	1.19	0.73	1.33	0.84	1.71	1.42

[i], the round-lipped sound [o] with the greatest impact, and the smallest impact is [i], which is consistent with the acoustic coarticulation coefficient of vowels. Similarly, [o] is the most affected, and [i] is the least affected, which should be related to the high tongue position and small opening degree when [i] pronounces, and the pronunciation position of consonants will be affected when consonants are combined. No matter from the perspective of acoustics or physiology, or from the perspective of coarticulation coefficient or dispersion, it can be fully proved that the vowel [i] has the strongest anti-synergy ability.

5 Conclusion

This paper mainly studies the acoustic characteristics and pronunciation airflow characteristics of vowel vowels in Xiahe dialect. First, the relative and absolute duration of vowel pronunciation is studied from the perspective of pronunciation duration. The results show that the duration of vowel [a] is the longest, but from the perspective of the whole, the duration difference between vowel vowels is not significant. From the acoustic point of view, it has been proved that [i] and [e] in Xiahe dialect are similar in tongue position when pronouncing, and the similarity of the two vowels is as high as 25%. The two vowels with the greatest similarity are [o] and [u]. The air flow pattern of vowel pronunciation is not consistent with that of vowel resonance peak, but from the perspective of air flow, the air flow stability of vowel [i] is the best. From the perspective of synergetic pronunciation, [i] is also the vowel with the strongest ability to resist synergism, while [o] is the vowel with the weakest ability to resist synergism. This paper mainly carries out a comparative study on the acoustics and pronunciation airflow of the vowels and vowels of the Xiahe dialect. This research method is not widely used in the phonetics research at present, because the collection of airflow has certain instability, and the data obtained need to be corrected accordingly. At present, it is only a preliminary study, and more exploration will be carried out in the future.

Acknowledgment. This work was financially supported by Excellent doctoral program in Gansu Province (No. 22JR5RA179) and the Fundamental Research Funds for the Central Universities (No.31920220019).

References

1. Kan Hua, "Comparison of consonant consonants in Tibetan Languages of Xiahe and Maqu in Gannan," Journal of Northwest University for Nationalities (Philosophy and Social Sciences Edition), pp. 35–42, 1984.
2. Mocuo Wan, Nan La Cairang, Yufeng Chen. The phoneme distribution and acoustic performance of the vowels in Ando dialect of Tibetan [J]. Journal of Tibet University (Social Science Edition), 2021,36 (04): 116-122. https://doi.org/10.16249/j.cnki.1005-5738.2021.04.015
3. Wucuo Xia. A Study on the Aerodynamics of Tibetan Vowel Speech -- Taking Amdo dialect of Tibetan as an example [J]. Plateau Science Research, 2022,6 (01): 97-103. DOI: https://doi.org/10.16249/j.cnki.2096-4617.2022.01.013
4. Xiujuan Shi. On the application of vowel pattern in the study of Chinese dialects [J]. Qilu Journal, 2016 (04): 156-160
5. Feng Shi. Pattern of intonation - the intersection of phonetics and phonology. The Commercial Press, Beijing 2008.046
6. Wu Zongji, Lin Maocan. Summary of Experimental Phonetics [M]. Peking University Press, 2014
7. Li Yonghong.Study on the Air Pressure of Putonghua Stope Based on PAS6600 [A] Information Engineering Research Institute, USA.
8. Lei Zhang. Acoustical study of synergetic pronunciation in Mandarin syllables [D]. East China Normal University, 2012.