Advances in Social Science, Education and Humanities Research, volume 673
2022 3rd International Conference on Language, Art and Cultural Exchange (ICLACE 2022)

Acoustic Analysis of Compound Vowels in Ali Tibetan Language

Lingzhen Li [1], Yonghong Li[2*]

[1] China National Information Technology Research Institute, Northwest Minzu University
[2] China National Information Technology Research Institute, Northwest Minzu University
[*]Corresponding author. Email: 825760876@qq.com

ABSTRACT
By using the method of acoustic experiment, this paper analyzes the acoustic characteristics of Ali Tibetan compound vowels, such as duration, tone and formant. It is concluded that there are 15 kinds of compound vowels in Ali Tibetan language, which are divided into oral compound vowels and nasal compound vowels; They are true compound vowels, and they all come from the reduction of syllables; Compound vowels are all long vowels, which are matched with long tones; The keys are 44 and 14, The condition of high flat tone 44 is that the initial consonant is clear, and the condition of low rising tone 14 is that the initial consonant is turbid. The study of compound vowels in Ali dialect in this paper can provide a certain reference for the further study of compound vowels in Weizang dialect in the future.

Keywords: *Ali Tibetan, Compound vowels, Experimental phonetics, Acoustic analysis.*

1. INTRODUCTION

Ali is located in the west of the Tibet Autonomous Region and takes Tibetan as the main language. Traditionally, the Tibetan region is divided into three dialect regions: Weizang, Ando and Kangba. The dialect points Pulan taken in this article belongs to Weizang dialect.

Its phonetic characteristics are that the initial consonants of complex consonants fall off, the turbid consonants outside the nasal sound are cleared, resulting in a tone system. There are a large number of complex vowels with high frequency in the post Tibetan dialect group [1]. As for compound vowels, Qu Aitang [2] believes that there are mainly two compound vowels in Tibetan, which are divided into true compound vowels and false compound vowels according to their nature. In terms of Ali compound vowels, Tan Kerang [3] proposed that there are a large number of Ali Tibetan compound vowels, and most of them are the result of syllable reduction.

Study uses the method of acoustic experiment to analyze the acoustic characteristics of compound vowels in Pulan County by analyzing the spectrogram, extracting formant, duration and other data, hoping to provide data support for the further study of compound vowels in Tibetan dialect in the future. Tibetan is a major national language in China. It has a long history and a large number of users. Using the method of experimental phonetics to analyze the acoustic parameters of Tibetan compound vowels is of great value in phonetics and speech engineering.

2. EXPERIMENTAL DESCRIPTION

2.1. Pronunciation Person and Material

The pronunciation partner is a college student of Tibet University for nationalities, with clear speech and no obvious pronunciation defects and can speak local dialect without being affected by other dialects. The pronunciation material used in this study comes from the Tibetan dialect questionnaire [4], which requires the pronunciation person to select the commonly used monosyllabic and disyllabic compound vowels, and finally select a total of 40. Examples of pronunciation materials are shown in Figure 1 below.

Tibetan	Chinese meaning	IPA	Tonal category	Tone pitch
ཨ	Tibetan script	ˀa	HA	D51
དབང	right	oŋ	HC	D51
སྦར	already lighted	bar	LB	D131
སྦོམ	coarse	bɔm	LC	D131
སྦོང	soak	boŋ	LC	D131
དབུགས	breath	ʊʔ	HE	D55

Figure 1. Yushu Tibetan pronunciation vocabulary

2.2. Voice Signal Acquisition

In order to ensure the accuracy of voice signals, this study recorded in a professional recording room. The recording equipment includes laptop, microphone ECM-44B Lavalier microphone, EURORACK UB1204FX-PRO mixer and Blaster X-Fi Surround5.1Pro external sound card, etc; The recording software is Adobe Audition3 0, single channel recording is adopted, the sampling accuracy is 16 bits, the sampling frequency is 22050hz, and the recording samples are stored in (*. WAV) format.

2.3. Experimental Data Extraction

Praat speech analysis software is used to extract all acoustic parameters in this study. Firstly, the Praat script program is used to mark the voice file to distinguish the initial segment, the final segment and the two vowels before and after the final segment; Then extract the time length, formant and other data respectively for comparative analysis; When extracting the fundamental frequency data, we should accurately deal with the bend and drop tail. If there is a special phenomenon, we need to deal with it in combination with the dialect phonology, and finally normalize it to 20 points.

2.4. Five Degree Conversion

In order to make the data more universal, it is necessary to normalize the F0 data and finally convert it into the traditional five degree value. The five degree value conversion adopts the T-value method used by Shi Feng, etc., and its formula is: $T = [(lgf0 lgmin) / (lgmax lgmin)] * 5$. Where Max is the upper limit frequency of frequency domain, min is the lower limit frequency of frequency domain, F0 is the frequency of measurement point, and the corresponding relationship between T value and five degree value is shown in Table 1.

Table 1. Comparison of T value and five degree value.

T value	0-1.1	0.9-2.1	1.9-3.1	2.9-4.1	3.9-5
Five degree value	1	2	3	4	5

3. PHONEMES AND SOURCES OF COMPOUND VOWELS IN PULAN DIALECT

As for the source of Ali Tibetan compound vowels, Mr. Tan Kerang believes that most of the compound vowels come from the reduction of disyllabic words such as root and word formation suffix [3]. According to the results of investigation and previous studies, the sources of Pulan Tibetan language can be summarized into the following categories:

ao、 *au*、 *iu*、 *ia*、 *ea*、 *oa*、 *ua*、 *eo*: from the reduction of root vowels and word formation suffixes of ancient words; *iã*、 *iõ*、 *eõ*、 *aũ*、 *aõ*、 *oã*: it comes from the reduction of the word formation suffix of the root vowel and nasal sound of ancient words.

4. DURATION ANALYSIS

Although the true compound vowel also has two sound points, there is almost no interval between the two vowels or the interval is very short, so it is not easy to detect [5]. We get the total duration of 15 kinds of compound vowels according to Praat, then get the average respectively, and finally make a cylindrical diagram of the total duration, as shown in Table 2.

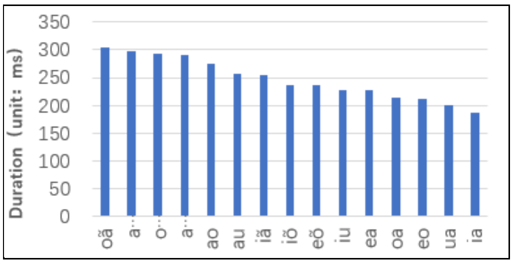

Figure2. Total duration of compound vowels in Pulan dialect

Through analysis, we conclude that the compound vowels in Pulan area are long vowels. Among the 15 types of compound vowels, the longest is *oã*, which is 304ms, and the shortest is *ia*, which is 188ms, with a difference of 116ms.

In order to further investigate the duration relationship between the front and rear vowels in the compound vowel, we extract the duration of the front and rear vowel stable segments respectively according to the Praat annotation, and draw a histogram, as shown in Figure 3. In Figure 3, blue represents the front stable section and orange represents the rear stable section.

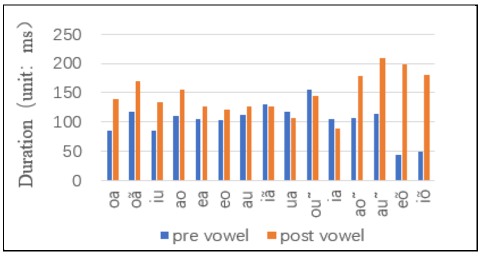

Figure3. Duration of pre vowel and post vowel of compound vowel in Pulan

As can be seen from Figure 3, The duration of the first stable segment of *iã*、 *ua*、 *oũ*、 *ia* are longer than that of the second stable segment, and the other compound vowels are on the contrary, but the first vowel still has an obvious stable segment, so it is a true compound vowel. In addition, in the 14th and 15th types of compound vowels, although the duration of the first vowel is significantly shorter than that of the second vowel, there is still obvious perception in hearing, and there is an obvious peak on the sound intensity curve. Therefore, it is still defined as true compound vowel here.

5. TONE ANALYSIS

The relationship between compound vowels and tone is first reflected in the tone range. Tan Kerang [3] pointed out that the reason why the true vowel is a syllable is that this syllable, like the unit sound, has only one tone range. True compound vowels are a kind of equivalent structure, that is, several vowels stand side by side without primary and secondary [2]. According to the acoustic analysis of this paper, it is found that the fundamental frequency curve of Pulan compound vowel is coherent without interruption, and the front and rear vowels have obvious ringing points, so we think it is a true compound vowel.

After normalizing the extracted fundamental frequency value into five degree value, we get the specific tone value and occurrence conditions of Pulan Tibetan compound vowel: (1) high flat tone 44, and the occurrence condition is that the initial consonant is clear consonant (the base word is clear base word); (2) Low rising tone 14 occurs when the initial consonant is turbid (the base word is turbid). See Figure 4 for details:

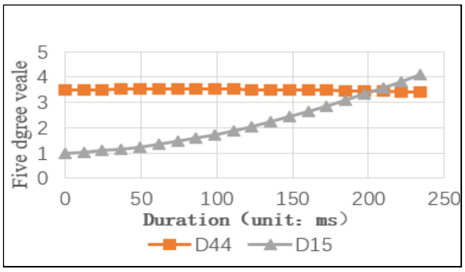

Figure 4. tone curve of Pulan Tibetan language

From Figure 4, we can see that the T value curve of high level adjustment is stable, there is no obvious upward or downward trend, and the adjustment value is 44; The T value curve of low rising tone rises obviously, starting from the first tone area and finally the fourth tone area, with the tone value of 14. The duration of the two types of tones is the same, both of which belong to long tone.

6. FORMANT ANALYSIS

We select a representative sound from all kinds of compound vowels in Ali Pulan Tibetan, mark it with Praat and draw a picture to show the formant characteristics of this kind of compound vowels. "-" in the label represents the transition segment.

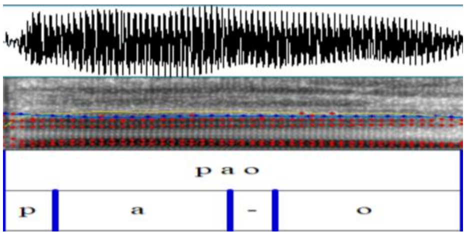

Figure5. "pao" (hero)

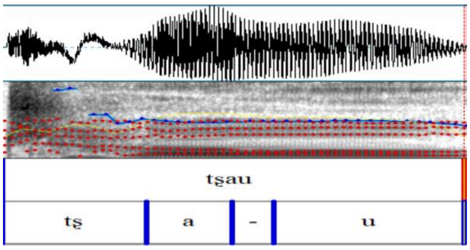

Figure6. "tʂau" (big copper pot)

As shown in Figure 5, the example word "hero" from the compound vowel *ao*. By averaging, F1 and F2 from [a] to [o] show a downward trend, F1 decreases from 835hz to 642hz, and the average distribution range of F2 changes from 1311hz to 1026hz; Figure 6 shows the example word "big copper pot" of compound vowel *au*. The average value of F1 is changed from 551hz to 431hz, and the average value of F2 is 1101hz to 828hz. Both of them slide from one stable state to another. It can be found that after different vowels, F2 of [a] has an obvious upward and downward trend.

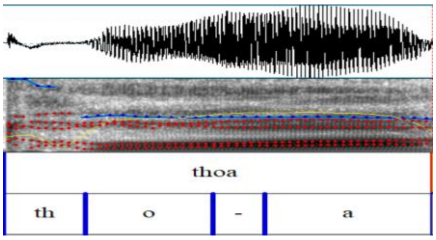

Figure7. "t^hoa" (hammer)

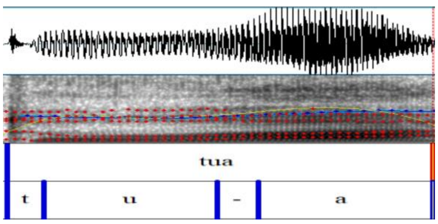

Figure8. "tua" (smoke)

Figure 7 is a diagram of compound vowel *oa*. It is concluded that the distribution range of F1 changes from 534-749hz to 851-1052hz, and the distribution range of F2 changes from 1082-1097hz to 1321-1548hz, showing an upward sliding trend. In the speech diagram of compound vowel *ua* in Figure 8, the distribution range of F1 value changes from [u] average 417hz to [a] 816hz, the average value of F2 increases from 936hz to 1372hz, and the overall formant frequency increases.

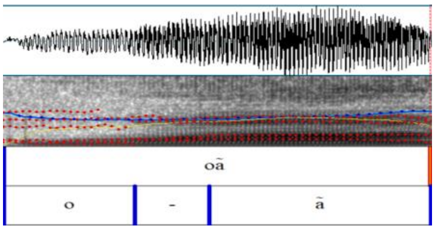

Figure9. “oã” (breast)

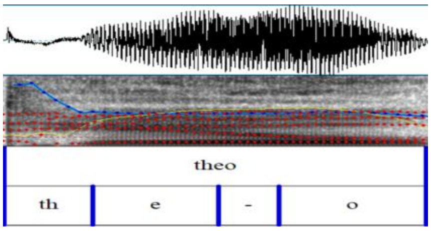

Figure10. “t^{h}eo” (thumb)

Fig. 9 is a diagram of compound vowel *oã*. F1 value rises from the average value of [o] 493hz to the average value of [ã] 911hz, and F2 slides up from 899hz to 1511hz. Figure 10 example word "thumb" from compound vowel. The average distribution range of F1 of all *eo* example words [e] is 523hz and F2 is 2158hz; [o] The average distribution value of F1 is 536hz and the average value of F2 is 1113hz. F2 has an obvious transition trend.

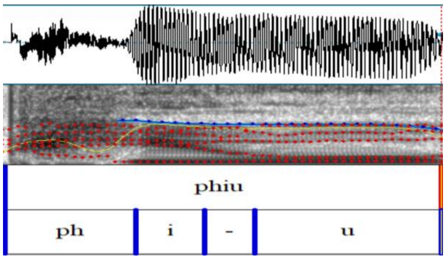

Figure11. “p^{h}iu” (night)

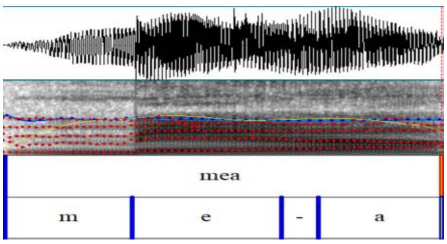

Figure12. “mea” (nevus)

The vowel *iu* of the “night” is in Figure 11. The value range of F1 transits from [i] 296hz to [u] 338hz, and the value range of F2 decreases from [i] 2559hz to [u] 951hz. Figure 12 shows the *ea* language of "nevus". The F1 value range slides up from [e] 607hz to [a] 949hz, and the F2 value range drops from [e] 2181hz to [a] 1809hz.

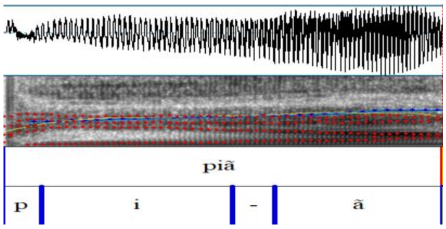

Figure13. “piã” (sand)

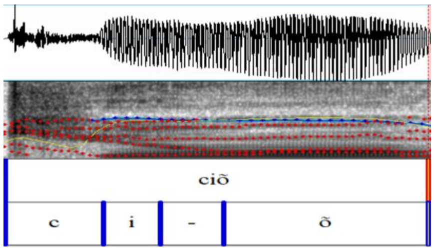

Figure14. “ciõ” (bitch)

Figure 13 example words from compound vowel *iã*. After averaging the data, the value range of F1 rises from [i] 407hz to [ã] 899hz, and the value range of F2 decreases from [i] 26321hz to [ã] 1182hz. Figure 14 shows the "bitch" *iõ* language. F1 value range slides up from [i] 407hz to [õ] 899hz, and F2 value range drops from [i] 2632hz to [õ] 1182hz.

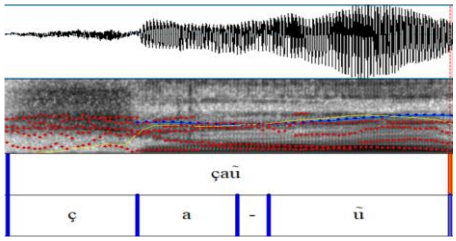

Figure15. “çaũ” (hat)

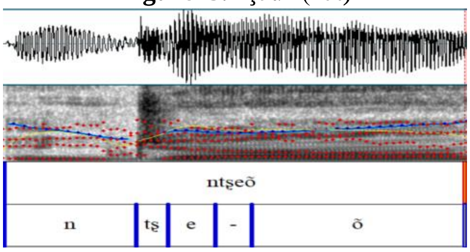

Figure16. “ntʂeõ” (female yak)

Figure 15 from compound vowel *aũ* of example word "hat". After averaging the data of all example words, it is obtained that the average value of [a] F1 is 430hz and F2 is 1215hz; [ũ] The average value of F1 is 605hz, F2 is 799-1501hz, and the average value is 1319hz. Fig. 16 is a diagram of compound vowel *eõ*. After averaging the data of all *eõ* example words, the average value of [e] F1 is 552hz and F2 is 1276hz; [õ] the average value of F1 is 613hz, f2113-1387hz, and the average value is 1182hz.

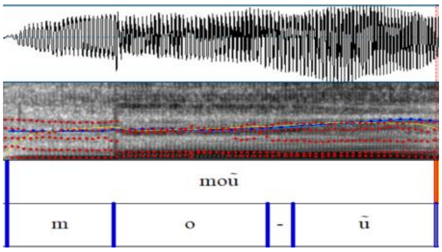

Figure17. “moũ” (female sheep)

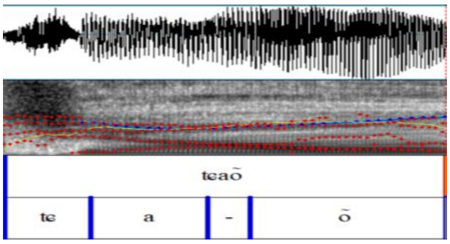

Figure18. “tɕaõ” (hen)

As shown in Figure 17 of *oũ* spectrogram, the obvious upward and downward trends of F1 and F2 can be seen in the language map. F1 transition from [o] average 608hz to[ũ] The average value is 462hz, and F2 decreases from 1075hz to 812hz. In Figure 18, after averaging the data of all *aõ* example words, it is concluded that the F1 distribution range of [a] is 522-958hz, the average value is 781hz, and the F2 distribution range is 1637-2322hz, the average value is 1866hz; The distribution range of F1 of [õ] is 711-780hz, with an average value of 738hz, and the distribution range of F2 is 768-976hz, with an average value of 853hz. The transition trend is obvious.

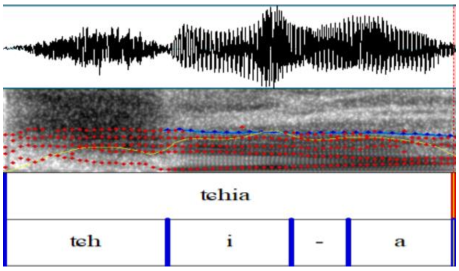

Figure19. “tɕ^{h}ia” (canine teeth)

As shown in Figure 19, it comes from the compound vowel *ia* example word "canine tooth". After averaging the data of all words, it is concluded that the distribution range of F1 of [i] is 330-555hz, the average value is 483hz, and the distribution range of F2 is 2129-2380hz, the average value is 2232hz; [a] The distribution range of F1 is 821-990hz, with an average value of 895hz, and the distribution range of F2 is 1486-1671hz, with an average value of 1548hz.

7. CONCLUSIONS

Through the phonological analysis of 15 kinds of compound vowels in Ali Pulan dialect, we draw the following conclusions:

a. Compared with other Tibetan areas, aripulan compound vowels have a large number and complex forms. They can be divided into oral compound vowels and nasal compound vowels. Both of them belong to true compound vowels and come from the reduction of syllables.

b. Aripulan compound vowels have stable duration before and after vowels, and there are obvious perceptual segments in auditory perception.

c. All compound vowels only produce growth tone, with tone values of 44 and 14, and tone types of low rising tone and high flat tone. High flat tone 44, the condition is that the initial consonant is clear; Low rising tone 14, if the initial consonant is turbid.

By analysing the relationship between compound vowels and tone evolution in Pulan dialect, this experiment comes to the conclusion that the generation of compound vowels is inseparable from the evolution of the whole language system. Starting from experimental phonetics, the methods of acoustic analysis and speech signal processing can effectively supplement the traditional phonetics.

ACKNOWLEDGMENTS

This work was financially supported by NSFC grant fund (No.11964034) and Research and innovation Projects (No.2021CXZX-674).

REFERENCES

[1] Gesang Jumian, Gesang Yangjing. Introduction to Tibetan dialect [M]. Ethnic Publishing House, 2002.

[2] Qu Aitang. Complex vowel in Tibetan [J]. Journal of Central University for Nationalities: Philosophy and Social Sciences Edition, 1987 (01): 75-83.

[3] Tan Kerang. Ali Tibetan compound vowel [J]. National languages, 1980 (3): 32-40.

[4] Kong Jiangping. Tibetan dialect questionnaire [M]. Commercial Press, 2011.

[5] Wang Shuangcheng. Authentic compound vowel sound in Amdo Tibetan [J]. Oriental linguistics, 2009 (2).

[6] Wang Shuangcheng. Characteristics of complex vowels in Amdo Tibetan [J]. National language, 2004 (03): 34-37.

[7] Qu Aitang. Tone of Tibetan language and its development [J]. Language studies, 1981 (1): 178-19.