Journal of Sociology and Ethnology (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/jsoce.2024.060304
ISSN 2616-2318 Vol. 6 Num. 3

Analysis of the Acoustic Characteristics of the Stops in Northern Yi Dialects

Shunqiang Chen[1,*], Renyu Gan[1]

[1]*Southwest Minzu University, Chinese Language and Literature College, Chengdu, Sichuan Province, China*

[*]*Corresponding author*

Keywords: Yi language; plosive; Acoustic characteristics

Abstract: This study primarily investigates the plosives of the northern Yi dialect, using the phonetic experiment as the research method to collect their physical signals and analyze their acoustic characteristics. In the study, MATLAB, Praat, and other software are employed to process the speech signal. The findings reveal that: (1) The energy distribution of plosives in the Yi language gathers "straight bars" from low frequency to high frequency, and the root tongue sounds /ga/, /ka/, /gga/, /mga/ exhibit double or even triple straight bars; (2) According to the positive and negative VOT parameters and the length of the VOT parameters, the VOT of /ba/, /pa/, /da/, /ta/, /ga/, /ka/ is positive and all are clear. In contrast, the VOT of /bba/, /nba/, /dda/, /nda/, /gga/, /mga/ is negative, implying voiced articulations. Moreover, /pa/, /ta/, /ka/ are found to be aspirated, while /ba/, /ga/, /da/ are not. (3) /nba/, /nda/, /mga/ all have nasal formants, which are proved to be nasal crown voiced plosives.

1. Introduction

In the acoustic analysis of experimental phonetics, consonants are proved to be more complex than vowels because of their articulatory movements, parts of speech, and methods of articulation. Traditional phonetics depicts the articulatory aspects and methods related to consonants according to the theory and method of "Oral transmission," which is the cornerstone for the subsequent understanding and study of phonetics. However, the lack of empirical evidence results in a somewhat subjective description of phonetic features in traditional phonetics, leading to an insufficiently detailed description of articulatory organs, articulatory parts, articulatory movements, and articulatory patterns.

With the continuous development of computer information technology, especially in the area of experimental phonetics research and related fields such as computer science, physics, physiology, medicine, and psychology, it is now possible to accurately interpret the intrinsic properties of speech. This marks a substantial stride in the comprehension of speech and the study thereof. In the early stages, phoneticians used wave pattern meters, oscilloscopes, and speech mappers to study speech. As research progressed, they began using electromagnetic articulators to monitor the movement of human articulatory organs in real time, electronic glottographs (EGGs) to record the vertical laryngeal movement patterns and vocal folds during vocalization, aerodynamic systems for speech production (APS) to evaluate articulation, and airflow barometers (PAS) to study whether the speech delivery is with or without air. The use of an electronic palatal positionograph (EPG) facilitated the real-time study of consonants. What's more, spiral CT and magnetic resonance imaging (MRI) were used to study both vocal tract shape and vocal tract movement, and high-speed photography allowed for the study of vocal fold vibration. Also, respiratory tapes were used to study respiratory patterns and rhyme. When it comes to electromyography (EMG), it provides insights into articulatory muscle movement, while electroencephalography (ERP) is used to study speech perception.

An overview of research on the phonetics of the Yi language: Mr.Su Lianke's *Research on the Basic Theory of the Yi Audible Language and its Digitization Methods* (2017) focused on the fundamental theory of the physiology and acoustics of the Yi language, as well as the study of digitization methods including the phonetics, script, and vocabulary of the audible language. As for Zhou Xuewen, the study *Acoustic Characteristics of Loose-Tight Laryngeal Vowels in Liangshan Yi* (2013) was based on the "Database of Acoustic Parameters of the Yi Language," and conducted an analysis to study the vowel durations, strengths, pitches, and resonance peaks of the loose-tight laryngeal vowels /u/, /y/, /ur/, and /yr/ in the CV syllables. Lan Zhenqun and Wu Xiyu's *A Study of the Generation Mechanism of Loose-Tight Vowel Oppositions in the Yi Language* (2017) comprehensively examined the generative mechanism of loose and tight vowel oppositions with a focus on the northern dialects of the Yi language. In her study of *The Effects of Clear and Turbid Plug Tones on the Realization of Tones in Yi* (2011), Wang Bei primarily analyzed the effects of clear and turbid plug tones on pitch across high, middle, and low tones of the Shengzha dialect of the Yi. Peng Chunfang's *Phonological Analysis of the All-Turbid Tones of Yi* (2010) conducted an acoustic analysis of the raw vowels of the turbid tones in the Yi language. In their work *Acoustic Performance of Loose and Tight Vowels in Southern Yi* (2005), Shi Feng and Zhou Decai mainly examined the acoustic performance of loose and tight vowels in the southern dialects of the Yi language from the perspective of the fundamental frequency energy of the sound, resonance peak values, as well as the pitch, length, and strength of these sounds. Chen Shunqiang's *Acoustic Characteristics of Vowels of the Northern Dialects of Yi and Their Related Techniques* (2021), applied the methodology of speech experimentalism to obtain physical signals of vowels in the northern dialects of the Yi language. The research investigates the connection between tongue height and vowel laxity. These scholars have transformed the era of introspective perception study of the Yi language into a scientific empirical study of phonetics by intersecting the study of the Yi language phonetics with computer science, physiology, psychology, brain science, and many other fields.

2. Research techniques and methods

In the acoustic analysis of consonants, it is not sufficient to merely interpret, describe, and categorize them based on the place and the method of articulation; It is also necessary to delve into the relationship between "the overtonal characteristics of consonants and the tracks, the resonance peaks of vowels, and the tongue position on the speech map."[1] Also, it is crucial to analyze the voicing and aspiration based on the Voice Onset Time (VOT), which refers to the time from the release of the obstruent to the onset of the vowel, as well as whether or not the de-blocking is associated with a sender segment to determine aspiration. For instance, "the vocal folds begin to vibrate immediately after de-blocking with zero VOT, vocal folds vibration begins after a certain period following de-blocking with positive VOT, the vocal fold vibration occurs before de-blocking with negative VOT, and the turbulent VOT is generally negative, whereas the clarity of the vocal folds exhibits a positive VOT."[2] The presence or absence of the spectrum of impulses produced by the rupture of a stop consonant in the speech map, the straightening bar (spike), is used to determine whether it is a stop consonant. In addition, the presence or absence of turbulence in the mouth during articulation, the turbulence (fills), is used to identify fricative consonants.

2.1 Corpus design

The northern dialects of Yi include 43 consonants, which are also phonemes. “Consonants are categorized into five types according to the different methods of articulation: stops, stop-fricatives, nasals, affricates and fricatives.”[3] According to the movement state of the vocal cords in the process of pronunciation, i.e., whether the vocal cords remain steady or exhibit vibration, consonants are clear when they are pronounced without trembling; Conversely, consonants are turbid when pronounced with trembling of the vocal cords. This theoretical study concludes that the consonants of the Yi language exhibit a well-defined contrast between clear and turbid articulation during pronunciation. There are two kinds of consonants, and they can also be categorized as air-feeding or non-air-feeding according to the strength of the airflow. This paper focuses on and analyzes the bilabial consonants,as shown in Table 1: “/b/, /p/, /bb/, /nb/,” the medial consonants “/d/, /t/, /dd/, /nd/,” the root consonants “/g/, /k/, /gg/, /mg/,” and the root consonants “/g/, /k/, /gg/, /mg/.” These consonants were extracted for acoustic parameters and experimental analysis.

Table 1: Plosives of the northern dialects of the Yi Language (Yi corpus used for this study)

Place of articulation	Unaspirated	Aspirated	Voiced consonant	Nasal crown fricative
Bilabial consonant	ba	pa	bba	nba
Blade-alveolar	da	ta	dda	nda
Velar	ga	ka	gga	mga

Note: The Yi stops /b/, /p/, /bb/, /nb/, /d/, /t/, /dd/, /nd/, /g/, /k/, /gg/, and /mg/ used in this study are spelled with the vowel /a/.

Consonants are articulated when airflow is blocked by the articulatory organs, which is one of the biggest differences between consonants and vowels. Stop consonants are a type of bursting sound, which is produced by the collection of airflow and bursting of the sound after a complete blockage is formed during the three articulatory phases: blocking phase, holding phase, and de-blocking phase. The acoustic characteristic of the stop consonant is represented by the appearance of a straight bar.

Pronunciators were selected with two main criteria in mind. Firstly, people with the standard pronunciation of the language were chosen. Secondly, individuals with pure, unblemished pronunciation and no voice lesions were selected. As such, two female and two male voices were selected for this project of study.

Female 1, a 35-year-old native of Xide County, Liangshan, Sichuan Province. She works as a teacher, as well as a doctoral student, of Chinese minority language and literature (Yi language and literature) at the College of Chinese Language and Literature, Southwest Minzu University. Native from the Shengzha dialect area, where people speak the northern dialect of the Yi language, she has a pure pronunciation, which is suitable for the collection of speech signals.

Female 2, a 22-year-old native of Zhaojue County, Liangshan, Sichuan Province. She is a student majoring in Chinese Minority Language and Literature (Yi Language and Literature) in the College of Chinese Language and Literature, Southwest Minzu University. Native from the Shengzha dialect area, she also has a pure pronunciation, which is suitable for the collection of speech signals.

Male 1, a 24-year-old native of Xide County, Liangshan, Sichuan Province. He is a student majoring in Chinese Minority Language and Literature (Yi Language and Literature) in the College of Chinese Language and Literature, Southwest Minzu University. Native from the Shengzha dialect area, he has a pure pronunciation, which is suitable for the collection of speech signals.

Male 2, a 23-year-old native from Zhaojue County, Liangshan, Sichuan Province. He is a student majoring in Chinese Minority Language and Literature (Yi Language and Literature) at the College of Chinese Language and Literature, Southwest Minzu University. Native from the Shengzha dialect area, he also has a pure pronunciation, which is suitable for the collection of speech signals.

2.2 Voice Signal Acquisition

The sampling rate of the high and low frequencies determines the sound quality. Typically for vowel analysis, a rate of around 10k is sufficient, while for consonant analysis, a higher rate, ranging from 16k to 22k, is acceptable. Note that the research voice sampling rate of 44k-48k is optimal. Considering that the voice is acoustically analyzed later, the sampling rate of this experiment is set at 48k, and the voice data is stored in the *.wav format.

Nowadays, there are more software technologies to generate speech maps, analyze speech spectrograms, and conduct precise speech analysis, such as MATLAB, Praat, and AS. In this experiment, the speech signal processing software MATLAB is used to pre-process the captured speech data. Then Praat and AS are used to analyze the sampled information and extract the duration and pitch of the consonants, the VOT duration, and the voiceless segment duration (GAP).

2.3 Speech Map Production

Drawing upon the characteristics of speech signals, speech maps are classified into narrowband and wideband speech maps. According to the goals of speech acoustic analysis research, wideband speech maps can be used to analyze signals such as resonance peak crossbars, turbid crossbars of vowels, fricative frequencies, and impulse straight bars, while narrowband speech maps can be used to analyze harmonic signals. Figures 1 and 2 show the consonant /ba/ speech maps.

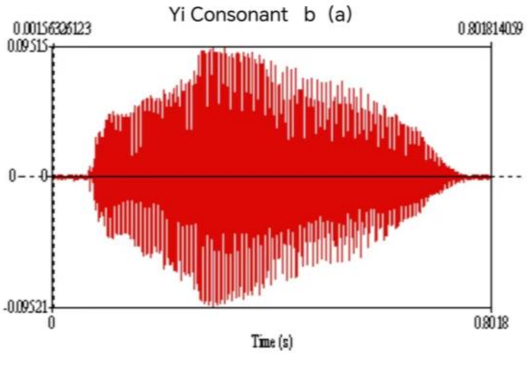

Figure 1: Waveform of the Yi consonant /b/

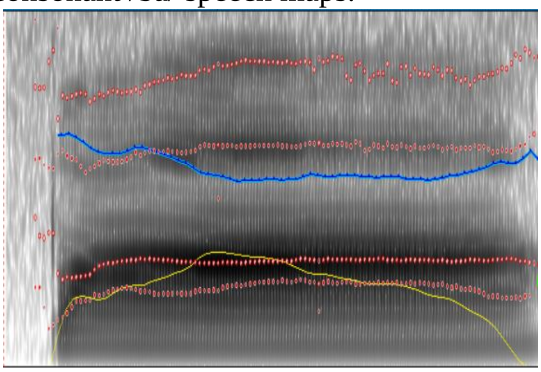

Figure 2: Spectrogram of the Yi consonant /b/

3. Acoustic analysis of plug consonants

3.1 Consonant /b/ acoustic parameter analysis

It can be seen in Figure 3: the speech map for the consonant /ba/ starts with a vertical bar of energy from low to high frequencies. The straight bar corresponds to the burst sound of /b/ when the articulatory organ is de-blocked. The whole chart starts with the energy map of the burst phase of the consonant /b/, followed by the energy map of the resonance peak of the vowel /a/. The duration of the deblocking phase is relatively short, and there is an energy concentration in the low-frequency region, which is between 2000Hz-3000Hz. The overall energy levels in other frequency ranges are weaker and boast a positive VOT.

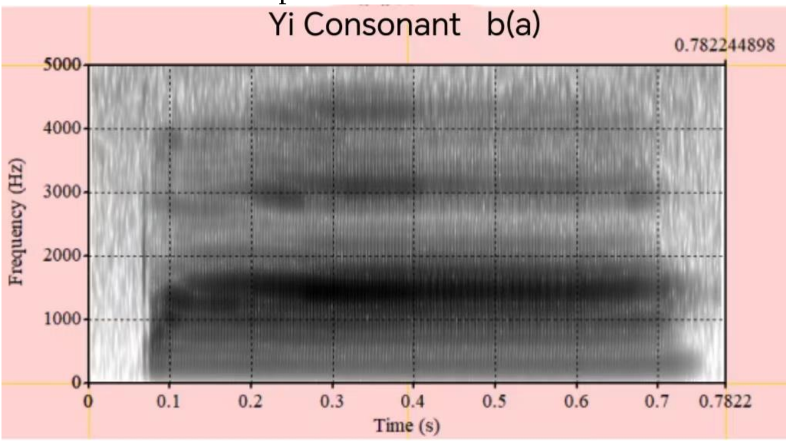

Figure 3: Spectrogram of the Yi consonant /ba/

3.2 Consonant /p/ acoustic Parameter Analysis

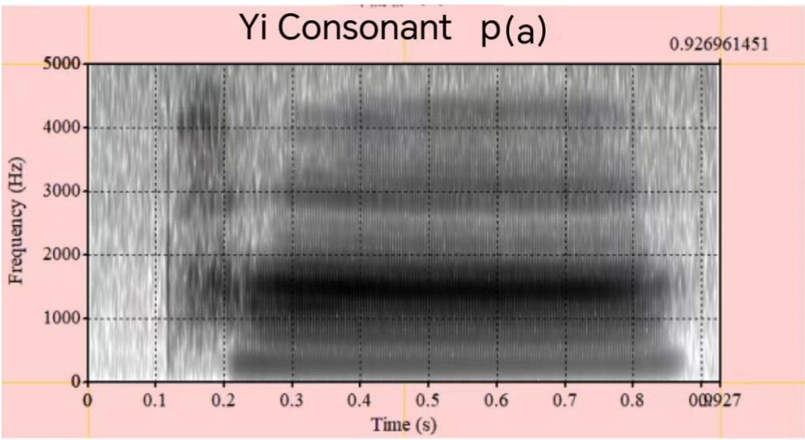

Figure 4: Spectrogram of the Yi consonant /pa/

It can be seen in Figure 4: the beginning of the consonant /pa/ spectrogram shows a straight bar with energy ranging from low frequency to high frequency. The whole starting section of the spectrogram represents the energy map of the burst tone of the consonant /p/. Following this, there is a longer duration of the airflow delivery phase. In this section, there is a concentration of energy in the low frequency at 1000Hz to 2000Hz and in the high frequency at 3000Hz-4000Hz. Therefore, the VOT is positive and the duration is longer. Subsequently, the spectrogram displays the energy map of the resonance peak of the vowel /a/.

3.3 Consonant /bb/ acoustic parameter analysis

It can be seen in Figure 5: the consonant /bba/ speech spectrogram begins with a section of pre-turbulence, indicated by a distinct voicing bar. Following this, there are a series of vertical bars with energy spanning from low to high frequencies. Energy concentration occurs at 500Hz and 1500Hz in the lower frequency range, as well as at 3200Hz in the higher frequency range. This is followed by the onset of the voice pronouncing the vowel /a/ and the VOT is negative. There is a representation of the resonance peaks for the vowel /a/.

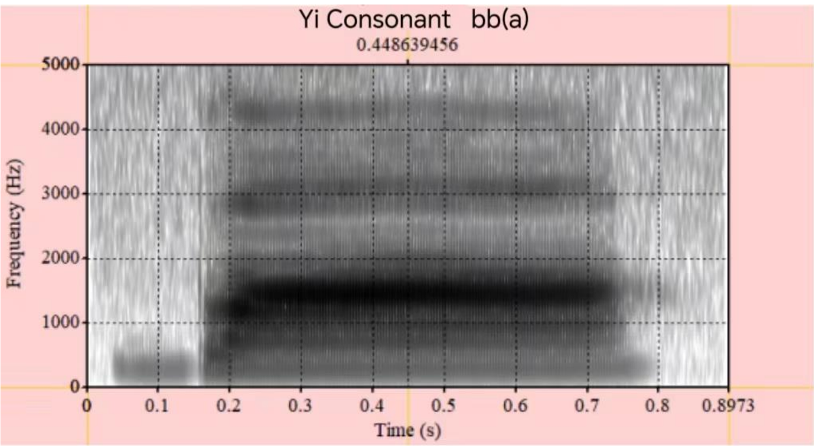

Figure 5: Spectrogram of the Yi consonant /bba/

3.4 Consonant /nb/ Acoustic Parameter Analysis

It can be seen in Figure 6: that the consonant /nba/ has a pre-turbulence at the beginning of the spectrogram, and there is an obvious turbulence bar. Due to the bifurcation of the vocal tract into the oral and nasal passages, the resonance of the vocal tract appears as an "anti-resonance peak", and there are obvious nasal resonance peaks in the pre-turbulence at the low-frequency band of 1000Hz and the high-frequency band of 3200Hz. Then there is a straight bar of energy spanning from low frequency to high frequency, and the energy of the straight bar is concentrated at 1500Hz in the lower frequency range, followed by the beginning of the voice of the vowel /a/, resulting in a negative VOT. After that, there is the resonance peak of the vowel /a/ energy map.

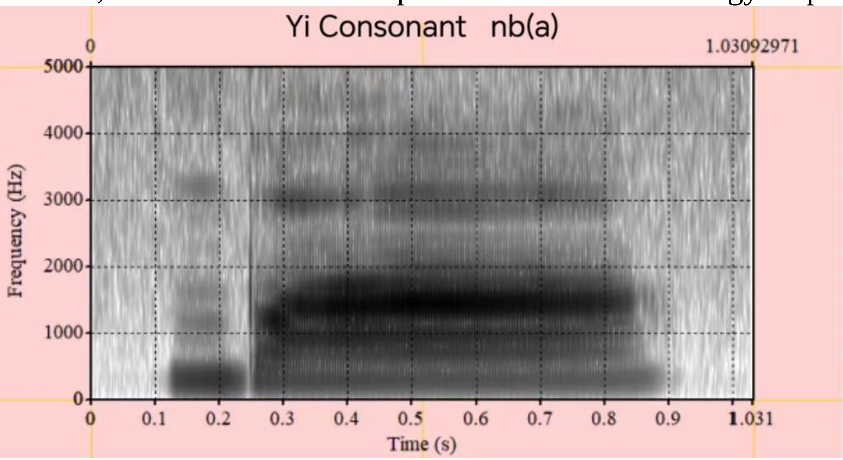

Figure 6: Spectrogram of the Yi consonant /nba/

4. Experimental Data and Analysis

Table 2: Acoustic Parameters of Plug Consonants (Bilabial, Labiodental) in the Northern Yi Dialect

Consonants	GAP(s)	VOT(s)	CD(s)	CA(dB)	F1(Hz)	F2(Hz)	F3(Hz)	F4(Hz)
b	0.041	0.0042	0.0452	67	1070	1540	3050	4120
p	0.115	0.0819	0.1969	64	1160	1570	3020	4160
bb	0.126	-0.0021	0.1239	66	937	1630	2920	4070
nb	0.141	-0.0019	0.1391	63	900	1620	3020	4080

Table 3: Acoustic Parameters of Plug Consonants (Mid-Tongue, Upper Alveolar) in the Northern Yi Dialect

Consonants	GAP(s)	VOT(s)	CD(s)	CA(dB)	F1(Hz)	F2(Hz)	F3(Hz)	F4(Hz)
d	0.039	0.0038	0.0428	63	1060	1560	2740	3780
t	0.178	0.0786	0.2566	64	1070	1750	3000	3800
dd	0.128	-0.0023	0.1257	66	840	1620	2640	3690
nd	0.152	-0.0021	0.1499	64	890	1720	2710	3860

Table 4: Acoustic Parameters of Plug Consonants (Velar Root and Velum) In the Northern Yi Dialect

Consonants	GAP(s)	VOT(s)	CD(s)	CA(dB)	F1(Hz)	F2(Hz)	F3(Hz)	F4(Hz)
g	0.040	0.0040	0.0440	65	1040	1810	3010	3960
k	0.182	0.0778	0.2598	63	1030	1740	2950	4000
gg	0.131	-0.0022	0.1288	65	860	1750	2960	4100
mg	0.156	-0.0019	0.1541	63	920	1720	3050	4210

Tables 2, 3, and 4 above are the acoustic experimental analysis data of the stops in the northern dialect of Yi language.

"The GAP has a stop and affricate consonants before their release, which is a manifestation of the consonant-forming and block-holding periods, causing the effect of tenuis. Although this phase appears as a gap, it is indispensable for stop consonants. Resistance-forming section= pre-resistance-forming section + post-resistance-forming section, GAP = post-resistance-forming section + resistance-holding section, and consonant length CD = GAP + VOT."[4]

5. Conclusion

From the tables 2, 3, and 4 above acoustic analysis parameters and experimental data, it can be seen that in the Yi language, stop consonants will have energy gathering of varying strengths and weaknesses from the low frequency to the high frequency when they are articulated in a burst. Moreover, for uvular sounds such as "ga", "ka", "gga", and "mga," they have a large obstruction area during the process of articulation, and the airflow is not as strong as it should be. Due to the relatively large area of the root of the tongue in the process of articulation, the airflow out of the blast will appear as double-impulse straight bars. In the case of male speakers producing "gga" "mga," three bursts can even occur. Such as Figure 7 "ga" was shown to have double bursts, and in Figure 8 "mga" shows three bursts.

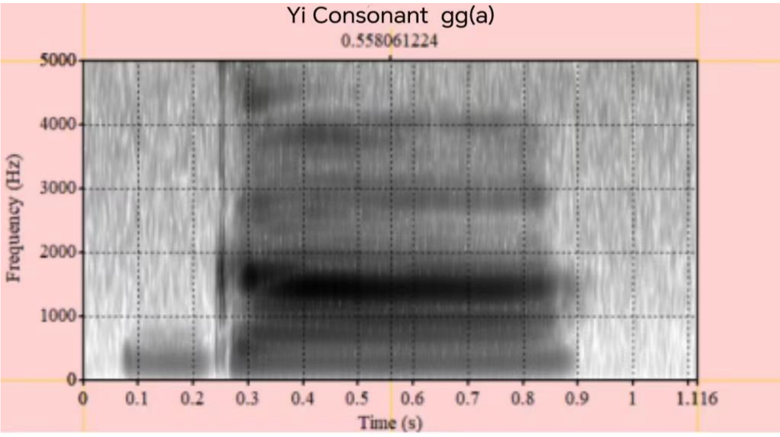

Figure 7: Spectrogram of the Yi consonant /ga/

From the above speech map and acoustic parameters, it can be learned that the consonants /ba/, /pa/, /bba/, /nba/, /da/, /ta/, /dda/, /nda/, /ga/, /ka/, /gga/, and /mga/ all have distinctive straight bars. They can be considered as stop consonants according to Mr. Kong Jiangping's statement, which "The acoustic nature of the stop consonants are mainly reflected in the straightening bars."[5] Based on the above VOT data, we can judge that the positive VOT values of /ba/, /pa/, /da/, /ta/, /ga/, /ka/ are all voiceless consonants, and the negative VOT values of /bba/, /nba/, /dda/, /nda/, /gga/, /mga/ are voiced consonants. Further analysis is conducted on the length of the VOT parameter to determine whether it is aspirated or not. The size of the VOT parameter of the above consonants is in the order of /p/>/t/>/k/>/b/>/g/>/d/. Together with the characterization of the spectrograms, the consonants /p/, /t/, /k/ are aspirated consonants, and /b/, /g/, and /d/ are unaspirated consonants. Thus, it empirically proves the viewpoints proposed by Mr. Su Lianke's *Research on the Basic Theory of Yi Voiced Language and its Digitization Method* (2017) as shown in Table 5.

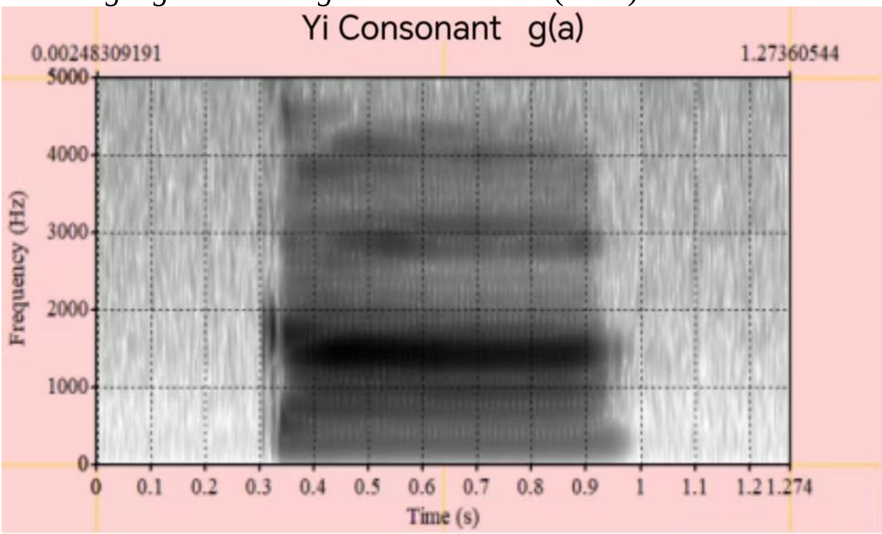

Figure 8: Spectrogram of the Yi consonant /gga/

Table 5: Plug consonants of the northern dialects of Yi

Place of articulation	Clear plosive		Turbidity stop sound	
	Unaspirated	Aspirated	Voiced consonant	Nasal crown fricative
Bilabial,labiodental	b	p	bb	nb
Alveolar,upper alveolar	d	t	dd	nd
Velar Root,soft Velum	g	k	gg	mg

Chart information from Mr. Su Lianke's Research on the Basic Theory of Yi Voiced Language and its Digitization Method (2017)

According to Mr. Kong Jiangping (2015), "the acoustic nature of nasal sounds is reflected in nasal resonance peaks," combined with the above speech map and acoustic parameters, it can be observed that the three sounds /nba/, /nda/, /mga/ all have "anti-resonance peaks (the phenomenon of nasal resonance peaks.)" This allows us to conclude and further confirm that /nba/, /nda/, and /mga/ are nasal crown turbid stops. Spectral map of Yi consonants are shown as Figure 9-11.

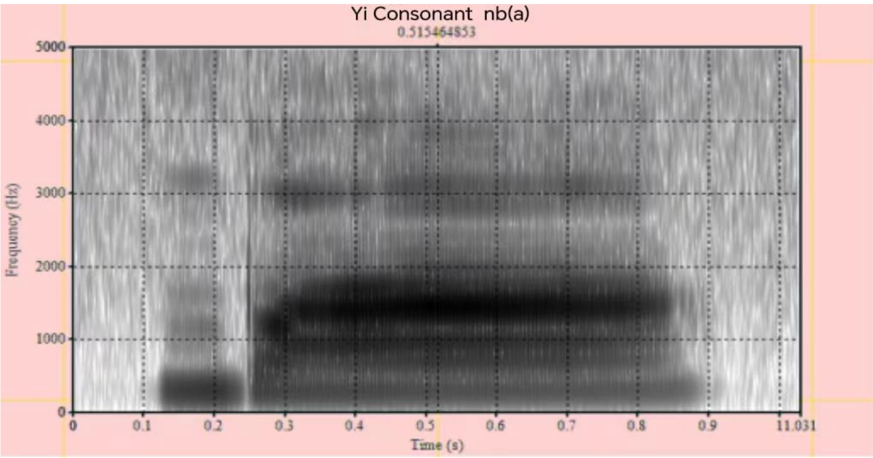

Figure 9: Spectrogram of the Yi consonant /nba/

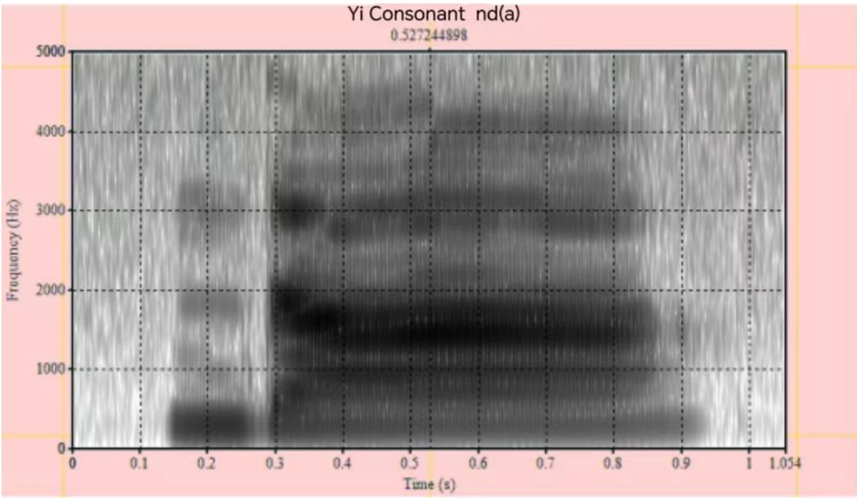

Figure 10: Spectrogram of the Yi consonant /nda/

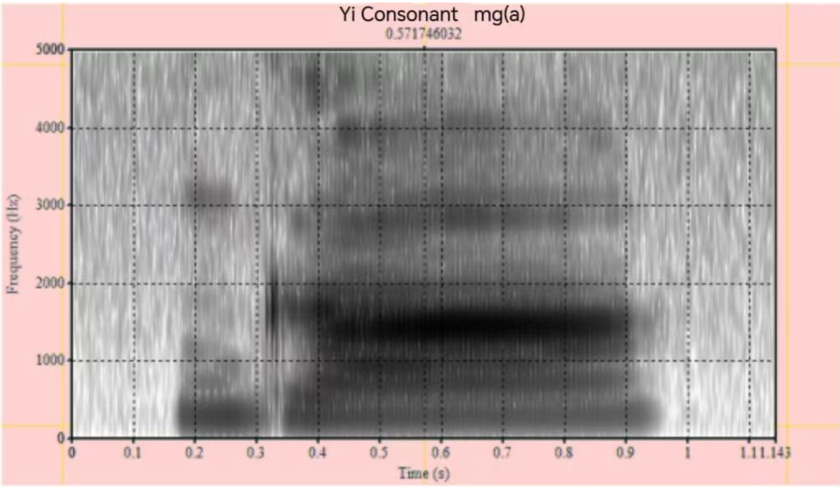

Figure 11: Spectrogram of the Yi consonant /mga/

Acknowledgement

Project Fund: General projects of the National Social Science Foundation: Research on the

Acoustic Speech Database of Yi Voice and its Vocalization Types and Rhythmic Characteristics (22BYY175).

References

[1] Gegen Tana and Hu A Xu, A Study on the Acoustics of Speech, Voice and Dynamic Palatal Position of Alashan Mongolian, Science Publishing House, September 2020. P 10.
[2] Kong Jiangping, On Speech Vocalization, Minzu University of China Press, November 2001. P 73-74.
[3] Su Lianke, Wu Bing, Chen Shunqiang, DaXi Aqie. Research on the Basic Theory of Yi Audible Language and Its Digitization Methods, Minzu Publishing House, August 2017. P 25.
[4] Bao Huaiqiao, "An Introduction to the Physiological and Acoustic Analysis of Mandarin Speech (Continued from 2)", Journal of Audiology and Speech Diseases, October 2004 (No.5), P 357-359.
[5] Kong Jiangping, A Basic Tutorial on Experimental Phonetics, Peking University Press, August 2015. P 106.