Computational Research on the Recitation Intonation of Hezhou Huaer

-- A Case Study of "Hezhou Sanling"

Li Ma[a], Yonghong Li[b*]

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

[a]2813808819@qq.com, [b*]lyhweiwei@126.com

Abstract. Based on the research method of intonation pattern, this paper quantitatively analyzes the data of pitch, duration, and intensity of the two chapters and four sentences of "Hezhou Sanling" recited in Hezhou dialect through experimental calculation, and uses statistical data to create tables and graphs to visually demonstrate the recitation intonation of "Hezhou Sanling""and discover its recitation rules, illustrating the strong rhythmic matching that exists when reciting Huaer Quling in Hezhou dialect.

Keywords: Intonation patterns; pitch fluctuation scale; duration ratio; volume ratio.

1 Introduction

By leveraging the powerful computing capabilities and intelligent algorithms of computers, the study of phonetics can be made more efficient, transforming the inherently invisible language into a visually intuitive research ontology. Hezhou Huaer is a kind of music song, including two parts of music and literature. The chapters and sentences of Hezhou Huaer are restricted by the melody of Huaer, and its basic style is four lines per song, and there are also new styles which called "broken waist" type, six or five lines per Huaer, has developed and evolved on the basis. Some singers will recite the lyrics before singing in order to make it easier for the audience to understand the content. The lyrics of Hezhou Huaer follow a certain rhyme and meter, making them catchy and enjoyable to read. They are filled with rhythm and beats, and sounds rhythmic as well. While Hezhou Huaer is not strictly defined as poetry in the traditional sense, its basic format consists of four lines per song, divided into two chapters.[2][3]

From the perspective of experimental phonetics, this paper utilizes speech software to extract suprasegmental speech information data such as pitch, duration, and energy. In the framework of intonation pattern, parameters such as pitch fluctuation scale, duration ratio, volume ratio are employed to quantitatively analyze the reciting intonation of Hezhou Huaer. We try to use data and intuitive graphics to show the unique phonological features when reciting the text of Hezhou type flower song in Hezhou dialect,

Y. Kuang et al. (eds.), *Proceedings of the 2024 5th International Conference on Education, Knowledge and Information Management (ICEKIM 2024)*, Atlantis Highlights in Computer Sciences 22,
https://doi.org/10.2991/978-94-6463-502-7_121

and scientifically prove that reciting the text of Hezhou type flower song in Hezhou dialect is regular and worthy of attention and research. [5][12][13]

2 The Sound and Tone System of Hezhou Dialect

Based on the investigation and research of the predecessors, this paper summarizes the sound and tone of Hezhou dialect in Gansu Province.

Table 1. Statistical table of phonological tones of 6 main representative points of Hezhou dialect in Gansu

Serial number	Representative point	Number of consonants	Number of vowel	Number of monosyllabic tones
1	Chengguan dialect of Guanghe County	23	32	3
2	Chengguan dialect of Hezheng County	24	31	3
3	Old Chengguan district dialect of Linxia	24	31	3
4	Hui dialect of Linxia eight square	24	31	3
5	Hanji dialect of Linxia county	24	31	3
6	Chui Ma Tan dialect of Jishishan county	24	29	3

Based on the above statistical tables, it can be seen that the phonological system of Hezhou dialect is relatively simple, typically consisting of 24 initials and 31 finals. Moreover, the finals tend to have a high degree of monophthongization. The dialect in Hezhou generally has three tones which are distributed in most dialect areas of Hezhou: level tone, rising tone, and falling tone. There is no zigzag tone. However, when the singers recite the Hezhou huaer in Hezhou dialect, the intonation is melodious and catchy, which has the aesthetic feeling of poetry. This paper will use detailed data, from the Hezhou dialect of the suprachedral features, to reveal its inherent law. [4]

3 Research Method

3.1 Pronunciation Corpus and Speakers

The obvious phonetic feature of reciting Hezhou Huaer in Hezhou dialect is "prolonged sound". Typically, in two chapters and four sentences of Hezhou Huaer, the prolonged sound is present in the first sentence, the third sentence, or simultaneously in both the first and third sentences. In order to better demonstrate the rhythmic characteristics of Hezhou Huaer melody, this article selects “Hezhou Sanling”as the experimental subject.

When reciting “Hezhou Sanling”in the Hezhou dialect, the pronunciation of the final words in the first and third sentences is elongated and rhymed, while the second and fourth sentences are read rhythmically and also rhyme.

The recorded corpus used in this paper is “Hezhou Sanling”, reciting in Hezhou dialect by Ma Yonghu, the inheritor of Huaer in Linxia, which was conducted in a professional recording studio with a sampling rate of 11025Hz.

3.2 Processing of Experimental Data

This article utilizes the Praat software to conduct acoustic measurements on the corpus, obtaining pitch (Hz), duration (ms), and intensity (amplitude product) for each character. Through computational formulas, these three parameters are converted into corresponding pitch fluctuation scale, duration ratio, volume ratio respectively. Based on these measurements, a single-item cross-matching analysis is conducted, and the Pearson correlation coefficient is calculated using the SPSS software program.[5][7]

4 Calculation of Pitch Fluctuation Scale, Duration Ratio and Volume Ratio

4.1 Pitch Fluctuation Scale Result

Pitch fluctuation scale measurement is an important quantitative indicator for intonation analysis, which can effectively reflect the change of pitch and the degree of fluctuation in speech. In this paper, Praat software is used to extract the fundamental frequency value of each word. Subsequently, these fundamental frequency data are converted into semitone values, and then the tone domain data is obtained according to the percentage data, that is, the range of pitch change. Semitone value is a kind of measuring unit suitable for intonation research, which is closer to people's psychological auditory perception.[1][11]

To convert the Hz value to a semitone value, the conversion formula is:

$$St=\frac{12*lg\frac{f}{fr}}{lg2} \tag{1}$$

In computer language, input = 12*LOG (f/fr, 2). In this equation, "f" represents the Hz value that needs to be converted, while "fr" represents the reference frequency, which is set to 55 Hz for males and 64 Hz for females.

After converting the Hz value to a semitone value, the semitone value is then converted to a percentage value. The calculation of the percentage value is a relative normalization method, and the computation method is as follows:

$$K=100*\frac{G-Smin}{Smax-Smin} \tag{2}$$

Here, G represents the semitone value of a certain point on a certain character, Smax is the upper limit semitone value of the sentence's pitch range, Smin is the lower limit semitone value of the sentence's pitch range, and K is the percentage of the pitch range for the character that we are calculating.[6]

After conducting measurements, Table 2 was obtained.

Table 2. Percentage values for each syllable

	First	Second	Third	Fourth	Fifth	Sisth	Seventh	Eighth	Nineth
Upper boundary of pitch range	73.1%	82.6%	57.7%	60.8%	100%	69.1%	43.2%	38.3%	38.3%
Lower boundary of pitch range	18.3%	17.8%	17.8%	28.6%	34.2%	29.9%	8%	0%	30.8%

As can be seen from Table 1, the key domains of each word go online and offline, and the key domains of word 1, word 2, and word 5 have a large difference. According to Table 1, we can get a graph of the fluctuation of the recitation of "Hezhou Sanling".

By looking at Table 1, we can see the upper and lower boundaries of the pitch range for each character. Among them, the pitch range differences of the first, second, and fifth characters are relatively large. Based on Table 1, we can obtain a graph of pitch fluctuation scale when reciting "Hezhou Sanling".

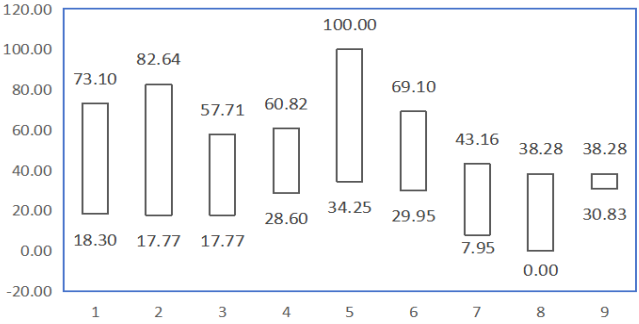

Fig. 1. The pitch fluctuation scale graph.

Through Figure 1, we can visually observe the Pitch fluctuation scale when the speaker recites ""Hezhou Sanling"." Among them, the pitch range of the first, second, and fifth characters is relatively large. Furthermore, the fluctuation amplitude of the upper line is greater than that of the lower line. Overall, the trend is downward. Therefore, when reciting Hezhou huaer, the tones exhibit distinct characteristics of rising and falling cadences.

4.2 Duration Ratio Result

The measurement of duration ratio is the key quantitative indicator in the duration dimension of intonation analysis. Through the measurement of duration ratio, we can better understand the intonation characteristics of speakers.[1][8]

Duration ratio calculation formula is:

$$Dx=\frac{Sx+Gx}{S\#} \tag{3}$$

The quantitative indicator, where Dx represents the duration ratio of a specific syllable, Sx stands for the duration of syllable x within a combination of syllables, Gx

denotes the pause that occurs after the syllable (within a single sentence, syllables are usually prolonged rather than paused, so this value is typically zero), and S# refers to the average duration of the syllable combination. If the duration ratio is greater than 1, it is considered as a segmental prolongation. Essentially, this is a normalization of duration.

The duration ratio of a specific character = the duration of that syllable / the average duration of all syllables in the sentence.

After measuring the duration of all syllables in the four sentences of the upper and lower chapters of the recitation of "“Hezhou Sanling”," we can obtain the duration ratio of each syllable within each sentence.

Table 3. Duration ratio

	First syllable	Second syllable	Third syllable	Fourth syllable	Sixth syllable	Seventh syllable	Eighth syllable	Ninth syllable	Tenth syllable
First sentence	1.19	0.83	0.28	1.04	0.72	0.61	0.40	0.67	3.26
Second sentence	1.50	1.23	0.44	0.98	0.81	1.01	0.94	1.09	
Third sentence	0.92	0.80	0.39	0.76	0.76	0.66	0.82	0.77	3.12
Fourth sentence	1.63	0.90	0.64	1.65	0.54	0.92	1.06	0.66	

Based on the duration ratio data in Table 2, we can clearly observe that the duration ratios of the final syllables in the first and third sentences are 3 to 5 times longer than the duration ratios of the other syllables, indicating that there is no prolongation of sound segments in those syllables. However, the second and fourth sentences exhibit varying degrees of sound segment prolongation. This analysis provides a deeper understanding of the rhythmic structure and pronunciation patterns within the recitation of “Hezhou Sanling”.

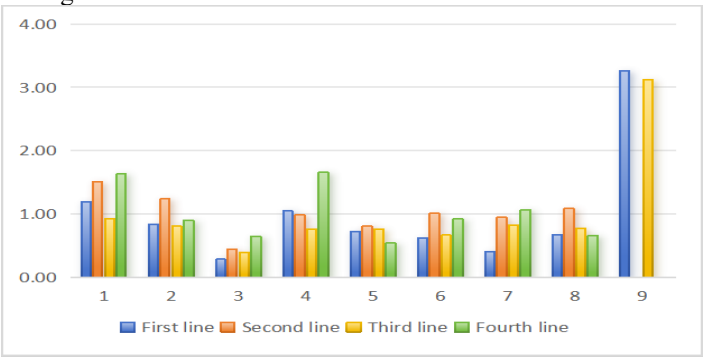

Fig. 2. The duration ratio graph.

From Figure 2, we can see that the duration ratio of the first and third sentences fluctuates more significantly, indicating that the speech speed is also faster. The duration ratio of the second and fourth sentences is relatively stable, suggesting a more moderate speaking pace compared to the first and third sentences.

4.3 Volume Ratio Result

Volume ratio measurement is an important quantitative indicator in intonation analysis, specifically related to the intensity aspect of speech sounds. It allows us to gain a deeper understanding of the intensity characteristics of speech sounds. The Praat software can extract amplitude product, which is directly proportional to the intensity and duration of the selected syllable.[1][9][10]

The formula for calculating the volume ratio is:

$$\text{A word volume ratio} = \frac{\textit{the amplitude product}}{\textit{the average amplitude product of all syllables}} \quad (4)$$

The resulting value represents the relative magnitude of the volume for each syllable. In this way, we can convert the amplitude product, which is influenced by multiple factors, into comparable and relative volume ratio values. Through measurement, we can obtain a chart of volume ratios.

Table 4. Volume ratio

	First syllable	Second syllable	Third syllable	Fourth syllable	Sixth syllable	Seventh syllable	Eighth syllable	Ninth syllable	Tenth syllable
First sentence	1.19	0.72	0.13	1.00	0.43	0.53	0.09	0.37	4.55
Second sentence	1.20	1.29	0.13	0.94	0.41	0.78	1.96	1.30	
Third sentence	1.35	0.97	0.21	0.76	0.91	0.36	0.33	0.34	3.76
Fourth sentence	2.19	1.83	0.60	1.44	0.21	1.17	0.44	0.12	

According to the volume ratio data in Table 3, the volume ratio of the last syllable in the first the third sentence is 3 to 5 times higher than that of other words, while the volume ratio of other syllables is basically less than 1. In terms of the volume ratio of the second and fourth sentences, the energy is relatively average. We use the data in Table 3 to create a more intuitive line chart of volume ratio.

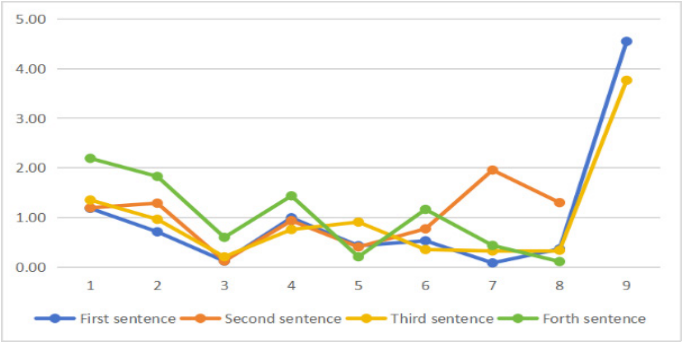

Fig. 3. The volume ratio graph.

Through Figure 3, it can be visually observed that the volume ratio of the last syllable in the first and third sentences is significantly greater than that of the other characters.

The volume of the other characters is relatively smaller compared to the second and fourth sentences. The volume of the first, third, and fourth sentences exhibits a clear downward trend compared to the initial volume, while the second sentence is the opposite.

4.4 Correlation Analysis Among Pitch Fluctuation Scale, Duration Ratio and Volume Ratio

From the data analysis of the recitation of the two chapters and four sentences of the song of "Hezhou Sanling" from the aspects of pitch fluctuation scale, duration ratio and volume ratio. It can be observed that there are specific patterns in reciting the Hezhou Huaer in Hezhou dialect. Below, we will conduct a correlation analysis among pitch fluctuation scale, duration ratio and volume ratio, using the Pearson correlation test in SPSS.[5] The results are presented in Table 5.

Table 5. Correlation among pitch fluctuation Scale (PFS), duration ratio (DR) and volume ratio (VR)

Category			PFS	DR	VR
First sentence	PFS	Pearson correlation	1	-0.197	-0.165
		Significance (Two-tailed)		0.611	0.671
	DR	Pearson correlation	-0.197	1	.996**
		Significance (Two-tailed)	0.611		0
	VR	Pearson correlation	-0.165	.996**	1
		Significance (Two-tailed)	0.671	0	
Second sentence	PFS	Pearson correlation	1	-0.02	0.028
		Significance (Two-tailed)		0.963	0.947
	DR	Pearson correlation	-0.02	1	.730*
		Significance (Two-tailed)	0.963		0.025
	VR	Pearson correlation	0.028	.730*	1
		Significance (Two-tailed)	0.947	0.025	
Third sentence	PFS	Pearson correlation	1	-0.073	0.414
		Significance (Two-tailed)		0.863	0.308
	DR	Pearson correlation	-0.073	1	.965**
		Significance (Two-tailed)	0.863		0
	VR	Pearson correlation	0.414	.965**	1
		Significance (Two-tailed)	0.308	0	
Fourth sentence	PFS	Pearson correlation	1	-0.004	0.188
		Significance (Two-tailed)		0.993	0.655
	DR	Pearson correlation	-0.004	1	.734*
		Significance (Two-tailed)	0.993		0.038
	VR	Pearson correlation	0.188	.734*	1
		Significance (Two-tailed)	0.655	0.038	

Based on the correlation coefficients in Table 4, it can be observed that there is a negative correlation between pitch fluctuation scale and duration ratio. Except for the first sentence, there is a positive correlation between pitch fluctuation scale and volume ratio, although the correlation coefficients are all less than 0.5. Additionally, there is a positive correlation between duration ratio and volume ratio, with correlation

coefficients greater than 0.7 indicating a strong correlation. Notably, the correlation between duration ratio and volume ratio is significantly strong for the first and third sentences. These findings suggest that when reciting the Hauer lyrics in Hezhou dialect, there exists an inherent unity and strong correlation among pitch fluctuation scale, duration ratio, and volume ratio, resulting in a high degree of rhythmic matching.

5 Conclusion

This article quantitatively studies and analyzes the "Hezhou Sanling" recited in Hezhou dialect from the perspectives of pitch, duration, and intensity, using acoustic parameters such as pitch fluctuation scale, duration ratio, and volume ratio. It visually demonstrates the intonation characteristics of the Hezhou Huaer melody during recitation. The pitch fluctuation scale reflects the cadence and modulation in recitation, the pause rate reveals the rhythmic sense of recitation, and the volume ratio can reflect the emotional investment and intensity changes during recitation. Therefore, it can be seen that the Hezhou Huaer melody recited in Hezhou dialect has a strong degree of rhythmic matching.

By utilizing experimental phonetic methods and intelligent, efficient computer tools, we can more scientifically reveal the inherent patterns of Hezhou Huaer melodies. This approach not only aids in a deeper understanding of the artistic characteristics of Hezhou Huaer, but also provides clues and evidence for subsequent research.

Acknowledgment

This work was financially supported by the Innovation Team of Northwest Minzu University and NSFC grant fund (No.11964034).

References

1. Shi Feng, Intonation Pattern: The Foundation of Experimental Linguistics [M], The Commercial Press, 2013.
2. Xi Huimin. 1989. Study of Huaer in Northwest China [M]. Lanzhou University Press, Lanzhou.
3. Wang Pei. 1992. A Study on Hezhou Huaer [M]. Lanzhou University Press, Lanzhou.
4. Jianjun Zhang. Study on Phonology of Hezhou Dialect [D]. Shaanxi Normal University,2009.
5. Shi Feng and Wang Ping. A Preliminary Discussion on the Prosodic Hierarchy System of Chinese Language. [J] Journal of Nankai Linguistics, Volume 1, 2014.
6. Shi Lin, Wen Baoying, and Xie Chenwei. The Prosodic Matching of Mandarin Chinese Intonation. [J] Journal of Nankai Linguistics, Volume 2, 2015.
7. Shi Feng, Liang Lei, and Wang Ping. The Fluctuation of Intonation in Mandarin Chinese Declarative Sentences. [J] Journal of Nankai Linguistics, Volume 2, 2009.
8. Wang Ping and Shi Feng. On the Research Methods of Intonation Patterns [J]. Contemporary Foreign Language Studies, Issue 5, 2011.

9. Wang Yi. An Experimental Study on the Intonation of Nanjing Dialect [D]. Nanjing Normal University, 2021.
10. Zhang Siqi. A Study on the Loudness of Mandarin Intonation in Shanghai [D]. Tianjin Normal University, 2018.
11. Liang Lei, Shi Feng. Analysis of Volume Ratio in Mandarin Two-Character Groups [J]. Nankai Linguistics, Issue 2, 2010.
12. LADD D Robert. Stylized intonation[J]. Language，1978.
13. Bolinger Dwight L,1989. Intonation and its Uses[M]. Stanford: Stanford University Press.