A study of the resonance peak characteristics of Mongolian long-key ode

Qina Zhou[1, a], Shu Liu[1, b], Chen Dai[1,c], Yaxin Xi[1,d], and Axu Hu*[,1, e]

[1] Key Research Base of Humanities and Social Sciences of the State Ethnic Affairs Commission; Key Laboratory of Language and Culture Computing, Ministry of Education, Northwest University for Nationalities.

[a]1749489610@qq.com, [b] liushuzkw@foxmail.com, [c] 327032091@qq.com, [d] 2986078238@qq.com, [*,e] axuhu@xbmu.edu.cn

Abstract. The lyrics of the ode are mainly praising ancestors, princes, and nobles, singing about parents, hometowns, and horses, with a solemn style. In 2005, the Mongolian long tune was selected as a world-class intangible cultural heritage, and the study of Mongolian long tunes has become a worldwide issue. At present, domestic, and foreign research on long-key odes mainly focuses on music theory, artistic expression, historical development, cultural connotation, performance methods, etc. The research contents and fields are relatively rich. However, little research has been done on the generation principle and articulation mechanism of the Mongolian long-key ode due to the limitation of research techniques and means. By extracting some physiological acoustic parameters of articulation, establishing an orally transmitted cultural parameter model, and analyzing the relationship between parameters and articulatory actions, this modern phonological method based on computer and language science to study the fundamental mechanism of long-tune vocalization can deepen the theory of speech vocalization, explore the mechanism of speech vocalization. The study has found a new application for modern phonology accumulated over the years, which can help to preserve culture digitally and promote the presentation of cultural multimodality.

Keywords: Mongolian Long -key ode; Cultural connotation.

1. Introduction

The oral culture of any ethnic group has gradually developed over millions of years of human evolution. Therefore, the process of the emergence and development of this “cultural species” is not inferior to the formation of natural species, and it is a precious wealth of human wisdom. In 2005, the Mongolian long song was selected as a world-class intangible cultural heritage, making the study of the Mongolian long song a global issue. “Ode” in Mongolian is called “Sasiltiir Dawaa,” and the lyrics of the song mainly praise the ancestors, and the merits of nobles and royals, sing about parents, hometown, and horses, and have a solemn and restrained style.[1] Due to their solemnity, long songs generally have time and space limitations and cannot be sung at will. Representative works include “The Rich and Vast Alashan,” “The Mountains and Rivers of the State,” “The Rising Sun of Alashan,” and others. Regarding the study of Mongolian long songs, whether domestic or foreign, it all started with the collection and compilation of Mongolian folk songs.

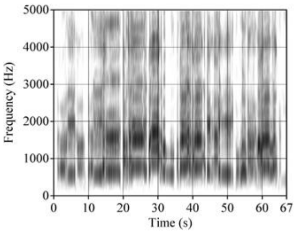
“Singing resonance peak phenomenon,” also known as “resonance peak,” is defined as the pulsation effect of the vocal tract. If the vocal tract is viewed as a resonant cavity, the resonance peak is the resonant frequency of this cavity. In computer music, the resonance peak is an important parameter that determines the timbre and quality of sound. Currently, the language description of the “singer's resonance peak” can be summarized as follows: the singer’s resonance peak refers to a resonance peak that appears in the frequency range of 2200Hz-3200Hz. Its existence can enhance the brightness and penetration of the singer's voice, preventing it from being overshadowed by the accompaniment or other sounds.[2] Studying Mongolian long songs from the perspective of singing acoustics is a new research trend in the field of long song research over the past decade, belonging to interdisciplinary research. This provides more possibilities and feasibility for long songs research and protection, especially with the application of scientific instruments, which provide scientific data as a basis for long songs research. For example, Bascom in the United States summarized in “The Physical Definition of Beautiful Singing Voices for Men and Women” that “the best resonance effect in singing is the resonance peak of 2800Hz for male voices and 3200Hz for female voices.” Mongolian long song is one of the most representative forms of Mongolian oral culture, expressing the deep feelings and wisdom of the ethnic group. The unique and distant, majestic qualities of Mongolian long songs are always emotional sustenance for the people, regardless of when and where.

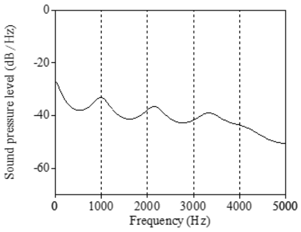
2. Techniques and Methods Used

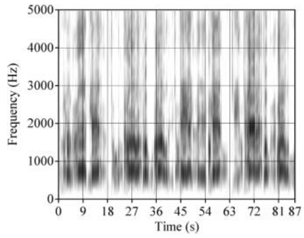
Based on the previous years’ research results, this study has achieved the design of experiments on Mongolian long songs by studying some relevant knowledge of speech theory and referring to the experience of other academic papers. Firstly, in the preprocessing stage, the corpus was selected, recorded, and segmented, the rules for file naming and saving were clarified, and a database was established. Energy, resonance peak, duration, fundamental frequency, opening quotient, speed quotient, and other parameters were extracted, and a database was established. The signal acquisition hardware equipment used in this experiment mainly includes a laptop computer, Audition recording software, external sound card, mixing console, microphone, etc. The selected long song representative works include “The Rich and Vast Alashan,” “The Mountains and Rivers of the State,” “The Rising Sun of Alashan,” and others. The first character of the long song name is used to replace the full name, for example, “The Rich and Vast Alashan” is abbreviated as ‘R’. The recorded speech files were saved, classified according to different themes, and saved in the same folder to form a corpus. The Cool Edit software was then used to segment the recordings. The entire song was segmented into sections, phrases, and words for later analysis. The parameters of the speech signal were mainly extracted and analyzed using Praat speech analysis software. The resonance peak parameters of the long song were extracted, and the differences were analyzed by comparing the resonance peak parameters of three representative works.

3. Acoustic analysis of resonance peak characteristics of musical sections

The acoustic characteristics of the resonance peak in singing can be seen from the characteristics of resonance. Resonance is an important factor in singing, and it is also a very important means of expression in singing art. It can make the voice bright, clear, and more penetrating. [3] The resonance peak curve can indirectly indicate the presence or absence of a singing resonance peak in a Mongolian long song. Using Praat software to extract the resonance peak of the long song sections and conduct statistics, we obtained the resonance peak spectrogram and envelop of three representative long songs. By observing the presence or absence of a singing resonance peak in Mongolian long songs, we can conclude the characteristics of the resonance peak in long song sections.







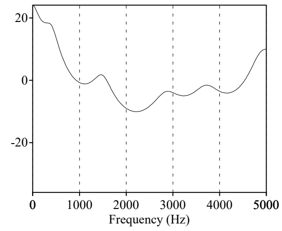


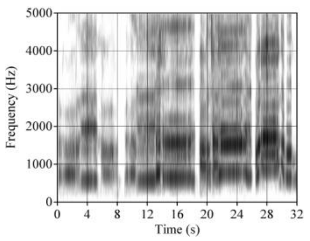
Fig.3-1 Singing Resonant Peak Spectrograms of “Fu”
Fig.3-2 Resonant Peak Spectrum Envelope of “Fu”
Fig. 3-3 Singing Resonant Peak Spectrograms of “Jiang”
Fig. 3-4 Resonant Peak Spectrum Envelope of “Jiang”

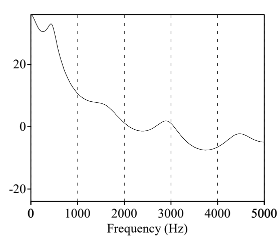
From Fig 3-1, the long song “Fu” has two dense and high-energy resonance peaks, which are distributed between 800Hz and 2500Hz and are included in the most sensitive sound frequency band of the human ear, between 1000Hz and 4000Hz. Sounds with such resonance peaks are prone to cause auditory excitement and an increase in loudness. By observing the spectrogram in Fig 3-1, a shallow resonance peak F3 can be found around the frequency of 3200Hz. The four resonance peaks are distributed between 700Hz and 5000Hz. F1 is below 1000Hz, F2 is around 1500Hz-2000Hz, F3 is around 2500Hz-3500Hz, and F4 is around 3500Hz-5000Hz. From the resonance peak spectrum envelope of “Fu” in Fig 3-2, several high-energy peaks, including F1 and F3, appear in the low-frequency region, and a relatively shallow resonance peak appears around 3200Hz.

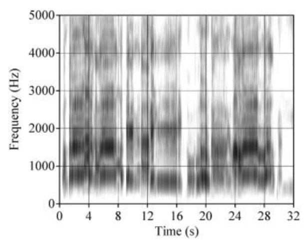
From Fig 3-3, the long song “Jiang” has two dense and high-energy resonance peaks, which are distributed between 800Hz and 2500Hz and are included in the most sensitive sound frequency band of the human ear, between 1000Hz and 4000Hz. Sounds with such resonance peaks are prone to cause auditory excitement and an increase in loudness. By observing the spectrogram in Fig 3-3, a shallow resonance peak F3 can be found around the frequency of 3200Hz at 72s. The four resonance peaks are distributed between 700Hz and 5000Hz. F1 is below 1000Hz, F2 is around 1500Hz-2000Hz, F3 is around 2500Hz-3500Hz, and F4 is around 3500Hz-5000Hz. From the resonance peak spectrum envelope of “Jiang” in Fig 3-4, several high-energy peaks, including F1 and F3, appear in the low-frequency region, and a relatively shallow resonance peak appears around 2800Hz.

4. Analysis of Resonance Peak Characteristics in Long Song Phrases

Excellent singers tend to have a high-energy resonance peak around 2400-3200Hz, centered at 2800Hz, which is produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram. This high-energy resonance peak is both penetrating and loud, and combined with the characteristics of the singer’s low-energy resonance peaks (F1, F2), it is known as the singer's resonance peak. Regardless of the voice part, pitch, vowel type, or tension of the sound, they always exist in the voices of excellent singers. Ultimately, the determining factor for the quality difference in voices is resonance, and “getting resonance in the voice” means acquiring 2800Hz. [4] In this study, the Praat software was used to extract the resonance peaks of the long song phrases, and some resonance peak spectrograms and spectrum envelopes were obtained through statistical analysis of three long songs, to observe whether there are singing resonance peaks in Mongolian long songs.







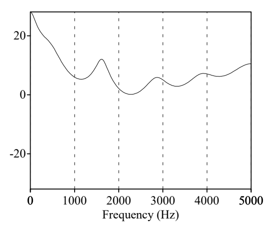
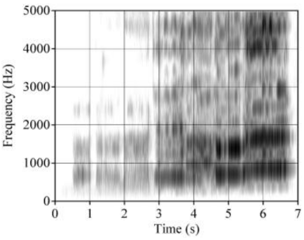
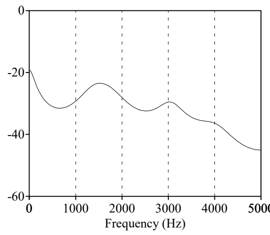


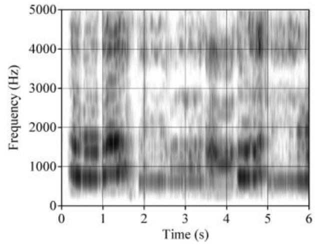
Fig. 4-1 Resonant Peak Spectrogram of the First Sentence of the Song “Fu”
Fig. 4-2 Resonant Peak Spectrum Envelope of the First Sentence of the Song “Fu”
Fig. 4-3 Resonant Peak Spectrogram of the Second Sentence of the Song “Fu”
Fig. 4-4 Resonant Peak Spectrum Envelope of the Second Sentence of the Song “Fu”

From the spectrogram of the first phrase of “Fu” in Fig 4-1, it can be observed that there are dense resonance peaks at around 1000Hz and 2000Hz, but the resonance peaks near 3000Hz and 4000Hz are shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third, fourth, and fifth formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the first phrase of “Fu” in Fig 4-2, there is a third resonance peak near 2800Hz.

Similarly, in Fig 4-3, it can be observed that there are dense resonance peaks at around 1000Hz and 2000Hz, but the resonance peaks near 3000Hz and 4000Hz are shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the second phrase of “Fu” in Fig 4-4, there is a third resonance peak near 2800Hz.







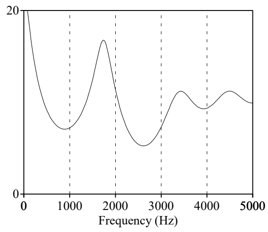
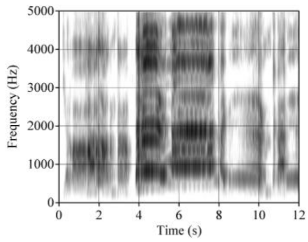
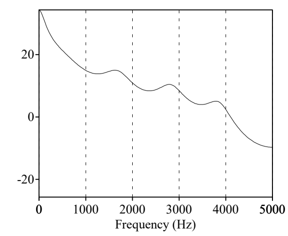


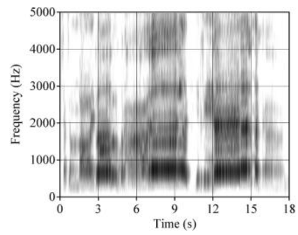
Fig. 4-5 Resonant Peak Spectrogram of the First Sentence of the Song “Xu”
Fig. 4-6 Resonant Peak Spectrum Envelope of the First Sentence of the Song “Xu”
Fig. 4-7 Resonant Peak Spectrogram of the Second Sentence of the Song “Xu”
Fig. 4-8 Resonant Peak Spectrum Envelope of the Second Sentence of the Song “Xu”

From Fig 4-5, it can be observed that there are dense resonance peaks at around 1000Hz, 2000Hz, 4000Hz, and 5000Hz, but the resonance peak near 3000Hz is shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the first phrase of “Xu” in Fig 4-6, there is a third resonance peak near 3000Hz.

From Fig 4-7, it can be observed that there are dense resonance peaks at around 1000Hz, 2000Hz, and 4000Hz, but the resonance peak near 3000Hz is shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the second phrase of “Xu” in Fig 4-8, there is a third resonance peak near 3500Hz.







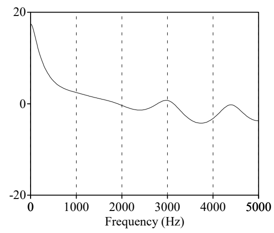
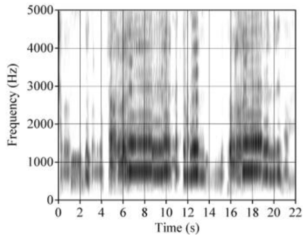
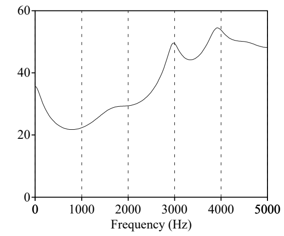


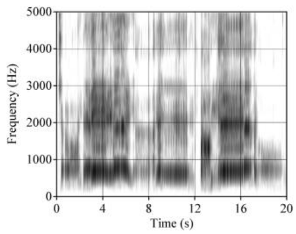
Fig. 4-9 Resonant Peak Spectrogram of the Seventh Sentence of the Song “Xu”
Fig. 4-10 Resonant Peak Spectrum Envelope of the Seventh Sentence of the Song “Xu”
Fig. 4-11 Resonant Peak Spectrogram of the First Sentence of the Song “Jiang”
Fig. 4-12 Resonant Peak Spectrum Envelope of the First Sentence of the Song “Jiang

From the spectrogram of the seventh phrase of “Xu” in Fig 4-9, it can be observed that there are dense resonance peaks at around 1000Hz, 2000Hz, 3000Hz, and 4000Hz within the time range of 4s to 8s, but the resonance peaks within the time range of 0s to 4s and 8s to 12s are shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the seventh phrase of “Xu” in Fig 4-10, it can be seen that there is a third resonance peak near 2800Hz.

From the spectrogram of the first phrase of “Jiang” in Fig 4-11, it can be observed that there are dense resonance peaks at around 1000Hz and 2000Hz, but the resonance peaks near 3000Hz and 4000Hz are shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the first phrase of “Jiang” in Fig 4-12, it can be seen that there is a third resonance peak near 3000Hz.







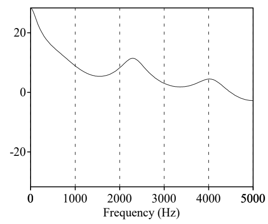


Fig. 4-13 Resonant Peak Spectrogram of the Third Sentence of the Song “Jiang”
Fig. 4-14 Resonant Peak Spectrum Envelope of the Third Sentence of the Song “Jiang”
Fig. 4-15 Resonant Peak Spectrogram of the Second Sentence of the Song “Jiang”
Fig. 4-16 Resonant Peak Spectrum Envelope of the Second Sentence of the Song “Jiang”

From the spectrogram of the second phrase of “Jiang” in Fig 4-13, it can be observed that there are dense resonance peaks at around 800Hz and 1800Hz, but the resonance peaks near 3000Hz and 4000Hz are shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the second phrase of “Jiang” in Fig 4-14, it can be seen that there is a third resonance peak near 3000Hz.

From the spectrogram of the third phrase of “Jiang” in Fig 4-15, it can be observed that there are dense resonance peaks at around 1000Hz and 2000Hz, but the resonance peaks near 3000Hz and 4000Hz are shallow. This is consistent with the characteristics of the high-energy resonance peak produced by multiple resonance peaks at the third (F3), fourth (F4), and fifth (F5) formants within the bright frequency band of the spectrogram in the voices of excellent singers. From the resonance peak spectrum envelope of the third phrase of “Jiang” in Fig 4-16, it can be seen that there is a third resonance peak near 2500Hz.

The above provides an in-depth look at the science of what makes a singer's voice sound good and how resonance peaks play a key role in creating that sound. Resonance refers to the amplification of sound waves through the interaction of sound waves with the natural cavities of the body, such as the mouth, throat, and nasal cavity. Through this interaction, certain frequencies are amplified, resulting in a unique sound quality unique to each singer. Studies point out that great singers tend to have a high-energy resonance peak around 2400-3200 Hz, centered at 2800 Hz. This peak is produced by multiple resonance peaks in the third (F3), fourth (F4), and fifth (F5) vowels in the bright band of the spectrogram. This high-energy resonance peak is both penetrating and loud, giving great singers their signature sound. The article emphasizes that this resonance peak is always present in the voices of great singers, regardless of voice part, pitch, vowel type, or vocal tension. To study the resonance peaks in Mongolian long songs, the researchers used Praat software to extract resonance peaks from long song phrases. They then analyzed spectrograms and spectral envelopes to determine if the songs contained resonance peaks consistent with those in the voices of good singers. The results of the study showed that the resonance peaks in the Mongolian long songs were consistent with the characteristics of the high-energy resonance peaks in the voices of the best singers.

Resonance is a determinant of differences in voice quality, and "getting resonance in the voice" means getting 2800 Hz. This means that singers who want to improve their voices must focus on developing their resonance peaks in the third, fourth, and fifth consonants in the bright frequency band of the spectrogram. By doing so, they can create a sound that is both powerful and pleasing to the ear. In summary, this study provides valuable insight into what makes a great singer's voice unique and how resonance peaks play a key role in creating that voice. A study of resonance peaks in Mongolian long songs demonstrates the universality of this concept and highlights the importance of developing resonance peaks for singers who want to improve their voices.

5. Analysis of the Vocabulary Features of Hymns

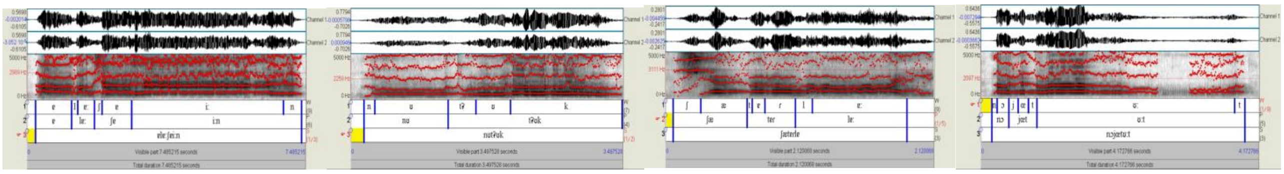


Fig.5-1 The 3-level labeling of the language diagram for the third word of the ode“Fu”, “ɑlɑšɑ-yin”

Fig.5-2 The 3-level labeling of the language diagram for the 7th word of the ode“Fu”, “ nutuɤ”

Fig.5-3 The 3-level labeling of the language diagram for the 9th word of the ode “Jiang”, “sitar lɑ”

Fig.5-4 The 3-level labeling of the language diagram for the 7th word of the ode “Jiang”, “nɔjɑtʊ:t”

Fig 5-1 is an example of a three-layer annotation of the spectrogram of the female speaker “ɑlɑšɑ -yin”. The F1 average value of the first vowel /ɐ/ is 874Hz, and the F2 average value is 1446Hz. The F1 average value of the second vowel /ɐ/ is 772Hz, and the F2 average value is 1408Hz. The F1 average value of the third vowel /ɐ/ is 848Hz, and the F2 average value is 1578Hz. The F1 average value of the final vowel /i/ is 821Hz, and the F2 average value is 1589Hz. The lower the tongue position, the higher the resonance peak, and vice versa. According to the Mongolian vowel tongue position chart, the tongue position of the vowel /ɐ/ is very low, so compared with other vowels, the resonance peak will be very high. This word is the one with the highest average resonance peak value in the hymn “Fu” because it has three /ɐ/ vowels, which are rounded, low, and back vowels. The average value of the second vowel /ɐ/ is lower than that of the first vowel /ɐ/ because the second vowel /ɐ/ is affected by the preceding consonant /l/.

Fig 5-2 is an example of a three-layer annotation of the spectrogram of the female speaker “nutuɤ”. The F1 average value of the first vowel /ʊ/ is 601Hz, and the F2 average value is 1056Hz. The F1 average value of the second vowel /ʊ/ is 578Hz, and the F2 average value is 1257Hz. The lower the tongue position, the higher the resonance peak, and vice versa. According to the Mongolian vowel tongue position chart, the tongue position of the vowel /ʊ/ is very high, so compared with other vowels, the resonance peak will be very low. This word has the smallest average resonance peak value in the hymn “Fu” because it has two /ʊ/ vowels: rounded, semi-high, and back vowels. The average value of the second vowel /ʊ/ is lower than that of the first vowel /ʊ/ because the second vowel /ʊ/ is affected by the preceding consonant /t^h/.

Fig 5-3 is an example of a three-layer annotation of the spectrogram of the female speaker “sitar lɑ”. The F1 average value of the first vowel /æ/ is 667Hz, and the F2 average value is 1764Hz. The F1 average value of the second vowel /ɐ/ is 718Hz, and the F2 average value is 1391Hz. The F1 average value of the final vowel /ɐ:/ is 772Hz, and the F2 average value is 1285Hz. The lower the tongue position, the higher the resonance peak, and vice versa. According to the Mongolian vowel tongue position chart, the tongue position of the vowels /æ/ and /ɐ/ is very low, so compared with other vowels, the resonance peak will be very high. This word is the one with the highest average resonance peak value in the hymn “Jiang” because both the /ɐ/ and /æ/ vowels have very high resonance peaks, which leads to this word having the highest average resonance peak value in the hymn “Jiang”.

Fig 5-4 is an example of a three-layer annotation of the spectrogram of the female speaker “nɔjɑ tʊ:t”. The F1 average value of the first vowel /ɔ/ is 635Hz, and the F2 average value is 1287Hz. The F1 average value of the second vowel /øe/ is 611Hz, and the F2 average value is 1521Hz. The F1 average value of the final vowel /ʊ:/ is 611Hz, and the F2 average value is 1521Hz. The lower the tongue position, the higher the resonance peak, and vice versa. According to the Mongolian vowel tongue position chart, the tongue position of the vowels /ɔ/ and /ʊ/ is very high, so compared with other vowels, the resonance peak will be very low. The /ɔ/ vowel is a rounded, mid-low, back vowel. This word is the one with the lowest average resonance peak value in the hymn “Jiang” because both the /ɔ/ and /ʊ/ vowels have very low resonance peaks, which leads to this word having the lowest average resonance peak value in the hymn “Jiang”.

6. Summary

This article delves into the acoustic properties of Mongolian long-tone hymns and explores how the extraction of resonance peaks can provide insights into the vocal techniques used by singers. By analyzing the spectrograms and spectral envelopes of three hymns, the researchers were able to draw several conclusions: all three hymns have two dense and high-energy resonance peaks between 800Hz-2500Hz, which fall within the most sensitive frequency range of the human ear (1000Hz-4000Hz). Additionally, all three hymns exhibit several high-energy peaks centered around F1 in the low-frequency range, including F3, which can increase auditory excitement and loudness. The spectral envelopes of the hymns indicate that many of the third resonance peaks are near 2400-3200Hz, with 2800Hz being the center of the singer's resonance peak, suggesting that these hymns are performed by excellent vocalists. Furthermore, the article discusses the relationship between tongue position and resonance peak frequencies, with higher tongue positions leading to lower F1 frequencies and lower tongue positions leading to higher F1 frequencies. The article also explores the relationship between F3 and tongue tip movement, as well as the effects of lip rounding on resonance peak frequencies, and how these factors relate to the pronunciation of Mongolian vowels. Overall, this research provides quantitative data for the ontology of Mongolian long-tone hymns and offers a more intuitive way to understand existing teaching methods. Real-time analysis of vocal cord vibrations and tongue movements during long-tone singing can be performed using this research. By extracting several physiological acoustic parameters of articulation, establishing a parametric model of oral transmission culture, and analyzing the relationship between parameters and articulatory actions, this modern phonological method based on computer and language science to study the fundamental mechanism of long tune vocalization can deepen the theory of speech vocalization and explore the mechanism of speech vocalization, which provides new insights for the research and conservation of oral transmission culture such as Mongolian long tune chant, and also provides a new application for modern phonology The study has found a new application for modern phonology accumulated over the years, and the interpretation of vocalization skills and teaching difficulties in long-toned chants will be more imaginative, which can open up a new field of oral culture research and conservation, help to preserve culture digitally and promote the presentation of cultural multimodality.

7. Acknowledgement

This study was supported by the Fundamental Research Funds for the Central Universities (31920230123) and the National Social Science Foundation (22BYY165). We thank the reviewers for their comments.

References

[1] "A singing throat nine blessings - the famous Mongolian long folk song inheritor O. Eriegejdema Alashan folk song singing album", [M/CD] Hohhot, Inner Mongolia Culture Audio and Video Publishing House, 2010: p. 448.

[2] Wang Shiqian, Voice mechanism and the resonance peak of singers - with an answer to Sundberg [J], Music Art, Journal of Shanghai Conservatory of Music, 1993, 01.

[3] Wang Fang, "Research on the Acoustic Rhythm of Tianzhu Tibetan Folk Songs" [D], Master's thesis, Northwest University for Nationalities, May 2016, p. 12.

[4] J.Sundberg.Level and Center Frequency of the Singer`s Formant, Journal of Voice2001(2), P176-186.