International Linguistics Research; Vol. 8, No. 2; 2025
ISSN 2576-2974 E-ISSN 2576-2982
https://doi.org/10.30560/ilr.v8n2p22

A Praat-Based Study on the Vowel Features in Chongqing Hip-Hop Lyrics

Jiang Xin[1]

[1] School of Foreign Studies, Jinan University, Guangzhou, Guangdong, China

Correspondence: Jiang Xin, School of Foreign Studies, Jinan University, Guangzhou, Guangdong, 511400, China. E-mail: jx17358469715@163.com

Received: April 16, 2025 Accepted: May 2, 2025 Online Published: May 7, 2025

Abstract

Chongqing dialect, having experienced the waves of ethnic migration, has given birth to distinctive linguistic features. Since the 1980s, academic research on the Chongqing dialect has gradually increased. However, most studies have only scratched the surface in terms of vocabulary, semantics, and phonological characteristics, with few attempts to delve into its internal phonetic structure. This study addresses this gap by using Praat software and theories of tongue position and formants to analyze rhyme vowels in Chongqing hip-hop lyrics. It examines formant frequencies (F1, F2, F3), pitch, and duration of vowels from local hip-hop artists' works. Results reveal that these vowels have distinct regional acoustic characteristics, differing from standard Mandarin, with lower formant frequencies aligning with the dialect's low-pitched nature. In addition, the study confirmed the phonetic changes that occurred in the 1970s, such as /ɤ/ → /ɛ/, /aʊ/, /uɔ/. This research enriches our understanding of Chongqing dialect's acoustic features and offers new insights for interdisciplinary studies in dialect phonetics and music linguistics.

Keywords: acoustic analysis, Chongqing hiphop, vowels, dialect studies

1. Introduction

Chongqing, located in Southwest China, is home to a unique dialect variant of Southwestern Mandarin. The Chongqing dialect's tonal contours and phonetic features distinguish it from Standard Mandarin and make it conducive to hiphop music. Additionally, the Chongqing dialect's phonetic structure is inherently rhythmic and highly adaptable for rhyming, a critical component of hiphop artistry. The dialect's alignment with the assertive and straightforward temperament of the Sichuan and Chongqing region's populace resonates deeply with the ethos of hiphop culture, fostering a natural synergy between the dialect and the genre.

This study employs a Praat-based acoustic analysis to delve into the vowel features of Chongqing Hip-Hop lyrics. By applying Praat, we aim to uncover the subtleties of the dialect as expressed in this modern musical context. In this study, a corpus of 300 Chongqing hiphop music was built, from which words with obvious vowel features were selected for recording, and then analyzed with praat software, focusing on the characteristics of their formants and analyzing the differences between Chongqing dialect and Putonghua.

2. Literature Review

Hip-hop music has become a global phenomenon, transcending borders and resonating with young people worldwide. Its critical and defiant nature, combined with its "political consciousness," has led to its adaptation and evolution in various cultural contexts. This process has enriched the genre, turning it into a dynamic and diverse mode of artistic and social commentary that reflects the struggles and aspirations of a new generation worldwide. (George, 2002) begins with the origins of hiphop, tracing its evolution from the post-soul era to the 21st century.[1] Furthermore, George examines its impact on society, economy, and politics.

Gradually, Alonzo Westbrook provides a comprehensive guide to hip-hop slang, laying the groundwork for linguistic research in this field. Neate explores how hip-hop transcends racial and national boundaries, seeking identity recognition in an unequal world.

In China, research on hip-hop has increased significantly since the debut of "The Rap of China" in 2017. Studies have focused on subcultural perspectives, fan communities, and communication studies. (Wu &Jiang,2017) takes the program phenomenon of "The Rap of China" as the research object based on narrative discourse theory, and analyzes its narrative discourse characteristics, while (Wang, 2018, p.113-123) highlights hip-hop's significance in

providing a means for youth expression. (Yang and Feng,2020, p.99-104) [2][3] Also, there are some other studies. (Chen,2012; Li, 2013; Zhang&Yan,2015; Song, 2017)

The study of the Chongqing dialect initially focused on its lexicon, with phonetic research emerging later. (Fan, 1979) explored the common lexicon of the Chongqing dialect. [4] (Huang, 1986) observed the reduplication of nouns with "er" suffixes.[5] (Zhu,1987) highlighted the importance of dialectal lexicon differences. [6]

Following this, focus turned to the nuances of phonetics and intonation. (Yang, 1995) and (Zeng, 2013) examined the phonological evolution of the Chongqing dialect. [7][8](Xu, 2001) analyzed the "er" suffix in Chongqing speech. (Yang, 2022) explored the "v" consonant in the Pengshui dialect. (Li, 2021) investigated phonetic features in local chronicles.[9]

Scholars use advanced tools like Praat software to measure and analyze formants in different languages and dialects. (Zhong, 2019, p.61-71; Huang, 2018, p.p. 4-5; Cao, 2018, p.352-356) have conducted such analyses. [10-12] (Joos, 1948, p.4-15) identified correlations between tongue position and formants, while (Delattre, 1951, p.4-15) expanded on this research.[13] (Jones, 1948) standardized the International Phonetic Alphabet.[14] A.C. Gimson and J.C. Wells documented the English vowel system. Leonard Bloomfield and Noam Chomsky emphasized the dynamic nature and innate structures of vowel changes. Their work provides a comprehensive perspective on the evolution and diversity of vowel sounds across languages.

3. Research Design

3.1 Research Question

Given the current state of research, which has made significant strides in understanding the broader dynamics of hiphop music and the Chongqing dialect, there remains a notable gap in the acoustic analysis of rhymes in Chongqing hiphop music. This research aims to bridge this gap by conducting a systematic acoustic analysis of the rhymes in Chongqing hiphop lyrics.

Hence, this study aims to tackle the following questions:

RQ1: What are the distinctive acoustic characteristics of the vowels in Chongqing hiphop music as compared to Standard Mandarin?

RQ2: Based on Praat software, what are the pronunciation characteristics of these vowels in lyrics?

RQ3: How do the formant frequencies (F1, F2, F3), pitch, and duration of the vowels in Chongqing hiphop lyrics?

Research Objects

The research subjects include a diverse selection of Chongqing-based hip-hop lyrics from both established and emerging artists, focusing on their representative works that showcase the dialect's phonetic characteristics.

3.2 Research Tools

The primary tool is a self-created Chongqing hip-hop music corpus, comprising carefully selected songs from various local artists to reflect the dialect's phonetic diversity. It serves as the main data source for acoustic analysis. The other research tool is Praat software, which measures formant frequencies (F1, F2, F3) of rhymes, analyzes vowel pitch and duration, and generates spectrograms to visualize acoustic properties.

3.3 Methods

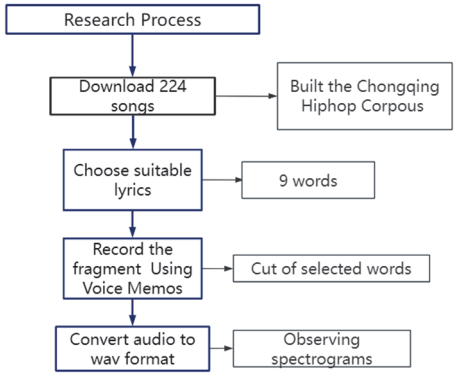

Figure 1. Procedures of data collection

In this study, a mixed-methods approach was used to analyze the vocalic characteristics of rhymes in Chongqing hip-hop music. (Figure.1) A corpus of Chongqing hip-hop songs was created by downloading diverse tracks from various artists. Excerpts of selected lyrics were recorded using Voice Memos and converted to wav format for compatibility with Praat software. Praat was then used to measure formant frequencies (F1, F2, F3), pitch, and duration, documenting key parameters for each rhyme. In the quantitative segment, statistical analysis was performed on the acoustic parameters of the rhymes, including calculating means, standard deviations, and conducting t-tests to identify significant differences between the Chongqing dialect and Standard Mandarin. The qualitative segment involved analyzing the rhymes' acoustic characteristics and their interaction with rhythm, melody, and other lyrical elements to understand their function and impact on the song's overall expression.

Ultimately, the results of the quantitative and qualitative research were integrated to provide a comprehensive interpretation of the acoustic characteristics of the vocalic rhymes in the Chongqing dialect. The study explored how these characteristics reflect the uniqueness of the Chongqing dialect and their artistic expression in hip-hop music.

4. Results and Analyses

This section focuses on using Praat software to generate spectrograms of isolated word pronunciations in hip-hop music and comparing them with Standard Mandarin pronunciations to directly reflect the phonetic characteristics of rhymes in the Chongqing dialect.

Using advanced corpus retrieval software, researchers selected 9 representative sets of words from a large Chongqing hip-hop corpus, covering both everyday and regionally specific vocabulary. The standard Mandarin and authentic Chongqing dialect pronunciations of these words were recorded. Praat was then used to generate spectrograms, which revealed differences in pitch, duration, and formants through comparative analysis.

4.1 Monophthong

1) /ɤ/→/ɛ/、/aʊ/、/uɔ/

Figure 2 shows spectrograms of the phrase "le shi wu du（勒是雾都）" in Standard Mandarin (left) and Chongqing dialect (right). Blue lines indicate pitch. Denser formant clusters in Standard Mandarin may result from tongue position adjustments, lip rounding variations, or faster speech rates compressing vowels.

Focusing on the character "le（勒）" the pitch of the Standard Mandarin pronunciation /lɤ/ is notably higher than that of the Chongqing dialect /lɛ/. This aligns with the characteristic lower pitch often associated with the Chongqing dialect.

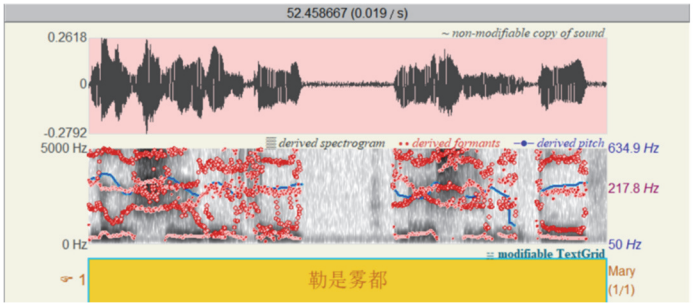

Figure 2. Spectrogram of le shi wu du(勒是雾都)

Figure 3 shows the spectrogram of Mandarin "le（勒）" with a short distance between F1 and F2, indicating a high tongue position and retroflex action. The vowel [e] has a wider opening, lower pitch, and nearly adjacent F2 and F3, suggesting minimal variation in tongue position or lack of labial rounding.

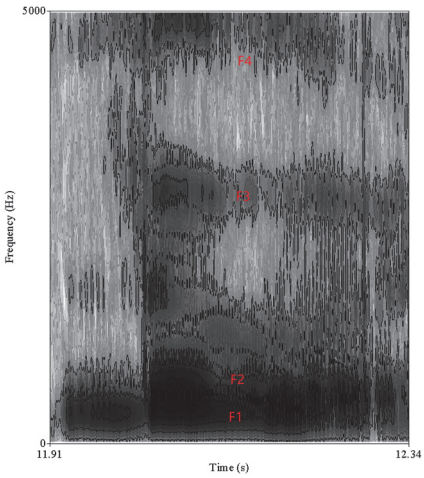

Figure 3. Formant pattern of /lɤ/

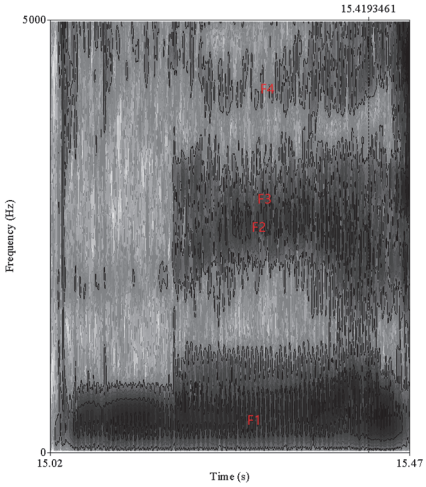

Figure 4. Formant pattern of /lɛ/

Figure 4 shows the lyrics from "2END" by Blinder Tang Jingkun and A Lin. The blue line indicates pitch, which is higher in Standard Mandarin than in the Chongqing dialect. For "le," the short distance between F1 and F2 in /lɤ/ suggests a higher tongue position. The closeness F1 and F2 in /ɤ/ suggest a smaller oral cavity opening and higher tongue position. For the diphthong /laʊ/, the spectrogram shows fluctuations in F1, F2, and F3, reflecting the transition from /a/ to /ʊ/. The distance between F2 and F3 varies with the pronunciation of /aʊ/. The vowel /a/ has a low tongue position and wide oral aperture, while /ʊ/ has a high tongue position and rounded lips. The transition from /a/ to /ʊ/ is shown in the spectrogram as a gradual increase in F2 and F3 frequencies.

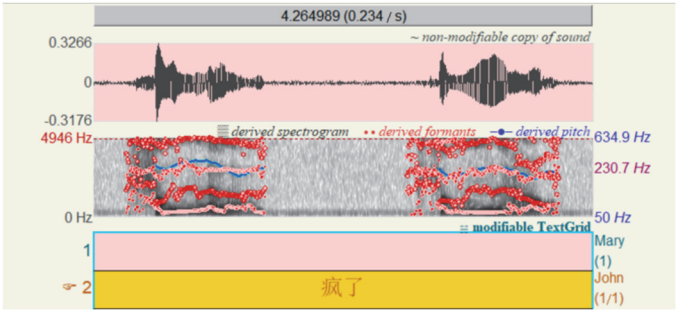

Figure 5. Spectrogram of feng le(疯了)

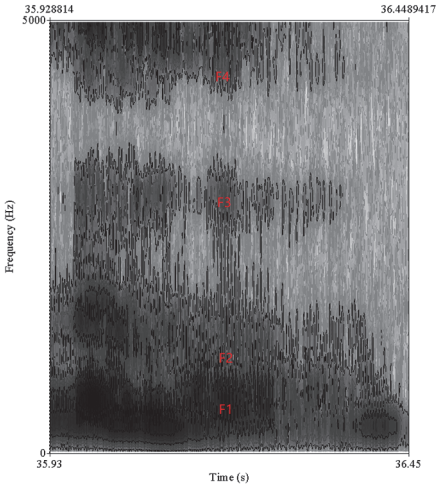

Figure 6. Formant pattern of /lɤ/

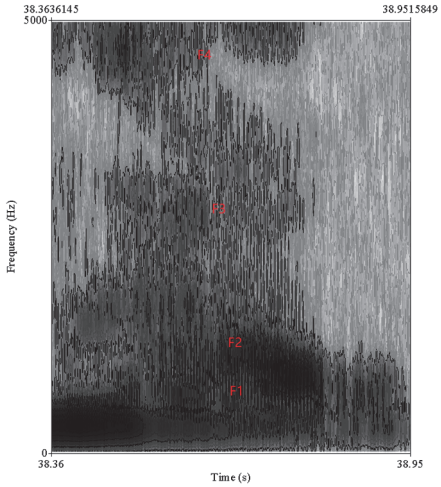

Figure 7. Formant pattern of /laʊ/

To summarize, the findings of this research reveal that the vowel /e/ has the most obvious change in Chongqing dialect. To take another example, ke(瞌) in "ke shui(瞌睡)" in Mandarin Chinese is /kɤ/, but in Chongqing dialect is /kuɔ/ named; the lip becomes rounded and the pitch becomes higher.

In wrapping up, the study has shown the phonetic shift from /ɤ/ to /ɛ/, /aʊ/, and /uɔ/. In fact, this series of changes was already completed before the early 17th century, as evidenced by the use of the Latin letter 'e' to represent characters such as "格," "革," and "德" in "Xiru Ermu Zi." (Wang, 1985)[15] In modern Northern and Southwestern Mandarin dialects, characters with the rhyming patterns of Mo Mai De in the first and second tones, as well as characters with the Zhi rhyme and Zhang system, have not undergone centralization of the tongue position and remain at the stage of pronouncing [e]. In today's Beijing dialect, characters from the ancient Zeng She first-class De rhyme with entering tone are pronounced as / ɤ/, /uo/, while in the Sichuan dialect, they are pronounced as/ɛ/, /uɛ/. For characters with the ancient Zeng She open three-class Zhi rhyme and initial consonants like "ce(测)" and "se(色)," the rhyme is pronounced as /ɤ/ in Beijing dialect, and with the tongue position moved forward to /ɛ/ in the Sichuan dialect. Characters from the Geng She second-class Mo Mai rhyme with entering tone, such as "bai(柏)" "mai(麦)" "ze(泽)" and "ge(革)" guo(国) are pronounced with the rhymes /ai/, /ɤ/, /uo/ in Beijing dialect, and as /ɛ/, /uɛ/ in the Sichuan dialect. (Li,2021).

4.2 Diphthong

1) /ei/→/ɛ/

Figure 8 "hei an(黑暗)" is the lyric of the song 4REAL. Observing the spectrographs on both sides, it can be seen that the pitch of Putonghua is still higher than that of Chongqing dialect. This is in line with the characteristics of Chongqing dialect bass.

Figure 9 illustrates the pronunciation of /xei/ in Standard Mandarin, where it can be observed that the formant scatter plot of the latter part, /ei/, exhibits an upward trend and increased density. This likely indicates that as time progresses, the resonant frequencies of the vowels are increasing. Such a pattern is typically associated with changes in the tongue position, possibly due to the transition from /x/ to /ei/, where the tongue moves from a higher to a lower position. In contrast, Figure 10 suggests that /xɛ/, unlike /xei/, may not have the transition characteristic of a diphthong, thus potentially displaying a distinct formant pattern with a more constricted oral opening.

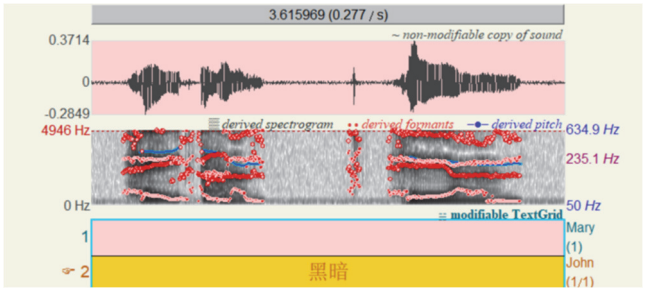

Figure 8. Spectrogram of hei an(黑暗)

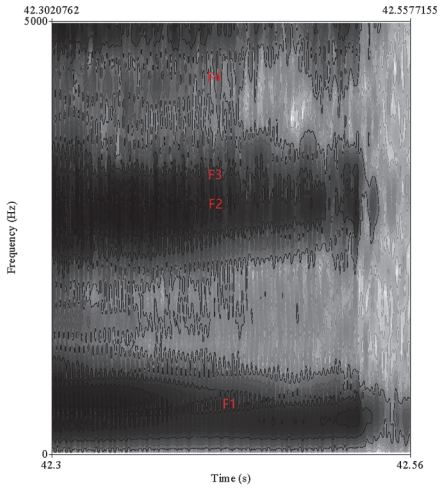

Figure 9. Formant pattern of /xeɪ/

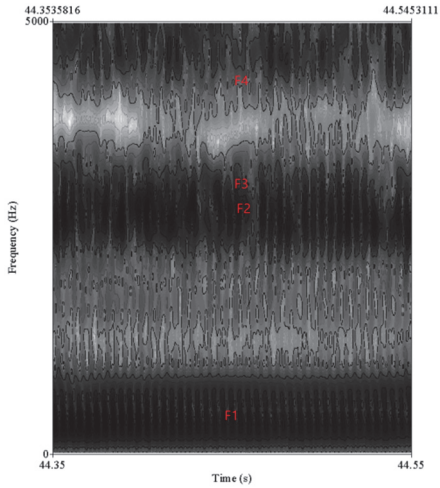

Figure 10. Formant pattern of /xɛ/

2) /ai/→/ɛ/ 、/iɛ/

"bai mi（百米）" is a lyric from the 2017 Chongqing Teahouse Cypher, produced by rappers (Tory/Bridge/GAI). Notably, the spectrogram on the left, which represents the Standard Mandarin pronunciation, has a longer duration, and the formant scatter plot is more concentrated. (Figure 11) The concentration of the formant scatter plot indicates that there is minimal variation in the shape of the vocal tract during articulation, or in other words, the quality of the vowels is relatively stable. This likely reflects the typical characteristics of vowel pronunciation in Standard Mandarin, where the adjustments of tongue position and lip shape during articulation are more consistent.

Figures 12 and 13 depict the acoustic characteristics of the word "bai（百）" in Standard Mandarin and Chongqing dialect respectively. Notably, in Standard Mandarin /pai/, the distance between F1 and F2 is shorter, indicating a more centralized vowel space or a higher tongue position for the /ai/ diphthong. Additionally, the longer duration of the sound may be related to the diphthong's articulatory features, which could involve a more extended transition between the two vowel elements. In contrast, for the Chongqing dialect /pɛ/, the distance between F2 and F3 is shorter, suggesting a potentially different quality or a more constricted articulation of the vowel, possibly due to the dialect's specific phonetic realization. The comparison highlights the phonetic variations that can occur in the pronunciation of the same word across different dialects, offering insights into the unique phonetic and phonological rules of each linguistic variety.

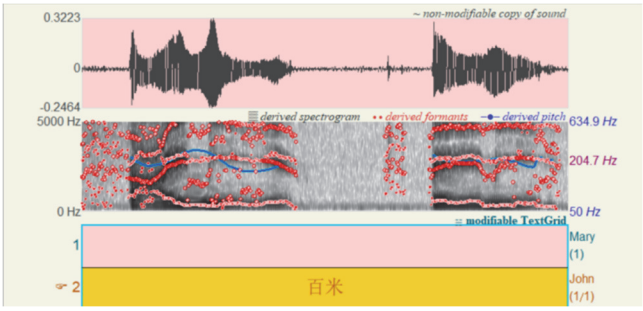

Figure 11. Spectrogram of bai mi(百米)

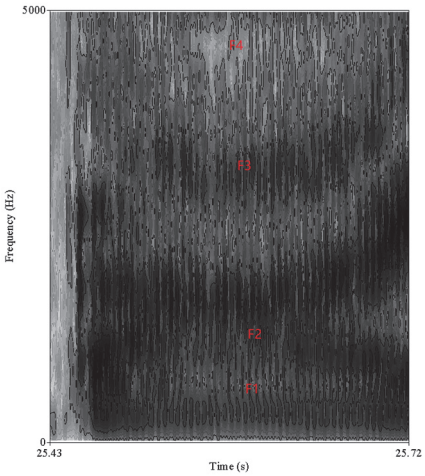

Figure 12. Formant pattern of /pai/

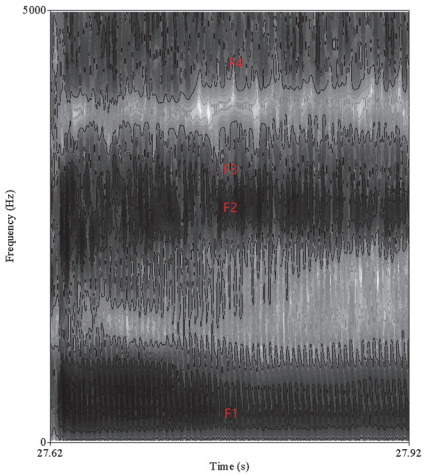

Figure 13. Formant pattern of /pe/

Next, figure 14 captures the lyrics from the hiphop song 2END, highlighting "mai(麦)" in "mai xiang（麦香）" which is under scrutiny. It consists of a consonant [m] paired with a diphthong [aɪ]. In Standard Mandarin, a more elevated pitch is distinctly noticeable, particularly for the consonant [m], likely due to its nature as a nasal consonant, which allows for greater freedom in vocal cord vibration, resulting in a higher pitch.

Furthermore, the formant scatter plot in Standard Mandarin is more compact. As seen in Figure 15, the F1 formant trace is more undulating, indicating a rise and fall in a fragmented pattern, which may reflect swift alterations in tongue positioning and vocal tract configuration, particularly during the transition within the diphthong /aɪ/. Conversely, Figure 16 illustrates a more continuous F1 formant, appearing as a smooth, unbroken line. In the Chongqing dialect, the pronunciation of /mai/ has mutated into /miɛ/, characterized by a more constricted oral

aperture and a steadier tongue position, maintaining a consistent vocal tract shape that may enhance the fluidity and connectedness of the sound. These disparities could be linked to the varying phonetic practices between the two dialects, with Standard Mandarin placing greater emphasis on the transition between consonants and vowels, whereas the Chongqing dialect may focus on the stability of pronunciation as a whole.

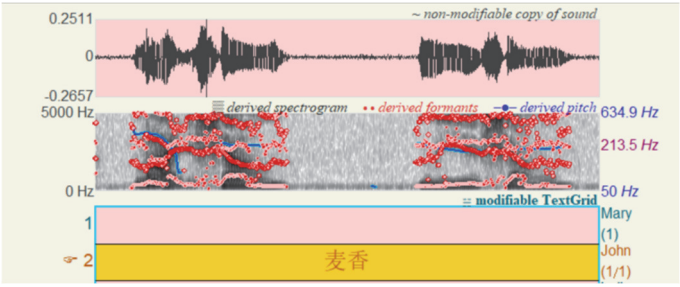

Figure 14. Spectrogram ofmai xiang (麦香)

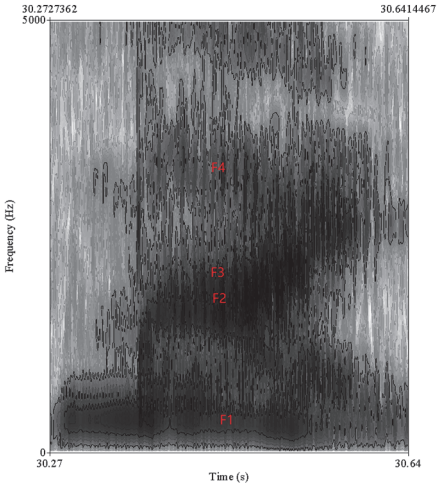

Figure 15. Formant pattern of /maɪ/

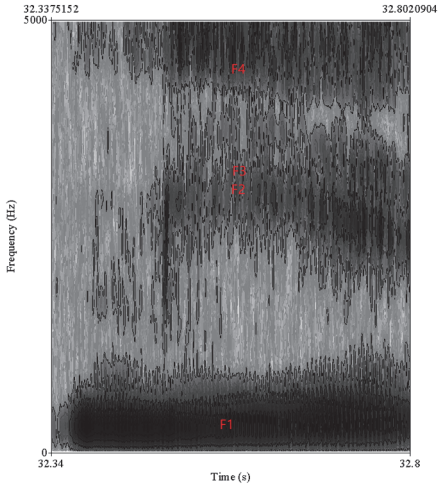

Figure 16. Formant pattern of /miɛ/

5. Conclusion

The study's findings can be summarized in three key points:

1) Chongqing dialect vowels have distinct acoustic features from Standard Mandarin, particularly lower pitch frequencies, aligning with its low-pitched nature and benefiting hiphop's popularity.

2) Phonetic shifts are observed, such as /ɤ/ → /ɛ/, /aʊ/, /uɔ/, and /ai/ → /ɛ/, /iɛ/, /ei/ → /ɛ/.

3) The Chongqing dialect prefers the sound /ɛ/, with these changes solidified in the 1970s, highlighting its dynamic and historical richness.

The phonological changes may reflect shifts in cognitive mechanisms during language development, possibly driven by language simplification, the principle of least effort, or social communication needs. In music, especially hip-hop, artists adjust and innovate rhymes based on rhythm, melody, and style, showing how cognitive mechanisms adapt in the interplay between language and music.

Despite its achievements in analyzing the acoustic characteristics of rhyme vowels in the Chongqing dialect, the study has limitations. First, background music and complex tones in hip-hop may cause interference. Second, due to technological constraints, the study relied solely on Praat spectrograms, though the relationship between tongue position and formants is not definitive (Delattre, 1951, p.4-15).[13] Future research could benefit from using techniques like EMMA and X-ray imaging to enhance acoustic analysis.

References

[1] George, N. (2002). *Hip hop America* (Vol. 3). City Press Bloc.

[2] Wang, J. J. (2018). The Rap of China and the cultural politics of hip-hop. *Literary and Art Studies*, (6), 113–123.

[3] Yang, Y. L., & Feng, Y. Q. (2020). "Authenticity" construction: Cultural production of symbols and meanings in vertical music variety shows. *Modern Communication (Journal of Communication University of China)*, (5), 99–104.

[4] Fan, J. Y. (1979). The differentiation of the character "Xia (下)" in Chongqing dialect. *Dialects*, (7).

[5] Huang, X. Z. (1986). The classification of Southwestern Mandarin. *Dialects*, (12).

[6] Zhu, Y. K. (1987). On lexical differences in Chinese dialects. *Journal of Shantou University*, (10).

[7] Yang, H. M. (1995). The evolution of Chongqing phonology over the past forty years. *Journal of Chongqing Education College*, (3).

[8] Zeng, X. Y. (2013). A comparative analysis of Chongqing dialect system over the past 70 years. *Journal of Chongqing Radio and Television University*, (2).

[9] Li, Y. H. (2021). *A study on dialect phonetics in Bashu local chronicles* [Master's thesis, Southwest Jiaotong University].

[10] Zhong, C. S. (2019). Cross-linguistic forensic speech comparison based on vowel formants. *Foreign Languages (Journal of Shanghai International Studies University)*, (1), 61–71.

[11] Huang, Z. H. (2018). The impact of formants on the intelligibility of Uyghur speech. In *Proceedings of the 2018 National Acoustics Conference* (pp. 4–5). College of Information Science and Engineering, Xinjiang University.

[12] Cao, C. Y., Xie, Y. L., & Zhang, J. S. (2018). The influence of vowels on tone perception under different formant distribution. *Journal of Tsinghua University (Science and Technology)*, (4), 352–356. https://doi.org/10.16511/j.cnki.qhdxxb.2018.04.008

[13] Delattre, P. (1951). The physiological interpretation of sound spectrograms. *PMLA*, *66*(5), 864–875. https://doi.org/10.2307/459529

[14] Joos, M. (1948). Acoustic phonetics. *Language*, *24*(2), 5–136. https://doi.org/10.2307/522229

[15] Wang, L. (1985). *A history of Chinese phonology*. China Social Sciences Press.

[16] Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, *63*, 75–86. https://doi.org/10.1016/j.wocn.2017.05.002