An Acoustic Study on the Vowel Features of Ordos Mongolian Based on a Speech Database

Naqin Zhao[1,a], Nanzi Liu[1,b], Gegen Tana[2,c], Axu Hu[2,*]

[1]Northwest Minzu University, Department of Chinese Language and Literature; Key Laboratory of Language and Cultural Computing of Ministry of Education, Lanzhou 730030, China
[2]Northwest Minzu University, Key Laboratory of Language and Cultural Computing of Ministry of Education, Lanzhou 730030, China

[a]2947225741@qq.com, [b]liunanzi@foxmail.com, [c]87354683@qq.com,
*huaxu84@foxmail.com
*Corresponding author.

Abstract. This study presents an acoustic comparative analysis of word-initial and non-word-initial short vowels in Ordos Mongolian and Standard Mongolian based on a speech database constructed from field recordings[1]. A corpus drawn from an Ordos Mongolian dialect dictionary was recorded, annotated in Praat, and processed to extract formant, duration, intensity, and pitch parameters. The results show that Ordos Mongolian has 11 word-initial short vowels: /ɐ/, /e/, /i/, /ɪ/, /ɔ/, /ʊ/, /ɵ/, /ʉ/, /u/, /æ/, and /œ/. Its non-word-initial syllables also contain 11 short vowels: /ɐ/, /e/, /i/, /ɪ/, /ɔ/, /ʊ/, /ɵ/, /ʉ/, /æ/, /œ/, and [ə]. Compared with Standard Mongolian, the overall tongue position of Ordos Mongolian short vowels is more retracted, while non-word-initial vowels show a lower and more retracted pattern. In addition, Ordos Mongolian short vowels generally have shorter duration and stronger intensity, whereas pitch differences between the two varieties are small. These findings provide database-supported acoustic evidence for describing the vowel system of Ordos Mongolian and for comparing regional Mongolian varieties.

Keywords: Ordos Mongolian, Acoustic analysis, Speech database, Articulatory features.

1 Introduction

Ordos Mongolian is one of the most representative regional varieties of Mongolian and is spoken primarily in the Ordos region of the Inner Mongolia Autonomous Region, China. Having developed for an extended period in a relatively isolated linguistic environment, it has preserved numerous archaic features in its phonology, lexicon, and grammar while also developing distinctive local characteristics. It therefore constitutes an important object of inquiry for research on the historical evolution, dialectal differentiation, and development of Mongolian. Although previous studies have provided relatively detailed descriptions of the phonological and dialectal features of Ordos

T. Chamaratana et al. (eds.), *Proceedings of the 2026 12th International Conference on Digital Humanities and Frontiers in Social Sciences (DHFSS 2026)*, Advances in Social Science, Education and Humanities Research 1031,
https://doi.org/10.2991/978-2-38476-611-6_19

Mongolian, most have relied primarily on traditional linguistic methods, and analyses of its vowel system have lacked systematic evidence from experimental phonetics. With advances in speech database construction and acoustic analysis technologies, the investigation of linguistic phenomena through objective acoustic measurements has become an important methodological approach in modern linguistics. Against this background, the present study constructs a speech database of Ordos Mongolian and employs Praat to quantitatively analyze the formant frequencies, duration, intensity, and fundamental frequency of short vowels occurring in word-initial and non-word-initial syllables. These acoustic parameters are further compared with their counterparts in Standard Mongolian. The study aims to identify the acoustic characteristics and phonetic realization patterns of the Ordos Mongolian vowel system, thereby providing a scientific basis and empirical evidence for research on Mongolian dialects, the development of speech databases, and the preservation and transmission of minority-language resources.

2 Speech Database Construction

2.1 Corpus Design

Corpus design constituted a crucial part of this study. The Ordos Mongolian linguistic corpus was selected based on Cao Namu's Ordos Dialect Dictionary[2]. The corpus comprised 376 monosyllabic words, 1,551 disyllabic words, 359 trisyllabic words, and 32 polysyllabic words. The fieldwork was conducted in Ejin Horo Banner, Ordos, with four informants. Each informant contributed approximately five hours of recordings, yielding a total of about 2,318 words.

Following data acquisition, the Praat software, developed by the University of Amsterdam, was used for annotation, resulting in the establishment of the Ordos Mongolian Speech Annotation Database. A three-tier annotation scheme was adopted.

The first tier, the segment tier, employed a phonetic transcription method based on actual pronunciation, capturing phonetic variations and segment durations. The second tier, the syllable tier, adopted a phonemic transcription method aligned with the phonological system of the target language. The third tier, the word tier, also used phonemic transcription, reflecting the phonological system or orthographic form of the target language.

Since Praat does not support direct International Phonetic Alphabet (IPA) input, SAMPA (Speech Assessment Methods Phonetic Alphabet) symbols were used for annotation.

2.2 Automatic Extraction of Acoustic Parameters

Acoustic parameters were automatically extracted from the annotated files using a script provided by the Institute of Ethnology and Anthropology, Chinese Academy of Social Sciences. The extracted data were subsequently organized in Excel spreadsheets. The acoustic parameters examined in this study included VD, VA, VP, F1, F2, SF, BF, and EF. VD represents vowel duration, VA represents vowel intensity, and VP

represents vowel pitch. The value of F1 corresponds to tongue height, while the value of F2 corresponds to tongue advancement. SF, BF, and EF are prosodic parameters related to pitch contour, denoting the starting frequency, turning point frequency (if any), and ending frequency of the pitch contour, respectively[3]. As shown in Fig. 1, sample data from the Ordos Mongolian Acoustic Parameter Database and the Standard Mongolian Acoustic Parameter Database are presented below.

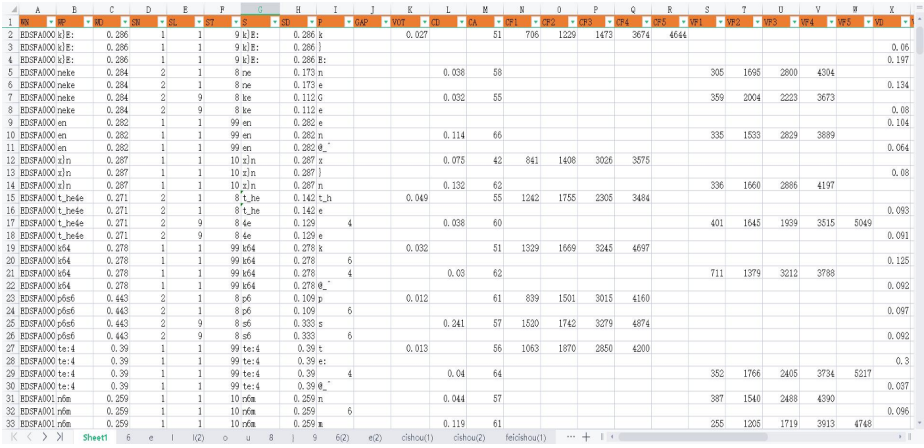

Fig. 1. Sample data from the Ordos Mongolian Acoustic Parameter Database and the Standard Mongolian Acoustic Parameter Database

3 Comparative Analysis of Formants in Word-Initial Short Vowels

Ordos Mongolian has 11 word-initial short vowels: /ɐ/, /e/, /i/, /ɪ/, /ɔ/, /ʊ/, /ɵ/, /ʉ/, [u], /æ/, and /œ/. As shown in Fig. 2, the acoustic vowel charts for word-initial short vowels in Ordos Mongolian (male and female speakers) present the formant distribution used for the following analysis.

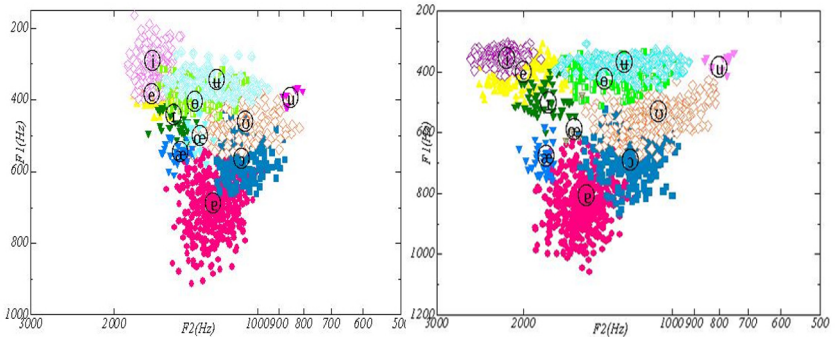

Fig. 2. Acoustic formant pattern chart of Ordos Mongolian

3.1 The vowel /ɐ/ in Ordos Mongolian

This is a near-low, central, unrounded vowel. In traditional phonetics, scholars generally considered this vowel to be a back, low vowel, transcribed as [ɑ], whereas Zhana identified it as a near-low, central vowel. The results of the present acoustic analysis align with Zhana's findings, indicating a near-low, central vowel (mean values: male speakers: F1 = 695 Hz, F2 = 1233 Hz; female speakers: F1 = 837 Hz, F2 = 1521 Hz). Therefore, transcribing this vowel as [ɐ] more accurately reflects its actual pronunciation. In polysyllabic words, when the vowel [ɪ] appears in a subsequent syllable, this vowel may shift to [æ] (mean values: male speakers: F1 = 532 Hz, F2 = 1431 Hz; female speakers: F1 = 651 Hz, F2 = 1766 Hz).

3.2 The vowel /e/ in Ordos Mongolian

This is a near-high, front, unrounded vowel. Scholarly opinions on this vowel have mainly diverged regarding its tongue height and transcription. Regarding tongue height, some scholars classify it as a near-high vowel, while others consider it a mid-high vowel. Regarding transcription, some argue it is a front vowel and should be transcribed as [e], whereas others suggest it lies between a front and a central vowel and should be transcribed as [ɘ]. The present experimental data show that this vowel is a near-high, front vowel (mean values: male speakers: F1 = 383 Hz, F2 = 1606 Hz; female speakers: F1 = 397 Hz, F2 = 2007 Hz), making [e] a fitting transcription.

3.3 The vowel /i/ in Ordos Mongolian

This is a high, front, unrounded vowel. There is a scholarly consensus that this vowel is a high, front vowel, and the present experimental results confirm this classification (mean values: male speakers: F1 = 315 Hz, F2 = 1655 Hz; female speakers: F1 = 352 Hz, F2 = 2206 Hz).

3.4 The vowel /ɪ/ in Ordos Mongolian

Acoustically, this vowel is a mid-high, central, unrounded vowel. In traditional phonetics, scholars have disagreed on whether it constitutes a distinct phoneme. Zhana considered it a near-high, front vowel, and other accounts have described it as a mid-high, front, unrounded vowel. The present experimental data, however, suggest a mid-high central vowel (mean values: male speakers: F1 = 457 Hz, F2 = 1484 Hz; female speakers: F1 = 490 Hz, F2 = 1817 Hz). The tongue position for female speakers is slightly more front.

3.5 The vowel /ɔ/ in Ordos Mongolian

This is a mid-low, back, rounded vowel. Traditional phonetics consistently identifies this vowel as mid-low and back, whereas Zhana classified it as a mid, back vowel. The present experimental results show it to be a mid-low, back vowel (mean values: male

speakers: F1 = 556 Hz, F2 = 1084 Hz; female speakers: F1 = 678 Hz, F2 = 1213 Hz). Female speakers exhibit a slightly lower and more front tongue position. In polysyllabic words, when the vowel [ɪ] appears in a subsequent syllable, /ɔ/ may be realized as [œ] (mean values: male speakers: F1 = 494 Hz, F2 = 1334 Hz; female speakers: F1 = 578 Hz, F2 = 1571 Hz).

3.6 The vowel /ʊ/ in Ordos Mongolian

This is a mid-high, back, rounded vowel. Traditional phonetics generally classifies this vowel as a near-high, back vowel, while Zhana identified it as a mid-high, back vowel. The present experimental results are consistent with Zhana's classification, confirming it as a mid-high, back vowel (mean values: male speakers: F1 = 462 Hz, F2 = 1086 Hz; female speakers: F1 = 552 Hz, F2 = 1107 Hz). Female speakers have a slightly lower tongue position.

3.7 The vowel /ɵ/ in Ordos Mongolian

This is a near-high, central, rounded vowel. Traditional scholarship has shown divergence regarding the tongue advancement and transcription of this vowel, with the main views being: (1) a front vowel transcribed as [ø]; (2) a central vowel transcribed as [ɵ]; or (3) a vowel between front and central, also transcribed as [ɵ]. Zhana considered it a near-high, central vowel. The present experimental data support Zhana's view, indicating a near-high, central vowel (mean values: male speakers: F1 = 378 Hz, F2 = 1328 Hz; female speakers: F1 = 407 Hz, F2 = 1313 Hz), and the transcription [ɵ] is more appropriate for its actual pronunciation.

3.8 The vowel /ʉ/ in Ordos Mongolian

This is a high, central, rounded vowel. Scholars consistently classify this vowel as high and central. The present experimental results likewise confirm it as a high, central vowel (mean values: male speakers: F1 = 355 Hz, F2 = 1267 Hz; female speakers: F1 = 374 Hz, F2 = 1255 Hz). Additionally, when the vowel [ʉ] appears in a syllable adjacent to the consonants /k/ or /x/, it may be realized as the vowel [u] (mean values: male speakers: F1 = 389 Hz, F2 = 849 Hz; female speakers: F1 = 367 Hz, F2 = 796 Hz).

3.9 The vowel [u] in Ordos Mongolian

This is a near-high, back, rounded vowel. Traditional phonetics has never discussed the vowel [u] as a distinct category. Based on the present experimental data, this vowel is identified as a conditional allophone of /ʉ/. When the vowel [ʉ] appears in a syllable adjacent to the consonants /k/ or /x/, it is realized as [u], which is a near-high, back vowel (mean values: male speakers: F1 = 389 Hz, F2 = 849 Hz; female speakers: F1 = 367 Hz, F2 = 796 Hz).

3.10 The vowel /æ/ in Ordos Mongolian

This is a mid, front, unrounded vowel. Scholars have disagreed on the existence of this vowel as a distinct phoneme. The present experimental data indicate that this vowel arises from [ɐ] when influenced by the vowel /ɪ/ in a subsequent syllable or by a preceding consonant /j/. It is characterized as a mid, front vowel (mean values: male speakers: F1 = 548 Hz, F2 = 1470 Hz; female speakers: F1 = 688 Hz, F2 = 1828 Hz). Male speakers show a slightly lower and more front tongue position.

3.11 The vowel /œ/ in Ordos Mongolian

This is a mid, central, rounded vowel. Scholarly opinions on this vowel are inconsistent. Some scholars consider it an allophone of /ɔ/, arising mainly under the influence of a following vowel /ɪ/, while others argue that it does not exist in this dialect. The present experimental data confirm its existence. Its articulatory features are those of a mid, central, rounded vowel (mean values: male speakers: F1 = 494 Hz, F2 = 1334 Hz; female speakers: F1 = 578 Hz, F2 = 1571 Hz).

4 Comparative Analysis of Acoustic Parameters: Duration, Intensity, and Pitch

4.1 Comparison of Vowel Duration

Duration refers to the length of a sound, with its acoustic parameter being length of time, which depends on the duration of articulatory movement[4]. Duration is measured in milliseconds (ms). In experimental phonetic studies of Mongolian, considerable research has been conducted on vowel duration. Scholars such as Quejingzhabu, Baiyinmende, and Hasiqimuge have proposed that vowel duration is related to syllable type, number of syllables, presence or absence of a preceding consonant, and the type of preceding consonant.

The present experimental data show that the order of word-initial short vowel duration in Ordos Mongolian from longest to shortest is as follows: for male speakers, [u] > /æ/ > /e/ > /œ/, /ɐ/ > /ɔ/ > /ɵ/ > /i/, /ʊ/ > /ʉ/, /ɪ/; for female speakers, [u] > /ɐ/ > /ɔ/ > /æ/ > /ɵ/ > /i/, /ʊ/ > /e/ > /ʉ/ > /œ/ > /ɪ/. Among these, the vowel [u] has the longest duration (male speakers: 83 ms; female speakers: 92 ms), while the vowel /ɪ/ has the shortest duration (male speakers: 56 ms; female speakers: 56 ms).

4.2 Comparison of Vowel Intensity

Intensity corresponds to the amplitude of sound waves[5]. Greater amplitude results in stronger sound, while smaller amplitude results in weaker sound. Intensity is measured in decibels (dB). The present experimental data show that the order of word-initial short vowel intensity in Ordos Mongolian from strongest to weakest is as follows: for male speakers, /æ/ > /e/, /œ/ > /ɐ/, /ɵ/, /ɔ/ > /ʉ/, [u] > /ʊ/, /ɪ/ > /i/; for female

speakers, /ɔ/, /ʊ/, /ɵ/, /ʉ/, [u] > /e/, /œ/ > /ɐ/, /æ/, /i/ > /ɪ/. The vowel /æ/ for male speakers (68 dB) and the rounded vowels for female speakers (70 dB) show the strongest intensity. The vowel /i/ for male speakers (62 dB) and the vowel /ɪ/ for female speakers (66 dB) show the weakest intensity.

4.3 Comparison of Vowel Pitch

Pitch is determined by the fundamental frequency of a sound wave. Within a certain range, higher vibration frequency results in higher pitch, while lower frequency results in lower pitch. Frequency refers to the number of cycles per second and is measured in hertz (Hz). The present experimental data show that the order of word-initial short vowel pitch in Ordos Mongolian from highest to lowest is as follows: for male speakers, /e/, /æ/ > /ɵ/ > /œ/ > /ɔ/, /ʉ/ > /i/ > /ɐ/, /ɪ/ > /ʊ/ > [u]; for female speakers, /e/ > /ɵ/ > [u] > /ʉ/ > /i/, /ɪ/ > /æ/ > /ɐ/, /ʊ/ > /ɔ/ > /œ/. Among these, the vowel /e/ has the highest pitch (male speakers: 125 Hz; female speakers: 179 Hz). The vowel [u] for male speakers (113 Hz) and the vowel /œ/ for female speakers (160 Hz) show the lowest pitch.

5 Comparative Analysis of Acoustic Features Between Ordos Mongolian and Standard Mongolian

This study applied the theories and methods of experimental phonetics to conduct an acoustic analysis and comparison of vowels in Ordos Mongolian with respect to formants, duration, intensity, and pitch. The following conclusions were drawn.

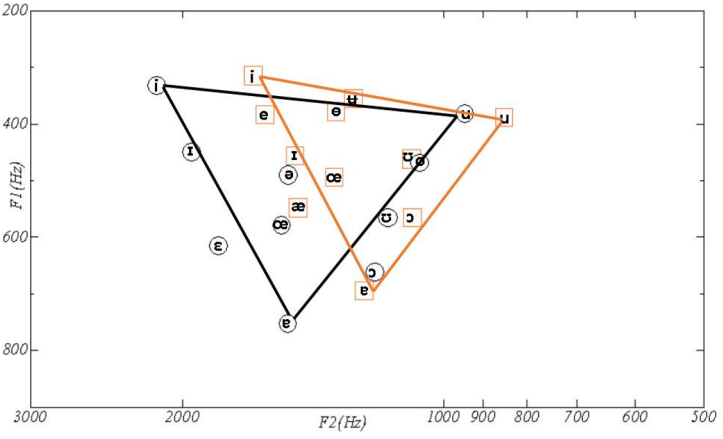

Fig. 3. Acoustic vowel space comparison between Ordos Mongolian and Standard Mongolian

Word-initial short vowels in Ordos Mongolian comprise 11 vowels: /ɐ/, /e/, /i/, /ɪ/, /ɔ/, /ʊ/, /ɵ/, /ʉ/, [u], /æ/, and /œ/. Based on the formant parameters and acoustic vowel space of the word-initial short vowels, the overall tongue position of

Ordos Mongolian word-initial short vowels is higher and more retracted than that of Standard Mongolian. The acoustic vowel space comparison between the two varieties is shown in Fig. 3.

In terms of duration, the word-initial short vowels in Ordos Mongolian are shorter than those in Standard Mongolian. In terms of intensity, the word-initial short vowels in Ordos Mongolian are stronger than those in Standard Mongolian. In terms of pitch, the word-initial short vowels in Ordos Mongolian are similar to those in Standard Mongolian.

Non-word-initial short vowels in Ordos Mongolian consist of 11 vowels: /ɐ/, /e/, /i/, /ɪ/, /ɔ/, /ʊ/, /ɵ/, /ʉ/, /æ/, /œ/, and [ə]. In Standard Mongolian, there are six non-word-initial short vowels: [ɜ], /ə/, /i/, /ɪ/, [ɞ], and [ɵ]. The vowels /ɐ/, /e/, and /ɔ/ in Ordos Mongolian non-word-initial syllables correspond to [ɜ], [ə], and [ɞ], respectively, in Standard Mongolian.

Based on the formant parameters and acoustic vowel space of the non-word-initial short vowels in the two varieties, the overall tongue position of Ordos Mongolian non-word-initial short vowels is lower and more retracted than that of Standard Mongolian.

6 Conclusion

This study conducted an acoustic analysis of vowels in Ordos Mongolian using acoustic methods and theories, and compared them with those of Standard Mongolian. An Acoustic Parameter Database of Ordos Mongolian Speech was first established using experimental phonetic methods. Following statistical analysis of the database, acoustic charts and vowel-space plots for the short vowels examined in Ordos Mongolian and Standard Mongolian were generated using the plotting software Sma4win. Descriptive methods were then employed to characterize and compare each vowel feature between the two varieties. This research provides a more accurate and scientific representation of the similarities and differences in vowel articulation between the two varieties.

Acknowledgments

This work was supported by the National Social Science Fund of China (22BYY165); National Natural Science Foundation of China (62366045); Fundamental Operations Budget Project of Central Universities (31920230123); Fundamental Operations Budget Project of Central Universities (31920240027); Special Fund for Basic Scientific Research Operations of Central Universities (31920250001-9); Science Research Project Supported by Shaanxi Provincial Department of Education (23JK0329).

Disclosure of Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Wu, Z., Lin, M.: An Outline of Experimental Phonetics. Revised and expanded by Bao, H., Lin, M. Peking University Press, Beijing (2014). https://ling.cass.cn/keyan/xueshuchengguo/cgtj/202112/t20211209_5380367.html
2. Caonamu: A Dictionary of the Ordos Dialect, rev. ed. Inner Mongolia People's Publishing House, Hohhot (2008). https://www.jd.com/brand/1315abfabcc88b9bcaf0.html
3. Huhe: An Acoustic Study of Mongolian Speech. Social Sciences Academic Press, Beijing (2018) .https://doi.org/10.978.7520/123235
4. Reetz, H., Jongman, A.: Phonetics: Transcription, Production, Acoustics, and Perception. Translated by Cao, M., Li, A. China Social Sciences Press, Beijing (2018), p. 219. https://paslab.phonetics.org.cn/?p=5830
5. Xie, F.: Mandarin Pronunciation and Voice Production, 2nd ed. Beiyue Literature and Art Publishing House, Taiyuan (2015), p. 15. https://libspace.hubu.edu.cn/mspace/searchDetail-Local/m0dc12795e018ed5f269b56ca7d668fd4