

On the relation between the similarity of the acoustic distribution patterns of vowels and the language closeness



Huhe Harnud¹ · Zhou Xuewen¹

Received: 24 February 2021 / Accepted: 27 September 2021 / Published online: 30 October 2021
© The Author(s) 2021

Abstract

Based on the “Unified Platform for Speech Acoustic Parameters of Chinese Minority Languages”, this paper calculates and compares the acoustic distribution of vowels in Mongolian, Uyghur, and Ewenki and proposes a hypothesis that the relevance between the similarity of the acoustic distribution patterns of vowels and language closeness does exist. It indicates that the acoustic pattern implies clues of closeness and relevance among the three languages. The results demonstrate that, in terms of vowels, Mongolian and Ewenki are closely related. Both those languages and the Uyghur language are distant relatives, with only typological similarity. This paper provides a new perspective for the research methodology of language kindred. It proves that the comparison of acoustic pattern is of significance in studies in linguistics, linguistic typology, historical comparative linguistics, and anthropology.

Keywords Altaic languages · Acoustic distribution pattern · Similarity · Closeness

Introduction

Since the first hypothesis about the relationship between Altaic and some other languages was proposed by Swedish officer Philipp Johann von Strahlenberg in the first half of the eighteenth century, Altaic linguistics has undergone nearly 300 years of research. If *Einführung in die Altaische Sprachwissenschaft* written by G. J. Ramstedt

✉ Huhe Harnud
huhe@cass.org.cn

Zhou Xuewen
zhouxw@cass.org.cn

¹ The Institute of Ethnology and Anthropology, Chinese Academy of Social Sciences, Beijing, China

(1952, 2004), a Finnish scholar, laid the theoretical foundation of Altaics and established the “Altaic Theory”, which separated the “Altaic languages” from the “Ural Altaic language” and became an independent Altaic Theory, then it can be said that publications of N. N. Poppe’s (1897-1991, American scholar) *Altaiisch und Urtürkisch* (Ungarische Jahrbücher, Bd.VI, Berlin 1926), *Introduction to Mongolian Comparative Studies* (MS-FOU 110, Helsinki 1955), *Vergleichende Grammatik der Altaischen Sprachen* (Otto Harrassowitz, Wiesbaden, 1960), *Introduction to Altaic Linguistics* (Otto Harrassowitz, Wiesbaden, 1965, etc. have advanced research on the Altaic Theory. For many years, although historical comparative linguists have consolidated the Altaic Theory from aspects of phonetic features, word formation methods, syntactic structures, phonetic rules and cognate words, many scholars still suspect or oppose Altaicism. For example, the British scholar G. Clauson, the German scholar G. Doerfer et al. maintain that the Turkic language had a strong influence on Mongolian and that Mongolian language had a great influence on Tungusic (Clauson 1959, 1962; Doerfer 1966). The common components (similarities) among Turkic, Mongolian and Manchu-Tungusic languages have evolved from the original structural similarity and language contact (borrowing and influencing). These scholars argue that the Altaic languages have only typological similarity instead of kindred. One important proof is that there are no common numerals and cognate words between Mongolian and Turkic languages. Scholars such as J. Benzing, L. Ligeti, K. Grönbech, D. Sinor, A.Róna-Tas, etc., claim that it is too early to conclude that “Altaic languages” have relevance in etymology. Some scholars hold the opposite stand towards to the Altaic Theory and some scholars suggest more research before making conclusions.

Apparently, scholars realize that commonalities and similarities of phonetic features, word formation and syntactic structure are not sufficient to authenticate the features of original languages. These may be common features rather than original features, and cannot be used to illustrate that these languages have a common origin. This is a major difference between historical comparative linguistics and linguistic typology. For example, although there are some similarities in phonetic features, word formation and syntactic structure for Ural and Altaic languages, the commonalities in phonetic corresponding and cognate words are scarce. Therefore, most scholars do not support the theory of kinship between Ural and Altaic languages (Huhe 2013a, b).

Altaic languages are mainly spoken in the vast area of North and Central Asia and some areas in Europe. According to statistics, there are about 100 million people speaking an Altaic language, excluding Japan. Then, how many languages belong to the Altaic language family? Due to the differences in the classification of language family, language, and dialect, consensus has not yet been achieved. According to the latest viewpoints of most scholars, there are about 50 Altaic languages. The Altaic language family has been widely used in China, covering all the Altaic languages such as Uyghur, Kazakh, Uzbek, Western Yugu, Kirgiz, Tatar, Salar, Tuvan (Turkic language family); Mongolian, Monguor, Dawoer, Dongxiang, Eastern Yugu, Baoan (Mongolic language family); Manchu, Xibo, Ewenki, Oroqen, Hezhe (Manchu-Tungusic family), etc. Although China is the birthplace of Altaic languages, the contribution of Chinese scholars to the Altaic Theory is far less than that of foreign scholars. In recent years, some scholars, such as Qinggeltai, Geng Shimin, Litipu Tohuti and Chaoke have published some research works; however, most of them are

engaged in the descriptive study of a single language family or a comparative study within the same language family. Achievements in the comparative study of Altaic languages based on the perspective of historical comparative linguistics have rarely been reported (Hugjiltu 2004).

Since the establishment of the Permanent International Altaic Conference at the 24th International Conference of Orientalism held in Munich, Germany in 1957, until 2019, 62 international Altaic Conferences have been held, continuously advancing the research of international Altaics. However, due to the complexity of the Altaic Theory and the shortage of research resources all over the world, the Altaic Theory still remains at the stage of hypothesis. In order to make a breakthrough in Altaic research, (1) we should not only actively conduct the comparative study of Altaic languages (language history) using the theories and methodologies of historical comparative linguistics, (2) but also conduct the quantitative and qualitative research of language ontology based on modern science and technology, such as experimental phonetics, computational linguistics, statistics and so on, (3) more importantly, apply archaeology, anthropology, ethnology (folklore) and historiography to examine the living traces left by the Altaic ethnic groups; (4) investigate the genetic information and ethnic differences of these ethnic groups by applying anatomy, genetics, and especially DNA technology. Up to present, scholars have mainly focused on the first and the third methods (these two are based on living traces. However, most of the Altaic ethnic groups are nomadic, and very few written documents and relics exist), while the second and third methods are rarely used (these two are empirical and rich resources can be utilized). This paper applies the second method (experimental phonetics) to conduct the research.

Theory and methods

Since the early 1990s, we conducted some quantitative and qualitative studies on the segmental and suprasegmental phonetic features of Mongolic dialects and even Altaic languages by applying the theories and methods in experimental phonetics, made some new findings which are different from those found by traditional linguistic studies, and propose some new theories, methods, and viewpoints to solve problems in phonetics and prosody which cannot be resolved with traditional linguistic means. However, our previous research focuses mainly on the averaged values of the acoustic parameters of a single language, and less attention is paid to the distribution pattern and variation (range and trend) of phonetic segments in acoustic space. Generally, we pay too much attention to the description of the synchronic status (static state) of segments while ignoring the historical evolution (dynamic) of segments.

Since 2013, to facilitate research on the acoustic feature of speech, our research team has developed software tools to automatically label and retrieve phonetic feature of acoustic parameters. We also developed “Unified Platform Software” (Unified Platform for Speech Acoustic Parameters of Chinese Minority Languages), which accomplishes inquiry, output, and analysis of acoustic parameters of speech. So far, the Unified Platform contains acoustic parameters (vowel, consonant, segmental and suprasegmental features) of 10 minority languages. Each language contains a

word-list of 1000-2000 polysyllabic words which are pronounced at reading speed. Based on the Unified Platform (Huhe et al. 2009), by analyzing and studying “vowel acoustic dynamic distribution”, “voice acoustic distribution pattern” and “voice acoustic distribution type” of a single language, we found that the similarity of acoustic distribution pattern of segments is related to language closeness. By comparing the similarities of the “vowel acoustic triangle” and the “acoustic distribution pattern” between Mongolic and Altaic languages, we examine the closeness or relevance between these languages (Huhe 2013a, b, 2016, 2019). Through these empirical researches, we realize that our results and conclusions can be used to verify and correct the conclusions obtained in historical comparative linguistics and consolidate the Altaic Theory. At present, we mainly focus on this issue so as to verify the relevance among languages. The Research Roadmap is demonstrated in Fig. 1.

However, in terms of the relationship among languages, terms such as similarity, closeness, relevance, and kindred (or kinship) are used to evaluate distances among languages (the four terms are sorted by distances from far to near). With a high similarity, two languages are close to each other. With a high closeness, two languages may have relevance. But, it is impossible to conclude kinship of languages with just

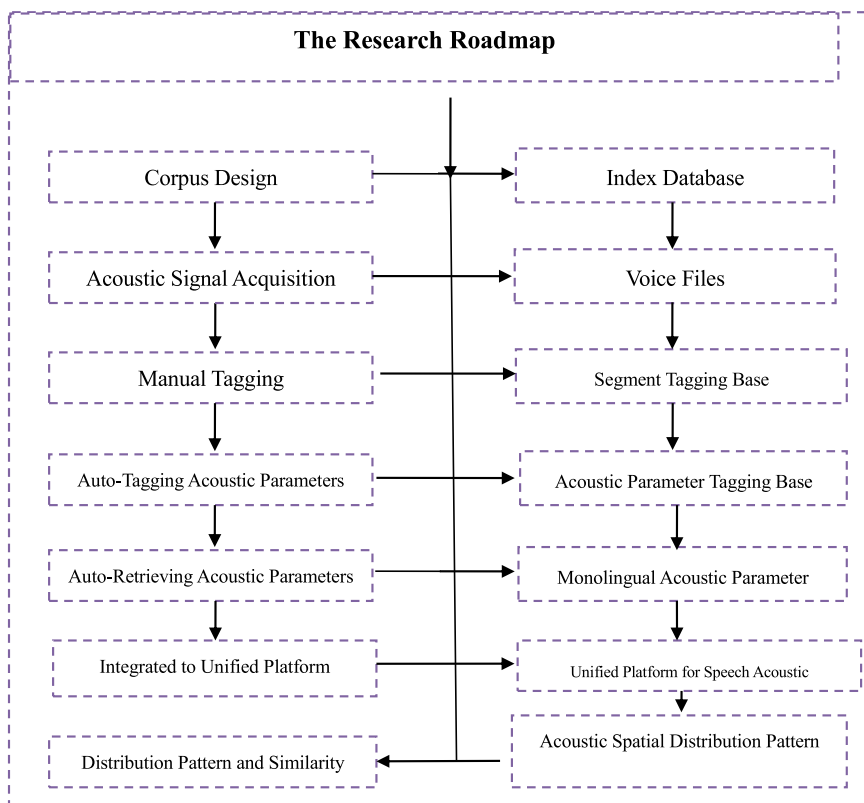


Fig. 1 The Research Roadmap

relevance because it involves multiple complicated factors which exceed our discussions in the paper.

Our proposed “acoustic distribution model of phonetic segments” is visible and measurable. In order to distinguish the individuality and generality of the acoustic distribution patterns of phonetic segments, we suggest that the acoustic distribution characteristics of monolingual speech be called “acoustic distribution pattern” (actual system), and the original model reconstructed by analyzing and comparing the multilingual “acoustic distribution pattern” be called “phonemic distribution pattern” (reconstructed system). The acoustic distribution pattern (modern and actual) of Altaic vowels is shown in Fig. 2. The phonemic distribution pattern (ancient and reconstructed) of Altaic vowels is shown in Fig. 3. In the two figures ellipses represent phonemes and allophones. The numbers and the scopes of phonemes and allophones in the two figures are different, indicating evolution trend of phonological systems. In Figs. 2 and 3, all data come from male informants (MGYM - Mongolian, UGYM - Uyghur, EWKY - Ewenki).

First, the first and second formants of all vowels in the first syllable of each word for each language were extracted by using the Unified Platform. Thus, the acoustic distribution pattern of vowels in each language could be computed. Second, based on the vowel acoustic distribution pattern, we demonstrated the acoustic distribution pattern of vowels in each language (Figs. 2 and 3). Thirdly, we calculated the similarity between the two languages with “histogram distance method” and “block histogram method” (please refer to Tables 1, 2 and 3 and notes [1]). Finally, the similarity and closeness were analyzed by using historical comparative linguistics and experimental phonetics and the vowel phonemic pattern of the original language was reconstructed.

The histogram distance calculation algorithm was applied to measure the similarity of two images. Firstly, the histograms, Hista and Histb, of the two images were calculated respectively. Then, the normalized correlation coefficients (Bhattacharyya distance, histogram intersection distance) of the two histograms were computed. Bhattacharyya distance refers to the similarity of two discrete or continuous probability distribution values. It is closely related to the Bhattacharyya coefficient, which was used to measure the overlapping between two statistical samples. Meanwhile, the Bhattacharyya coefficient can be used to measure the dispersion of class variables.

Results

Similarity comparison of vowel acoustic distribution patterns (actual)

As Table 1 shows, the similarity between Mongolian and Ewenki (the highest similarity value reaches 85%) is higher than that between Mongolian and Uyghur, or Uyghur and Ewenki. Here is the result:

“Histogram Distance” (similarity from large to small):85% (Mongolian — Ewenki) > 79% (Mongolian — Uyghur) > 76% (Uyghur — Ewenki).

“Block Histogram Distance” (similarity from large to small):67% (Mongolian — Ewenki) > 54% (Mongolian — Uyghur) > 52% (Uyghur — Ewenki).

Table 3 (calculated based on acoustic parameters) shows that the similarity between Mongolian and Ewenki (the highest similarity value reaches 69%) is higher

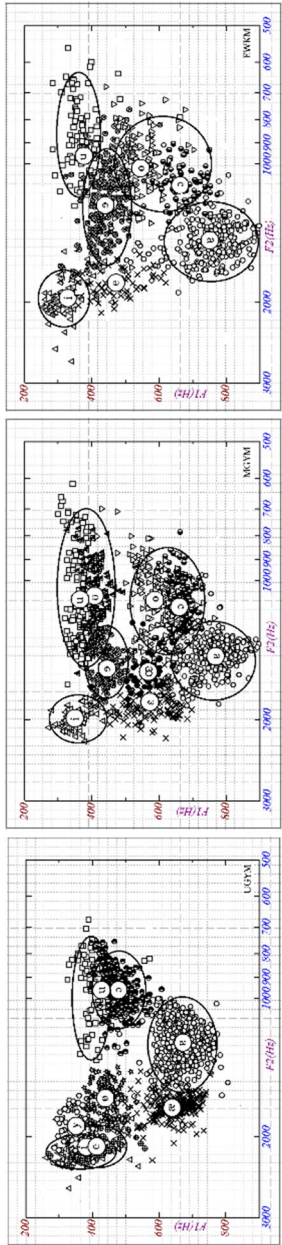


Fig. 2 Acoustic distribution patterns of vowels in Uyghur, Mongolian and Ewenki (actual)

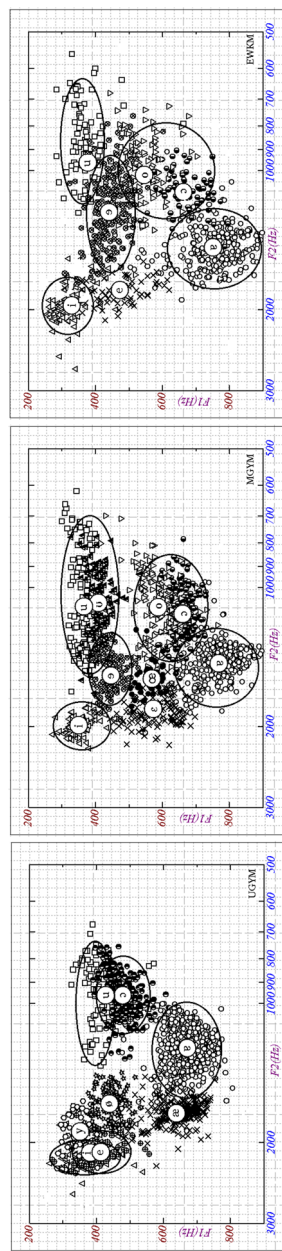


Fig. 3 Phonemic distribution pattern of vowels in Uyghur, Mongolian and Ewenki (reconstructed)

Table 1 Comparison of acoustic distribution patterns of vowels in three languages (actual)

Algorithm and result Languages	“Histogram Dis- tance”	“Blocked Histogram Distance”
Mongolian — Ewenki	85%	67%
Mongolian — Uyghur	79%	54%
Uyghur — Ewenki	76%	52%

Note: image similarity comparison based on Fig. 2

Table 2 Comparison of acoustic distribution patterns of reconstructed vowels in three languages (reconstructed)

Algorithm and result Languages	“Histogram Dis- tance”		“Blocked Histo- gram Distance”	
	Male	Female	Male	Female
Mongolian — Ewenki	71%	69%	57%	57%
Mongolian — Uyghur	69%	62%	54%	51%
Uyghur — Ewenki	60%	67%	52%	53%

Note: image similarity comparison based on Fig. 3

Table 3 Comparison of acoustic distribution patterns of Vowels in three languages (parameter)

Algorithm and result Languages	Based on acoustic parameter	
	Male	Female
Mongolian — Ewenki	69%	64%
Mongolian — Uyghur	65%	59%
Uyghur — Ewenki	57%	46%

Note: distance similarity comparison of F1 and F2 of vowels in the Unified Platform

than that between Mongolian and Uyghur, Uyghur and Ewenki. The following is the calculation results:

Male: 69% (Mongolian — Ewenki) > 65% (Mongolian — Uyghur) > 57% (Uyghur — Ewenki).

Female: 64% (Mongolian — Ewenki) > 59% (Mongolian — Uyghur) > 46% (Uyghur — Ewenki).

Similarity comparison of vowel acoustic distribution pattern (reconstructed)

Table 2 shows that the similarity values of reconstructed vowel patterns of the three languages are basically consistent with the above results. For example, the highest similarity between Mongolian and Ewenki is 71%.

“Histogram Distance” (similarity from large to small):

Male: 71% (Mongolian — Ewenki) > 69% (Mongolian — Uyghur) > 60% (Uyghur — Ewenki).

Female:69% (Mongolian — Ewenki) > 67% (Uyghur — Ewenki) > 62% (Mongolian — Uyghur).

“Block Histogram Distance” (similarity from large to small):

Male:57% (Mongolian — Ewenki) > 54% (Mongolian — Uyghur) > 52% (Uyghur — Ewenki).

Female:57% (Mongolian — Ewenki) > 53% (Uyghur — Ewenki) > 51% (Mongolian — Uyghur).

The similarity comparison of cardinal vowels of three languages by using logarithmic quotient model

The vowel normalization algorithm extracts the essence of features of vowels by eliminating pronunciation variance caused by speaker, context etc. After evaluating the performances of some classical vowel normalization models (Johnson 2005), Zhou Xuewen proposed a high-performance vowel normalization algorithm—“logarithmic quotient model” (Xuewen 2013; Xuewen and Long 2017). We apply this model to normalize the formant values of three cardinal vowels (a, i, u) of the three languages and compare their articulation distances. Table 4 shows the distances of averaged normalized values of three vowels in the three languages. The right-most column (sum of distances of three vowels) shows that the distance between Mongolian and Ewenki is the smallest (0.090), indicating that Mongolian and Ewenki are more similar.

The relevance between the similarity of vowel acoustic distribution pattern and language closeness

Some scholars maintain that the common elements (similarity) between Altaic languages originate from the result of primitive structural similarity as well as mutual contact and influence (borrowing or interaction). Similarity in typology in Altaic languages does exist. In this paper, we only focus on the closeness of the three languages. Whether in Figs. 2 and 3 (qualitative) or from the similarity value (quantitative) in Tables 1, 2 and 3, we find out that the similarity of vowel acoustic distribution patterns between Mongolian and Ewenki is higher than that between Mongolian and Uyghur. Therefore, we are convinced that, in terms of vowels, the closeness of the first two languages is higher than that of the latter. In addition, the similarity values in Tables 1, 2 and 3 show that the similarity between the three languages is more than 50%, indicating that their remarkable typological similarity.

Table 4 Distances among normalized values of the three vowels

distances	a	i	u	Sum
Mongolian — Ewenki	0.016	0.013	0.061	0.090
Mongolian — Uyghur	0.039	0.049	0.064	0.152
Uyghur — Ewenki	0.051	0.061	0.021	0.133

Discussion and conclusion

This paper proposes the hypothesis that the similarity of the acoustic distribution pattern of vowels in phonetic segments and language closeness are related. By calculating and comparing the similarity values of the distribution pattern of vowels in the first syllable in Mongolian, Uyghur and Ewenki (Altaic language family), based on the Unified Platform, we examined the closeness and relevance among the three languages and have made the followings preliminary conclusions:

Mongolian and Ewenki languages are close relatives. They and the Uyghur language are distant relatives and share typological similarity.

We maintain that languages in the same language family share original and common “language genetic information” (abbreviated as “language DNA”). This “language DNA” in acoustic space is reflected as “acoustic parameter model of speech and prosody” (abbreviated as “speech acoustic model”). Like the DNA of organisms, “language DNA” contains the “linguistic genetic information” from the origin languages of contemporary languages’ origins. Although languages have undergone varying degrees of variation, change, and evolution in their long history, the original and common “language DNA” of the same language family is stable. By comparing the similarity found in the “acoustic parameter model” among languages, we can find out the original common “DNA” of the same language family.

Although many problems remain to be further clarified and solved, as interdisciplinary research, this study possesses pioneering implications and addresses some important issues. For example, (1) the acoustic distribution pattern of vowels includes both modern and historical sounds (evolution clues). How to accurately examine historical phonetics from a synchronic perspective? How to understand and explain the relevance between synchronic and diachronic phonetics? (2) Although the phonetic system is stable, it is also evolutionary. The synchronic model cannot fully reflect the diachronic model. What is the relationship between phonetic stability and change? (3) To what extent do modern phonetic patterns reflect historical patterns? To what extent do they reflect the change? (4) How to apply the modern experimental phonetic theory to assist the studies of historical origins of phonetics? (5) As an assessment of closeness or relevance of languages, the similarity value (index) needs to be further quantified. As the research continues, it is certain that these problems can be solved by calculating and comparing the acoustic distribution patterns and pattern similarity values of segmental and suprasegmental acoustics, thus advancing the development of linguistics, linguistic typology, historical comparative linguistics, and anthropology.

At present, the above discussions and results remain at an experimental and exploratory stage; thus, no final conclusions have been made. In addition, features of consonants and syllables are not included in the paper. Because the results of this research rely closely on image recognition and vowel normalization technologies, however, and with advanced technologies and increased acoustic data, it is expected that our research will be further expanded and deepened and the resulting conclusions will be more accurate and convincing.

Note.

[1] Bhattacharyya distance and Bhattacharyya coefficient are named after A. Bhattacharyya, a statistician who worked at the Indian Institute of statistics in the 1930s.

(1) Bhattacharyya Distance

For discrete $D_B(p, q) = -\ln(BC(p, q))$ probability p and q , which reside in the same domain X , Bhattacharyya Distance.

Among which, $BC(p, q) = \sum_{x \in X} \sqrt{p(x)q(x)}$ is Bhattacharyya coefficient.

For continuous probability p and q , Bhattacharyya coefficient is: $BC(p, q) = \int \sqrt{p(x)q(x)} dx$.

In case of $0 \leq BC \leq 1$ and $0 \leq D_B \leq \infty$, D_B does not accord with trigonometric inequality (in addition, Hellinger distance does not accord with trigonometric inequality $0 \leq D_B \leq \infty D_B \sqrt{1 - BC}$).

For Multi-variable Gauss distribution, the sum of $p_i = N(m_i, P_i)$ and $D_B = \frac{1}{8} (m_1 - m_2)^T P^{-1} (m_1 - m_2) + \frac{1}{2} \ln \left(\frac{\det P}{\sqrt{\det P_1} \sqrt{\det P_2}} \right)$ is distribution of means and covariance $p = \frac{P_1 + P_2}{2}$.

In this case, Bhattacharyya distance in the first item is related to Mahalanobis distance.

(2) Bhattacharyya coefficient

Bhattacharyya coefficient is the approximate measurement of overlapping between two samples a and b . Overlapping area is divided into sub-zones (the number is n).

$$\text{Bhattacharyya} = \sum_{i=1}^n \sqrt{(\Sigma a_i \cdot \Sigma b_i)}$$

This algorithm is based on measuring image similarity by computing mathematical vector differences. It demonstrates two advantages. First, it is easy to normalize a histogram. Second, similarity between two images with different resolution can be computed with a histogram easily and efficiently.

Acknowledgements Not applicable.

Authors' contributions Huhe is responsible for analyzing data, proposing theory and writing most of the manuscript in Chinese. Zhou verifies the theory by applying a vowel normalization algorithm, supplements part of manuscript and is the major contributor in writing the manuscript in English. All authors read and approved the final manuscript.

Funding Not applicable.

Availability of data and materials All data generated or analysed during this study are included in this published article.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Clauson, G., 1959. The Turkish elements in XIV-th century Mongolian, *Central Asiatic Journal* (The Hague--Wiesbaden), IV, 3.
- Clauson, G. 1962. *Turkish and Mongolian studies*. London: Royal Asiatic Society.
- Doerfer, G. 1966. *Zur Verwandtschaft der altaischen Sprachen*. Vol. 71, 12. Wiesbaden: Ural Altaische Jahrbücher.
- Hugjiltu (呼格吉勒图), 2004. *A comparative study of vowels of Mongolic languages* (蒙古语族语言基本元音比较研究), Huhehote: Inner Mongolia Education Press (呼和浩特:内蒙古教育出版社).
- Huhe (呼和), 2013a. "A preliminary study on the relevance of Altaic languages based on acoustic model" (基于语音声学模式的阿尔泰语系语言亲属关系初探), Beijing: Minority Languages of China (北京:民族语文), 3, pp.73—81.
- Huhe (呼和), 2013b. "Phonetics-acoustic model and relevance between languages" (语音声学模式与语言之间的亲属关系问题), Hong Kong: International Conference on Phonetics of the Languages in China (ICPLC) (香港:中国语言的国际语音学会议).
- Huhe (呼和), 2016. *A comparative study on acoustic distribution of vowels in Altaic languages* (阿尔泰语系语言元音声学空间分布特征比较研究), Indiana University (USA): *The 8th International Symposium on evolutionary linguistics* (美国印第安纳大学:第八届演化语言学国际讨论会).
- Huhe (呼和), 2019. "On the acoustic distribution types of speech acoustics" (语音声学的分布类型), Beijing: Minority Languages of China (北京:民族语文), 2019 vol. 4.
- Huhe (呼和), Hasqimeg (哈斯其木格), Zhou Xuewen (周学文), 2009. "Development approach of Chinese minority language acoustic parameter database" (中国少数民族语音声学参数数据库的开发方法), Urumqi: NCMMSC (National Conference on Man-Machine Speech Communication) 2009 (乌鲁木齐:全国人机通讯学术会议2009).
- Johnson, K. 2005. Speaker normalization in speech perception. In *The handbook of speech perception*, ed. D.B. Pisoni and R. Remez. Oxford: Blackwell Publishers.
- Ramstedt, G.J. 1952. *Einführung in die Altaische Sprachwissenschaft*. Helsinki: Pentti Aalto.
- Ramstedt, G.J., translated by Zhou Jianqi (周建奇译), 2004. *The introduction to Altaics* (阿尔泰语言学导论), Huhehote: Inner Mongolia Education Press (呼和浩特:内蒙古教育出版社).
- Zhou Xuewen (周学文), 2013. "Vowel normalization algorithm- logarithmic quotient model" (元音归一算法--对数商模型), Guiyang: NCMMSC (National Conference on Man-Machine Speech Communication) 2013 (贵阳:全国人机通讯学术会议2013).
- Zhou Xuewen, Congjun Long (周学文, 龙从军), 2017. "An efficient vowel normalization algorithm-logarithmic quotient model" (一个有效的元音归一算法--对数商模型), Seoul: COCOSA (Committee for the Co-Ordination and Standardization of Speech Databases and Assessment) 2017 (首尔:语音数据库标准和评测协调委员会2017学术研讨会).

Comments

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.