

# AN INVESTIGATION INTO MODERN MONGOLIAN VOWEL HARMONY USING REAL-TIME MAGNETIC RESONANCE IMAGING

Yoshio Saitō<sup>a</sup>, Yurong<sup>b</sup> & Kikuo Maekawa<sup>c</sup>

<sup>a</sup>Takushoku University, Japan; <sup>b</sup>Inner Mongolia University, China;

<sup>c</sup>National Institute for Japanese Language and Linguistics, Japan  
 emailsaito@gmail.com; umyurong@yahoo.co.jp; kikuo@ninjal.ac.jp

## ABSTRACT

The principle which underlies modern Mongolian vowel harmony has been a matter of dispute. Many recent studies insist that the harmonic sets of vowels are phonetically distinguished in terms of the tongue root position (i.e., the size of the pharynx), while the distinction has been traditionally described by backness of the tongue.

As no crucial articulatory evidence has been provided so far, we investigated this problem using real-time magnetic resonance imaging. In this paper we present robust physiological evidence that the two harmonic sets of vowels can be distinguished in terms of size of the pharynx and degree of tongue height in modern East Mongolian.

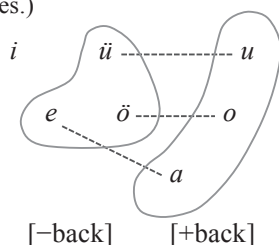
**Keywords:** Mongolian, vowel harmony, Expanded, ATR, MRI

## 1. INTRODUCTION

### 1.1. Vowel harmony

Vowel harmony is a phonological process that restricts the occurrence of vowels within a morphological unit such as a word. In the typical “Altaic-type” vowel harmony, the main principle that underlies the process had been considered to be backness of the tongue, and this harmony process is called palatal harmony, which only allows either [+back] vowels or [–back] ones within a morphological unit. For example, Middle Mongolian has seven vowels, and six of them, excluding the neutral vowel *i*, harmonize within a word or a word plus suffixes (Figure 1).

**Figure 1:** Harmonic sets of Middle Mongolian vowels (Three harmonic pairs are shown with dotted lines.)



In suffixation, for example, the harmony is realized by choosing either member of each pair of vowels for suffixes, which is determined by the category of vowels in a stem. E.g.,

[–back] vocalic stem *üje-gde-gsen-dür*  
 to see-PASS-PST-LOC

[+back] vocalic stem *yabu-gda-gsan-dur*  
 to go-PASS-PST-LOC

### 1.2. Modern Mongolian vowel harmony

The modern Mongolian language dealt with here is the Bārin dialect, which is a variety of East Mongolian spoken in the south-eastern part of the Inner Mongolian Autonomous Region in China. The Mongolian vowels in Figure 2 employ the traditional Roman transcription system widely used in Western scholarship along with the corresponding modern Bārin vowels in parentheses transcribed with IPA symbols<sup>1</sup>.

**Figure 2:** Mongolian vowels

*i* [i]    *ü* [u]    *u* [ʊ]  
*e* [ɛ]<sup>2</sup>    *ö* [ɔ]    *o* [ɔ]  
*a* [a]<sup>3</sup>

Modern Mongolian has vowel length contrast and all seven vowels have their long counterparts with the same phonetic values. As can be seen from the phonetic symbols in Figure 2, modern Mongolian has undergone a great vowel shift [12][13]. The vowel harmony pairs, however, have been maintained irrespective of the change in phonetic value.

It is difficult to describe the so-called “palatal harmony” in terms of backness of the tongue in modern Mongolian. Many recent studies insist that advanced tongue root (ATR) is the feature that governs the harmony process. Based on X-ray photography, Chinggeltei and Sinedke [2] found as early as 1959 (i.e., several years before similar studies on vowels of West African languages such as Ladefoged [6] and Stewart [11]), that the position of

the tongue root is the major difference between the two vowel harmony sets in Mongolian. Unfortunately, they did not, or could not, include X-ray photos in their article, and the original X-ray films seem to be lost. According to Svantesson et al. [13], similar results were reported by Buraev [1] for Buriad, a Mongolic language in South Siberia, and by Möömöö [10] for the Khalkha dialect of the Mongolian language spoken in Mongolia. Svantesson et al. [13], as well as Svantesson [12], supports the tongue root theory for East Mongolian including Khalkha and Inner Mongolian dialects from an acoustic phonetic viewpoint, stating that “the second member of each pair [u ~ ʊ], [o ~ ɔ], and [e ~ a]<sup>4</sup> has consistently higher F1 and normally lower F2 than the first member.” Jōo conducted some acoustic, physiological (phonolaryngographic), and auditory (electroencephalographic) experiments on modern Mongolian vowel harmony mainly from the 1990s. Jōo [5] calls it “radial harmony” from an acoustic phonetic point of view, drawing spoke-like lines radiating from one point by linking members of each harmonic pair on F1-F2 diagram. He says that his acoustic data suggest articulatory difference in backness and height of the tongue, but provide no positive evidence for pharyngeal harmony.

The purpose of our study is to investigate modern Mongolian vowel articulations to find out if there is/are any common feature(s) which could be a phonetic basis for the distinction of the vowel harmony sets.

## 2. METHOD

### 2.1. Data

MRI movies of 694 words uttered by a native speaker of the Bārin dialect<sup>5</sup> (194 words uttered three times each and the remaining 500 uttered once) were recorded at the Brain Activity Imaging Centre (BAIC) of ATR-Promotions, an affiliate company of ATR (Advanced Telecommunications Research Institute International) using 3T MRI system (Siemens MAGNETOM Prisma fit 3T). Recordings of articulatory movements in the midsagittal plane were made with FLASH sequence with acceleration factor 3. Spatial resolution was 256×256 pixels, and 1 pixel corresponds to 1 mm in length. Slice thickness was 10 mm, and temporal reconstruction rate was 14 frames per second.

A total of 166 vowels in minimal and quasi-minimal pairs of words were selected for the present study. Short vowels in non-initial syllables, which are all reduced in modern Mongolian, were excluded. The vowels investigated were short vowels in initial syllables and long vowels in initial and second

syllables (Table 1). A frame from each movie at a point closest to the middle of each vowel, which should show vowel articulation least affected by adjacent consonants, was chosen for measurements.

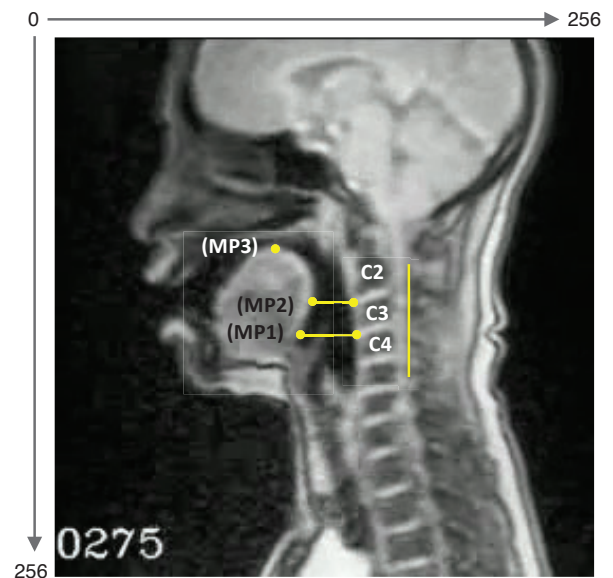
**Table 1:** Minimal and quasi-minimal pairs of words selected for the present study<sup>6</sup>

Vowels	a/e		o/ö		u/ü	
Words	am	em	om	öm	ul	ül
	gar	ger	xor	xör	udlā	üdlē
	damjix	demjix	tosol	tösöl	ujūr	üjūr
	xalax	xelax	nogō	nögō	ūr	ūr
	galīg	gerīg	ō	ō	dū	dū
	xatā	xerē	ōč	ōč	nūr	nūr
	ālǰ	ēlj	tō	tō	bīlū	bīlū
	nāx	nēx	tōl	tōl		
	yīrād	irēd				
	udlā	üdlē				

### 2.2. Measurements

Measurements of x- and y-coordinate values were made at the following points for each vowel using an image processing programme, ImageJ (Ferreira and Rasband [3]). These measured points are shown with dots in Figure 3.

**Figure 3:** Measured points



(MP1) The point on the pharyngeal wall at the intervertebral disc between the third and fourth cervical vertebrae (C3-C4) and its opposite point across the pharynx on the surface of the tongue root with the same y-axis value

(MP2) The point on the pharyngeal wall at the intervertebral disc between the second and third cervical vertebrae (C2-C3) and its opposite point across the pharynx on the surface of the tongue root with the same y-axis value

### (MP3) The highest point of the tongue

The “backness” of vowels in articulatory phonetics is not defined by physiological criteria alone, but to see the differences in shape of the oral cavity, we measured the highest point of the tongue for each vowel. In our MRI pictures the rear surfaces of the third and fourth cervical vertebrae are nearly vertical (horizontal in supine position), and, therefore, the tongue’s highest point was tentatively determined to be the closest point on the surface of the tongue to the zero point on the y-axis.

## 3. RESULTS

Independent two-tailed *t*-test was performed on the average values of x-axis at MP1 and MP2 and of x- and y-axis at MP3 for each vowel. The results are shown in Table 2.

**Table 2:** Average x- and y-axis values and statistics

Measured point Vowel	MP1		MP2		MP3	
	Lx	Px	Lx	Px	Lx	Ly
<b>a</b>	127.6	140.3	127.7	138.6	104.4	104.6
<b>e</b>	118.2	142.1	123.4	140.6	106.1	100.7
Diff.	9.4	-1.7	4.3	-2.0	-1.6	3.9
Sign.	****	****	****	****	*	****
<b>o</b>	129.4	141.7	130.5	139.9	110.6	103.7
<b>ö</b>	121.2	142.2	126.9	140.8	110.3	101.1
Diff.	8.3	-0.5	3.6	-0.9	0.4	2.6
Sign.	****	*	****	****	n.s.	****
<b>u</b>	127.4	140.9	129.1	139.6	110.4	103.0
<b>ü</b>	115.6	142.1	123.3	140.8	107.4	98.3
Diff.	11.8	-1.2	5.8	-1.2	3.0	4.8
Sign.	****	****	****	****	****	****

Lx: X-value of the tongue surface

Ly: Y-value of the tongue surface

Px: X-value of the pharyngeal wall

Significance levels:

\*\*\*\**p* < .0001, \*\*\**p* < .001, \*\**p* < .01, \**p* < .05

### 3.1. Size of the pharynx

The distance between the pharyngeal wall and the tongue root in the midsagittal plane was calculated from the measurements at two positions, MP1 and MP2.

The large positive values of the difference in Lx at MP1 and MP2 (3.6-11.8 mm) between members of each harmonic pair in Table 2 show that the tongue root is advanced in *e*, *ö*, and *ü*. In addition, as the small negative values of the difference in Px at MP1 and MP2 (0.5-2.0 mm) indicate, the pharyngeal wall is slightly pulled backward at the same time in their articulation. This means that the pharynx is expanded in *e*, *ö*, and *ü*, but it is not so in *a*, *o*, and *u*. Therefore, following Lindau [8], etc.,

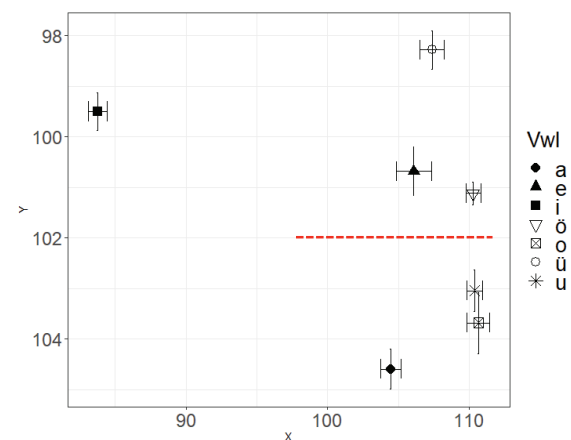
we employ the term [±Expanded], which is more appropriate than the term [±ATR] for the description of modern Mongolian vowel harmony in terms of pharynx size, although the expansion of the pharynx is mostly made by the forward movement of the tongue root.

### 3.2. Highest point of the tongue

The average x- and y-axis values of the highest point of the tongue (Lx and Ly at MP3 in Table 2) for each vowel and their 95% confidence intervals are plotted on a graph in Figure 4.

**Figure 4:** Mean location of vowels with 95% confidence interval

(A dotted line drawn parallel to x-axis on the graph separates the two sets of vowels.)



The members of each harmonic pair are significantly different in tongue height at the significance level of .0001, and the two harmonic sets can be distinguished in terms of tongue height. As for backness of the tongue, the difference between *o* and *ö* is not significant, and the two groups of vowels cannot be separated on the basis of backness of the tongue.

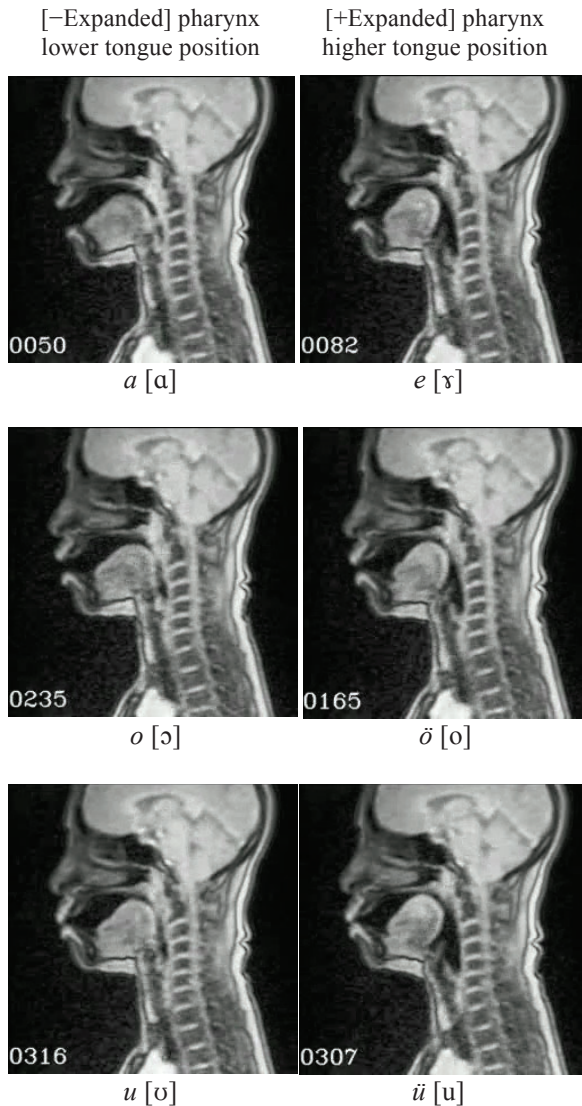
## 4. CONCLUSION

Our data show that the two harmonizing sets of modern Mongolian vowels can be distinguished by two articulatory features: the existence and non-existence of expansion of the pharynx by advancement of the tongue root (ATR) and retraction of the pharyngeal wall; and degree of tongue height (See Figure 5 for MRI pictures).

From the results of the above measurements, our view at this stage of the study is that the size of the pharynx is the major factor and the difference in tongue height is an accompanying effect of tongue root movement, taking the degree of difference in vocal tract shape into consideration.



**Figure 5:** MRI photos of the articulations of three harmonic pairs of vowels



How much acoustic difference is made by pharyngeal expansion and by raising of the tongue should be examined to determine the phonetic basis for the distinction of the two harmonic sets. We will investigate this problem in the next step of our study by synthesizing vowels using three-dimensional vocal tract shape data that we obtained in an independent experiment. Larynx height will also be taken into account.

## 5. REFERENCES

- [1] Buraev, I. D. 1959. *Zvukovoj sostav burjatskogo jazyka* [The sound structure of the Buriat language]. Ulan-Ude.
- [2] Chinggeltei, Sinedke. 1959. Monggol kelenü ündüsün egesigüdün tukai [On the basic vowels of the Mongolian language]. *Öbör monggolun yeke surgaguli erdem sinjilegenü sedkül* [Journal of Inner

Mongolia University], 97–114. Reprinted in: *Chinggelteiyin jokiyal bütiigelün tegübüri 2* [Collected works of Chinggeltei Vol. 2]. 2010. Chifēng: Öbör monggolun sinjilekü ukagan teknik mergejlün keblelün koriya, 125–146.

- [3] Ferreira, T., Rasband W. 2012. *ImageJ User Guide (IJ 1.46r)*. <https://imagej.nih.gov/ij/docs/guide/index.html>
- [4] Janhunen, J. ed. 2003. *The Mongolic Languages*. London: Routledge.
- [5] Jōo, H. 2005. *Mongoru-go boin-chōwa no kenkyū* [Studies on Mongolian vowel harmony]. Tōkyō: Bensei-shuppan.
- [6] Ladefoged, P. 1964, 1968<sup>2</sup>. *A Phonetic Study of West African Languages*. London: Cambridge University Press.
- [7] Lindau, M. 1975. *Features for vowels*. Ph.D. dissertation, UCLA.
- [8] Lindau, M. 1978. Vowel features. *Language* Vol. 54, No. 3, 541–563.
- [9] Lindau, M. 1979. The feature expanded. *Journal of Phonetics* 7, 163–176.
- [10] Möömöö. 1977. *Sistema fonem sovremennogo mongol'skogo jazyka* [The phonemic system of the modern Mongolian language]. Warszawa.
- [11] Stewart, J. M. 1967. Tongue root position in Akan vowel harmony. *Phonetica* 16, 185–204.
- [12] Svantesson, J.-O. 1985. Vowel harmony shift in Mongolian. *Lingua* 67: 283–327.
- [13] Svantesson, J.-O., et al. 2005. *The phonology of Mongolian*. Oxford: Oxford University Press.

## ACKNOWLEDGEMENTS

This study was supported by a grant from the Japan Society for the Promotion of Science [Grant-in-aid for Scientific Research (B) “Elaboration of articulatory phonetics by means of real-time MRI and WAVE data” (Project No.17H02339, Principal investigator: Kikuo Maekawa)].

We thank Tatsuya Kitamura, Yukiko Nōta, and Hironori Takemoto for their advice and help, Takuya Asai for development of an analysis and annotation system for real-time MRI movies, and anonymous reviewers for their comments and suggestions.

<sup>1</sup> As a result of umlauting, etc., some additional vowels with different phonetic values emerged in modern dialects. However, they do not concern us here in the present study.

<sup>2</sup> The symbol [ə] is used in Inner Mongolian academic literature.

<sup>3</sup> The symbol [ɛ] is used in Inner Mongolian academic literature.

<sup>4</sup> The phonetic value of the original *e* is not [ɛ] but [e] in Khalkha.

<sup>5</sup> A female speaker born in 1968 in Ulānxada (Chifēng), Inner Mongolia, and brought up there.

<sup>6</sup> Thirty-two, 27, and 24 pairs of words were selected for *a/e*, *o/ö*, and *u/ü* respectively. All words were uttered three times, but two utterances each were used for the words *nāx* and *nēx*.