

# **In search of the acoustic correlates of tongue root contrast in three Altaic languages: Western Buriat, Tsongol Buriat, and Ewen\***

Hijo KANG (Stony Brook University, USA)  
Seongyeon Ko (Cornell University, USA)

**KANG, Hijo & Seongyeon Ko (2012). In search of the acoustic correlates of tongue root contrast in three Altaic languages: Western Buriat, Tsongol Buriat, and Ewen. *Altai hakpo* 22. 179-203. The Altaic Society of Korea.**

This paper is an investigation of the acoustic characteristics of vowels in three Altaic languages: two Mongolic languages, Western Buriat (WB) and Tsongol Buriat (TB), and a Tungusic language, Ewen. Based on the formant frequencies of simple vowels in initial syllables, we first demonstrate that all the three languages can be best described as having a tongue root (TR) contrast-based vowel system. Using various acoustic measures such as F1, F2, normalized A1-A2, B1, and center of gravity, we then compare the two vowels in each harmonic pair in search of the acoustic correlates of the RTR contrast in Altaic languages. The results show an overall resemblance to the previous acoustic findings of the ATR contrast in West African languages. F1 successfully distinguishes all vowel pairs in all the three languages. B1 and the center of gravity also differentiate most vowel pairs. However, the

---

\* Earlier versions of this paper were presented at the 85th Annual Meeting of LSA (January 6-9, 2011), the CUNY Conference on the Phonology of Endangered Languages (January 12-14, 2011), and the 10th Seoul International Altaic Conference (July 15-16, 2011). We are grateful to the audiences at those events, two anonymous referees, and Prof. Juwon Kim for their comments and suggestions. All remaining errors are ours.

We use “three’ Altaic languages” in the title since, despite its name ending with “Buriat,” Tsongol Buriat is generally regarded as a variety of Mongolian and thus will be treated as such herein. See §3.1 for further details.

distinction in the values of normalized A1-A2 is observed only in WB, not in the other two languages TB and Ewen.

**Keywords:** tongue root contrast, ATR, RTR, Altaic, Buriat, Ewen, F1, normalized A1-A2, B1, center of gravity

## 1. Introduction

Tongue root harmony (Ladefoged 1964, Stewart 1967, Lindau 1979, Casali 2008 among others), once believed to exist exclusively in West African languages (Stewart 1967), has also been found in many languages in non-Turkic branches of the proposed Altaic family (see Novikova 1960, Ard 1981, Kim 1989, Li 1996, Zhang 1996 for Tungusic; Cenggeltei 1959, Svantesson 1985, Svantesson et al. 2005 for Mongolic). However, compared to the relatively well-studied West African languages (Lindau 1979, Hess 1992, Tiede 1996, Fulop et al. 1998, Hudu 2010 for Niger-Congo languages; Jacobson 1980, Guion et al. 2004 for Nilo-Saharan languages), no consensus has been reached on the nature of vowel harmony in Altaic languages, partly due to a paucity of instrumental studies. As our first step toward investigating vowels in Altaic languages, this study aims to confirm the tongue root-based nature of the vowel systems in Buriat (Mongolic) and Ewen (Tungusic) on the ground of acoustic data.

There are two tongue root features proposed in the phonological literature: [Advanced Tongue Root] and [Retracted Tongue Root] ([ATR] and [RTR] hereafter). It is highly controversial whether [ATR] and [RTR] are two distinct features or two opposing values of a single feature. However, following Steriade (1995), we assume as a working hypothesis that they are two distinct features which involve “two opposing gestures on the same or related articulatory dimensions” (pp. 149-152). It is generally accepted that the African tongue root systems utilize [ATR] whereas the Altaic systems utilize [RTR] (Clements & Rialland 2008, Kim 1989, Kim 1999, Li 1996).<sup>1)</sup> With this in background, this

---

1) Evidence is mostly phonological: for example, Li (1996: 318ff.) notes the phonological differences between African and Tungusic vowel systems in

paper attempts to identify the acoustic correlates of the hypothesized [RTR] contrast in Western Buriat, Tsongol Buriat, and Ewen by replicating the methods adopted in the previous studies of African languages. It is expected that the comparison between Altaic and African tongue root systems will contribute to the typology of tongue root harmony. This is especially so when we consider the endangered status of the three Altaic languages.

The remainder of this paper is structured as follows: Section 2 provides an overview of the previous instrumental studies of the tongue root contrast, including articulatory and acoustic studies. Section 3 describes the data and method of our acoustic study and Section 4 shows the results. Section 5 concludes the paper with a general discussion.

## 2. Previous instrumental phonetic studies

Since Stewart (1967) proposed the feature [ATR] (“root-advanced” in his terminology) for Akan vowel harmony, it has been attempted with the development of the technical methodology to unveil the phonetics of the proposed tongue root feature. This section reviews the major findings in the previous instrumental studies of tongue root contrast, which can be divided roughly into two types: articulatory and acoustic studies.

### 2.1 Articulatory studies

#### 2.1.1 Cineradiography (X-ray tracings)

In earlier studies, the articulatory mechanism of the tongue root contrast was mostly investigated by means of X-ray photography. First, Ladefoged’s (1964) X-ray tracings of Igbo vowels show that the primary difference between the two vowel sets in Igbo is the advancement vs. retraction of the *body* of the tongue. Stewart (1967) reinterpreted Ladefoged’s tongue *body* advancement/ retraction as the tongue *root* advancement/retraction. Lindau (1974, 1979) further utilized X-ray to find that the primary gesture for the [+ATR] vowels in Akan (a Niger-Congo language spoken in Ghana)

---

terms of vowel inventories, typical merger patterns, and typical neutral vowels. However, no rigorous instrumental research has been conducted to compare the phonetics of [ATR] and [RTR] contrast.

involves a change in the size of the pharyngeal cavity accomplished by lowered larynx as well as advanced tongue root. Similarly, Jacobson (1980) reports the pharyngeal cavity expansion for the [+ATR] vowels in three Nilo-Saharan languages, although the expansion is not uniformly achieved across all the languages.

Buriat and Ewen have also been investigated using cineradiography. Using X-ray tracings, Buraev (1959) rejects a palatal analysis of Buriat vowel contrast, characterizing the “soft” vowels as “raising of the central part of the tongue blade” (Svantesson et al. 2005: 220). Novikova’s (1960) X-ray images show that for the “pharyngealized vowels” of (Ola dialect of) Ewen (Ladefoged & Maddieson 1996: 306–310) the size of the pharyngeal cavity decreases as a result of pharyngeal passage narrowing and larynx raising triggered by tongue root retraction.<sup>2)</sup>

### 2.1.2 Magnetic Resonance Imaging (MRI)

In addition to the sagittal expansion reported in the previous X-ray studies, Tiede’s (1996) magnetic resonance imaging (MRI) data of Akan vowels show the lateral expansion of the pharyngeal cavity in the [+ATR] vowels. Tiede’s articulatory data also show that the ATR contrast in Akan is distinct from the English tense vs. lax contrast in spite of the similarity of the vowel distribution in the acoustic space.<sup>3)</sup>

### 2.1.3 (Transnasal) endoscopy

Recently, Edmondson & Esling’s (2006: 175–9) transnasal endoscopic (or laryngoscopic) study shows that the tongue root contrast between the two vowel sets in Somali (a Cushitic language) and Kabiye (a Gur language) involves different laryngeal valve settings. For example, the contrast between

---

2) However, as pointed out by Aralova & Grawunder (2011), the settings of Novikova’s experiment are not clearly described. More crucially, it is noticed in Ladefoged & Maddieson (1996) that all vowels in Novikova’s X-ray images and tracings have a lowered velum, which means that they are all nasalized vowels. This undermines the validity of Novikova’s X-rays.

3) In English, the change of pharyngeal region is concomitant with that of tongue body, whereas, in Akan, the pharynx is adjusted independently of the tongue body. This means Akan speakers make an active use of the pharynx to make the tongue root contrast.

“non-constricted” vowels /i e æ ö ʊ/ (traditionally described as [+ATR] vowels) vs. “constricted” vowels /ɪ ɛ ɑ ɔ u/ ([−ATR] vowels) in Somali is characterized as involving different arytenoid-epiglottal apertures.

#### 2.1.4 Ultrasound imaging

More recently, ultrasound imaging started being used as a harmless and non-invasive way to visualize the tongue root position (Hudu et al. 2009, Hudu 2010). For example, with a hypothesis on the direct mapping between articulatory gestures (the tongue root position) and phonological features (the dominance of [ATR] or [RTR] feature), Hudu et al. (2009) identify the relative advancement of the tongue root for the [+ATR] vowels in Dagbani compared to the inter-speech posture (ISP).<sup>4)</sup>

#### 2.1.5 Summary

The various image technologies have shown in common that the tongue root contrast is realized by the advancement/retraction of the tongue root and the expansion/contraction of the pharyngeal cavity from the perspective of production. These findings raised a question from the perspective of the other side of linguistic activity – perception. How do the hearers distinguish the advanced ([+ATR]) vowels from the retracted ([−ATR]) vowels or how is the contrast acoustically realized? The next section gives the answers that have been presented in the previous phonetic studies of African languages.

## 2.2 Acoustic studies

### 2.2.1 F1, F2, and F3

The first two formants have been extensively used in phonetics to characterize vowels. There are widely accepted correlations between F1/F2 and vowel height/frontness. The F1-height correlation is negative: the higher the vowel, the lower its F1 frequency. The F2-frontness correlation is positive: the fronter the vowel, the higher its F2 frequency. However, F1 and F2 are also affected by other factors such as lip rounding and pharyngeal cavity expansion.

---

4) To our best knowledge, there has been no MRI-, endoscopy-, or ultrasound-based researches conducted on Altaic languages.

In African languages, it has been pointed out that F1 is the most reliable acoustic cue of the [ATR] feature: e.g., Akan (Hess 1992), Degema (Fulop et al. 1998), Maa (Guion et al. 2004), and Yoruba (Przedziecki 2005). [+ATR] vowels have lower F1 than their [-ATR] counterparts very consistently. This has effect that [+ATR] vowels appear to be raised in the acoustic space (Ladefoged & Maddieson 1996: 305), although the lower F1 frequency of [+ATR] vowels is better associated with the pharyngeal cavity expansion rather than the actual tongue body raising, as the aforementioned articulatory studies suggest.

By contrast, F2 does not show a consistent effect on the contrast. In many languages, front [+ATR] vowels have higher F2 values than their front [-ATR] counterparts, while back vowels show the opposite pattern. This indicates that [+ATR] vowels are more peripheral than [-ATR] vowels in general. Also, F2 does not always differentiate the two series of vowels. For example, although /e/-/ɛ/, /i/-/ɪ/, and /o/-/ɔ/ pairs in Degema were successfully distinguished by F2, /ə/-/a/ and /u/-/ʊ/ pairs were not (Fulop et al. 1998).

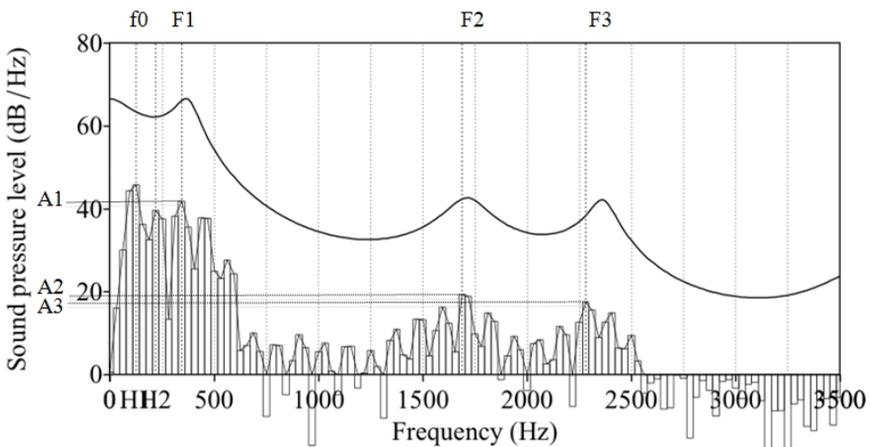
In addition to F1 and F2, the lowering of F3 has also been noticed as a possible acoustic cue for the “pharyngealized” vowels as opposed to plain ones in Caucasian languages (Catford 1994: 59). Note that Ladefoged and Maddieson (1996: 306–310) relate this vowel contrast in Caucasian to that in Ewen described in Novikova (1960).

Even though formant frequencies (F1 and F2 in particular) work as effective cues within pairs, they might fail to differentiate vowels across pairs. For instance, the high [-ATR] vowel /ɪ/ was not always separated from the mid [+ATR] vowel /e/ in terms of F1 and F2 frequencies depending on speakers (Hess 1992, Guion et al. 2004). This suggests that there might be other acoustic cues than formant frequencies.

### 2.2.2 The normalized A1-A2

On the basis of perceptual impression, Halle & Stevens (1969) suggest that [ATR] contrast is related to phonation, with [+ATR] vowels having the characteristics of “breathy” voice while [-ATR] vowels having those of

“creaky” voice.<sup>5)</sup> Thus, “spectral tilt,” the representative acoustic cue of phonation (Ladefoged & Maddieson 1996, Gordon & Ladefoged 2001), has been investigated to find an acoustic cue of [ATR]. In the acoustic phonetics literature, spectral tilt (or spectral flatness) has been measured in various forms such as H1-H2, H1-A1, H1-A2, and H1-A3<sup>6)</sup> (Garellek & Keating 2011 and references therein). Figure 1 below shows how formants, harmonics, and amplitudes are measured on a short-term spectrum.



**Figure 1.** Measurement of harmonics, formants, and amplitudes on the FFT (Fast Fourier Transform) spectrum of the middle of [ə] in *tɔŋgəri* ‘sky’ in WB

More recently, the *normalized A1-A2* was first proposed as a measure of the spectral tilt in Degema (Fulop et al. 1998) and also applied to Maa (Guion et al. 2004). The results showed that [+ATR] vowels have higher normalized A1-A2 values (breathy voice) than their [-ATR] counterparts (creaky voice)

5) Other impressionistic descriptions include “dull,” “deep,” or “hollow” for [+ATR] vowels and “bright,” “tight,” or “choked” for [-ATR] vowels (Starwalt 2008: 3).

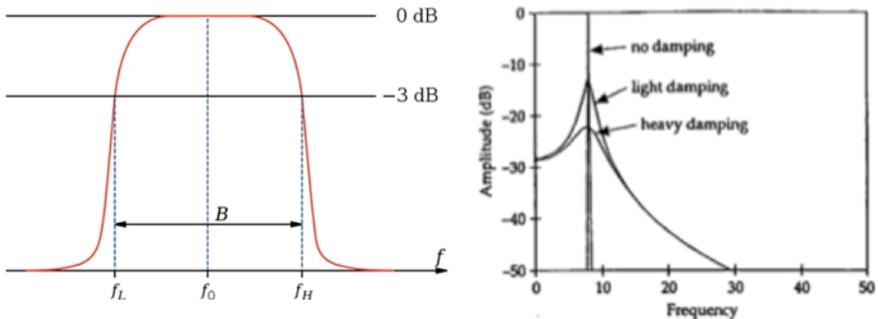
6) Here “H” stands for “harmonic” and “A” for “amplitude.” H1 coincides with f0 (fundamental frequency) and other harmonics (H2, H3, and so on) are basically integral multiples of f0. A1 is the amplitude of F1, A2 the amplitude of F2, and so on, as shown in Figure 1.

in both languages. This means that, energy is more concentrated on relatively higher formants in [-ATR] vowels. Thus, [-ATR] vowels tend to have a gentler slope in the spectrum and a relatively “brighter” impression.

However, two things should be noted about the spectral tilt as an acoustic correlate of the tongue root feature. First, spectral tilt does not present a consistent difference between the two series of vowels. Depending on language, speaker, or vowel pair, it often fails to distinguish the [+ATR] and [-ATR] vowels (Fulop et al. 1998, Guion et al. 2004). Second, as pointed out by Casali (2008: 510), the difference is not as drastic as in the distinction between actual breathy vs. creaky voice vowels in, e.g., San Lucas Quiavini Zapotec (Gordon & Ladefoged 2001), Gujarati (Fischer-Jørgensen 1967), and Hmong (Huffman 1987).

### 2.2.3 B1

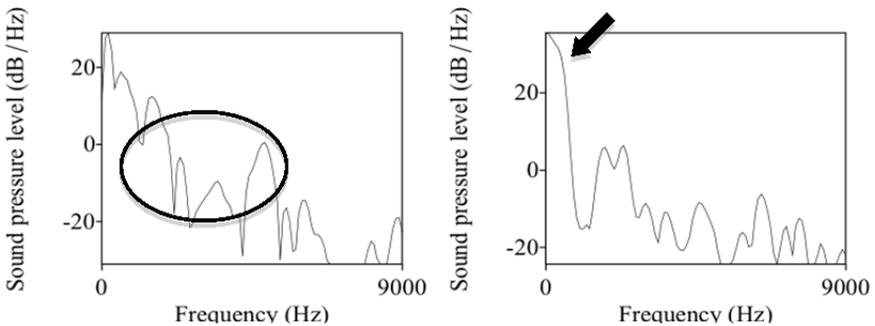
The bandwidth of F1 (B1, cf. Figure 2) has also been investigated in Akan (Hess 1992), Yoruba (Przedziecki 2005), LuBwisi, Ifè, and nine other languages (Starwalt 2008). [+ATR] vowels were found to be narrower in B1 than [-ATR] vowels. Although B1 is affected by F1, the difference in B1 was bigger than that predicted by the difference in F1 (Starwalt 2008). However, the results varied depending on languages and speakers, which means B1 may not be as consistent as F1.



**Figure 2.** Measurement of bandwidth (on the left), from Wikipedia (<http://www.wikipedia.org>): Each bandwidth is measured under 3dB from the peak of formant; bandwidth differences (on the right), from Johnson (1997: 143): Bandwidth is related to the degree of damping.

### 2.2.4 The center of gravity<sup>7)</sup>

The *center of gravity* (also known as *spectral mean*, cf. Figure 3) was recently tested as an acoustic cue of the tongue root contrast. Starwalt (2008) found in languages such as Ifè, Mboosi, and Kwa a tendency that [+ATR] vowels have lower center of gravity values.



**Figure 3.** The spectra of [a] in *naraj* ‘sun’ (left) and [ə] in *təŋgəri* ‘sky’ (right) in WB. The energy is concentrated on low frequency range in [ə] (made by arrow) but is distributed on a relatively high frequency range in [a] (made by an oval). As a result, the center of gravity is higher in [a] than in [ə].

### 2.2.5 Summary

Thus far, we have seen that F1 is the most reliable acoustic correlate of the tongue root feature. The other acoustic features introduced here are not so robust and, more importantly, not independent of each other. One acoustic feature is somehow related to another feature presumably because they are caused by the same articulatory mechanism. In spite of this incompleteness of the contemporary acoustic knowledge on the [ATR] contrast in African languages, the aforementioned acoustic studies will provide the methodological ground for the current study and make it possible to acoustically compare the

---

7) This is defined as “the measure of the mean of the frequencies of the sound’s spectrum over a specific domain” (Starwalt 2008: 94). This reflects where energy is relatively concentrated.

(hypothetically) two distinct tongue root features, [ATR] and [RTR], in African and Altaic languages, respectively.

### 3. Data and method

#### 3.1 Target languages

Buriat is a Northern Mongolic language spoken in Russia (Buriatia), Mongolia, and Inner Mongolia by 400,000 or fewer speakers (Skribnik 2003). Most Buriat varieties including Western Buriat (WB) are known to have seven short vowels /i, ə, a, u, ʊ, o, ɔ/. By contrast, Eastern Buriat has only six short vowels due to the merger of short /o/ with /u/ (Poppe 1960, Skribnik 2003, Svantesson et al. 2005).

(1) Buriat vowels (Poppe 1960, Skribnik 2003: 104)

|                 |     |
|-----------------|-----|
| i               | u   |
|                 | ʊ   |
| ə <sup>8)</sup> | (o) |
| a               | ɔ   |

As its name implies, Tsongol-Sartul Buriat (TB) is often regarded as a variety of Buriat, i.e., Southern Buriat or South Selenge Buriat. However, from both historical and linguistic perspectives, it may well be viewed as a variety of Mongolian proper for the following reasons (Skribnik 2003, Svantesson et al. 2005 and references therein): First, it is spoken by the descendants (estimated at about 20,000) of Mongolian immigrants who moved to the South Selenge area of Buriatia in the 17th century and,

---

8) This vowel has been normally rendered as /e/ in the literature. However, it has been noted that it is pronounced as a central vowel (Poppe 1960: 6) at least in some varieties. Thus, we will use /ə/ and, accordingly, present data from other sources with necessary modifications.

second, it is known not to share the same phonological developments as Buriat. Since Mongolian proper has seven vowels as (Western) Buriat, TB is expected to retain the same number of vowels. However, as we will see later, there seems to be an ongoing merger of short /o/ and /u/, which is undoubtedly a Buriat feature.

Like many other Altaic languages, Buriat displays a vowel harmony process based on which vowels can be divided into the following three groups:

(2) Harmonic sets in Buriat

- |                      |   |   |   |
|----------------------|---|---|---|
| a. Set A (= [-RTR]): | ə | u | o |
| b. Set B (= [+RTR]): | a | ʊ | ɔ |
| c. Neutral:          | i |   |   |

The Buriat vowel harmony pattern is illustrated in the following examples (Kaun 1995, drawn from Poppe 1960):

(3) Buriat vowel harmony

a. [-RTR] vowels

- əxə-də: 'mother-DAT.REFL'  
 xul-də 'foot-DAT'  
 o:r-do 'self-DAT.REFL'

b. [+RTR] vowels

- axa-da: 'elder brother-DAT.REFL'  
 xʊŋ-da 'swan-DAT'  
 mɔɖɔn-dɔ: 'tree-DAT.REFL'

(4) As the only vowel in a stem, /i/ acts like a [-RTR] vowel

- xi:-də 'dung dust-DAT'  
 ti:g-ə:d 'to do that way-GERUND'

Ewen is a Northern Tungusic language spoken by only 7,000 speakers scattered in Eastern Siberia (Russian census, 2002). It has eight vowel phonemes as presented in (5), which are grouped into the two harmonic sets as in (6).

(5) Ewen vowels (Novikova 1960, Kim 1989, Kim 2011, Li 1996)

- |   |   |
|---|---|
| i | u |
| ɪ | ʊ |
| ə | o |
| a | ɔ |

## (6) Harmonic sets in Ewen (Kim 2011, Kaun 1995)

|                      |   |   |   |   |
|----------------------|---|---|---|---|
| a. Set A (= [-RTR]): | i | ə | u | o |
| b. Set B (= [+RTR]): | ɪ | a | ʊ | ɔ |

The Ewen vowel harmony pattern is illustrated in the following examples:

## (7) Suffixal vowel harmony in Ewen (data drawn from Kim 2011)

|                  |                  |                  |                     |
|------------------|------------------|------------------|---------------------|
| a. [-RTR] vowels |                  | b. [+RTR] vowels |                     |
| hor-li           | 'go-IMP_2SG.IMP' | hilkat-li        | 'rinse-IMP_2SG.IMP' |
| tugəñi-du        | 'winter-DAT'     | jugan-du         | 'summer-DAT'        |
| toŋər-duk        | 'lake-ABL'       | bazar-dok        | 'market-ABL'        |
| hupkučək-lə      | 'school-LOC'     | dɔlbañ-la        | 'night-LOC'         |

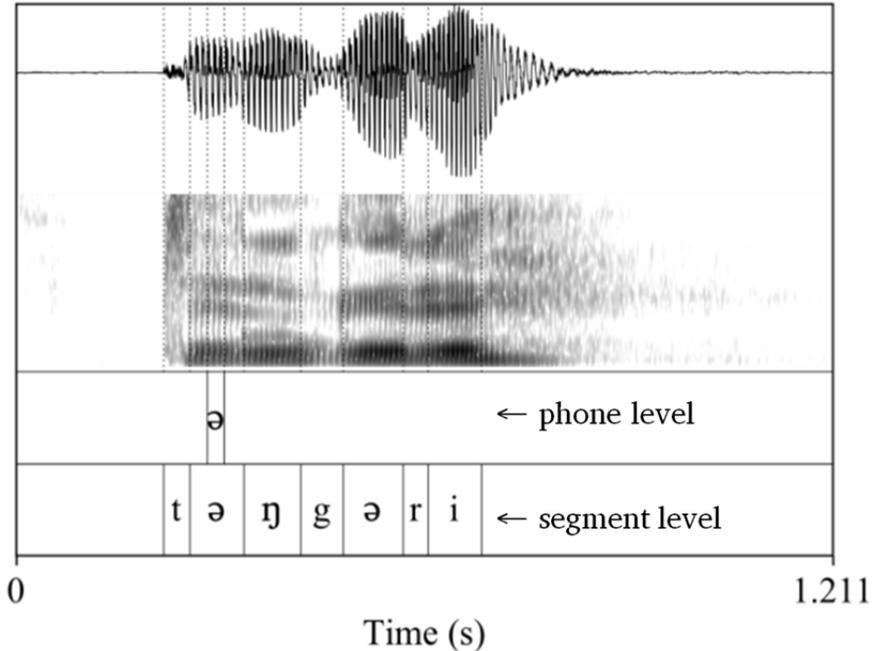
### 3.2 Data and analysis

The data were collected by the ASK REAL (the Altaic Society of Korea, Researches on Endangered Altaic Languages) project team (Kim et al. 2008).<sup>9)</sup> The WB data were recorded in Ust-Orda Buryatia, Irkutsk Oblast, Russia in January 2007 by a male speaker (born in 1943) of the Ekhirit-Bulagat variety. The TB data were recorded in Tashir, Buriatia in August 2007 by a female Tsongol speaker (born in 1950). Lastly, the Ewen data were obtained in Khabarovsk, Russia in February 2006. A female speaker of Eastern Ewen, who was born in a remote village of Magadan Oblast in 1940, participated in the recording.<sup>10)</sup>

After provided to the authors, the data were transcribed into IPA in consultation with Bosson (1962) for Buriat and Robbek & Robbek (2005) for Ewen. Only short vowels in initial syllables were segmented from the 1,271 (WB), 1,233 (TB), and 899 (Ewen) tokens. The onset of a vowel was marked on the first peak of a periodic wave and the offset was on the last peak. Figure 4 exemplifies the segmentation process.

9) The authors are deeply indebted to the principal investigator Prof. Juwon Kim who allowed us to use the ASK REAL materials.

10) See Kim et al. (2008) or go to the ASK REAL Digital Archive ([http://altaireal.snu.ac.kr/askreal\\_v25/](http://altaireal.snu.ac.kr/askreal_v25/)) for more information on the recordings or the fieldwork in general.



**Figure 4.** Segmentation of *təŋgəri* 'sky' in WB.

After the segmentation, the following acoustic features were measured at the middle of each vowel phone<sup>11</sup>: the fundamental frequency (f0), the first three formants (F1, F2, F3), the amplitudes of the three formants (A1, A2, A3), the bandwidths of the three formants (B1, B2, B3), the first two harmonics (H1, H2), and the center of gravity. Then, H1-H2, H1-A1, H1-A2, H1-A3, and the normalized A1-A2 values were calculated as a measure of spectral tilt. All the processes from segmentation to measurement were carried out using Praat (Boersma 2001). Then a series of ANOVAs (Analysis of Variance) were performed to the acoustic values for the comparison of vowels in each pair in each language. The results of the potential correlates of the tongue root

11) The phone of a vowel was defined as a period where the first three formants are stable, to say, where the vowel formants are not affected by adjacent segments (cf. Hertz 1991).

feature (F1, the normalized A1-A2, B1, and center of gravity), as well as the overall results, will be presented in the next section.

## 4. Results

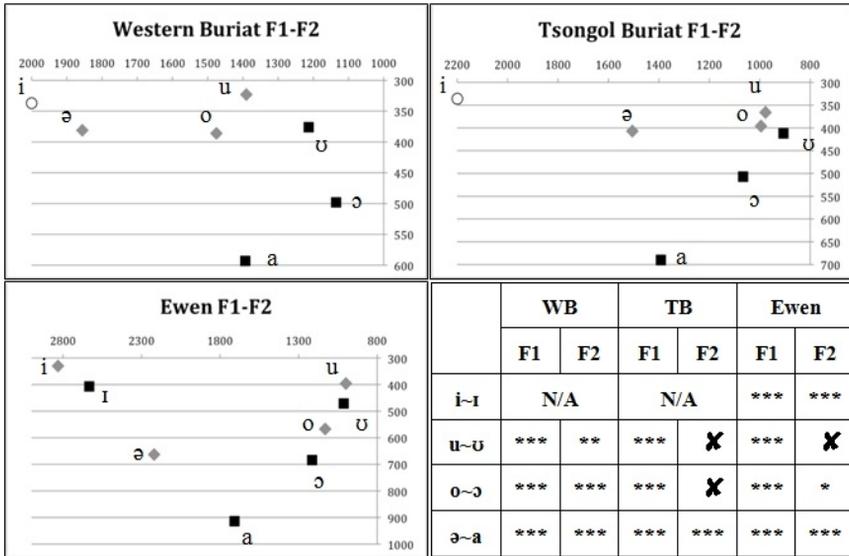
The table below shows the results of all the measured and calculated acoustic features. Above all, it should be noted that F1 is the only acoustic cue that is significantly different in every harmonic pair in every language. In WB and TB, A2, H1, H1-A2, and the center of gravity are also distinguishing acoustic cues in every harmonic pair. B1 and the center of gravity are found to be relatively robust cues, though they are not as reliable as F1. Overall, the results are very similar to those observed in West African languages.

|      | f0   | F1   | F2   | F3   | A1   | A2   | A3   | B1   | B2   | B3   | H1   | H2   | H1-A2 | H1-A3 | N.<br>A1-A2 | CoG  |
|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|-------|-------------|------|
| WB   | +2/3 | -3/3 | +3/3 |      | +3/3 | -3/3 |      | -3/3 |      | +1/3 | +3/3 | +3/3 | +3/3  |       | +3/3        | -3/3 |
| TB   | +3/3 | -3/3 |      | -1/3 |      | -3/3 | -3/3 | -2/3 | +2/3 |      | +3/3 | +2/3 | +3/3  | +3/3  | +1/3        | -3/3 |
| Ewen | +2/4 | -4/4 |      |      |      |      |      | -3/4 |      | -2/4 |      |      | +1/4  |       | +2/4        | -3/4 |

**Table 1.** Overall results (+: ATR greater, -: non-ATR greater, x/y: significant difference in x pair(s) out of y pairs)

### 4.1 F1, F2, and F3

The average F1 and F2 frequencies are marked on the three vowel formant charts in Figure 5 below. Of the two formant frequencies, F1 is a robust cue of the vowel contrast, whereas F2 is not. As the table in the lower right corner of Figure 5 shows, [+RTR] vowels were always distinguished from their [-RTR] counterparts by F1. They had higher F1 values (lower in the vowel space) than their [-RTR] counterparts consistently throughout all the pairs in the three varieties (for each pair,  $p < .01$ ).



**Figure 5.** Vowel distribution in F1(y-axis, Hz)-F2(x-axis, Hz) space (\*\*\*:  $p < .001$ , \*\*:  $p < .01$ , \*:  $p < .05$ , ✗:  $p > .05$ )

In contrast, F2 is not as consistent as F1 in terms of the direction and the significance of difference. In WB, the [-RTR] vowels consistently had higher F2 values (fronter in the acoustic vowel space) than their [+RTR] counterparts in all harmonic pairs. However, in TB, the harmonic pair of /o/ and /ɔ/ as well as the pair of /u/ and /ʊ/ were not distinguishable by means of F2. In Ewen, front [-RTR] vowels were fronter than their [+RTR] counterparts and back [-RTR] vowels were backer than their [+RTR] counterparts, though F2 difference was not always significant. In this respect, the vowel system of Ewen resembles those of African tongue root languages in which [+ATR] vowels are more peripheral than their [-ATR] counterparts in the vowel space.

Note that F3 turned out to be irrelevant to the vowel contrast.

All the means and standard deviations are presented in Table 2.

Before moving on to the other acoustic features, it should be noted that in TB /o/ marginally failed to be distinguished from /u/ in terms of both F1 ( $F = 3.76$ ,  $p = .054$ ) and F2 ( $F = 3.66$ ,  $p = .057$ ). With the fact that /o/ was not

fronter than /ɔ/, this suggests that /o/ is on the verge of merging into /u/ in TB.

| WB   | F1   |       | F2   |       | F3   |       |
|------|------|-------|------|-------|------|-------|
|      | Mean | SD    | Mean | SD    | Mean | SD    |
| i    | 337  | 43.2  | 2000 | 143.3 | 3244 | 409.3 |
| ə    | 381  | 50.6  | 1856 | 194.3 | 3216 | 602.0 |
| a    | 593  | 103.8 | 1393 | 223.2 | 3197 | 570.4 |
| u    | 323  | 53.7  | 1390 | 475.0 | 2890 | 751.9 |
| ɔ    | 376  | 75.5  | 1213 | 276.6 | 2986 | 710.4 |
| o    | 386  | 42.6  | 1475 | 528.1 | 2883 | 743.7 |
| ɔ    | 498  | 77.5  | 1135 | 628.4 | 2747 | 752.0 |
| TB   | F1   |       | F2   |       | F3   |       |
|      | Mean | SD    | Mean | SD    | Mean | SD    |
| i    | 336  | 52.3  | 2199 | 284.7 | 2894 | 328.8 |
| ə    | 407  | 53.4  | 1505 | 270.0 | 2825 | 292.6 |
| a    | 690  | 100.4 | 1392 | 185.2 | 2839 | 284.1 |
| u    | 366  | 71.8  | 977  | 457.8 | 2859 | 372.5 |
| ɔ    | 412  | 73.2  | 906  | 283.0 | 2954 | 300.2 |
| o    | 396  | 91.7  | 995  | 478.4 | 2845 | 348.4 |
| ɔ    | 507  | 58.5  | 1066 | 291.4 | 2866 | 280.8 |
| Ewen | F1   |       | F2   |       | F3   |       |
|      | Mean | SD    | Mean | SD    | Mean | SD    |
| i    | 330  | 57.7  | 2832 | 208.8 | 3398 | 266.9 |
| ɪ    | 407  | 88.9  | 2633 | 262.4 | 3276 | 289.4 |
| ə    | 663  | 91.5  | 2220 | 205.1 | 3158 | 176.7 |
| a    | 914  | 127.0 | 1708 | 248.3 | 3454 | 322.0 |
| u    | 396  | 58.7  | 999  | 250.3 | 3063 | 181.5 |
| ɔ    | 471  | 89.9  | 1013 | 180.9 | 3304 | 309.3 |
| o    | 567  | 75.4  | 1131 | 166.3 | 3149 | 229.4 |
| ɔ    | 684  | 106.1 | 1214 | 177.6 | 3375 | 267.5 |

**Table 2.** The means and standard deviations of F1, F2, and F3 in the three languages

### 4.2 The normalized A1-A2

The result of normalized A1-A2 was not consistent across all the three languages (Figure 6). In WB, on the one hand, it proved to be a reliable acoustic cue for the tongue root contrast. In each harmonic pair, the value was significantly higher in a [+RTR] vowel than in its [-RTR] counterpart, which was consistent with the results of African languages (Fulop et al. 1998, Guion et al. 2004). On the other hand, however, the same consistency was not found in the other two languages. As Figure 6 shows, only one pair in TB and two pairs in Ewen were successfully distinguished. These results, by and large, coincide with those of African languages where the significance of difference varied depending on vowel pairs and speakers. The results of Ewen are also somewhat different from those in Aralova & Grawunder (2011) who found consistent higher values in “advanced” vowels than in “retracted” vowels in Ewen using simple A1-A2 without normalization. Another thing to be noted is that the overall values were much higher than those in African languages.<sup>12)</sup>

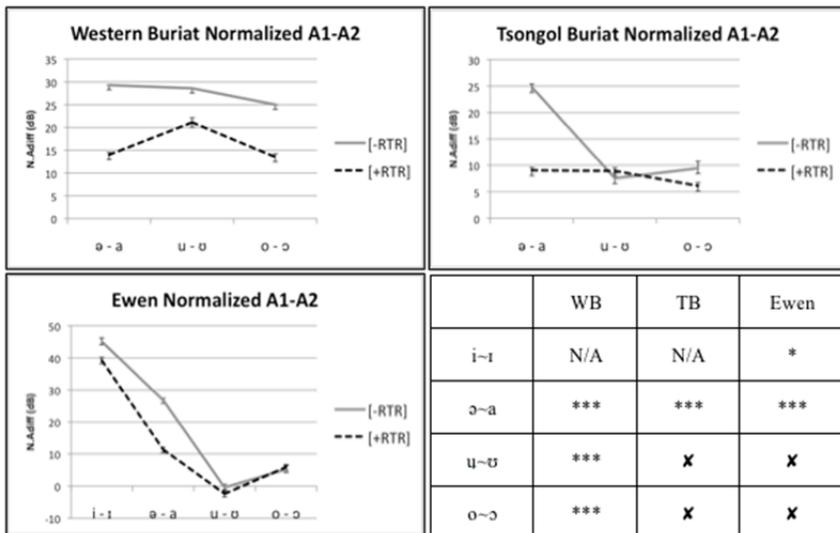


Figure 6. The normalized A1-A2 (\*\*\*: p < 0.001, \*\*: p < .01, \*: p < .05, ✗: p > .05 )

12) In African languages, the range is roughly between -10 and 10.

4.3 B1

The bandwidth of F1 (B1) was more reliable than the normalized A1-A2 (Figure 7). It distinguished a [-RTR] vowel from its [+RTR] counterpart in all harmonic pairs except for the high back vowel pair (/u/ and /ʊ/) in TB and the mid back vowel pair (/o/ and /ɔ/) in Ewen. In general, as shown in the charts below, [-RTR] vowels had lower B1 values than [+RTR] vowels. This is consistent with the results of African languages (refer back to §2.2.3).

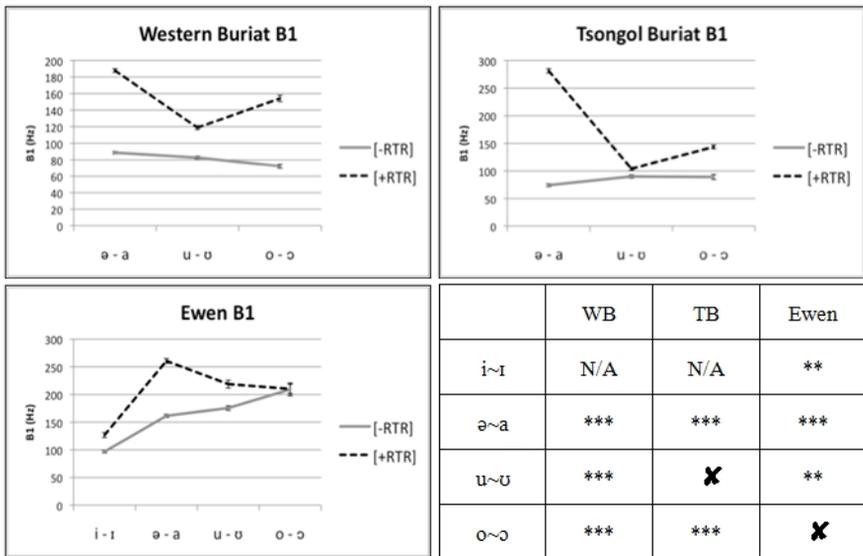


Figure 7. The bandwidth of F1 (\*\*\*: p<001, \*\*: p<.01, \*:p<.05, ✗: p>.05 )

4.4 The center of gravity

The center of gravity was found to be another acoustic cue for the tongue root contrast in the three languages (Figure 8). The only exception was the front high vowel pair in Ewen, which did not show any significant difference. As was the case with F1 and B1, center of gravity was higher in [+RTR] vowels than in [-RTR] vowels.

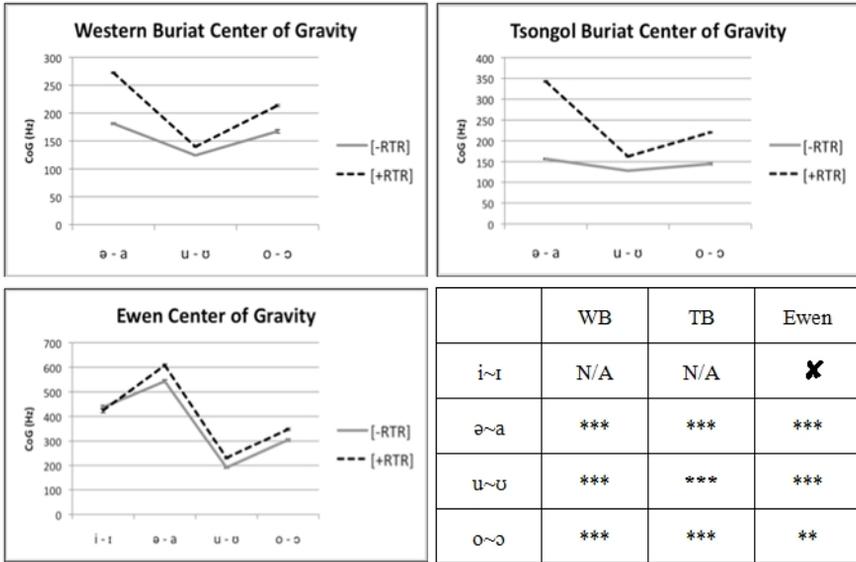


Figure 8. The center of gravity (\*\*\*:  $p < 0.001$ , \*\*:  $p < .01$ , \*:  $p < .05$ , ✗:  $p > .05$  )

### 4.5 Summary

F1 was the only acoustic cue to distinguish [-RTR] vowels from their [+RTR] counterparts in every pair and in every language. B1 and the center of gravity also distinguished one set of vowels from the other, to slightly varying degrees as well as with a few exceptions. The normalized A1-A2 was the least successful out of the four acoustic features tested in this paper in distinguishing the two series of vowels. Overall, the correlations between these measures and the [RTR] feature were consistent with those found in West African languages.

## 5. Discussion

### 5.1 Vowel systems in WB, TB, and Ewen

The acoustic data we have shown so far, especially the average F1 and F2 frequency data shown in Figure 5, indicate that the vowel systems in WB, TB, and Ewen are all best characterized as a tongue root system. By contrast, a

palatal or a height harmony analysis would not be able to explain the acoustic distribution of vowels.

The Buriat vowel system has been traditionally assumed to be based on a palatal contrast (Poppe 1960 among others). However, note in the WB data in Figure 5 that, despite the relative F2 differences between the harmonic counterparts, all rounded vowels surface as back vowels. Similar results are reported in Bayarmend's (2006) acoustic study of Khori Buriat (Eastern Buriat, EB) which has 6 short vowel phonemes /i, ə, e, u, ʊ, ɔ/ (no /o/) in the inventory. Although Bayarmend does not attempt a phonological feature analysis, the fact that all the three rounded vowels /u, ʊ, ɔ/ are acoustically realized as back vowels indicates that the EB vowel system is also based on a tongue root contrast rather than a palatal contrast. These results are largely consistent with Buraev's (1959) X-ray studies as well: Now that we have convincing acoustic data from both WB and EB, Buraev's X-ray images can be reinterpreted to show the tongue root-based nature of the Buriat vowel system.

The TB data also reveal that its vowel system is based on a tongue root contrast. All vowels other than /i/ (including /ə/) can be viewed as back vowels. This characteristic is basically very similar to that of EB explored by Bayarmend (2006). Note that, compared to the WB /ə/ which is realized as [e], the TB /ə/ is realized as "high-mid, central, unrounded vowel" as in EB (Bayarmend 2006: 107). In addition, there is another contrast between WB vs. TB, the retention of /o/ in WB vs. the merger /o/ > /u/ in TB and EB.<sup>13)</sup> This change of TB into a 6-vowel system may tell us that TB has been strongly influenced by EB, although TB is originally a Mongolian dialect and thus expected to have a 7-vowel system like other Mongolian dialects.

The Ewen data are particularly significant since there has been no acoustic study of Ewen vowels until recently (cf. Aralova & Grawunder 2011). In the meantime, Novikova's (1960) X-ray tracings have served as a basis of the "RTR" hypothesis (Kim 1989, Li 1996) of the entire Tungusic languages or the "pharyngeal" hypothesis (Ladefoged & Maddieson 1996) of the languages

---

13) Note the TB merger seems to be ongoing, with the two merging vowels not completely overlapping yet.

with the Ewen type vowel contrast such as some Caucasian languages. In our new acoustic data, the overall distribution of Ewen vowels in Figure 5 confirms that the Ewen vowel system is based on a tongue root contrast. There is no observed lowering of F3 in [+RTR] vowels. This implies that the vowel contrast in Ewen may not be of the same type as the contrast between plain vs. pharyngealized vowels in the Caucasian languages.

## 5.2 Acoustic correlates of tongue root contrast

After all, the results show that the acoustic realization of the tongue root contrast in Altaic languages is not different from that of African languages. In both groups of languages, [+ATR]/[-RTR] vowels have lower F1, B1, and center of gravity and higher normalized A1-A2 than their [-ATR]/[+RTR] counterparts. However, as Starwalt (2008: 85, 100) points out, B1 and center of gravity are seriously affected by F1. That is to say, the differences in B1 and center of gravity may be attributed to that in F1. For this reason, it is premature to conclude that all these acoustic features are the characteristics of the tongue root contrast. Another thing Starwalt (2008: 416) points out is that the normalized A1-A2 was not as robust as other acoustic features such as B1. We reached the same conclusion. With the discrepancy of the absolute values between Altaic and African languages, this might cast a doubt on the appropriateness of the normalized A1-A2 as the acoustic correlate of the tongue root contrast.

In this study, we found no acoustic grounds for distinguishing [ATR] and [RTR] as two distinct features. However, needless to say, we need more acoustic data and analyses to confirm one way or the other. In addition, articulatory studies based on MRI (Tiede 1996), laryngoscopy (Edmondson and Esling 2006), ultrasound imaging (Hudu 2010) may reveal the physiological causes for the acoustic characteristics. We hope that we will find finer data and methods in the future and obtain fruitful results on the articulatory and acoustic nature of the tongue root contrast in Altaic languages.

## References

- Aralova, Natalia & Sven Grawunder. 2011. Does Even have ATR? Handout presented at the CUNY conference on the phonology of endangered languages, CUNY Graduate Center. <http://www.cunyphonologyforum.net/endan.php>. (March 15, 2012)
- Ard, Josh. 1981. A sketch of vowel harmony in the Tungus languages. In Bernard Comrie (ed.), *Studies in the languages of the USSR*, 23–43. Edmonton-Carbondale, Alberta: Linguistic Research.
- Bayarmend, Borjigin. 2006. Acoustic analysis of the Buriat short vowels. *Altai hakpo* 16. 105–116. The Altaic Society of Korea.
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10). 341–345. Blackwell.
- Bosson, James E. 1962. *Buriat reader* (Uralic and Altaic series 8). Bloomington, IN: Indiana University Publications.
- Buraev, Ignatij Dmitrievich. 1959. *Zvukovoj sostav burjatskogo jazyka* [*The sound structure of Buriat*]. Ulan-Ude: Burjatskoe knizhnoe izdatel'stvo.
- Casali, Roderic F. 2008. ATR harmony in African languages. *Language and linguistics compass* 2(3). 496–549. Blackwell. (doi:10.1111/j.1749-818X.2008.00064.x)
- Catford, J. C. 1994. Vowel systems of Caucasian languages. In Howard Isaac Aronson (ed.), *Non-Slavic languages of the USSR: Papers from the fourth conference*, 44–60. Columbus, OH: Slavica Publishers.
- Cenggeltei. 1959. Monggol kelen-ü Bagarin ayalgun-u abiy-a-yin jüi ba üges-ün jüi [Phonology and morphology of the Baarin dialect of Mongolian]. *Öbör Monggol-un Yeke Surgaguli-yin erdem sinjilegen-ü sedgüil: Gün uqagan neyigem sinjilekü uqagan*. 1–96. Inner Mongolia University.
- Clements, G. N. & Annie Rialland. 2008. Africa as a phonological area. In Bernd Heine & Derek Nurse (ed.), *A linguistic geography of Africa*, 36–85. New York: Cambridge University Press.
- Edmondson, Jerold A. & John H. Esling. 2006. The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology* 23(2): 157–191. Cambridge, UK: Cambridge University Press. (doi:10.1017/S095267570600087X)
- Fischer-Jørgensen, Eli. 1967. Phonetic analysis of breathy (murmured) vowels in Gujarati. *Indian Linguistics* 28. 71–139. Linguistic Society of India.
- Fulop, Sean A., E. Kari & Peter Ladefoged. 1998. An acoustic study of the tongue root contrast in Degema vowels. *Phonetica* 55(1-2). 80–98. International Society of Phonetic Sciences. (doi:10.1159/000028425)

- Garellek, Marc & Patricia Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41(2). 185–205. Cambridge, UK: Cambridge University Press. (doi:10.1017/S0025100311000193).
- Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of phonetics* 29(4). 383–406. Elsevier. (doi:10.1006/jpho.2001.0147)
- Guion, Susan G., Mark W. Post & Doris L. Payne. 2004. Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of phonetics* 32(4). 517–542. Elsevier. (doi:10.1016/j.wocn.2004.04.002)
- Halle, Morris & K. N. Stevens. 1969. *On the feature “advanced tongue root.”* (MIT Research Laboratory of Electronics Quarterly Progress Report). Cambridge, MA: MIT.
- Hertz, S. R. 1991. Streams, phones, and transitions: Toward a phonological and phonetic model of rormant timing. *Journal of phonetics* 19. 91–109. Elsevier.
- Hess, S. 1992. Assimilatory effects in a vowel harmony system: An acoustic analysis of advanced tongue root in Akan. *Journal of phonetics* 20. 475–492. Elsevier.
- Hudu, Fusheini Angulu. 2010. *Dagbani tongue-root harmony: A formal account with ultrasound investigation.* Vancouver, Canada: University of British Columbia Ph.D. dissertation. <https://circle.ubc.ca/handle/2429/27777>. (March 1, 2012)
- Hudu, Fusheini Angulu, Amanda Miller & Douglas Pulleyblank. 2009. Ultrasound imaging and theories of tongue root phenomena in African languages. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds.), *Proceedings of conference on language documentation and linguistic theory 2*, 153–163. London: SOAS. <http://www.hrlep.org/publications/ldlt2/papers/ldlt2proceedings.html>. (March 1, 2012)
- Huffman, Marie K. 1987. Measures of phonation type in Hmong. *The journal of the Acoustical Society of America* 81(2). 495–504. American Institute of Physics for the Acoustical Society of America.
- Jacobson, Leon Carl. 1980. Voice-quality harmony in Western Nilotic languages. In Robert Michael Vago (ed.), *Issues in vowel harmony: Proceedings of the CUNY linguistics conference on vowel harmony, 14th May 1977*, 183–200. Amsterdam: John Benjamins Publishing Company.
- Johnson, Keith. 1997. *Acoustic and auditory phonetics.* 1st edn. Cambridge, MA: Blackwell.
- Kaun, Abigail Rhoades. 1995. *The typology of rounding harmony: An optimality theoretic approach.* Los Angeles, CA: University of California Ph.D. dissertation.
- Kim, Juwon. 1989. *Mancwu thwungkwusuceyey moumcohwa yenkwu [A study on vowel*

- harmony in the Manchu-tungus languages*]. Seoul, Korea: Seoul National University Ph.D. dissertation.
- Kim, Juwon. 1999. Selkunchwuchwukkwa selkuncencin [RTR and ATR]. In *Eneuy yeksa* [History of languages: in honor of Professor Baeg-in Seong], 311–341. Seoul, Korea: Taehaksa.
- Kim, Juwon. 2011. *A grammar of Ewen* (Altaic Languages Series 6). Seoul, Korea: Seoul National University Press.
- Kim, Juwon, Jae-Il Kwon, Dongho Ko, Yoon-Shin Kim & Soonhwan Jeon. 2008. *Salacye kanun Althaienelul chacase* [Documentation of endangered Altaic languages]. 2nd edn. Phacwu, Korea: Taehaksa.
- Ladefoged, Peter. 1964. *A phonetic study of West African languages: An auditory-instrumental survey* (West African Language Monographs 1). Cambridge, UK: Cambridge University Press.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford, UK: Blackwell.
- Li, Bing. 1996. *Tungusic vowel harmony: Description and analysis* (HIL Dissertations 18). Dordrecht: Holland Institute of Generative Linguistics.
- Lindau, Mona. 1974. The feature advanced tongue root. In Erhard Friedrich Karl Voeltz (ed.), *Third annual conference on African linguistics, 7-8 April 1972*, (African Series 7), 127-133. Bloomington: Indiana University.
- Lindau, Mona. 1979. The feature expanded. *Journal of phonetics* 7. 163–176. Elsevier.
- Novikova, Klavdija A. 1960. *Ocherki dialektov evenskogo jazyka: Ol'skij govor* 1. Leningrad: Nauka.
- Poppe, Nicholas N. 1960. *Buriat grammar*. Bloomington, IN: Indiana University.
- Przedzdiecki, Marek Andrzej. 2005. *Vowel harmony and coarticulation in three dialects of Yoruba: Phonetics determining phonology*. Ithaca, NY: Cornell University Ph.D. dissertation.
- Robbek, V. A. & M. E. Robbek. 2005. *Evensko-russkij slovar'* [Ewen-Russian Dictionary]. Novosibirsk: Nauka.
- Skribnik, Elena K. 2003. Buryat. In Juha Janhunen (ed.), *The Mongolic Languages* (Routledge Language Family Series 5), 102–128. London: Routledge.
- Starwalt, Coleen Grace Anderson. 2008. *The acoustic correlates of ATR harmony in seven- and nine-vowel African languages: A phonetic inquiry into phonological structure*. Arlington, TX: University of Texas Ph.D. dissertation.
- Steriade, Donca. 1995. Underspecification and markedness. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 114–174. Oxford: Blackwell.
- Stewart, John M. 1967. Tongue root position in Akan vowel harmony. *Phonetica* 16: 185–204. International Society of Phonetic Sciences. (doi:10.1159/000258568)

- Svantesson, Jan-Olof. 1985. Vowel harmony shift in Mongolian. *Lingua* 67(4) (December). 283–327. Elsevier. (doi:10.1016/0024-3841(85)90002-6)
- Svantesson, Jan-Olof, Anna Tsendina, Anastasia Mukhanova Karlsson & Vivan Franzén. 2005. *The phonology of Mongolian*. Oxford, UK: Oxford University Press.
- Tiede, Mark K. 1996. An MRI-based study of pharyngeal volume contrasts in Akan and English. *Journal of Phonetics* 24. 399–421. Elsevier.
- Zhang, Xi. 1996. *Vowel systems of the Manchu-Tungus languages of China*. Toronto, ON: University of Toronto Ph.D. dissertation.

Hijo KANG

Department of Linguistics, Room S201 Social & Behavioral Sciences Building  
Stony Brook University  
Stony Brook, NY 11794-4376, USA  
<hikang@ic.sunysb.edu >

Seongyeon Ko

Department of Linguistics, 203 Morrill Hall, 159 Central Avenue  
Cornell University  
Ithaca, NY 14853-4701, USA  
<sk484@cornell.edu >

Received 30 March 2012;  
accepted 31 May 2012.