

Vowels of Beryozovka Ewen: An acoustic phonetic study*

KANG Hijo (Chosun University)

YUN Jiwon (Stony Brook University)

Ko Seongyeon (CUNY Queens College)

KANG, Hijo, Jiwon YUN & Seongyeon Ko. 2017. Vowels of Beryozovka Ewen: An acoustic phonetic study. *Altai Hakpo* 27. 1-23. The Altaic Society of Korea.

This study acoustically analyzes 13,540 vowel tokens of Beryozovka Ewen with the aid of automated post-transcriptional processing technique. The focus of the analysis is on the acoustic correlates of [RTR], which is the harmonic feature of the language. In addition to the first three formants, acoustic values representing spectral tilt such as H1 – H2, H1 – A2, and B1 are measured as potential acoustic cues of [RTR]. The results show that F1, F3, and B1 are the most reliable cues of the feature and that H1 – H2 and H1 – A2 are nearly reliable. These acoustic cues are also shown to interact with length and position. In general, the acoustic distance between [–RTR] and [+RTR] vowels are farther in long and word-initial vowels than in short and non-initial vowels, respectively. We claim that [RTR] is more appropriate for the harmonic feature of Ewen than [ATR], and that the greater perceptibility of word-initial vowels is understood as a means to facilitate the lexical access in a language with vowel harmony.

Keywords: Beryozovka Ewen, vowel, [RTR], formant, spectral tilt, position

* We are grateful to Professor Juwon Kim for giving us permission to use the Beryozovka Ewen materials. We also thank the reviewers for their comments on the earlier draft. This work was supported by Laboratory Program for Korean Studies through the Ministry of Education of Republic of Korea and Korean Studies Promotion Service of the Academy of Korean Studies (AKS-2016-LAB-2250004).

1. Introduction

In this paper, we investigate the acoustic properties of vowels in Beryozovka Ewen (an Eastern Ewen dialect, Northern Tungusic), with emphasis on the acoustic correlates of the tongue root contrast based on the feature [Retracted Tongue Root (hereafter RTR)].

We first present the statistical analysis of the frequencies of the first three formants (F1, F2, F3) and show how these values do (or do not) distinguish the two groups of vowels. The analysis suggests that F1 and F3 are the two most reliable acoustic correlates of tongue root contrast in Beryozovka Ewen. Second, we also show that some measurements of spectral tilt (such as H1 – H2, H1 – A2, and B1) may also serve as acoustic cues for the differentiation of non-RTR and RTR vowels, although there are some exceptions. We also show how the length and the position of vowels affect these values. For example, we find that high front vowels (/i/ vs. /ɪ/) and low vowels (/ə/ vs. /a/) are better distinguished in initial positions due to greater differences. We also observe some significant—but inconsistent—effect of position and length on spectral tilt.

This paper is structured as follows. Section 2 provides background information on Beryozovka Ewen and a brief literature review on previous phonetic studies. Section 3 describes our language materials and research methods. Section 4 presents the results of various acoustic measurements. Then, section 5 provides a discussion on our findings and concludes the paper.

2. Background

2.1 Beryozovka Ewen

The Ewen language (ISO 639-3 code: eve) is a severely endangered Northern Tungusic variety spoken in the Russian Far East by 5,660 native speakers out of 21,800 ethnic Ewens (Simons & Fennig 2017). There are at least two dialect groups: The Eastern group includes the Kamchatka dialect, the Chukotka dialect, the Okhotsk Sea shore dialect, and the Srednekolymskiy dialect of the Sakha (Yakutia) Republic, while the Western group includes all other dialects of Sakha Republic (except for the Srednekolymskiy dialect).

Beryozovka Ewen, as a sub-dialect of the Srednekolymskiy dialect, belongs to the Eastern group (Kim 2011:3). Due to lack of thorough descriptions, it is difficult to tell the phonetic and phonological differences among the dialects and sub-dialects of Ewen.

The consonant and vowel phonemes of Beryozovka Ewen are presented below:

	Bilabial	Labio-dental	Alveolar	Palatal	Velar	Pharyngeal
Plosive	p b		t d		k g	
Affricate				tʃ ɕ		
Fricative			s			(h)
Nasal	m		n	ɲ	ŋ	
Trill			r			
Lateral			l			
Approximant		w		j		

Table 1. Consonant phonemes in Beryozovka Ewen (Kim 2011:23)

	Front		Central		Back	
High	i ɪ	i: ɪ:			u ʊ	u: ʊ:
Mid			ə	ə:	o ɔ	o: ɔ:
Low			a	a:		

Table 2. Vowel phonemes of Beryozovka Ewen (Kim 2011:11–22)

In the rest of the paper, we will categorize the vowels into “high front” (/i, ɪ/), “high back” (/u, ʊ/), “mid back” (/o, ɔ/), and “low” (/ə, a/) vowels.¹⁾

Like most of the other Tungusic languages, Ewen exhibits vowel harmony operating on the feature [RTR]:

1) The *phonetically* mid vowels /ə, o, ɔ/ should be viewed as *phonologically* low vowels from a contrastivist’s viewpoint (Ko 2012; Ko 2013). However, we treat only the mid central vowel /ə/ as a low vowel to form a harmonic pair with another low vowel /a/.

(1) Harmonic sets in Beryozovka Ewen (Kim 2011:39–41)

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

(2) Vowel harmony in Beryozovka Ewen (Kang & Ko 2012:190, data drawn from Kim 2011)

[–RTR] vowels		[+RTR] vowels	
hor-li	‘go-IMP’	hilkat-lɪ	‘rinse-IMP’
tugəŋi-du	‘winter-DAT’	ʝogani-du	‘summer-DAT’
toŋər-duk	‘lake-ABL’	bazar-dok	‘market-ABL’
hupkučək-lə	‘school-LOC’	dəlbəŋi-la	‘night-LOC’

As illustrated in the above examples, all vowels in a stem/word must bear the same value for the feature [RTR].

2.2 Previous phonetic studies of Ewen

Early phonetic studies of the vowels of Ewen include Novikova (1960 for Ola dialect) and Lebedev (1978 for Okhotsk dialect). In particular, Novikova’s X-ray images provide articulatory evidence that the “pharyngealized vowels” of Ola Ewen (which correspond to the Set B vowels (1b) of Beryozovka Ewen) are articulated with narrower pharyngeal passage and a raised larynx. There has been a general consensus in the phonology literature that Novikova’s “pharyngealized vowels” are indeed [+RTR] vowels, i.e., that the vowel harmony in Ewen is based on the tongue root contrast (Ard 1981; Kim 1989; Kaun 1995; Li 1996; Kim 2011; Ko 2012; Ko, Joseph & Whitman 2014).

In recent years, several acoustic studies have been conducted for Ewen vowels. Building upon the findings in the acoustic phonetic studies conducted on African tongue root harmony languages (Hess 1992; Fulop, Kari & Ladefoged 1998; Guion, Post & Payne 2004; Przedziecki 2005; Starwalt 2008), Aralova, Grawunder & Winter (2011 Bystraia and Sebian dialects) and Kang & Ko (2012 an Eastern dialect) attempt to find the acoustic correlates of the tongue root contrast in Ewen.²⁾ Both studies confirm that all non-retracted vowels have lower F1 (the frequency of the first formant) values than their retracted counterparts with the tendency that

2) See Kang & Ko (2012) for an overview of the instrumental studies of tongue root harmony in both African and Altaic languages.

the difference is bigger for low vowels. Aralova, Grawunder & Winter (2011) further claim that F2 (the frequency of the second formant) is higher for non-retracted vowels, but Kang & Ko (2012) report no such effect in F2 values. In Aralova, Grawunder & Winter (2011), $A1 - A2$ (the amplitude of the first formant *minus* the amplitude of the second formant) values—measured to see the difference in spectral slope—are found to be consistently smaller for retracted vowels. In contrast, Kang & Ko (2012) observe that the *normalized* $A1 - A2$ values (Fulop, Kari & Ladefoged 1998) do not show a consistent difference between the two vowels in all harmonic pairs. Based on Catford's (1994) observation on Caucasian languages, Ladefoged & Maddieson (1996:306–310) suggest that F3 (the frequency of the third formant) could be markedly lower for “pharyngealized” vowels (i.e., [+RTR] vowels) in languages like Ewen. Aralova, Grawunder & Winter (2011) report this effect in one of the two Ewen dialects they investigated (Sebian dialect), but not in the other (Bystraia dialect). Kang & Ko (2012) find no consistent effect of F3 differences between the two groups of vowels in the Ewen variety they investigate. Kang & Ko (2012) further examine B1 (the bandwidth of the first formant) (Hess 1992) and the center of gravity (Starwalt 2008) to notice that B1 and the center of gravity are wider and higher in [+RTR] vowels, respectively, in general.

3. Language Materials and Methods

The language materials used in this study were collected by a team of fieldwork researchers of the Altaic Society of Korea—under the project title “Researches on Endangered Altaic Languages” (Principal Investigator: Professor Juwon Kim at Seoul National University)—in 2007 in Yakutsk, Sakha Republic, Russia. The language consultant, Ms. Mariya Ivanovna Buldukina, was a reindeer herder from Beryozovka, Srednekolymskiy District, and used Ewen regularly in her daily life. For the audio recordings, Ms. Buldukina repeated twice the Ewen words, phrases, or sentences corresponding to the Russian expressions in the questionnaire. See Kim (2011:5–10) for more details of the fieldwork, the language consultant, and the collected materials.

From the audio files (11 hours 32 minutes) recorded to collect Ewen vocabulary items, we extracted 2,186 words (or phrases in some cases) into separate sound files and then ran Prosodylab-aligner (Gorman et al. 2011)

vowel	short			long			sum
	initial	medial	final	initial	medial	final	
i	328	644	364	116	10	12	1,474
ɪ	246	454	352	34	10	0	1,096
u	352	474	124	52	10	0	1,012
ʊ	348	330	148	68	6	14	914
o	230	62	0	158	30	22	502
ɔ	268	38	12	148	28	0	494
ə	694	1,316	1,574	280	148	48	4,060
a	698	1,268	1,450	346	164	62	3,988
sum	3,164	4,586	4,024	1,202	406	158	13,540

Table 3. Tokens of short and long monophthongs of Ewen measured in this study

to automatically segment and label them using the phonemic transcriptions in Kim (2011).³⁾ Table 3 shows how many tokens of each vowel quality (short and long) in three different positions were prepared for acoustic measurements through the above-mentioned process. Vowels in monosyllabic words were counted as an initial vowel. Note that there was no token of /ɪ:/, /u:/, /o:/, /ɔ:/ in final position.

4. Results

To find the acoustic effect of [RTR], position, and length, the annotated audio data (totally 13,540 vowels included in 2,186 sound files) went through acoustic measurements,⁴⁾ which were all automatically carried out by means of Praat scripts (Boersma & Weenink 2014).

Independent factors for statistical analyses such as vowel, position, length, and [RTR] were coded within the result file, which was subject to ANOVA

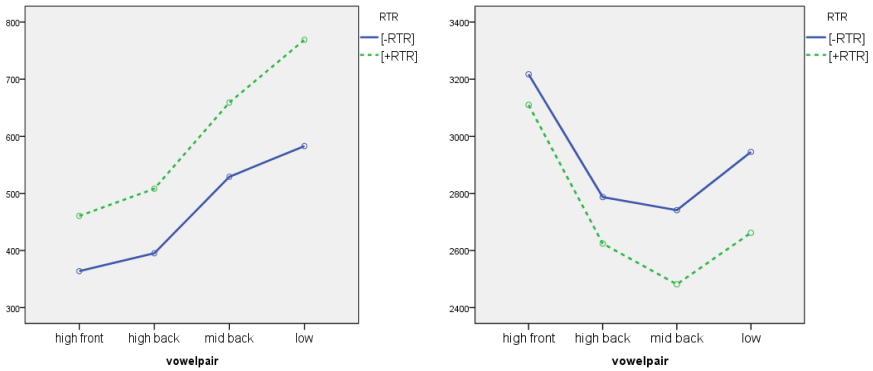
3) See Yun, Kang & Ko (2016) for the details on this semi-automated post-transcriptional processing technique applied to the analysis of Najkxin Nanai (Tungusic).

4) Note that Aralova, Grawunder & Winter (2011) and Kang & Ko (2012) measure in total 3,336 and 899 tokens of vowels, respectively, with the latter investigating only short vowels in initial syllables.

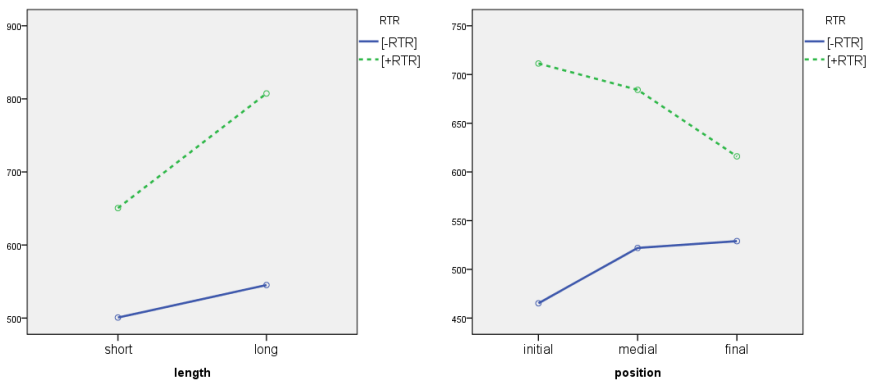
(analysis of variance, SPSS 23). This section presents the results of statistics, focusing on the effect of [RTR], position, and length. In each sub-section, we first examine the effect of [RTR] and then move on to the effects of position and length.

4.1 Formant structures and quadrilaterals of vowels

F1 turns out to be one of the most reliable cues for the [RTR] feature in that all the [+RTR] vowels have significantly higher F1 values than their [-RTR] counterparts ($F(7,13532) = 1052, p < .001$). This is consistent with the results of previous studies on other tongue root harmony languages in Africa and



Figures 1 (LEFT) and 2 (RIGHT). F1 (LEFT) and F3 (RIGHT) values of non-RTR and RTR vowels in all four harmonic pairs.



Figures 3 (LEFT) & 4 (RIGHT). F1 values of non-RTR and RTR vowels with the effect of vowel length (LEFT) and positions (RIGHT).

Northeast Asia (Hess 1992; Fulop, Kari & Ladefoged 1998; Guion, Post & Payne 2004; Aralova, Grawunder & Winter 2011; Kang & Ko 2012; Yun, Kang & Ko 2016). In addition, F3 systematically distinguishes [+RTR] vowels from [-RTR] vowels in every pair ($F(7,13532) = 600.3, p < .001$). The two acoustic cues show that low vowels (/ə/ vs. /a/) have bigger differences than high vowels (/i/ vs. /ɪ/ & /u/ vs. /ʊ/) as shown in Figures 1 and 2 above. However, the [RTR] feature is not realized as F2 in back vowels.

The main effects of length and position are also confirmed in two-way ANOVA's with another main effect of [RTR], as in Figures 3 and 4 above.

The interactions of length \times [RTR] and position \times [RTR] are significant for all the formant values. The differences between [-RTR] and [+RTR] vowels are bigger in long vowels than in short vowels ($F_1: F(3,13536) = 110.7, p < .001$; $F_2: F(3,13536) = 19.7, p < .001$; $F_3: F(3,13536) = 47.6, p < .001$) and in initial positions than in non-initial positions ($F_1: F(5,13534) = 153.7, p < .001$; $F_2: F(5,13534) = 5.2, p < .01$; $F_3: F(5,13534) = 90.3, p < .001$). This means that long and initial vowels are more distinguishable than short and non-initial vowels, respectively.

In order to tease apart the effect of position from that of length,⁵⁾ the

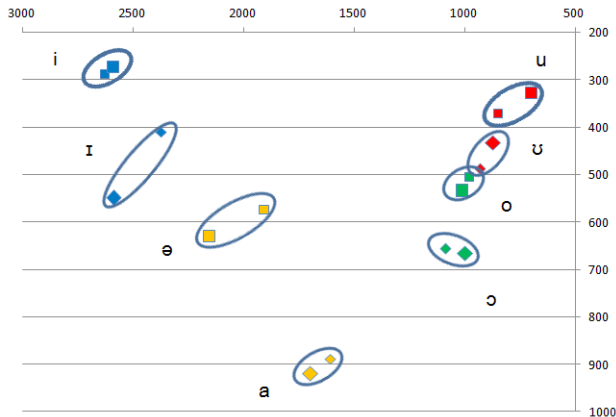


Figure 5. Vowel quadrilateral (F1 & F2 in Hz on x- and y-axis, respectively) for long vs. short vowels in monosyllabic words. (■: [-RTR] & ◆: [+RTR], Bigger icons stand for long vowels.)

5) For example, vowels in monosyllabic words are in initial and final positions at the same time, which is problematic for the analysis of the “position effect.”

data were divided into four groups, depending on the length of words: monosyllabic, disyllabic, trisyllabic, and quadrisyllabic. Here we show the F1-F2 charts of the first three groups, one by one.

Figure 5 shows the length effect in monosyllabic words: long vowels in monosyllabic words are more peripheral than their short counterparts—with the exception of /i/ and /o/, which show a reverse peripherality.⁶⁾

The position of vowels has a similar but greater effect. Basically, vowels in initial positions take larger vowel space than those in other positions. Figure 6 shows that, in disyllabic words, initial vowels are more peripheral than final vowels.

Figure 7 shows that, in trisyllabic words, the size of vowel space decreases as the vowels go towards the final positions of words. These all suggest that the distinctiveness of vowels is not consistently preserved throughout all the positions in words.

So far we have seen that [RTR], length, and position affect the formant values of vowels generally. Now let us take a look at each pair to see whether they are effective in all the non-RTR and RTR vowel pairs.

In line with the overall results presented above, F1 and F3 effectively distinguish [–RTR] vowels from their [+RTR] counterparts in every pair.

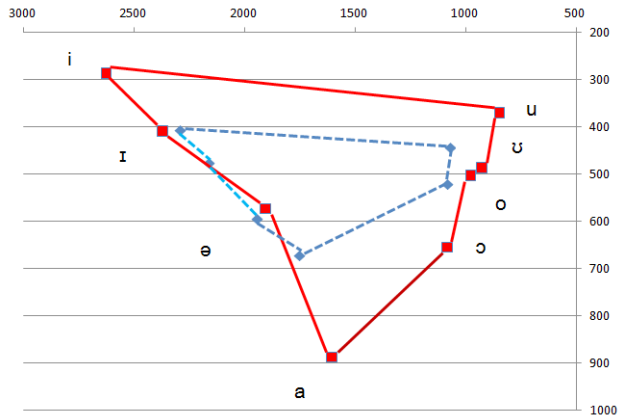


Figure 6. Vowel quadrilateral (F1 & F2 in Hz on y- and x-axis, respectively) for initial (■) vs. final (◆) vowels in disyllabic words. Mid back vowels, /o/ and /ɔ/, do not appear in the final positions of disyllabic words.

6) However, the short and long vowel pairs of /i/ and /o/ do not show meaningful differences.

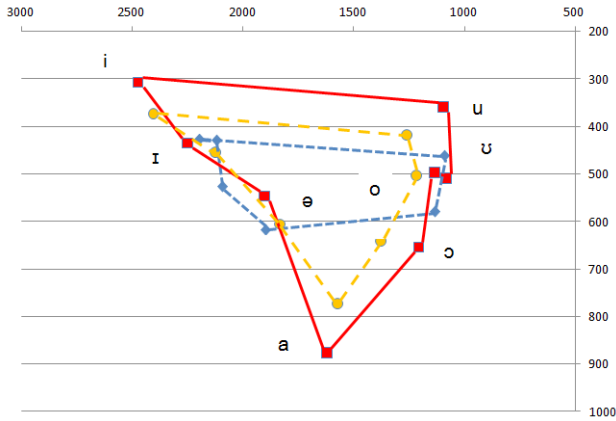


Figure 7. Vowel quadrilateral (F1 & F2 in Hz on y- and x-axis, respectively) for initial (■) vs. medial (●) vs. final (◆) vowels in trisyllabic words. Mid back vowels, /o/ and /ɔ/, do not appear in finals, and /ɔ/ does not appear in medials.

		/i/	<	/ɪ/	/u/	<	/ʊ/	/o/	<	659	583	<	767
F1	M	365	<	461	395	<	510	530	<	659	583	<	767
	SD	136	***	172	111	***	137	143	***	157	185	***	216
F2	M	2350	>	2163	1107	>	1106	1122	<	1134	1936	>	1703
	SD	552	***	473	410		333	319	**	294	389	***	347
F3	M	3213	>	3108	2786	>	2623	2742	>	2479	2945	>	2806
	SD	339	***	293	219	***	262	210	***	295	380	***	438

Table 4. Formant values of 4 non-RTR and RTR vowel pairs and the results of ANOVA (M: Mean, SD: Standard Deviation, ***: $p < .001$, **: $p < .01$, *: $p < .05$)

On the contrary, F2 does not show a significant difference in the high back vowel pair and the direction is the opposite in the mid back vowel pair.

Lastly, we present the effect of position and length by the results of interaction (RTR × position and RTR × length) in Table 5 below. A significant interaction here means that differences between [−RTR] and [+RTR] vowels are affected by position or length. For example, the difference may be bigger in a specific position, i.e., in the initial position.

Though the effects of position and length were confirmed in the whole data, it is not the case that every vowel is influenced. By and large, the

acoustic feature	vowel pair	position		length	
		Interaction	bigger difference	Interaction	bigger difference
F1	/i/ vs. /ɪ/	***	initial	***	long
	/u/ vs. /ʊ/		initial		long
	/o/ vs. /ɔ/		initial		short
	/ə/ vs. /a/	***	initial	***	long
F2	/i/ vs. /ɪ/		initial		short
	/u/ vs. /ʊ/		initial	*	long
	/o/ vs. /ɔ/	***	final	*	long
	/ə/ vs. /a/	***	initial	***	long
F3	/i/ vs. /ɪ/	*	initial		short
	/u/ vs. /ʊ/		final	*	short
	/o/ vs. /ɔ/	***	initial		long
	/ə/ vs. /a/	***	initial	***	long

Table 5. Interactions of [RTR] × position and [RTR] × length (***: $p < .001$, **: $p < .01$, *: $p < .05$)

acoustic distance between [−RTR] and [+RTR] vowels is greater in initial positions or when the vowels are long, in line with Figures 5 to 7. Interestingly, F1 is affected by the two factors only in the high front and low pairs. In other words, the F1 values of round vowels are not sensitive to position nor to length. When it comes to F2 and F3, low and mid vowels seem to vary depending on position and length, but high vowels do not.

To sum up, F1 and F3 values of Beryozovka Ewen vowels are differentiated by [RTR]. [−RTR] vowels have lower F1 and higher F3 values than their [+RTR] counterparts in all four pairs. F1 and F3 values are also affected by the position and length of vowels in general. When it comes to position, high front vowels (/i/ vs. /ɪ/) and low vowels (/ə/ vs. /a/) are more distinguishable in initial positions due to greater differences. Vowel length exerts its influence only on the formant structures of low vowels.

4.2 Spectral tilt

Since Halle & Stevens (1969) suggested that the feature [Advanced Tongue Root (ATR)] is related to phonation, spectral tilt (or slope) has been investigated as a possible acoustic correlate of tongue root contrast (Hess 1992; Fulop, Kari & Ladefoged 1998; Guion, Post & Payne 2004; Przewdziecki

		/i/	/ɪ/	/u/	/ʊ/	/o/	/ɔ/	/ə/	/a/
H1 – H2	M	9.92	> 6.92 ***	4.39	< 5.83 ***	3.81	< 7.22 ***	6.32	< 9.01 ***
	SD	8.81	8.52	5.86	6.65	5.46	6.53	6.62	6.47
H1 – A1	M	9.2	> 8.41	10.55	> 8.11 ***	6.54	> 5.03	6.39	< 7.07 **
	SD	11.93	10.82	12.47	8.75	8.4	6.19	8.14	7.79
H1 – A2	M	33.48	> 31.31 ***	22.1	> 19.5 **	14.95	> 14.48	24.84	> 19.42 ***
	SD	8.4	9.71	10.61	11.81	9.17	11.33	10.17	11.6
H1 – A3	M	33.59	> 32.91	46.54	< 43.33 ***	44.24	> 41.9	33.88	> 32.04 ***
	SD	8.17	9.33	8.95	9.63	8.65	8.8	8.55	9.46
A1 – A2	M	24.27	> 22.89 **	11.56	> 11.39	8.40	< 9.44	18.45	> 12.34 ***
	SD	15.04	12.56	14.87	12	10.61	11.14	11.05	9.85

Table 6. Spectral tilt values of 4 non-RTR and RTR vowel pairs and the results of ANOVA (M: Mean, SD: Standard Deviation, ***: $p < .001$, **: $p < .01$, *: $p < .05$)

2005; Starwalt 2008). In this study, we measured and calculated 5 different values (H1 – H2, H1 – A1, H1 – A2, H1 – A3, and A1 – A2) to see whether they distinguish [–RTR] and [+RTR] vowels. Two-way ANOVA's with the factors of [RTR] × length and [RTR] × position revealed that the main effect of [RTR] is significant in H1 – H2 ($F(1,13327) = 100.1$, $p < .001$), H1 – A2 ($F(1,13327) = 336.7$, $p < .001$), H1 – A3 ($F(1,13327) = 73.1$, $p < .001$), and A1 – A2 ($F(1,13327) = 345.829$, $p < .001$). However, the effect is not consistent through all the pairs, as presented in Table 6.⁷⁾

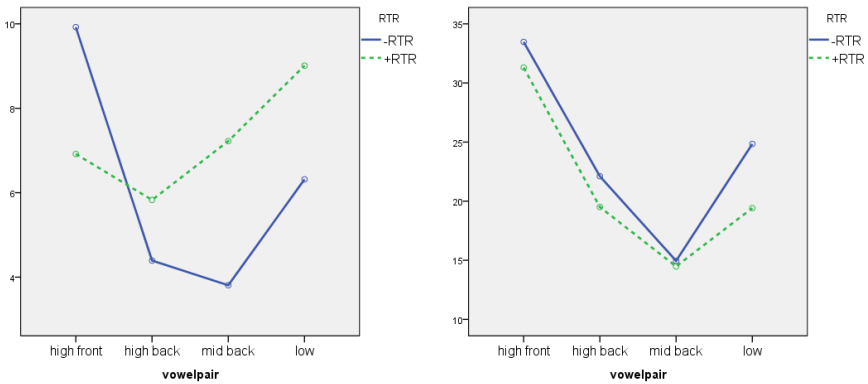
H1 – H2 turns out to effectively distinguish [–RTR] vowels from their [+RTR] counterparts, but /i/ shows higher value than /ɪ/, while other [–RTR] vowels show lower values than their [+RTR] counterparts (Figure 8). H1 – A2 values are consistently higher in [–RTR] vowels than in [+RTR] vowels, but the difference is not significant between /o/ and /ɔ/ (Figure 9). The other three values (H1 – A1, H1 – A3, and A1 – A2) are not reliable cues: They do not show consistency in terms of direction and significance. For this

7) The 'spectral tilt' values were measured on 13,328 vowels with the loss of 212 vowels due to technical issues.

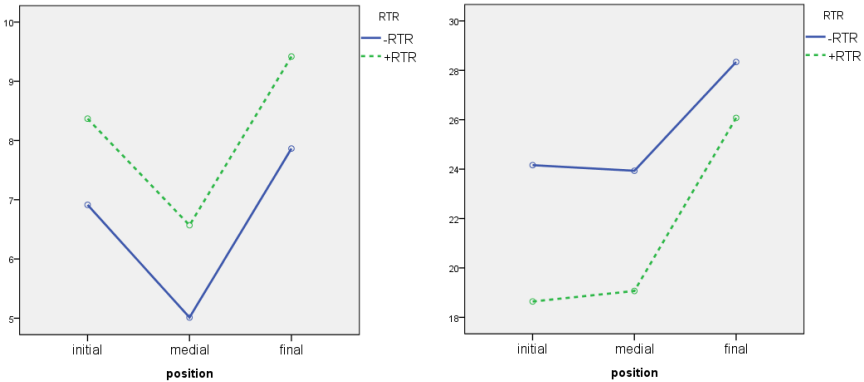
reason, we will omit them when we discuss the effect of position and length below.

As seen in Figures 10 and 11, H1 – H2 difference between [–RTR] and [+RTR] vowels does not change depending on the position of vowel, while H1 – A2 values differ more in initial and medial positions than in final positions.

However, the effect of position and length is not found in all the pairs, either. Table 7 shows the effect of position and length in each vowel pair. There is an overall tendency that initial positions and long vowels maximize



Figures 8 (LEFT) and 9 (RIGHT). H1 – H2 values (LEFT) and H1 – A2 values (RIGHT) of the four vowel pairs



Figures 10 (LEFT) and 11 (RIGHT). H1 – H2 values (LEFT) and H1 – A2 values (RIGHT) of non-RTR and RTR vowels in initial, medial, and final positions

acoustic feature	vowel pair	position		length	
		Interaction	bigger difference	Interaction	bigger difference
H1 – H2	/i/ vs. /ɪ/	***	initial	**	long
	/u/ vs. /ʊ/		final	***	long
	/o/ vs. /ɔ/	***	final	***	long
	/ə/ vs. /a/		initial		long
H1 – A2	/i/ vs. /ɪ/	**	medial	*	long
	/u/ vs. /ʊ/	*	initial		long
	/o/ vs. /ɔ/		final		long
	/ə/ vs. /a/	***	initial	***	long

Table 7. Interactions of [RTR] × position and [RTR] × length (***: $p < .001$, **: $p < .01$, *: $p < .05$)

the spectral tilt difference between [–RTR] and [+RTR] vowels. But in Table 7 there are also cases where this tendency is not observed. This may indicate that the position and length effect is not systematic in Beryozovka Ewen.

Lastly, we measured and compared B1 as a potential acoustic cue of [RTR] as previous studies did (Hess 1992, Kang & Ko 2012). In every pair, B1 turns out to be higher in [+RTR] vowels than their [–RTR] counterparts (high front: $F(1,2360) = 59.614, p < .001$; high back: $F(1,2363) = 151.583, p < .001$; mid back: $F(1,976) = 10.308, p < .01$; low: $F(1,7925) = 303.915, p < .001$), as seen in Figure 12.

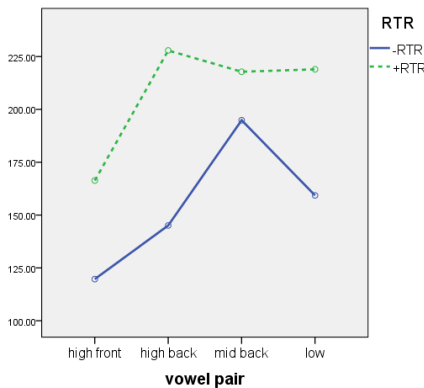


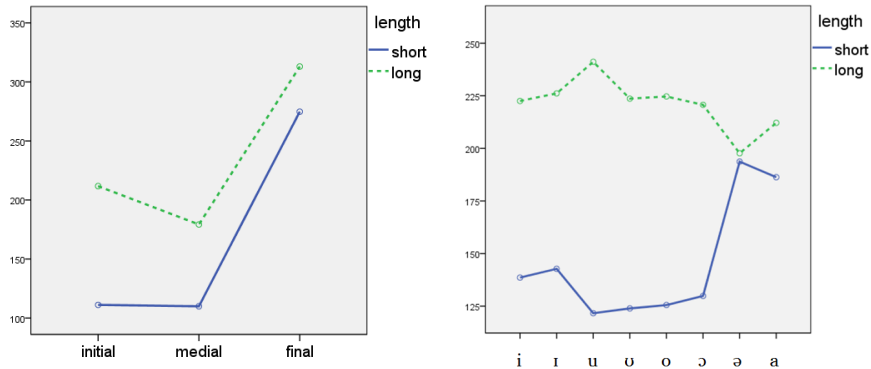
Figure 12. B1 values of 4 non-RTR and RTR vowel pairs

To summarize the results of spectral tilt, these acoustic values are not as consistent and reliable as F1 and F3, though H1–H2, H1–A2, and B1 can be regarded as plausible candidates for the acoustic cues distinguishing [–RTR] and [+RTR] vowels.

4.3 Duration

Though duration is not the focus of this study, we present the results of measurements and comparisons as empirical data.⁸⁾ First, we examine whether the (phonologically) long vowels are indeed realized as longer at the phonetic level than their (phonologically) short counterparts. The mean values are 213.3 ms vs. 160.6 ms, which are significantly different ($F(1,13539) = 416.697, p < .001$). Interaction with position is also significant ($F(5,13534) = 416.697, p < .001$), which means that the vowel length contrast is affected by the position of vowel. As Figure 13 shows, the gap between long and short vowels is maximized in initial positions and minimized in final positions. We also find the effect of final lengthening in the same figure.

Another interesting fact about duration is that /ə/ does not show the length contrast at the phonetic level (Figure 14). We speculate that this is due to the fact that short /ə/ (and /a/ as well) appears in word-final



Figures 13 (LEFT) and 14 (RIGHT). Durations of long and short vowels in initial, medial, and final positions (LEFT) and length contrast of each vowel (RIGHT)

8) Since we used the automated segmentation, it should be noted that the duration of vowels might not be as accurate as that processed by manual segmentation. See Yun, Hwang & Ko (2012) and Yun, Kang & Ko (2016) for discussions on this issue.

positions more frequently than other vowels, and thus often undergoes the final lengthening, which might mask the underlying length contrast at the surface.

5. Discussion and Conclusion

The acoustic analysis of the vowels of Beryozovka Ewen, which was assisted by the automated annotation, revealed that the most reliable acoustic cues of the feature [RTR] are F1 and F3 in formant structures. There are also other potential acoustic cues such as H1 – H2 (with an exception of the high front vowel pair), H1 – A2 (with an exception of the mid back vowel pair), and B1, among many values representing the spectral tilt of vowels. The position and length effects are not robust, though there are strong tendencies that the acoustic value differences between [–RTR] and [+RTR] vowels are bigger in initial and long vowels than medial and short vowels, respectively. In this section, we will discuss the acoustic cue of [RTR] first and the position effect next, in comparison with the previous studies.

5.1 The acoustic cue of [RTR]

One of our findings drawn from the results is that F1 is one of the most reliable acoustic cues of [RTR] in Beryozovka Ewen. This is not surprising because it has been proven that F1 values effectively distinguish [+RTR] and [–RTR] vowels from each other in other dialects of Ewen (Aralova, Grawunder & Winter 2011; Kang & Ko 2012), other Altaic languages (Kang & Ko 2012 [Western and Tsongol Buriat]; Lulich & Whaley 2012 [Oroqen]; Yun, Kang & Ko 2016 [Nanai]), and African languages (Hess 1992 [Akan]; Fulop, Kari & Ladefoged 1998 [Degema]; Guion, Post & Payne 2004 [Maa] among others). The empirical contribution of this study is the finding that F3 is also effective, supporting Ladefoged & Maddieson (1996) who suggest that “pharyngealized” vowels in Ewen may have lowered F3 values as in Caucasian languages. This point distinguishes the current study from the previous ones which have reported no (Kang & Ko 2012 [Western Buriat]) or only partial F3 distinction (Kang & Ko 2012 Tsongol Buriat; Lulich & Whaley 2012 [Oroqen]; Yun, Kang & Ko 2016 [Nanai]).

Another acoustic cue which has been taken into account is spectral tilt. Halle & Stevens (1969) relate tongue root feature ([ATR]) to phonation (which is possibly realized as different values for spectral tilt measurements

such as H1 – H2 and H1 – A1). Ladefoged & Maddieson (1996) also claim that [+ATR] vowels have higher energy at high frequency (around F3) than [–ATR] vowels, which may result in high values of spectral tilt measurements such as H1 – A2, H1 – A3, etc. Following these studies, we hypothesized that the [+RTR] and [–RTR] vowels in Beryozovka Ewen would be distinguished by spectral tilt too. This is partially borne out by the results, which are similar to those in our previous studies of the Buriat (Kang & Ko 2012) and Nanai vowels (Yun, Kang & Ko 2016). The rather inconsistent results of spectral tilt measurements found in these studies (including our current ones presented in §4.2) might be attributed to the way to measure the spectral tilt. As Kim (2016) points out, automated analysis of H1 – H2 may not be fully reliable. She measures the H1 – H2 values of Korean vowels, both manually and automatically, using Praat. The comparison reveals that manually measured values were more similar to the results of the previous studies. Reporting some cases of obvious errors found in the automatically measured values, she claims that acoustic analyses using automated measurement should be manually corrected to enhance the reliability. Of course, to process a large set of data, utilizing automated segmentation and measurement is necessary. Though it is assumed that the amount of data is negatively correlated with that of errors, the comparison of the two methods should be carried out in the future.

5.2 Position effect

The F1-F2 values of Ewen vowels reveal that the dispersion of vowels is

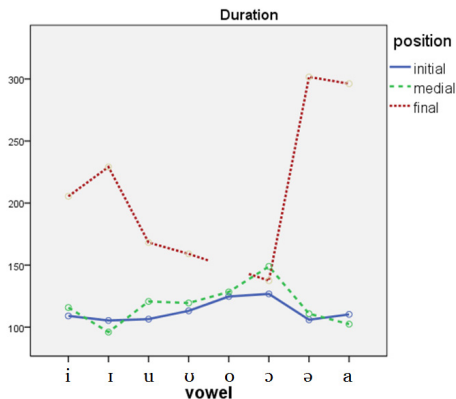


Figure 15. Duration of short vowels in initial, medial, and final positions.

smaller in non-initial positions, which is interpreted as phonetic reduction. Among others, stress is the best known linguistic factor which induces vowel reduction (Fourakis 1991 [English]; Padgett & Tabain 2005 [Russian]; Renwick & Ladd 2016 [Italian]). This leads us to a suspicion that the vowel reduction in Ewen might be related to stress. In other words, it could be claimed that stress regularly falls on initial syllables in Ewen, causing the reduction of vowels in non-initial positions. To attest this possibility, we measured three acoustic correlates of stress—duration, intensity, and pitch (Fry 1955; Fry 1958; Lieberman 1960, among others).

As shown in Figure 15, the duration of vowels does not differ in initial and medial positions ($p = .853$). Final vowels are significantly longer than the others, but this is attributed to the final lengthening. When it comes to intensity, the mean value is higher in medial than in initial ($p < .001$), which also indicates that initial vowels are not stressed in Ewen (Figure 16). Finally, though F0 is the highest in initial positions ($p < .05$) on average, the pattern is not systematic through all the vowels, as seen in Figure 17. All these results guide us to the conclusion that initial vowels in Ewen are not stressed.

Then, why are vowels more distinguishable in initial positions than in others? We believe that the answer should be sought by considering that initial vowels in Ewen must be perceptually prominent—although they are not prosodically prominent—given the rightward feature spreading in suffixal vowel harmony pattern presented in section 2.1. The crucial role

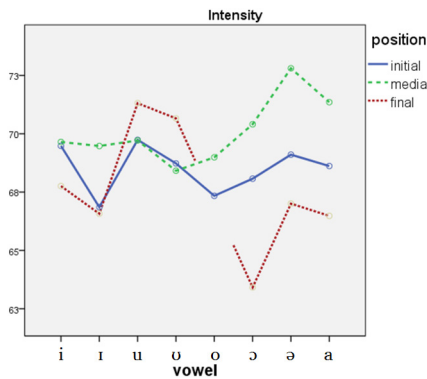


Figure 16. Intensity of short vowels in initial, medial, and final positions.

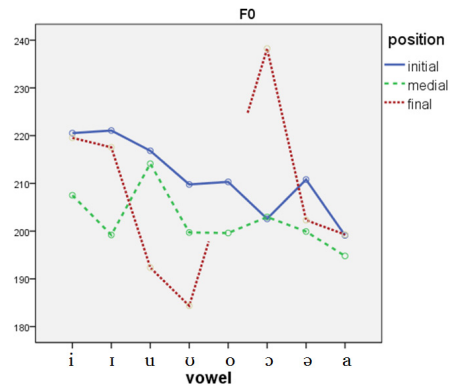


Figure 17. F0 of short vowels in initial, medial, and final positions.

that initial syllables play in lexical access has been noticed in the psychology literature. For example, initial portions provide more effective cues for word recognition or lexical retrieval (Horowitz, White & Atwood 1968; Nootboom 1981) and initial parts help subjects recall the target words better than other parts (Brown & McNeill 1966). This is presumably why non-initial vowels harmonize with the initial one, which is of great importance in lexical access, in many language families such as Turkic, Tungusic, Mongolian, Finno-Ugric, and Bantu (Trubetzkoy 1939; Kiparsky 1981; Beckman 1998; Rose & Walker 2011, among many others). In Ewen as well as in many other Tungusic languages, the initial vowel of a word determines the series of the following vowels in the word. This makes us infer that word-initial vowels should be perceptually salient whereas non-initial vowels do not have to be and, therefore, it might be the case that initial vowels resist reduction while non-initial vowels tends to be reduced at the phonetic level. We predict that similar acoustic observations on vowel reduction patterns—that are not governed by stress—will be made in many other Tungusic (as well as other Altaic) languages with similar vowel harmony patterns.

5.3 Concluding remarks

In this study, we have investigated the vowels of Beryozovka Ewen to find the acoustic correlates of its harmonic feature [RTR]. To this end, we applied a new technology (automated post-transcriptional processing technique) to process a large set of data in relatively short time. The results are basically in line with those of previous studies with some refined findings. It is shown that [RTR] in Beryozovka Ewen is systematically realized as F1 and F3. In addition, though not as reliable as F1 and F3, some acoustic measurements representing spectral tilt are also contributive to the distinction between [–RTR] and [+RTR] vowels. These acoustic cues are relatively fortified in initial positions and in long vowels. These results confirm that Ewen has [RTR] as the harmonic feature, and also suggest that the feature is acoustically better realized in initial positions than in non-initial positions, presumably in order to facilitate the lexical retrieval of Ewen words which comply with vowel harmony.

References

- Aralova, Natalia, Sven Grawunder & Bodo Winter. 2011. The acoustic correlates of tongue root vowel harmony in Even (Tungusic). *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)*. 240–243. Hong Kong, China: City University of Hong Kong.
- 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.
- Beckman, Jill N. 1998. *Positional faithfulness*. Amherst, MA: University of Massachusetts at Amherst Ph.D. dissertation.
- Boersma, Paul & David Weenink. 2014. Praat: doing phonetics by computer [computer program]. Version 5.4.03, retrieved December 18, 2014 from <http://www.praat.org/>.
- Brown, Roger & David McNeill. 1966. The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5(4). 325–337. Elsevier. doi:10.1016/S0022-5371(66)80040-3.
- 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.
- Fourakis, Marios. 1991. Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical Society of America* 90(4). 1816–1827. Acoustical Society of America. doi:10.1121/1.401662.
- Fry, Dennis B. 1955. Duration and Intensity as Physical Correlates of Linguistic Stress. *The Journal of the Acoustical Society of America* 27(4). 765–768. Acoustical Society of America. doi:10.1121/1.1908022.
- Fry, Dennis B. 1958. Experiments in the Perception of Stress. *Language and Speech* 1(2). 126–152. SAGE Publications. doi:10.1177/002383095800100207.
- 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. Karger. doi:10.1159/000028425.
- 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 & Kenneth N. Stevens. 1969. *On the feature “Advanced Tongue Root.”* (MIT Research Laboratory of Electronics Quarterly Progress Report). Cambridge, MA: MIT.
- 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.

- Horowitz, Leonard M., Margaret A. White & Douglas W. Atwood. 1968. Word fragments as aids to recall: The organization of a word. *Journal of Experimental Psychology* 76(2p1). 219–226. American Psychological Association. doi:10.1037/h0025384.
- 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.
- Kaun, Abigail Rhoades. 1995. *The typology of rounding harmony: an Optimality Theoretic approach*. Los Angeles, CA: University of California at Los Angeles Ph.D. dissertation.
- Kim, Juwon. 1989. *Mancwu thwungkwusuceyeuy moumcohwa yenkwu* [A study on vowel harmony in the Manchu-Tungus languages]. Seoul, Korea: Seoul National University Ph.D. dissertation.
- Kim, Juwon. 2011. *A grammar of Ewen* (Altaic Languages Series 6). Seoul, Korea: Seoul National University Press.
- Kim, Mi-Ryoung. 2016. A critical study on the accuracy and reliability of an automated analysis on H1-H2 for Korean phonation types. *The Linguistic Association of Korea Journal* 24(4). 103–127. The Linguistic Association of Korea.
- Kiparsky, Paul. 1981. Vowel harmony. Unpublished ms, MIT.
- Ko, Seongyeon. 2012. *Tongue root harmony and vowel contrast in Northeast Asian languages*. Ithaca, NY: Cornell University Ph.D. dissertation.
- Ko, Seongyeon. 2013. Vowel feature hierarchies in Tungusic. In Juwon Kim & Dongho Ko (eds.), *Current Trends in Altaic Linguistics: A Festschrift for Professor Emeritus Seong Baeg-in on his 80th Birthday*, 91–124. Seoul, Korea: The Altaic Society of Korea.
- Ko, Seongyeon, Andrew Joseph & John Whitman. 2014. Comparative consequences for a tongue root harmony analysis for proto-Tungusic, proto-Mongolic, and proto-Korean. In Walter Bisang & Martine Irma Robbeets (eds.), *Paradigm Change in Historical Reconstruction: The Transeurasian Languages and Beyond*. (Studies in Language Companion Series (SLCS)), Amsterdam, The Netherlands: John Benjamins.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Hoboken, NJ: Wiley-Blackwell.
- Lebedev, V. D. 1978. *Jazyk evenov Jakutii*. Leningrad: Nauka.
- Li, Bing. 1996. *Tungusic vowel harmony: description and analysis* (HIL Dissertations 18). Dordrecht: Holland Institute of Generative Linguistics.
- Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America* 32(4). 451–454. Acoustical Society of America. doi:10.1121/1.1908095.
- Lulich, Steven M. & Lindsay J. Whaley. 2012. An acoustic phonetic study of Oroqen vowels. In Andrei L. Malchukov & Lindsay J. Whaley (eds.), *Recent Advances in Tungusic Linguistics* (Turcologica 89), 59–78. Wiesbaden, Germany: Harrassowitz.

- Nooteboom, Sieb G. 1981. Lexical retrieval from fragments of spoken words: beginnings vs. endings. *Journal of Phonetics* 9(4). 407–424. Elsevier.
- Novikova, Klavdija A. 1960. *Ocherki dialektov Evenskogo jazyka: ol'skij govor*. Vol. 1. Leningrad: Nauka.
- Padgett, Jaye & Marija Tabain. 2005. Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica* 62(1). 14–54. Karger. doi:10.1159/000087223.
- Przedziecki, Marek Andrzej. 2005. *Vowel harmony and coarticulation in three dialects of Yoruba: phonetics determining phonology*. Ithaca, New York: Cornell University Ph.D. dissertation.
- Renwick, Margaret E. L. & D. Robert Ladd. 2016. Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 7(1). 1–29. Ubiquity Press. doi:10.5334/labphon.17.
- Rose, Sharon & Rachel Walker. 2011. Harmony systems. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The Handbook of Phonological Theory*, 240–290. Hoboken, NJ: Wiley-Blackwell.
- Simons, Gary F. & Charles D. Fennig (eds.). 2017. *Ethnologue: languages of the world*. 20th ed. Dallas, Texas: SIL International. <http://www.ethnologue.com> (May 2, 2017).
- 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, Texas: University of Texas at Arlington Ph.D. dissertation.
- Trubetzkoy, Nikolai Sergeevich. 1939[1969]. *Principles of phonology (Grundzüge der Phonologie)*. (Trans.) Christiane A. M. Baltaxe. Berkeley; Los Angeles: University of California Press.
- Yun, Jiwon, Hyun Kyung Hwang & Ko Seongyeon. 2012. Automatic Annotation for Korean Speech Corpus Analysis. Poster presented at the The International Workshop on Corpus Linguistics and Endangered Dialects. National Institute for Japanese Language and Linguistics, Tokyo, Japan.
- Yun, Jiwon, Hijo Kang & Seongyeon Ko. 2016. A phonetic study of Nanai vowels using automated post-transcriptional processing techniques. *Altai Hakpo* 26. 29–44. The Altaic Society of Korea.

KANG Hijo
 Department of English Education
 Chosun University
 309 Pilmun-daero Dong-gu
 GWANGJU 501-759 KOREA
 <hijokang@chosun.ac.kr >

YUN Jiwon

Department of Linguistics

Stony Brook University

STONY BROOK, NY 11794-4376, USA

<jiwon.yun@stonybrook.edu >

Ko Seongyeon (corresponding author)

Department of Classical, Middle Eastern, and Asian Languages and Cultures

Queens College of the City University of New York

65-30 Kissena Blvd

QUEENS, NY 11367-1597, USA

<Seongyeon.Ko@qc.cuny.edu >

Received 6 May 2017;

revision received 31 May 2017;

accepted 12 June 2017.

