

PRODUCTION AND PERCEPTION OF NORTH AND SOUTH KOREAN VOWELS: A PILOT STUDY

Jungah Lee, Charlotte Vaughn, Kaori Idemaru

University of Oregon

jlee27@uoregon.edu, cvaughn@uoregon.edu, idemaru@uoregon.edu

ABSTRACT

Previous studies have reported that some North Korean and South Korean vowels have undergone changes in different directions after decades of relative isolation. This paper presents a pilot study concentrating on three such pairs of vowels ([ɯ]-[u], [ʌ]-[o], and [e]-[æ]), with a particular interest in examining language standards in each variety by comparing newscasters and regular speakers from each country. Acoustic analyses of vowel spaces confirmed some North-South differences, and a perception test verified that North Korean [ʌ] and [o] are highly confusable to South Korean speakers. Noteworthy in the data, North Korean regular speakers who now reside in South Korea showed some vowel patterns that were different from both North Korean newscasters and South Korean speech.

Keywords: North Korean phonetics, vowel analyses, second dialect acquisition, sound change

1. INTRODUCTION

North Korean and South Korean are mutually intelligible varieties of Korean that have diverged along several dimensions since the separation of the two countries in 1957. Due to this long period of isolation, and different language policies, major lexical, morphological, phonological, and phonetic differences have been reported between the two varieties. For example, [e] and [æ] are now considered to have merged to a single category ([ɛ]) in South Korean [3, 12], but this merger has not been reported for North Korean. Meanwhile, in North Korea the back vowel [ɯ] has been reported to have merged to [u] [4, 5], and [ʌ] and [o] are reported to have merged in North Korean [4, 5], or possibly to have reversed height positions [6], compared to South Korean.

Owing to the difficulty in communication between North Koreans and those outside the country, these North Korean vowel patterns have been measured from the speech of speakers living in South Korea (North Korean defectors) [4, 5, 6]. It is still unclear whether these reported vowel mergers ([ɯ] and [u], [ʌ] and [o]) and non-mergers ([e] and

[æ]) represent patterns in North Korean speech, or something specific to North Korean defectors' speech. Further, how a North Korean standard language ideology would evaluate the status of these sound changes is unknown. It may be that pre-division language norms are highly esteemed, or that North Korean innovations hold prestige. In the 1960s, the North Korean regime decreed the speech of working class Pyongyang speakers to be the "cultured language" [9], but the current status of North Korean language ideology, particularly around more recent vowel changes, is unclear. As a way of investigating what may approximate standard North Korean language, we have examined samples of newscasters' speech available online. We also collected South Korean newscasters' speech to compare standard language of the two countries. We further examined speech samples from regular North Korean speakers (defectors) and South Korean speakers to examine the extent to which newscasters and regular speakers from each country are aligned in their vowel patterns.

Vowel formants (F1 and F2) were measured in speech samples from four groups of speakers (newscasters and "regular" speakers from North and South Korea), and analyses focused on 3 contrasts: [ɯ] vs. [u], [ʌ] vs. [o], and [e] vs. [æ]. A perception study was also conducted to further examine one of the reported North Korean mergers, between [ʌ] and [o], to see whether those two North Korean vowels are indeed confusable to South Korean listeners.

2. VOWEL PRODUCTION STUDY

2.1. Methodology

2.1.1. Speakers and speech samples

Speech samples were collected from nine speakers (Table 1). North and South Korean newscasters' speech was extracted from YouTube videos [1, 2, 11, 13]. Samples for all newscasters were selected to have the same topic, each reporting on events of the 2018 Winter Olympics held in Pyeong-chang, South Korea. The extracted samples were between 70 and 120 sec.

Samples for the "regular" speaker group came from 5 speakers' reading productions of the "North Wind and the Sun" passage and short lists of isolated words from [8]. Three North Korean speakers were

defectors living in South Korea. Of the three, the 2 female speakers were recent arrivals (6 months), and the male had lived in South Korea for about 30 years at the time of recording. These speakers were recorded in Seoul, Korea, in a sound booth using a flash digital recorder (Marantz PMD 670) and a microphone (Zoom H1) at a sampling rate of 44 kHz and 16-bit quantization. The male South Korean regular speech was from the male recording provided on [8] (recorded after several years away from South Korea), and the female speech came from a South Korean speaker recruited in the US recorded using the same equipment as the North Korean defectors (length of residence 2 months in US).

Table 1: Speakers.

Country	Speaker type	Gender	Age
North	Newscaster	F	60
		M	46
	Regular	F	38
		F	37
South	Newscaster	M	50
		F	45
	Regular	M	60
		F	38
		M	47

2.1.2. Tokens

Tokens of eight Korean vowels, [i], [e], [æ], [ɯ], [u], [ʌ], [o], [a], were segmented (Table 2) and their first and second formants (F1 and F2) were measured at the mid-point. The analysis here focuses on vowel contrasts of [ɯ]-[u], [ʌ]-[o], and [e]-[æ]. Raw F1 and F2 values were normalized using the Lobanov normalization method in the R Vowels package [7] in order to compare across speaker genders.

Table 2: Token counts.

Speaker type	Vowels									Total
		[i]	[e]	[æ]	[ɯ]	[u]	[ʌ]	[o]	[a]	
NN	F	35	20	24	40	22	41	63	38	283
	M	25	5	11	15	12	15	15	13	111
NR	F	134	59	40	113	92	121	58	235	852
	M	80	28	19	42	34	47	28	76	354
SN	F	55	24	28	46	22	29	40	71	315
	M	39	22	27	32	10	35	24	52	241
NR	F	75	31	17	41	45	49	28	89	375
	M	24	11	7	15	17	18	11	27	130
Total		467	200	173	344	254	355	267	601	2661

NN = North Korean newscasters, NR = North Korean regular speakers, SN = South Korean newscasters, SR = South Korean regular speakers

2.2. Results

Figure 1 presents the normalized F1 and F2 of the eight vowels for each speaker group.

2.2.1. Analysis of category merger

In order to explore the degree of vowel category mergers, group-wide Pillai statistics, a standard measure of vowel merger [10] were calculated for each of the three vowel pairs of interest. Lower values of Pillai indicate a higher degree of merger, and higher values indicate more distinction. Because of the low number of speakers per group, statistical testing on speaker-level Pillai is not possible. Thus, we give group-level Pillai scores here as an indication of the overall trends. With that caveat in mind, the numerical trends in Pillai in this study, shown in Table 3, do follow predictions from prior work, with North Koreans indeed showing more merger (lower Pillai) of [ɯ] and [u], and [ʌ] and [o], than South Koreans [4, 5], and South Koreans showing more merger of [e] and [æ] than North Koreans [3, 12]. However, we also note that newscasters and regular speakers did not always pattern together within countries.

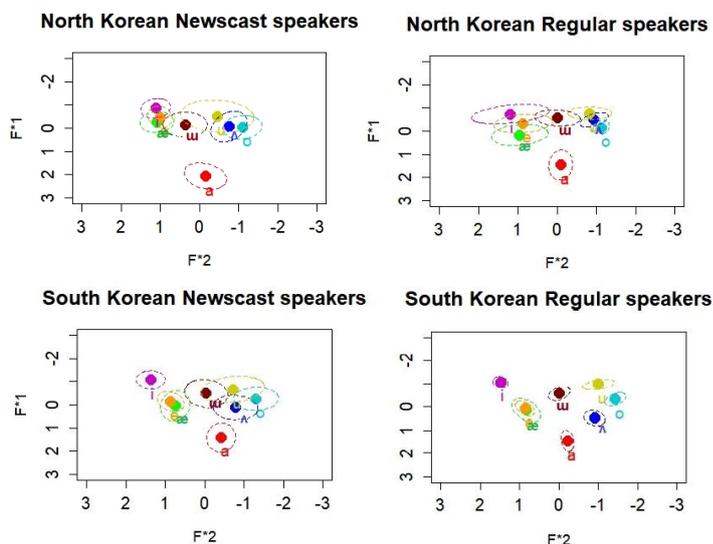


Figure 1: Vowel production results. Ellipses = 1 SD for each vowel category.

Table 3: Pillai statistics by country and speaker type

Country	Speaker Type	[ɯ]-[u] Pillai	[ʌ]-[o] Pillai	[e]-[æ] Pillai
SK	Regular	0.674	0.717	0.009
	Newscaster	0.206	0.255	0.050
	Total	0.440	0.486	0.029
NK	Regular	0.401	0.348	0.252
	Newscaster	0.294	0.118	0.065
	Total	0.348	0.233	0.158
Which country more merged?		NK	NK	SK

2.2.2. Analysis of F1 and F2

In order to understand the more fine-grained patterning of vowel height and backness dimensions separately for these speakers, normalized F1 and F2 of the vowel pairs of interest (i.e., [ɯ]-[u], [ʌ]-[o], and [e]-[æ]) were analyzed using linear mixed effect models with vowel, country, and speaker type

(SpkrType, newscaster vs. regular) as fixed effects and item as a random effect.

Here we report results involving the factor Vowel, which can indicate separation or non-separation of each vowel pair. Pillai statistics reported in Section 3.1 showed numerically more merger of [u] and [ʊ] for North Korean than South Korean speakers, supporting prior work [4, 5]. In these models of F1 and F2, however, country was not a significant predictor of [u] and [ʊ]. Rather, a significant main effect of vowel suggests that the F1 and F2 of the two vowels were statistically separate, across country ($B = 0.374$, $t = 4.156$, $p < 0.0001$ for F1, $B = 0.842$, $t = 6.850$, $p < 0.0001$ for F2). Thus, despite the trend in Pillai indicating more merger for North Korean speakers, [u] and [ʊ] were statistically distinct across speakers in terms of F1 and F2.

For [ʌ] and [o], Pillai statistics also showed more merger in North Korean than South Korean speakers. The F1 and F2 results for [ʌ]-[o] present a complex picture. For F1, there was a statistically significant Vowel x Country x SpkrType interaction ($B = -0.698$, $t = -4.799$, $p < 0.0001$). Post-hoc tests exploring the interaction found that whereas North Korean newscasters did not separate [ʌ] and [o] in terms of height, North Korean regular speakers separated them by producing [ʌ] higher than [o], and South Korean speakers, regardless of speaker types, separated them by producing [ʌ] lower than [o] ($p < 0.001$ for both). For F2, we found a statistically significant Vowel x SpkrType interaction ($B = 0.180$, $t = 2.010$, $p = 0.045$). Post-hoc tests indicated that both newscasters and regular speakers across country separated [ʌ] and [o] in terms backness ($p < 0.0001$ for both speaker types), but newscasters showed a greater degree of separation. These results are generally consistent with prior findings that [ʌ] and [o] show an inverse relationship in height between North and South Korean [6]. Our North Korean newscasters did not statistically follow this pattern, showing that the two vowels are not distinct in terms of height (F1). Further, it is noteworthy that North Korean regular speakers (defectors living in South Korea) produce these vowels in a manner different from both North Korean newscasters and South Koreans. In general, these results also support the idea that the height dimension is indeed more at work in differentiating between countries for these two vowels.

Finally, for [e] and [æ] our Pillai results were consistent with prior work reporting merger in South Korea [3, 12]. However, for F1 and F2 models, we found that these two vowels were not statistically separate either by F1 or F2 for any speaker groups. For all the speakers in this sample, it appears that [e] and [æ] are close together, though the vowel plots in Figure 1 and Pillai statistics suggest that the degree of

overlap between [e] and [æ] in North Korean speakers is indeed less.

3. VOWEL PERCEPTION STUDY

To investigate the perceptual consequence of the reported merger of [ʌ] and [o] in North Korean (and the complex pattern of [ʌ] and [o] across North and South Korean speakers in our production results), we tested whether South Korean listeners could identify these vowels produced by North Korean newscasters and North Korean regular speakers (defectors).

3.1. Methodology

3.1.1. Listeners and experiment conditions

Twenty native South Korean listeners participated. All were residing in the U.S. at the time of testing. Twelve listeners (5 female, mean age = 32) were instructed simply that they would hear syllables and that their task was to identify which vowel they heard (Not-Informed Condition). Eight listeners (5 female, mean age = 31) were given the same instructions and were additionally informed that the speech came from North Korean speakers (Informed Condition). A post-test debriefing survey revealed that participants were unfamiliar with North Korean speech, and with differences between South Korean and North Korean.

3.1.2. Stimuli and task

In addition to the vowels of interest, [ʌ] and [o], the vowel [a] served as a filler that should be easy for South Korean speakers to identify. To collect vowel tokens for use as perception stimuli, CV syllables that included the vowels [a], [ʌ], and [o] (78, 81, 61 tokens respectively) were identified in the North Korean speech samples (see Section 1.2) and extracted. The preceding consonant in the CV frame was [p], [t], or [s]. This resulted in 220 stimuli in total. The stimuli were amplitude normalized to 75dB using Praat.

The perception test was administered to participants individually in a sound booth. Seated in front of a computer, listeners heard an auditory presentation of a syllable through headphones, and were asked to indicate which vowel they heard in the syllable. The response choices—boxes containing [a], [ʌ], and [o] written in Korean orthography—appeared on the computer screen at each trial, and listeners indicated their response choice by using a mouse to click the box corresponding to the vowel they just heard. Listeners were allowed to listen to the stimulus syllable up to 7 times.

3.2. Results

The perception study investigated one of the reported North Korean mergers, between [ʌ] and [o], to see

whether South Korean listeners can accurately identify between the two vowels. Descriptive statistics are presented in Table 4 for the Not-informed condition (given no information about the speakers) and Table 5 for Informed condition (told the speakers were North Korean). In general, [a] was easy to identify across the conditions (92% correct and 81% correct), [o] was harder to identify (69% and 73%), and [ʌ] was hardest to identify (18% and 19%).

Response data were analyzed using a mixed effect logistic regression model with stimulus vowel ([o] as the reference) and condition (Informed as the reference) as the fixed effects, and item as a random effect.

Table 4: Perception results for not informed condition

Response \ Stimulus	[a]	[ʌ]	[o]
[a]	862 (92%)	70	4
[ʌ]	3	175 (18%)	794
[o]	4	223	505 (69%)

Table 5: Perception results for informed condition

Response \ Stimulus	[a]	[ʌ]	[o]
[a]	503 (81%)	103	18
[ʌ]	12	125 (19%)	511
[o]	4	128	356 (73%)

The results confirmed that correct identification of [ʌ] was lower than that of [o] ($B = -2.440$, $z = -17.029$, $p < 0.0001$), which in turn was lower than that of [a] ($B = 0.435$, $z = 3.021$, $p = 0.003$). Indeed, North Korean [ʌ] and [o] were confusable for South Korean listeners. Interestingly, we found a Condition x [a] interaction ($b = 0.225$, $z = 5.992$, $p < 0.001$). This indicates that listeners in the Informed condition were less accurate than listeners in the Not-informed condition at identifying North Korean [a]. Although [a] was included as a filler, predicted to be easy to identify and thus not of central interest, the significant interaction with condition reveals something about participants' approach to the task: South Korean listeners may have been less certain in their decisions when they were warned that they would be responding to unfamiliar North Korean speech compared to when no such warning was provided.

4. DISCUSSION

The current study investigated three vowel contrasts, namely, [ʉ]-[u], [ʌ]-[o], and [e]-[æ], in North Korean and South Korean speech provided by newscasters and regular speakers.

As for the production results, although [ʉ]-[u] did show higher degree of merger for North Korean speech than for South Korean speech, F1 and F2 did not statistically overlap across these vowel categories,

and this separation was the same across countries. These results raise a question regarding the status of reported [ʉ] and [u] merger [4, 5]. More production data is needed to examine [ʉ] and [u] in more detail. It will also be critical to conduct perception studies to determine the extent to which North Korean [ʉ] and [u] are perceptually confusable.

As for [ʌ]-[o], the production study revealed subtle differences in vowel height between North and South Korean, yet the perception study indicated that South Koreans found North Korean productions of these vowels highly confusable. Both North Korean newscasters and defectors produced [ʌ] higher than [o] (numerically), the opposite of the South Korean pattern, confirming the prior report [6]. However, our results revealed differences between North Korean newscasters (standard speakers) and defectors: newscasters showed a stronger degree of merger between [ʌ] and [o] than the defectors. These results suggest that more work in production and perception should be done to determine the best way to characterize these vowels across the two varieties, as a height reversal or a merger.

The [e] and [æ] merger has been well documented for South Korean [3, 12], and our data confirmed that the two vowels are quite close for all speakers. Although a potential [e]-[æ] merger in North Korean has attracted relatively little attention, in these data, the degree of [e]-[æ] overlap is stronger than that of [ʌ]-[o].

These results, particularly that of [ʌ]-[o], indicate the importance of using different types of speech samples when attempting to characterize North Korean speech. In particular, the speech of North Korean defectors is hard to interpret, having only speech from North Korean newscasters and South Korean speakers to serve as comparisons. Future research is needed to investigate whether North Korean defectors' patterns indicate a dissociation in stance from either speaker group as a form of identity marking, or an early stage in a process of second dialect acquisition. Future research incorporating a larger sample of speakers will help address these issues.

Although these results add data to a body of work attempting to characterize vowel variation and change in this interesting dialect situation, more work is needed to better understand the linguistic and social patterns observed. More studies of production, including other cues like duration, coupled with perceptual studies testing both North and South Korean listeners, will be informative.

5. REFERENCES

- [1] Choson TV: KCTV conghappoto 03wel 01il 107. 2018. [107th summarized news of 03/01/2018] *Online resource*: <https://www.youtube.com/watch?v=4hZhw-LoZss>
- [2] Choson TV: mochelem malyentoy nampwukkwankyeykaysenpwunwikilul hulye nohci malla namcosensimintancheyka yokwu- oy 1ken. [Do not break this peaceful relationship between South and North Korea in this Olympic season that South Korean civic group asked-and one more news] 2018. *Online resource*: <https://www.youtube.com/watch?v=GyEeGpRe0y0>
- [3] Eychenne, J., Jang, T. Y. 2015. On the merger of Korean mid front vowels: Phonetic and phonological evidence. *Journal of the Korean Society of Speech Sciences Vol, 7(2)*.
- [4] Ihyenpok. 1991. pwukhanuy wulimal umsenghak yenkwu hyenhwang. [North Korean phonetic study in the present situation] *Hankul, (213), 7-28*
- [5] Kangswunkyeng. 1999. pwukhan hwusel moumuy yunghap(merger) hyensang. [The mergers of North Korean back vowels] *Umsengkwahak, 5(2), 41-55*.
- [6] Kang, Y., Yun, S. 2018. Acquisition of second dialect features by North Koreans in Seoul. Poster presented at *LabPhon16*, University of Lisbon, Lisbon, Portugal.
- [7] Kendall, T., Erik R. 2010. Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package, version 1.1.
- [8] Lee, H. B. 2000. IPA Handbook Downloads | *International Phonetic Association*. Retrieved from <https://www.internationalphoneticassociation.org/content/ipa-handbook-downloads>
- [9] Lee, I. S., Ramsey, S. R. 2000. *The Korean Language*. Albany, New York: State University of New York Press.
- [10] Nycz, J., Hall-Lew, L. 2013, December. Best practices in measuring vowel merger. In *Proceedings of Meetings on Acoustics 166ASA* (Vol. 20, No. 1, p. 060008). ASA.
- [11] MBN News. Kimcwuhayuy 2wel 8il nyusuchocem wikisel pwuchwukinun ilpon. [Kim Ju-ha's February 8th News Focus, Japan thinks North Korea will make troubles in this Olympic period] 2018. *Online source*: <https://www.youtube.com/watch?v=Qn0lvxGb9iI>
- [12] Shin, J., Kiaer, J., Cha, J. 2012. *The sounds of Korean*. Cambridge University Press.
- [13] [Sonsekhuyuy ayngkhepuliphing] “ankyeng ssun yeca...” [Sohn Suk-hee's anchor briefing “Woman athlete with glasses”] 2018. *Online resource*: <https://www.youtube.com/watch?v=F3oCAhAaFzU>