CONTRASTIVE HYPERARTICULATION OF VOWELS IN TWO DIALECTS OF KOREAN

Yoonjung Kang[1,2], Na-Young Ryu[2], & Suyeon Yun[3]

[1]University of Toronto Scarborough, [2]University of Toronto, [3]Ewha Womans University
yoonjung.kang@utoronto.ca, nayoung.ryu@mail.utoronto.ca, suyeonyun@ewha.ac.kr

ABSTRACT

Previous studies on vocalic hyperarticulation generally found that vowels become longer and more dispersed, i.e., more peripheral in vowel space, under hyperarticulation, rather than more spectrally distinct from contrasting vowels. This study examines contrastive hyperarticulation of vowels in the speech of 57 speakers of two dialects of Korean, Seoul Korean (South Korea) and Hamkyoung Korean (North Korea). Participants produced 24 sets of lexical minimal pairs three times, once in isolation (casual), once carefully as if speaking to a non-native speaker (careful), and once directly following the speaker's production of the word's lexical minimal pair (contrastive). In both dialects, we found a statistically significant effect of speech condition (contrastive > careful > casual) on vowel duration, degree of dispersion, and spectral distance between contrasting vowel pairs. A closer inspection of the data shows, however, that the evidence for spectral contrast enhancement independent of dispersion is inconclusive.

Keywords: Korean vowels, dialects, hyperarticulation, minimal pairs

1. BACKGROUND

Speakers skillfully adjust their speech to adapt to the communicative demands of the speech context [15]. In particular, previous research examined if and how speakers adjust their speech when they are prompted to speak clearly, especially to distinguish target words from their lexical competitors [17, 22]. Under contrastive hyperarticulation, speakers tend to slow down their speech, producing longer segments, and exaggerate acoustic contrasts such as VOT/voicing [1, 20] and vowel formants [4]. Such hyperarticulation emerged as an important topic of research due to its implications for the adaptive nature of speech production and its purported interaction with the mental representation of words and speech sounds.

To probe how underlying phonological contrasts affect the pattern of hyperarticulation, previous studies compared clear speech production of identical or comparable target segments across different languages, dialects, or speaker groups. For example, Kang and Guion [10] examined contrastive hyperarticulation of Korean stops, which are undergoing tonogenetic sound change whereby lenis and aspirated stops are merging in VOT with F0 taking over as the primary distinctive cue in phrase-initial position. They found that older Seoul Korean speakers enhance the VOT distinction in clear speech, while younger Seoul Korean speakers exaggerate both VOT and F0, reflecting the changing nature of the contrast.

Unlike consonantal contrasts, evidence for contrast-specific enhancement of vowels remains elusive [13, 17, 20]. While some studies found evidence for "global" enhancement such as longer vowel duration [17] and vowel space expansion [4] in clear speech, evidence for contrast-specific spectral enhancement remains limited at best. For example, despite a large difference in vowel inventory size, speakers of Croatian and English were comparable in vowel peripheralization under clear speech conditions [21].

Clopper and Tamati [5] examined the low-front and low-back vowels in two dialects of American English and found that words with minimal pair competitors were produced with larger spectral differences than those without lexical competitors. Moreover, the lexical competitor effect was more pronounced in the dialect where the contrasting vowel pairs were acoustically more similar and hence in need of more contrast enhancement. However, it is possible that the attested contrast expansion is in fact an epiphenomenon of vowel peripheralization. Wedel, et al. [24] specifically distinguished these two possibilities by looking at vowel contrasts where contrast enhancement would result in centralization, rather than peripheralization of target vowels. However, these two studies that found evidence for spectral enhancement of lexical contrast are based on general speech corpus, rather than on speech data specifically designed to elicit contrastive clear speech in the context of lexical competitors. Hence, as far as we know, there has been no convincing demonstration of contrast-specific vocalic spectral enhancement in clear speech.

In this paper, we examine the contrastive hyperarticulation of vowels in Seoul Korean and Hamkyoung Korean, two dialects of Korean known to differ in their vowel structures [12, 14]. We examine the manifestation of contrastive hyperarticulation along three acoustical dimensions—duration, spectral dispersion, and contrast-specific spectral enhancement—across the two dialects. Table 1 provides an inventory of Korean monophthongs. The front rounded vowels are diphthongized in many dialects of Korean and are not included in our study.

Table 1: Inventory of Korean monophthongs

i	(y)	ɨ	u
e	(ø)	ʌ	o
ɛ		ɑ	

2. METHOD

2.1. Participants

Data collection took place in Seoul between December 2016 and January 2017. Participants included 36 North Korean defectors who are native speakers of the Northern Hamkyeong dialect of Korean, spoken in the northeastern region of North Korea, and 21 speakers of Seoul Korean. The participants were balanced for age groups in both dialects and balanced for gender in Seoul but not for Hamkyoung, due to the fact that the North Korean refugee population is predominantly female (http://www.unikorea.go.kr/). Table 2 summarizes the demographic information of the participants.

Table 2: Participant demographic information

	Old (above 40)	Young (40 or under)
Seoul	5F, 5M	5F, 6M
Hamkyoung	18F, 3M	13F, 2M

2.2. Speech materials and Procedure

A total of 24 minimal pairs were included, three for each of the eight pairs of neighbouring monophthongs (/i-e/, /e-ɛ/, /ɛ-ɑ/, /ɑ-ʌ/, /ʌ-o/, /o-u/, /u-ɨ/, and /ɨ-i/). For example, for the /ɛ-ɑ/ contrast, we included /sɛtɑ/ 'to leak' - /sɑtɑ/ 'to buy', /pɛm/ 'snake' - /pɑm/ 'night', and /ɛksu/ 'amount' - /ɑksu/ 'handshake'. Four words were used for two lexical pairs, resulting in a total of 44 words (8 vowel pairs * 3 minimal pairs * 2 words – 4 repeated words).

Each word was presented in standard orthography along with a picture depicting the target word, to disambiguate it from homophones and to reduce the monotonous nature of the reading task. The stimuli were presented to the speakers using *PsychoPy* [17] on a Microsoft Surface tablet. The task was self-paced and the participants advanced the words by touch screen. The participants were first asked to read the word on the screen comfortably ("casual" condition). After the casual reading, the participants were then asked to read the same words but this time as if they were speaking to a non-native speaker of Korean to help them learn Korean. For this phase, the words were presented in pairs. First, one word of the pair appeared on the left side of the screen along with its corresponding picture. After participants read the first word, its minimal pair appeared on the right. Each pair of words was presented twice in alternating order so that both words were presented once as the first item of the pair and once as the second item of the pair. The word presented first was intended to induce general hyperarticulation ("careful" condition) while the word presented second, right after its minimal pair, was intended to induce *contrastive* hyperarticulation ("contrastive" condition). So, in total, each word was read three times, once in the "casual" condition, once in the "careful" condition, and once in the "contrastive" condition. There were no repetitions. In total, each speaker produced 140 word tokens (24 word pairs * 2 words * 3 conditions – 4 words repeated words produced only once in the casual condition). 10 tokens were omitted due to error, mispronunciation, or background noise, and a total of 7970 tokens were analyzed (140 * 57 speakers – 10 omissions).

2.3. Acoustic analyses

The recorded speech was segmented and analyzed in *Praat* [3]. For each vowel, duration and the first two formants from the mid 10% of the vowel were automatically measured. Based on an exploratory analysis, the formant measurement settings that minimized errors (as indirectly measured by variance in formant values [7]) were determined. The formant ceiling was set at 4,000 Hz for males and 4,500 Hz for females, and the number of target formants was set at 5 for back vowels (/ɑ, ʌ, o, u, ɨ/) and 4 for front vowels (/i, e, ɛ/). The automatically measured values were further filtered by excluding 273 outliers that fall outside 2.5 standard deviations in either F1 or F2 for each vowel. This process removed all instances of visually identifiable outliers in the vowel plot.

2.4. Statistical analyses

Statistical analyses were conducted in R [19]. The formant values were z-scored transformed for each speaker (cf. [16]) to allow for comparisons across gender and speakers. We examined the effect of speech condition on hyperarticulation using three phonetic variables; (i) vowel duration, (ii) dispersion from the centre of the vowel space, and (iii) the distance between contrasting vowels in minimal pairs. Vowel duration is straightforward. The dispersion was calculated as the Euclidean distance of each vowel token from the centre of the respective speakers' vowel space, as shown in (1). The distance between contrasting vowel pairs was calculated as the

Euclidean distance of the two vowels in each lexical minimal pair produced for each speech condition, as shown in (2).

(1) $Dispersion_i = \sqrt{F1_i^{\,2} + F2_i^{\,2}}$

(2) $Distance_{ij} = \sqrt{(F1_i - F1_j)^2 + (F2_i - F2_j)^2}$

Three separate linear mixed-effects regression models were built using the *lmer*() function of the *lme4* package [2], with each of the three hyperarticulation measures as a dependent variable. The fixed effects predictors in the initial models included CONDITION (casual, careful, and contrastive), DIALECT (Seoul, Hamkyoung), VOWEL/VOWEL.PAIR, GENDER (Female, Male), AGE (Old, Young), and their full interactions. All predictors were simple coded (reference levels are underlined). The models included random intercepts for speakers and words/word.pairs, as well as a by-speaker random slope for CONDITION, if no convergence problem occurred. The initial full models were pared down by backward step-wise regression using the *step()* function. Post-hoc tests were done using the *testInteractions*() function of the *phia* package [6]. Below we report the results focusing on CONDITION, the main topic of our study, and its interactions with other predictors. Statistical analyses were also conducted to test for dialect- and age-based differences on vowel formants. The alpha value of 0.05 is assumed for significance reporting and statistical details are suppressed for space.

3. RESULTS

3.1. Overview of the vowel positions

Figure 1 summarizes the position of the eight monophthongs by dialect and age. We observed a number of statistically significant differences for dialect and age which are in line with previous related studies [8, 9, 11, 23]. The two front mid vowels (/e, ɛ/) are almost merged in Seoul, more so for younger than older speakers, while they remain distinct and stable across age groups in Hamkyoung. The high central vowel /ɨ/ is much more front in Seoul, with younger speakers fronting the vowel more than older speakers, while in Hamkyoung, the vowel is further back, although younger speakers show fronting toward a Seoul-like position. The positions of the two mid back vowels are also very different between the two dialects. In Seoul, /o/ is raising to a position close to /u/, with younger speakers raising more than older speakers, while /ʌ/ is in a lower mid vowel position. In Hamkyoung, the height of the two vowels is reversed in older speakers, while younger speakers again show movements toward the Seoul-like positions, raising /o/ and lowering /ʌ/.

3.2. Duration

Figure 2 summarizes the effects of speech condition and dialect on vowel duration. We found a significant main effect of CONDITION (contrastive > careful > casual), DIALECT (Hamkyoung > Seoul), and VOWEL (not shown). In other words, speakers produced longer vowels when asked to speak carefully, and the contrastive speech condition induced further lengthening over the careful condition. In terms of dialect, Hamkyoung speakers generally produced longer vowels than Seoul speakers. We also found a significant two-way interaction of CONDITION * VOWEL. A post-hoc test shows that while a significant durational difference between the casual and contrastive conditions was consistently found across all vowels, the careful condition does not differ from the other two conditions consistently across all vowels.

Figure 1: Korean vowels by dialect, age, and speech condition

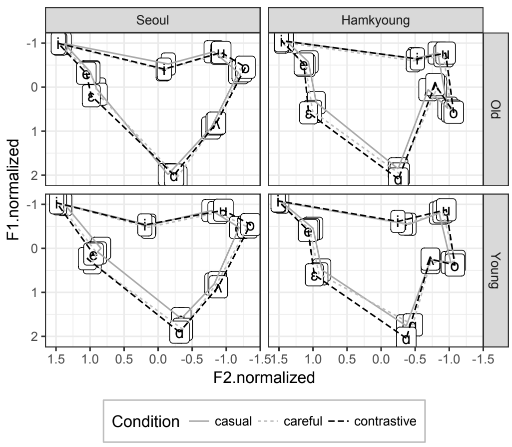
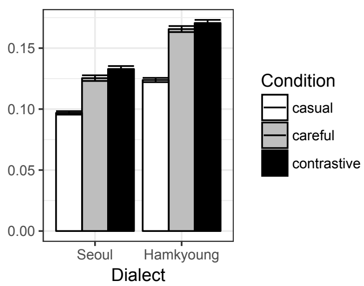


Figure 2: Vowel duration by speech condition



3.3. Dispersion

Figure 3 summarizes the effects of speech condition and dialect on vowel dispersion. There was a significant main effect of CONDITION (contrastive > careful > casual) and VOWEL (not shown). In other words, some vowels are farther away from the centre than others, as expected, and the contrastive condition induced the most peripheral vowel production followed by the careful and then casual speech condition. A significant two-way interaction was found for CONDITION * VOWEL. A post-hoc test shows that while a significant degree of dispersion difference between the casual and contrastive conditions was consistently found across all vowels (except for /ɨ/), the careful condition does not differ from the other two conditions consistently across all vowels. Also, a significant three-way interaction was found for CONDITION * DIALECT * AGE. Post-hoc tests show that this is due to the fact that speech condition has a much weaker effect in older Seoul speakers' speech, while the effect is more consistent for the other groups – younger Seoul speakers, older and younger Hamkyoung speakers.

Figure 3: Dispersion by speech condition

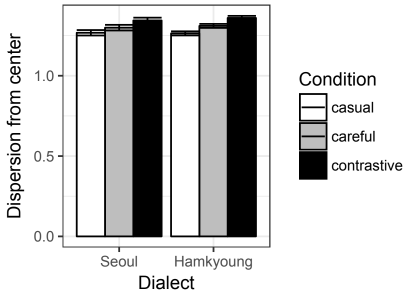
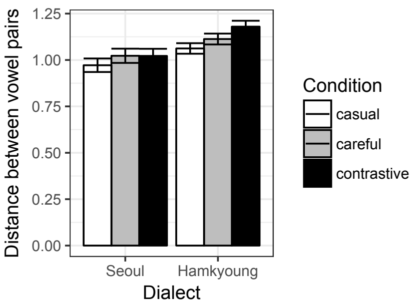


Figure 4: Vowel distance by speech condition



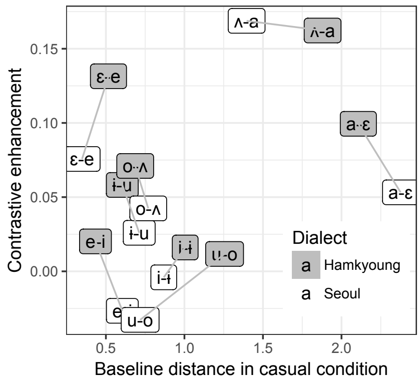
3.4. Distance between contrasting vowels

Figure 4 summarizes the effects of speech condition and dialect on the distance between neighboring vowels in lexical minimal pairs. There was a significant main effect of CONDITION (contrastive ~ careful > casual), DIALECT (Hamkyoung > Seoul), and VOWEL (not shown). The distance between neighboring vowels was smaller in casual speech compared to the careful and contrastive conditions. The difference between the careful and contrastive conditions was in the expected direction but did not reach significance. A significant two-way interaction was found for CONDITION * VOWEL. A post-hoc test shows that the effect of CONDITION was significant only for /ɑ-ɛ/ and /ɑ-ʌ/, and marginal for /ɛ-e/ and /ɨ-u/. This suggests that the significant main effect of CONDITION was likely mainly driven by the lowering and peripheralization of /ɑ/.

3.5. Similarity and enhancement

We also tested the hypothesis that the spectral contrast enhancement of vowel pairs is larger in dialects where the vowels are more acoustically similar [5]. Figure 5 plots the baseline mean acoustic difference, by dialect, between the target vowel and its competitor vowel in casual speech on the x-axis and the increase in distance from casual to contrastive speech conditions on the y-axis. For five of the eight vowel pairs, there is a predicted negative correlation between the baseline distance and contrastive enhancement across dialects, i.e., distance between vowels is enhanced more in the dialect where the vowels are acoustically more similar. However, in the other three vowel pairs /u-o/, /ɛ-e/, and /i-ɨ/, the Seoul dialect has a smaller baseline difference than Hamkyoung but the enhancement is smaller or negative (the distance from the constrasting vowel *decreased* in the contrastive condition). Interestingly, these three pairs are precisely the ones involved in change in progress in Seoul Korean that reduces the distance between the contrasting vowels (merger of /e-ɛ/, fronting of /ɨ/, and raising of /o/). Further analyses will investigate the correlation between sound change and the direction and degree of contrastive enhancement and its interaction with speaker age.

Figure 5: Baseline distance and contrast enhancement



4. CONCLUSION

This study examined contrastive hyperarticulation in two dialects of Korean. We found evidence of global enhancement of vowels in careful and contrastive clear speech, i.e., lengthening and peripheralization. However, the evidence for spectral enhancement of contrasting vowels seems marginal at best, in line with previous studies. We found a general trend of a negative correlation between baseline acoustic differences in vowel pairs and the degree of contrastive enhancement across dialects. Exceptions found are suggestive of its interaction with sound change.

ACKNOWLEDGEMENTS

The authors thank the Department of Linguistics at Seoul National University for lab use, Hana Centre for Refugee Adaptation and MS. Sunhee Moon for recruitment assistance, Mr. Hyoung Seok Kwon for assistance with acoustic analysis, and *the Social Sciences and Humanities Research Council of Canada* for funding.

5. REFERENCES

[1] Baese-Berk, M., Goldrick, M. 2009. Mechanisms of interaction in speech production. *Language and cognitive processes* 24, 527-554.

[2] Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G. 2017. lme4: Linear mixed-effects models using 'eigen' and S4. R package version 1.1-13. [Computer software] https://CRAN.R-project.org/package=lme4

[3] Boersma, P., Weenink, D. 2016. Praat: Doing phonetics by computer (Version 6.0.23). [Computer program] http://www.praat.org/.

[4] Cho, T., Lee, Y., Kim, S. 2011. Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics* 39, 344-361.

[5] Clopper, C. G., Tamati, T. N. 2014. Effects of local lexical competition and regional dialect on vowel production. *The Journal of the Acoustical Society of America* 136, 1-4.

[6] De Rosario-Martinez, H., 2015. phia: Post-Hoc interaction analysis. R package version 0.2-1. [Computer Software] https://CRAN.R-project.org/package=phia.

[7] Escudero, P., Boersma, P., Rauber, A. S., Bion, R. A. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America* 126, 1379-1393.

[8] Jung, M. S. 2002. Sound change of language used in broadcasting. *Kwukehak* 39, 221–249. [In Korean]

[9] Kang, J., Kong, E. 2016. Static and dynamic spectral properties of the monophthong vowels in Seoul Korean: Implication on sound change. *Phonetics and Speech Sciences* 8, 39-47. [In Korean]

[10] Kang, K. H., Guion, S. G. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America* 124, 3909-3917.

[11] Kang, Y. 2016. A corpus-based study of positional variation in Seoul Korean vowels. In: Kenstowicz, M., Levin, T., Masuda, R. (eds), *Japanese/Korean Linguistics* 23. Stanford: CSLI Publications, 3-22.

[12] Kang, Y., Schertz, J., Han, S. 2015. Vowels of Korean dialects. Poster presented at *the 169th Meeting of the Acoustical Society of America*, Pittsburgh, Pennsylvania, USA.

[13] Kirov, C., Wilson, C. 2012. The specificity of online variation in speech production. *Proceedings of the Annual Meeting of the Cognitive Science Society* 34.

[14] Kwak, C. G. 2003. The vowel system of Contemporary Korean and direction of change. *Kwukehak* 41, 59-91. [In Korean]

[15] Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, W. J., Marchal, A. (eds), *Speech production and speech modeling*. Amsterdam, The Netherlands: Kluwer Academic.

[16] Lobanov, B. M. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49, 606-608.

[17] Ohala, J. J. 1994. Acoustic study of clear speech: A test of the contrastive hypothesis. *Proceedings of the international symposium on prosody*, 75-89.

[18] Peirce, J. W. 2007. PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods* 162, 8-13.

[19] R Core Team 2016. R: A language and environment for statistical computing (Version 3.3.2) [Computer program]. Vienna, Austria: R Foundation for Statistical Computing.

[20] Schertz, J. 2013. Exaggeration of featural contrasts in clarifications of misheard speech in English. *Journal of Phonetics* 41, 249-263.

[21] Smiljanić, R., Bradlow, A. R. 2005. Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America* 118, 1677-1688.

[22] Smiljanić, R., Bradlow, A. R. 2009. Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and linguistics compass* 3, 236-264.

[23] Yoon, T. J., Kang, Y., Han, S., Maeng, H. S., Lee, J., Kim, K. 2015. A corpus-based approach to dialectal variation in Korean vowels. *Proceedings of the 18th International Congress of Phonetic Sciences*.

[24] Wedel, A., Nelson, N., Sharp, R. 2018. The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language* 100, 61-88.