Journal of the Korean Society of Speech Sciences Vol.7 No.2 (2015.06.30) www.speechsciences.or.kr
ISSN 2005-8063, pp. 119~129 http://dx.doi.org/10.13064/KSSS.2015.7.2.119

On the Merger of Korean Mid Front Vowels: Phonetic and Phonological Evidence

Eychenne, Julien[1) · Jang, Tae-Yeoub[2)

ABSTRACT

This paper investigates the status of the merger between the mid front unrounded vowels ㅔ[e] and ㅐ[ɛ] in contemporary Korean. Our analysis is based on a balanced corpus of production and perception data from young subjects from three dialectal areas (Seoul, Daegu and Gwangju). Except for expected gender differences, the production data display no difference in the realization of these vowels, in any of the dialects. The perception data, while mostly in line with the production results, show that Seoul females tend to better discriminate the two vowels in terms of perceived height: vowels with a lower F1 are more likely to be categorized as ㅔ by this group. We then investigate the possible causes of this merger: based on an empirical study of transcribed spoken Korean, we show that the pair of vowels ㅔ/ㅐ has a very low functional load. We argue that this factor, together with the phonetic similarity of the two vowels, may have been responsible for the observed merger.

Keywords: mid front vowels, merger, Korean monophthongs, functional load

1. Introduction

The vowel inventory of contemporary Korean needs intensive empirical investigation due to lack of agreement in phonological and phonetic studies. The number of monophthong phonemes of Seoul Korean varies, depending on scholars, from seven (namely /i, e, ɑ, ʌ, o, u, ɯ/) to ten (plus /ɛ, ø, y/). Traditional scholarship (Huh, 1965; Lee, 1996; Sohn, 1999, among others) consistently presents ten monophthongs as phonemes, due presumably to the influence of the Korean writing system (Hangeul), which uses different graphemes for all those ten vowels. More recently, phoneticians (Shin, 2000; Hwang & Moon, 2005) have argued for a reduced number of vowels (7 to 9, instead of 10) based on results of phonetic experiments. The three front vowels /ɛ, ø, y/ are categorized differently by different scholars. We will put aside the two front rounded vowels written ㅚ and ㅟ and transcribed as /ø/ and /y/, respectively, as they are regarded as diphthongs (/we/ and /wi/, respectively) by most researchers. But there remains a controversy about the proper treatment of non-high unrounded vowels: the vowel written ㅔ is described as a mid (or mid-high) vowel and is usually transcribed /e/; the vowel ㅐ is described as a mid low (or sometimes low) vowel, and is generally transcribed either /ɛ/ or /æ/. (We will use the symbol /ɛ/ in the remainder of this paper.) However, it has repeatedly been observed that these two vowels are often realized identically (see for instance Shin *et al*., 2013: 99-101), although the issue surrounding this has not been fully resolved yet.

The pair of Korean monophthongs /e/ and /ɛ/ thus raises two fundamental issues: (i) to what extent is the merger complete in contemporary Korean? (ii) What triggered the merger in the first place? In an attempt to answer the first question, the current study investigates the status of these vowels from production and perception data. As for production, relatively recent phonetic research generally agrees that no particular distinction between or

1) Hankuk Univ. of Foreign Studies, jeychenne@hufs.ac.kr
2) Hankuk Univ. of Foreign Studies, tae@hufs.ac.kr, Corresponding author

This work was supported by Hankuk University of Foreign Studies Research Fund of 2015.

Received: April 29, 2015
Revised: June 12, 2015
Accepted: June 15, 2015

within the two vowels is available except when spoken by speakers of different gender (Yang, 1996; Seong, 2004; Moon, 2007). While confirming this gender effect, the current approach also explores whether there is any cross-dialectal variation in the realization of these two vowels since, as is widely known, vowels produced by speakers of Southern provinces of Korea are different in quality as compared to the corresponding vowels produced by speakers in Seoul or other areas (*e.g.*, merger of /ʌ/ and /ɯ/ in Southeastern dialects. See Yeon, 2012:12 for examples).

Apart from the acoustic characteristics of Korean /e/ and /ɛ/, there have been few systematic phonetic studies that checked whether Koreans perceptually distinguish between them. Thus, we performed a perception experiment on this issue. Instead of using recorded tokens as stimuli, we conducted vowel identification tests with synthesized stimuli that cover the whole F1 and F2 space of Korean monophthongs. This approach is inspired by previous works in the study of vowel perception (Chládková & Escudero, 2012; Mannell, 2004, among others), although it varies in the details. This method has been found to be especially useful for comparing dialects or languages.

In order to account for various interactions between variables including dialect, gender as well as vowel category we employ the statistical analysis known as *linear mixed-effect modeling* (Pinheiro & Bates, 2000, Baayen *et al*., 2008). This model is recognized to be more advantageous for analyses with unbalanced data and non-independent parameters than the traditional techniques. Assuming that the two formants F1 and F2 are critical cues to distinguishing quality of Korean vowels, reflecting native speakers' articulatory gestures, the main statistical analyses are performed in such a way that each formant is placed as a dependent variable while the other variables such as vowel identity, dialect and gender are regarded as independent variables or predictors.

In addition, to address our second question, we investigated the role of *functional load* (Martinet, 2005 [1955]), a measure of the degree of contrastivity of a pair of phonemes in a given language, in order to understand how phonological factors may have played a role in the merger of /e/ and /ɛ/. Based on an analysis of the spoken part of the Sejong corpus, we show that this pair of phonemes indeed has a low functional load, which may have contributed to the merger of the two vowels.

The rest of this paper is organized as follows: section 2 presents and discusses data and results from our production test; section 3 describes the perception experiment that was carried out; section 4 provides an overview of the corpus-based investigation of functional load and section 5 provides a general discussion of the results, as well as other issues related to the distinction of Korean mid front vowels.

2. Production

2.1 Data

2.1.1 Subjects

To test the possibility of a DIALECT effect on vowel production, we selected three contrasted dialectal zones, namely Seoul (in the center, representative of standard Korean), Daegu (Southeast, Gyeongsang dialect) and Gwangju (Southwest, Jeolla dialect) (see Sohn, 1999 for an overview). A total of 30 subjects, equally balanced for gender across all survey points (5 males and 5 females per group) were recorded. We selected subjects who were all under 40 (mean = 22.4; s.d. = 3.8) not only because this study aims to examine the contemporary Korean pronunciation of young adult speakers, but because the perceptually normalized frequency scale (see §3.1 for details) used in the perception experiment is only assumed to be valid for young listeners at moderate sound level. Consequently, neither age effect nor diachronic change of the vowels in question is of interest in the current study.

Participants were all University students at HUFS at the time of recording. All of them were monolingual native speakers of Korean who had spent their primary school, middle school and high school in the city they were from, except for one Daegu male who spent his high school in Geochang (located 60 km southwest of Daegu) and one Gwangju male who spent his high school in Naju (located 20 km southwest of Gwangju); however, both cities belong to the same administrative province and dialectal area as Daegu and Gwangju, respectively. None of the participants had spent more than 6 months abroad in the last 3 years and none of them had any known speaking or hearing impairment.

2.1.2 Recording

For practical reasons, the experiment was conducted at the University's two campuses (in Seoul and Yongin): 18 participants were recorded in a sound-proof booth in Seoul, while the other 12 were recorded in a quiet room. The same hardware was used in all cases.

The recording prompt was not limited to the two relevant mid front vowels but included the eight Korean monophthongs (all the vowels mentioned in §1 except /ø/ and /y/). This setting was due to our broader research purpose of examining the perceptual correlates of all Korean monophthongs, but it was also useful in the current study for concealing from the subjects the purpose of experimentation.

Participants were asked to produce sentences containing each vowel in [h__dɑ] context. The series of 8 sentences was repeated 5 times in different random order each time. The data were acquired through a TASCAM US-155MK II audio interface, using a head-mounted AKG-C520 microphone. Sound was sampled at 16 kHz and stored in WAV format. The recorded chunk was tokenized into individual speech files so that each token includes one vowel. The beginning and ending point of each vowel was demarcated and annotated based on a phone alignment tool followed by manual adjustments.

The whole experiment, including the perception test reported in §3, took less than 30 minutes for each participant. Subjects were paid a small amount for their participation.

2.2 Formant extraction

A total of 300 tokens (2 vowels x 10 speakers x 3 dialects x 5 repetitions) were used for examination. For each vowel, values of F1 and F2 were extracted at its midpoint using the Praat program (Boersma & Weenink 2014) with processing parameters: 5500 Hz as maximum frequency and 6 as number of formants. After values were automatically extracted and examined, obvious errors (such as 750 Hz marked as F2) were manually corrected. It is not immediately clear why this type of formant detection errors occurred but they only appeared in the production of a few particular participants.

2.3 Results

The overall mean (and standard deviation) values (in Hz) for F1 and F2 were: 539 (90) Hz, 2078 (262) Hz for /e/ and 549 (90) Hz, 2054 (259) Hz for /ɛ/. <Figure 1> shows the distribution of F1-F2 coordinates for the two vowels without separate consideration of such variables as DIALECT and GENDER. The distributions are largely overlaid and vowels are not distinguishable from each other by either F1 or F2 as is confirmed statistically ($F(1, 298) = 0.77$, $p = 0.381$ for F1, $F(1, 298) = 0.64$, $p = 0.425$ for F2).

As comparing overall magnitudes provides nothing but a general picture of the two vowels, consideration of relevant variables and their interaction is necessary. <Table 1> and <Figures 2> and <Figure 3> provide a more analytical summary of magnitudes and their distribution on each subject group.

Table 1. Mean and standard deviation of F1 and F2 values on dialect and gender

		Seoul		Daegu		Gwangju	
		M	F	M	F	M	F
e	F2	1868 (115)	2261 (148)	1896 (243)	2246 (176)	1981 (271)	2210 (227)
	F1	531 (73)	573 (91)	525 (58)	557 (118)	482 (46)	566 (101)
ɛ	F2	1910 (200)	2268 (172)	1856 (223)	2209 (141)	1917 (231)	2159 (238)
	F1	539 (63)	573 (87)	534 (49)	578 (142)	494 (51)	571 (90)

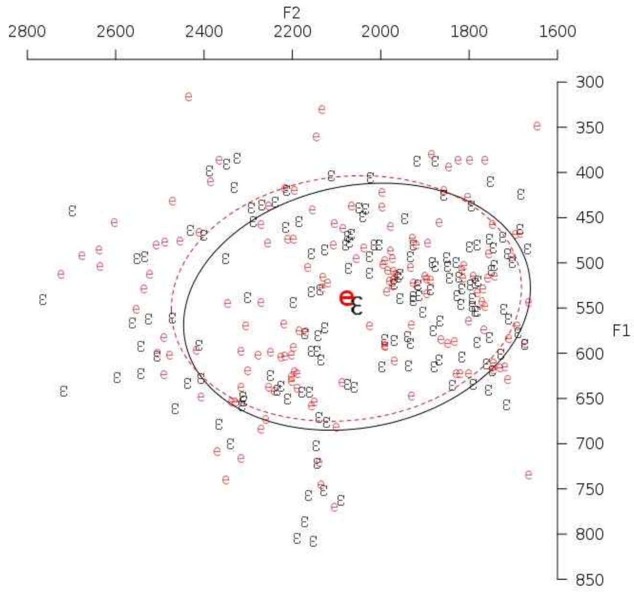

Figure 1. Overall distribution of F1 and F2 of /e/ and /ɛ/: Each ellipse encompasses 68.27 %, or one standard deviation, confidence limit.

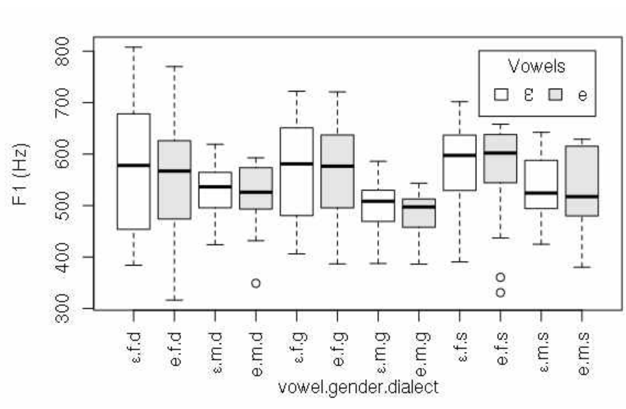

Figure 2. Distribution of F1 on speaker groups

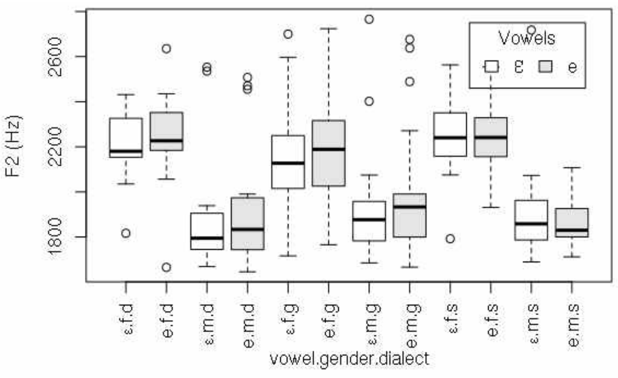

Figure 3. Distribution of F2 on speaker groups

No DIALECT effect has been statistically confirmed for either F1 ($\chi^2(2) = 0.81$, $p = 0.667$)) or F2 ($\chi^2(2) = 0.19$, $p = 0.908$). The result leads to the speculation that the two mid front vowels are not phonetically realized differently from each other in any of the three dialects of Korean.[3]

As for GENDER, female speakers produce vowels with distinctly higher frequency than male speakers for both F1 ($F(3, 296) = 9.4$, $p < 0.001$) and F2 ($F(3,296) = 61.5$, $p < 0.001$). When the data were fit to the *linear mixed-effects model* with subsequent *maximum likelihood ratio tests*, it was found that male voice has a lower F1 magnitude by about 52.1 Hz ± 26.1 ($\chi^2(1) = 4.27$, $p = 0.039$) and lower F2 by about 320.9 Hz ± 51.0 ($\chi^2(1) = 27.80$, $p < 0.001$). The observed GENDER effect that male speakers produce the vowels with lower formants than female speakers agrees with previous research on Korean vowel production (Moon, 2007; Seong, 2004; Yang, 1996). The cause of the effect is attributable to a language independent physiological property of human vocal organs rather than parameters to be verified by any linguistic examination. Corresponding gender differences in formant frequency can be easily found in reports of other languages (Hillenbrand *et al*., 1995 for English, Kriengwatana *et al*., 2014 for Dutch, Pépiot, 2012 for French).

To summarize, young Koreans who speak one of the three dialects do not produce the vowel /e/ and /ɛ/ differently from each other, while formant frequencies can be a cue to identifying speaker groups of different gender.

3) Although Daegu Korean is known to have lexical pitch accents unlike the other two dialects, there has been no reliable evidence suggesting that this prosodic feature is linked to the merger of /e/ and /ɛ/ in this dialect.

3. Perception

Much evidence has been provided to date on the discrepancy between production and perception of speech sounds. Thus, it is interesting to test whether the production of /e/ and /ɛ/ will be projected in the same pattern to the perception of those vowels. The main points to be checked are twofold: first, whether the gender difference which is found to be the only relevant factor in production will be maintained when the vowels are perceived, second, whether the two vowels are perceived distinctly from each other by any speaker group for a given dialect or gender. The same subjects who participated in the production tests were requested to participate in the perception tests as well.

3.1 Stimuli

Synthetic vowels were created using the Praat program (Boersma & Weenink, 2014). The F1/F2 acoustic space was sampled at psycho-acoustically equidistant points on the Equivalent Rectangular Bandwidth (ERB) scale. F1 values were sampled in 23 steps from 240 to 900 Hz, while F2 was sampled in 23 steps from 580 to 2,700 Hz. Extreme values falling outside of the hypothesized perceptual space were trimmed off using a step function similar to (Mannel, 2004). F3 was modeled as a function of F2: for vowels whose F2 was lower than 1500 Hz, F3 was assigned a fixed value of 2500 Hz; in vowels whose F2 was 1500 Hz or higher, F3 was assigned a value 1000 Hz higher than F2 (Chládková & Escudero, 2012). Higher formants (up to 10) were added using Praat's default settings.

The fundamental frequency was modeled in the following way: about half of the vowels had a slope falling linearly from 150 Hz to 100 Hz, while the other half had a slowly declining F0 through the first half of the vowel (from 150 to 140 Hz) and a more abrupt decline throughout the second half (from 140 to 100 Hz). In addition, two durations (200 ms and 300 ms) were used in order to create a more heterogeneous stimulus set. The F0 and duration criteria were assigned randomly (with a probability of 0.5 for each outcome) to each vowel, creating 4 distinct patterns. The final data set contained 343 unique stimuli.

The ERB scale was adopted for this perception test since it is considered to better approximate the frequency response of the basilar membrane than do other scales (such as the Bark scale), especially at low frequencies. The conversion from Hertz to ERB-rate units is based on (Glasberg & Moore, 1990):

$$ERB = 21.4 \times \log_{10}(0.00437 \times F_{Hz} + 1) \quad (1)$$

where F_{Hz} is the frequency in Hertz. One ERB unit is assumed to correspond to approximately 0.9 mm along the basilar membrane (Moore, 2013). <Figure 4> illustrates the correspondence between frequency in Hertz and ERB-rate units. Following (Rosner & Pickering, 1994), we will use E to represent the unit on the ERB-rate scale, and $E(F_i)$ to represent the ERB transform of formant F_i's center frequency (where $i \in \{1, 2\}$).

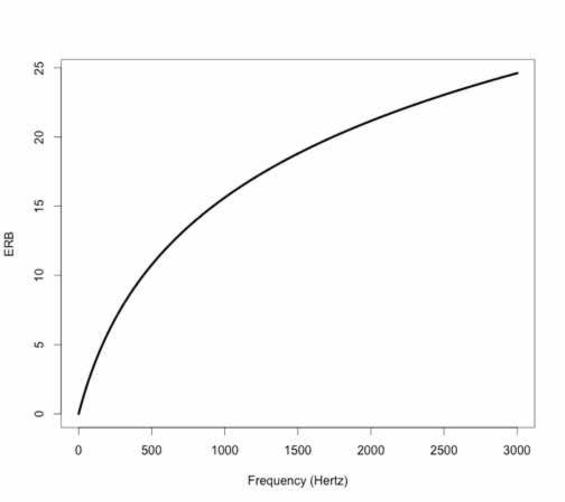

Figure 4. Correspondence between frequency in Hertz and cumulative equivalent rectangular bandwidth for frequencies up to 3,000 Hz

3.2 Procedure

All the subjects of the production experiment described in §2 also participated in the perception test. The task was a multiple forced-choice experiment where the synthesized vowels were presented to participants through circumaural headphones (using a Sony MDR-XD200 device). Stimuli were arranged in random order, and each of them was presented only once. Subjects first went through a short training session using a random subset of 7 items to make sure that they fully understood the task and to let them adjust the loudness to a comfortable level.

Participants were asked to categorize vowels using one of the 8 monophthongs of Korean including /e/ and /ɛ/, using standard Korean spelling: ㅣ, ㅐ, ㅔ, ㅗ, ㅜ, ㅡ, ㅓ, ㅏ corresponding to /i e ɛ u o ɯ ʌ ɑ/ respectively. Vowels were arranged on the screen in pseudorandom order, so as to make sure that vowels which were phonetically or orthographically close were not next to each other. <Figure 5> illustrates the results obtained for one subject.

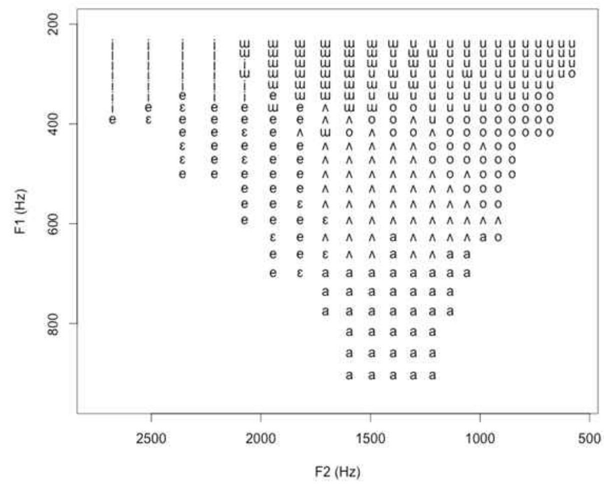

Figure 5. Perceptual map of one Seoul male subject

3.3 Statistical Analysis

Given the nature of the task, it is inevitable that some vowel tokens be miscategorized as a different vowel (*e.g.* a stimulus with a very low F1 and a very high F2, corresponding to the primary cardinal vowel [i], categorized as [ɑ]). To attenuate the effect of such random error, we excluded, for each vowel category, all tokens whose value for either F1 or F2 was more than 3 standard deviations away from the corresponding formant's mean value. A total of 45 tokens were discarded following this procedure, corresponding to 0.44% of the data set. A visual inspection of the data revealed that there were no false positives among the discarded tokens and we did not identify any pattern among the outliers (no particular vowel seemed to trigger more errors than others). The final data set contained 10,245 data points: for the purpose of this study, we restricted our attention to vowels labelled as ㅐ or ㅔ.

The data were analyzed by means of *linear mixed-effects models* with random intercept[4]), which allow modeling the combined effects of fixed and random predictors on a dependent variable. We examined F1 and F2 separately. For each formant, we built a model with the ERB-transformed formant (E(F1) or E(F2)) as the predicted value, the SUBJECT as a random effect and the DIALECT, GENDER and VOWEL as fixed effects. Besides accounting for the main effects, the model also takes

4) Mixed-effects models assume that the within-group residuals are independent and normally distributed, and that the random effects are normally distributed (Pinheiro & Bates, 2000: 174). The fitted models were systematically inspected by means of box plots of the residuals and quantile-quantile plots. We did not observe any deviance from the independence and normality assumptions.

into consideration the three possible two-way interactions (namely DIALECT × GENDER, GENDER × VOWEL and DIALECT × VOWEL) as well as the three-way interaction between DIALECT × GENDER × VOWEL, in order to investigate potential gender differentiation in any of the dialects.

3.4 Results

<Figure 6> provides a boxplot of the average perceived E(F1) for /e/ and /ɛ/, and the magnitudes of corresponding values are summarized in <Table 2>. We observe that Seoul females tend to have a lower F1 for /e/. Likewise, Gwangju males have a noticeably lower F1 than Gwangju females. Furthermore, Seoul females perceived /ɛ/ with a markedly higher F2: Seoul females' /e/ was on average 1.24 E lower than /ɛ/.

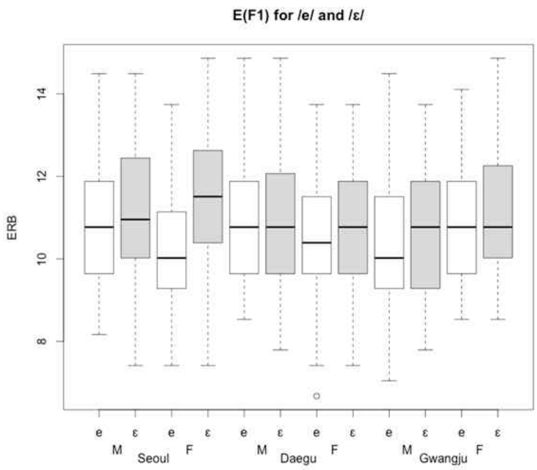

Figure 6. Box plot of ERB(F1) for /e/ vs /ɛ/ across dialects and genders. For each dialect, women appear first, males second. (white = /e/; gray = /ɛ/).

Table 2. Grand mean of F1 (in Hz) for /e/ and /ɛ/. The results were computed on ERB values and were converted to Hertz for readability only

	Seoul		Gwangju		Daegu	
	M	F	M	F	M	F
/e/	489	455	472	505	497	477
/ɛ/	521	553	479	526	521	502

The effect of the Subject random term in the model was estimated at 0.31 E (95% CI: [0.21, 0.46]), indicating that there was a significant amount of inter-subject variation. The mixed-effect model for E(F1) was further subjected to an Analysis of Variance (ANOVA). It did not reveal any main effect for DIALECT and GENDER, but we found a highly significant effect for VOWEL ($p < 0.0001$). Indeed, the mean difference for E(F1) across groups is 0.44 E, /e/ having a lower E(F1) than /ɛ/. The interaction between DIALECT × VOWEL ($p < 0.0001$), GENDER × VOWEL ($p < 0.05$) and the three-way interaction DIALECT × GENDER × VOWEL ($p < 0.0001$) were all significant. While the significance of main effects must be interpreted cautiously in the presence of a significant interaction term, it is clear from the above that there is a gender differentiation in terms of vowel height between these vowels, which is particularly noticeable for Seoul females.

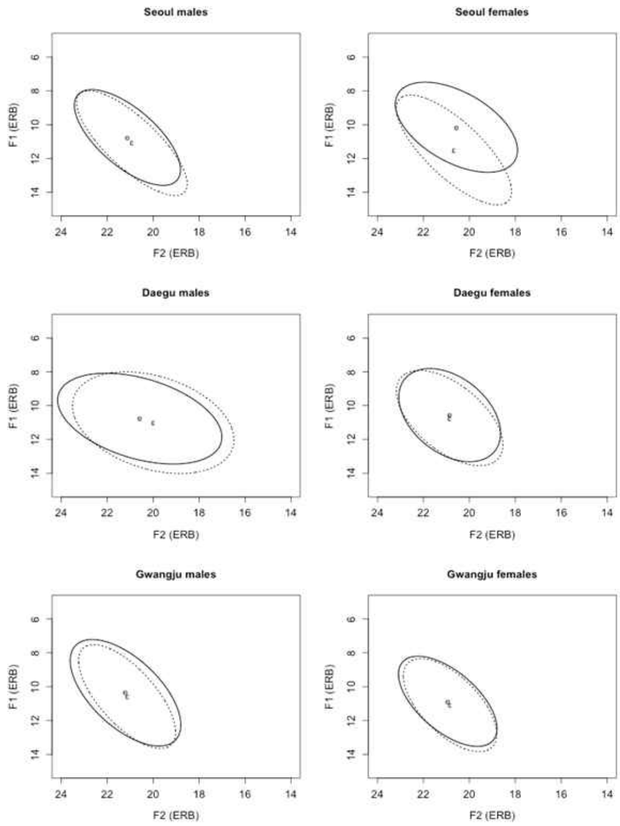

Figure 7. 95% confidence ellipses for the vowels /e/ and /ɛ/ across dialects and gender (/e/ = solid line; /ɛ/ = dotted line)

Regarding F2, the mixed-effects model revealed a significant effect of the SUBJECT, with an estimate of 0.47 E (95% confidence interval: [0.34, 0.64]). However, none of the predictors showed any significant effect (neither as a main effect nor as part of a factor interaction), which indicates that there is no significant difference in perceived backness between these two vowels.

<Figure 7> provides a visual summary of the dispersion of these two vowels in perceptual space in each dialect, for males and females separately. It is apparent that there is overall very

little difference except for Seoul females, who appear to have a much sharper distinction between a mid-high and a mid low vowel.

Interestingly, most subjects reported at the end of the experiment that they had had great difficulty discriminating between /e/ and /ɛ/, often observing that their choice must have been random. Our results suggest otherwise: Even though there is a significant overlap between the two vowels, there is, at least for some sub-populations (especially Seoul females), a greater-than-chance probability for a higher mid front vowel to be categorized as /e/ and for a lower mid front vowel to be categorized as /ɛ/. The difference in F1 between the two vowels was 1.023 E.

4. Functional factors

Since the results from both the production and perception experiments suggest that the contrast is largely lost, we set out to investigate possible causes for this merger. One of the oldest and most important phonological factors that have been argued to play a role in sound change is *functional load* (also called *functional yield*) (Martinet, 2005 [1955]: 35-39). In its most ‘naïve’ form, the functional load of a pair of phonemes represents the number of minimal pairs that these phonemes can distinguish in the lexicon of a given language. Such a definition has two obvious limitations: first, the conception of contrastivity that it assumes is too restrictive since, as is well known, what really matters in a phonological contrast is the fact that phonemes have an equivalent distribution, minimal pairs being only one special subcase where such a distribution is observed. Secondly, this definition does not take into account token frequency, and would attribute the same weight to rare and frequent words. But as Martinet himself pointed out (Martinet, 2005: 36), there are good reasons to assume that frequent words play a more important role than rare ones.

An outline of a more satisfactory definition of functional load was provided by (Hockett, 1955), who offered an information-theoretic formalization of the notion. He proposed the following equation:

$$FL(xy) = \frac{H(L) - H(L_{xy})}{H(L)} \tag{2}$$

where H represents the *entropy* (measured in bits) of a language, and L_{xy} represents the language L where the contrast x versus y has been neutralized. Under this interpretation, functional load represents the *information loss* corresponding to the merger of the two phonemes. While this equation provides a more robust way to quantify functional load, it cannot be used directly, since we cannot estimate the entropy of a language, which is an infinite set of utterances. However, as pointed by (Surendran & Nigoyi, 2006), we can approximate it provided that a sufficiently large corpus of the language is available.

4.1 Data

While we do not have access to a large sample of the spoken language when the merger took place (around the end of the 18th century, see Ahn & Iverson, 2007), we can obtain a reasonable approximation using a sample of contemporary Korean. To this end, we used the spoken part of the Sejong corpus (*a.k.a.* Korean National Corpus, see Kim, 2006), which contains orthographic transcriptions of contemporary spoken Korean. Word forms were extracted and converted to a linear phonological representation. In order to account (at least partially) for suprasegmental information, we added syllabic boundaries mirroring the graphical organization of letters in Hangeul. For example, the word form “있습니다” was phonetized as /.iS.sɯb.ni.dɑ./, where the dots represent syllabic boundaries. Note that we did not apply any phonological rule (such as regressive nasalization, obstruent tensification or resyllabification) to the resulting phonological forms: although using phonetic transcriptions could be preferable, it would not substantially alter the results since we are only analyzing the functional load of vocalic phonemes.

The final dataset that was used in this study contained, after cleanup, 805,514 tokens and 115,024 types.

4.2 Method

Our analysis of the data closely follows (Surendran & Nigoyi, 2006), who developed a computationally explicit method based on Hockett’s seminal idea[5]. For the sake of convenience, let *H(L)* denote the entropy of the dataset containing the full inventory of Korean phonemes (including, in our case, the two mid front unrounded vowels). Let $H(L_{xy})$ denote the entropy of the corpus where the contrast between vowels x and y has been neutralized. As the authors show, we can compute (an approximation of) the entropy of the dataset using equation (2)

5) For an analysis of functional load in Korean using a similar approach, albeit with a different focus and in a cross-linguistic perspective, see Oh *et al*., (2013).

above if we assume that the data were generated by a *k*-order Markov process (in other words, the probability of a phoneme *p* only depends on the *k* previous phonemes). The entropy of the dataset *L* can then be calculated with the following equation:

$$H(L) = \frac{1}{k+1}\left[-\sum_{x \in X} p(x)\log_2(x)\right] \quad (3)$$

In the current study, we used a second-order approximation ($k = 2$), which means that the sequences of phonemes (and/or syllabic boundaries) which were considered were trigrams[6].

The probability of each trigram *t* was estimated from normalized raw counts:

$$p(t) = \frac{c(t)}{\sum_{i=1}^{n} c(i)} \quad (4)$$

where *n* is the total number of trigrams.

4.3 Results

The results for all pairwise comparisons of vocalic phonemes are reported in <Table 3>. It must be emphasized that the functional load of these vowel pairs cannot be interpreted directly. The numerical values are sensitive to the number of phonemes in the language, to the dataset from which they were estimated and to the order of the Markov process assumed to have generated the data. Therefore, the functional load of a pair of phonemes is only meaningful when compared to that of other pairs from the same corpus.

As can be seen, the functional load of the /e ɛ/ pair is quite low compared to other pairs: with a value of 0.0082, it is the third lowest value, after /ɛ u/ and virtually identical to /ɛ ɯ/. To better appreciate the significance of these values, we provide in <Figure 8> a bar plot of the functional load of each pair, where pairs have been sorted in decreasing order (the pair with the highest functional load appears in the leftmost bin). As can be seen, the functional yield of /e ɛ/ is almost 4 times lower than that of /ɑ ʌ/, the most contrastive pair in the corpus.

Table 3. Functional load of all pairs of vowels in the spoken part of the Sejong corpus

	ɯ	u	e	o	ɛ	ʌ	ɑ
i	0.0206	0.0140	0.0149	0.0151	0.0115	0.0245	0.0294
ɯ		0.0121	0.0117	0.0144	0.0081	0.0208	0.0257
u			0.0094	0.0126	0.0074	0.0156	0.0189
e				0.0114	**0.0082**	0.0150	0.0180
o					0.0103	0.0193	0.0233
ɛ						0.0122	0.0169
ʌ							0.0317

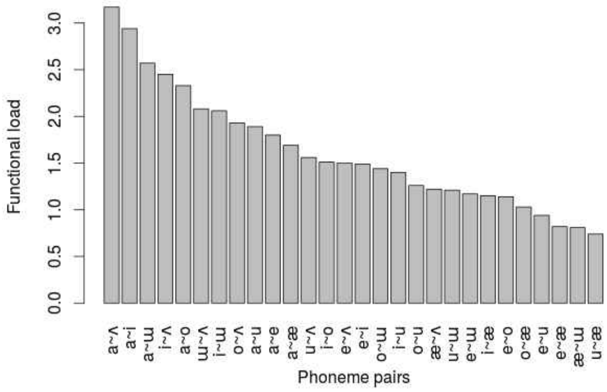

Figure 8. Functional load of all pairs of vowels

The fact that /ɛ/ merged with /e/ rather than /u/ or /ɯ/ can easily be explained on phonetic grounds, since the articulatory and acoustic distance between /e/ and /ɛ/ is far smaller than between /u/ and /ɛ/. Conversely, if it is indeed the case that /ɛ/ was a low front vowel (*i.e.*, closer to [æ] in the IPA sense) when it developed, as is often assumed (Ahn & Iverson 2007), the fact that /ɑ ɛ/ has a relatively high functional load may also explain why /ɛ/ merged with /e/ rather than with the low back vowel /ɑ/[7]. Determining how functional load and phonetic distance might interact in the development of mergers is an issue worth investigating, but which lies beyond the scope of this paper.

5. Summary and discussion

The results of the production experiment are mostly in line with results reported in previous studies. GENDER was the only

6) We repeated the experiment using a first-order Markov process. The results were numerically slightly different but consistent with those reported in the paper.

7) The merger of low vowels is a well-documented phenomenon (see for instance the ongoing merger of front /a/ and back /ɑ/ in Parisian French, Hansen & Juillard 2010).

statistically meaningful predictor that categorized subject groups. Both F1 and F2 were lower in male speech regardless of dialect of speakers or either of the two vowels spoken. The outcome appears to be reasonable considering naturally different shapes of articulatory and/or phonatory organs. Therefore, it can be stated that young native speakers of Korean do not make any phonetic differentiation among mid front vowels when they produce them. A caveat, however, is that the conclusion on the lack of DIALECT effect should be restricted to the three dialects examined in this paper without extended generalization across other major dialects of Korean (see Yeon, 2012: 168 for dialectal zones).

The results of the perception test did not consistently mirror the results of production. Above all, no predictor made any significant difference in F2. On the other hand, it was observed that F1 was meaningfully differentiated according to an interaction of predictors. Most prominently, the Seoul dialect is different from the others in that female subjects perceived the two vowels differently in terms of F1. In other words, they tended to perceive vowels with a lower F1 as /e/ rather than /ɛ/. It is hard to jump to a conclusion that the distinction of the two vowels based on the 'tongue height' parameter is systematic and sustainable in Seoul speakers' perception: we nevertheless cautiously observe that our results seem to be consistent with established results about gender differentiation in sociolinguistic research. As has been demonstrated by Labov, women generally show a higher rate of prestige variants associated with high social class or formal speech than do men (Labov, 2001: 266-7). The better distinction by Seoul female, who speak the standard variety, could reflect the fact that they have a higher sensitivity to this normative contrast. However, further refined experimentation including a diachronic approach with longitudinal settings and a much larger sample seems to be required for clarification[8].

The inconsistent results between production and perception bring up two related issues to consider: speaker normalization and sound change. First, the fact that such a distinct GENDER effect across all dialects in vowel production almost disappeared in perception supports the theory of listeners' capability of correcting what they hear in the context of speakers of different gender groups. Various linguistic and social strategies of talker normalization (see Johnson, 2005 for review) are likely to be employed by listeners at the perceptual level to eliminate acoustical differences in formant values of vowels between male and female voices.

Second, considering that Seoul female subjects perceptually distinguish between /e/ and /ɛ/ but are unable to produce the same vowels differently, they may have experienced hearing the two vowels as distinct sound units, produced by older speakers. If this is true, it can be hypothesized that the contrast between these two vowels is the manifestation of a change-in-progress, in such a way that it is being lost by the younger generation. It is also possible that Seoul females' better discrimination of these vowels is due, at least in part, to a higher sensitivity to this distinction, as conveyed by the media or the schooling system[9], for example. Solely based on the current phonetic experiments, it is however premature to draw any conclusion on the general status of Korean mid front vowels. Nevertheless, our results at least confirm the tendency that /e/ and /ɛ/, which are assumed to have been distinct at some point in the past, are not produced differently from each other any longer, at least by young speakers of Korean.

The last study reported on in this paper helps shed some light on the possible causes of this merger. There is a consensus among Korean linguists on diachronic changes of Korean vowels (Huh, 1952; Ahn & Iverson, 2007): It has been observed that the characters corresponding to the vowels /e/ and /ɛ/ were not even included in the list of basic vowels when Hangeul was first promulgated by King Sejong in 1446. Indeed, Late Middle Korean (15^{th} century) is reported to lack mid front vowels entirely, /i/ being the only front vowel (Ahn, 2002: 238; Ahn & Iverson, 2007). Evidence reveals that the two vowels had maintained their status as off-glide diphthongs, /ʌj/ and /aj/, respectively, until the late 17^{th} century when their monophthongized characters started to appear in literary works. Apart from an explanation based on the phonetic similarity of these vowels[10], the low functional load of the /e/ vs /ɛ/ contrast

8) An anonymous reviewer suggested that the Seoul females' better discrimination might be influenced by their better proficiency in English where /ɛ/ and /æ/ are contrastive. Unfortunately, since we have not documented our subjects' English proficiency in a systematic way, we are currently unable to shed light on this issue, but it is an interesting hypothesis that deserves to be investigated in future research.

9) To take but one example, some Korean high-school textbooks draw students' attention to minimal pairs such as 개 vs 게, 내 vs 네, insisting that they should be pronounced "accurately".

10) In the pronunciation of some speakers of Australian English, there is a case of merging /e/ and /æ/, so that the words *hell* and *hal* become homophonous (Cox & Palethorpe, 2007: 346).

appears to be a factor that may have facilitated, if not triggered the merger currently observed in contemporary Korean. If this analysis is correct, the dynamics of this change suggest that this merger is the outcome of a (long) process of stabilization in the mid front region of the vowel space, driven by a combination of articulatory, perceptual and functional factors.

Finally, although the variable AGE was not rigorously treated in this paper, it is worth re-considering. In fact, the age effect was assumed to have been controlled as all subjects of the experiments were 36 or younger. However, it is suspected that this restriction is not strict enough for full control of the effects. Thus, a supplementary statistical test was conducted to check if there is any sign of an age effect within subjects. When the production data was analyzed, no significant age effect was revealed for either F1 ($\chi^2(1)$ = 0.171, p = 0.680) or F2 ($\chi^2(1)$ = 3.717, p = 0.053). Consequently, there was no ground to argue for any systematic relation between subjects' age and mid front vowel production. With the perception data, it was found that F1 was not affected by age ($\chi^2(1)$ = 0.063, p = 0.802). Strikingly, however, F2 was found to vary upon age according to linear mixed-effects model fitting followed by maximum likelihood comparison ($\chi^2(1)$ = 11.99, p<0.001). As the data used in this study was not originally designed and established to analyze age effects, the age groups are not evenly distributed. Consequently, the current result does not necessarily imply that listeners' age influences Korean mid front vowel perception. Further investigation with data refined for age groups is desirable for more reliable inference. More generally, while the research reported in this paper offers new clues about the status of the merger in Korean, much work remains to be done in order to better understand the synchrony (across generations) and recent diachrony of the merger between /e/ and /ɛ/.

Acknowledgements

We feel especially grateful to Sulki Kim for her assistance in collecting and arranging data.

References

Ahn, S. & Iverson, G. K. (2007). Structured imbalances in the emergence of Korean vowel system. In Joseph C. Salmons & Shannon Dubenion-Smith (eds.), *Historical Linguistics 2005* (*Current Issues in Linguistic Theory,* Vol. 284), 275-293, Amsterdam & Philadelphia: John Benjamins Publishing Co.

Ahn, S. (2002). A dispersion account on Middle Korean Vowel Shifts, *Japanese/Korean Linguistics*. Vol. 10, 237-250.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items, *Journal of Memory and Language*. Vol. 59, 390-412.

Bae, J. (1996). *Introduction to Korean Phonology*. Seoul: Shingu Munhwasa. (in Korean)

Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program]. Version 5.3.66, http://www.praat.org.

Chládková, K. & Escudero, P. (2012). Comparing vowel perception and production in Spanish and Portuguese: European versus Latin American dialects, *Journal of the Acoustical Society of America*. Vol. 131, No. 2, 119-125.

Cox, F. & Palethorpe, S. (2007). Australian English, *Journal of International Phonetic Association*. Vol. 37, No. 3, 341-350.

Glasberg, B. R. & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data, *Hearing Research*. Vol. 47, 103-138.

Hansen, A. B. & Juillard, C. (2011). La phonologie parisienne à trente ans d'intervalle – Les voyelles à double timbre, J*ournal of French Language Studies*. Vol. 21, No. 3, 313-359.

Hockett, C. F. (1955). *A Manual of Phonology*. Baltimore: Waverly Press.

Huh, W. (1952). ay, ey, oy, ɔ uy umka ("on the sound quality of ay, ey, oy, o, ɔ"), *Kukeo Kukmunhak*, Vol. 1, 5-8. (in Korean)

Huh, W. (1991). *Korean Phonology*. Seoul: Sam Munhwasa. (in Korean).

Hwang, H. & Moon, S. (2005). An acoustic comparative study of Korean /에, 애/ and English /ɛ, æ/ pronounced by Korean young male speakers, *Malsori*. Vol. 56, 29-47. (in Korean)

Johnson, K. (2005). Speaker normalization in speech perception. In D. B. Pisoni & R. Remez (eds.), *The Handbook of Speech Perception*, 363-389, Oxford: Blackwell Publishers.

Kim, H. (2006). Korean national corpus in the 21st century Sejong project. In *Proceedings of the13th National Institute of Japanese Literature (NIJL) International Symposium*, 49-54.

Kim, M. (1992). *Korean Phonology*. Seoul: Hanshin Publishing Co. (in Korean)

Kriengwatana, B., Escudero, P. & Terry, J. (2014). Listeners cope with speaker and accent variation differently: Evidence from the Go/No-go task. In *Proceedings of the 15th Australasian International Conference on Speech Science and Technology*, 2-5 December 2014, Christchurch, 76-79.

Labov, W. (2001). *Principles of Linguistic Change. Volume 2:*

Social Factors. Malden: Blackwell.

Lee, H. (1996). *Korean Phonetics*. Seoul: Taehaksa. (in Korean)

Mannell, R. (2004). Perceptual vowel space for Australian English lax vowels: 1988 and 2004. In *Proceedings of the 10th Australian International Conference on Speech Science & Technology*, Sydney, 221-226.

Martinet, A. (2005 [1955]). *Economie des changements phonétiques. Traité de phonologie diachronique.* Paris: Maisonneuve & Larose.

Moon, S. (2007). A fundamental phonetic investigation of Korean Monophongs, *Journal of the Korean Society of Phonetic Sciences and Speech Technology*. Vol. 62, 1-17.

Moore, B. C. J. (2013). Aspects of auditory processing related to speech perception, In W. J. Hardcastle, J. Laver & F. E. Gibbon (eds), *The Handbook of Phonetic Sciences,* 2nd ed., West Sussex: Wiley-Blackwell. 454-488.

Oh, Y., Pellegrino, F., Coupé, C. & Marsico, E. (2013). Cross-language comparison of functional load for vowels, consonants, and tones. In *Proceedings of Interspeech 2013*, Lyon, France, 25-29.

Pépiot, E. (2013). Voice, speech and gender: male-female acoustic differences and cross-language variation in English and French speakers. XVèmes Rencontres Jeunes Chercheurs de l'ED 268, Jun 2012, Paris, France.

Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS.* New York: Springer.

Rosner, B. S. and Pickering, J. B. (1994). *Vowel Perception and Production*. Oxford: Oxford University Press.

Seong, C. (2004). An acoustic analysis on the Korean 8 monophthongs with respect to the acoustic variables on the F1/F2 vowel space, *Journal of the Acoustical Society of Korea*. Vol. 23, No. 6, 454-461.

Shin, J. (2000). *Understanding Spoken Language*. Seoul: Hankuk Munhwasa. (in Korean)

Shin, J., Kiaer, J. & Cha, J. (2013). *The Sounds of Korean*. Cambridge: Cambridge University Press.

Sohn, H. (1999). *The Korean Language.* Cambridge: Cambridge University Press. (in Korean)

Surendran, D., & Niyogi, P. (2006). Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In O. N. Thomsen (ed.), *Competing Models of Linguistic Change: Evolution and Beyond*. Amsterdam and Philadelphia: John Benjamins, 43-58.

Yang, B. (1992). An acoustical study of Korean monophthongs produced by male and female speakers, *Journal of Acoustical Society of America.* Vol. 91, No. 4, 2280-2283.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers, *Journal of Phonetics.* Vol. 24, 245-261.

Yeon, J. (2012). Korean dialects: a general survey. In N. Tranter (ed.), *The Languages of Japan and Korea*. Abingdon and New York: Routledge, 168-185.

• **Eychenne, Julien**
Dept. of Linguistics and Cognitive Science
Hankuk University of Foreign Studies
Mohyeon, Yongin, Gyeonggi 449-791, Korea
Tel: 031-330-4998
Email: jeychenne@hufs.ac.kr

• **Jang, Tae Yeoub, Corresponding Author**
Dept. of English Linguistics
Hankuk University of Foreign Studies
107 Imun-ro, Dongdaemun-gu, Seoul 130-791, Korea
Tel: 02-2173-3119
Email: tae@hufs.ac.kr