Frequency Effects on Vowel Length Contrast Merger in Seoul Korean

Yoonjung Kang, Tae-Jin Yoon, and Sungwoo Han

December 15, 2014

Abstract

This paper presents an apparent time study of vowel length contrast merger in Seoul Korean based on the duration measurements of over 370,000 vowels in word-initial syllables in a read speech corpus. The effects of word frequency on vowel duration and the lexical diffusion of long vowel shortening are also examined. The findings confirm the observation of previous literature that vowel length contrast is on its way out in the language and that this sound change is nearing completion. We also find a significant effect of frequency on long vowel duration: other things being equal, these vowels are shorter in high frequency words than in low frequency words. The rate of change did not differ significantly depending on the frequency of words apart from the high frequency words reaching the endpoint of change and bottoming out in the change earlier than mid and low frequency words. We do not find any general resistance of high frequency words to this sound change and therefore the finding does not provide any evidence for an analogical mechanism of sound change. The observed frequency effect is compatible with a model where the frequency effect on duration comes from on-line factors that affect phonetic implementation of speech sounds along with an across-the-board lenition bias that drives the sound change, not through stored tokens of word-specific variants.

1. Introduction

This study examines vowel length contrast in Seoul Korean. Seoul Korean is generally described as having contrastively long vowels, but many descriptions of the language also note that the length contrast is being lost or is already lost in younger speakers’ speech. The first goal of the study is to provide an empirical contribution to this topic by examining the vowel length contrast in The Reading-Style Speech Corpus of Standard Korean (The National Institute of the Korean
Language 2005), a speech corpus that contains read speech data from 118 speakers of Seoul Korean stratified for gender and age. Specifically, the study aims to examine change in long vowel realization across different age groups, which we interpret as a reflection of sound change in real time (Bailey et al. 1991).

The second goal of the study is to examine the effect of word frequency on this sound change, i.e., whether and how this sound change affects words of high and low frequency differently. Studies on lexical diffusion show that sound change may affect high or low frequency words differently depending on the nature of sound change (Bybee 2001; Phillips 2006). By examining the frequency effect, we can provide indirect evidence about the underlying mechanism of this sound change—whether the change is a phonetically gradual reduction or a structurally motivated phonological shortening—and also contribute to the general literature on frequency effects on sound change. It is also notable that most previous studies of frequency effects on lexical diffusion of sound change examine the final outcome of a sound change or a frequency effect on synchronic variation at a static point in time rather than track the effect of frequency on change over time.ii Our study contributes to filling this gap by examining the interaction of frequency with sound change as the sound change unfolds in (apparent) time.

2. Word frequency and lexical diffusion

Studies on lexical diffusion of sound change observe that changes that are physiologically motivated and phonetically gradient tend to affect high frequency words first while changes that are analogical in nature and phonetically abrupt tend to affect low frequency words first (Bybee 2001; Bybee 2002; Bybee & Hopper 2001; Hooper 1976; Phillips 1984; Pierrehumbert 2001; Schuchardt 1885/1972).iii An often cited example of phonetic sound change that affects high frequency words first is t/d deletion in English; other things being equal, t/d deletion is more frequent in high frequency words than in low frequency words in American English (Bybee 2002; Coetzee & Kawahara 2013; Jurafsky et al. 2001).iv Regularization of English past tense is an example of change that affects low frequency words first; high frequency words like *slept, left,* and *kept* retain their irregular past tense form while low frequency words like *wept, leapt,* and *crept* may regularize (Hooper 1976).
Such frequency effects on sound change have been put forth as supporting evidence for the *usage-based* model of phonology (Bybee 2001; Bybee & Hopper 2001) and the *exemplar-based* model of phonology (Kirchner 2012; Pierrehumbert 2001; Pierrehumbert 2002), as these models assume a direct link between the use of individual word forms and their representations. The resistance of high frequency words to a regularizing change follows from the assumption that high frequency words form a stronger and more independent mental representation and therefore can resist changes motivated by analogy to other forms. The propensity of high frequency words to undergo a reductive sound change follows from the assumption that the lexical representation of a word include phonetically detailed exemplars and this representation is constantly updated with each use of the word. In a leniting sound change, the more frequently a word is used, the more exemplars of lenited tokens accrue, which forms the basis for subsequent productions of the word. As a result, more frequent words are expected to undergo a reductive sound change at a faster rate than less frequent words.

But, it is worth noting that frequency-dependent reduction is found as a synchronic process even in the absence of sound change. Studies have shown that high frequency words are produced with a shorter duration than low frequency words. While some interpret such frequency effect as evidence for phonetically rich word-specific representation assumed by exemplar-based models (Gahl 2008), others note that such effect can be attributed to the higher resting activation level of frequent words which gives rise to faster lexical access and articulatory planning without these words necessarily having different phonetic representations (Bell et al. 2009; Pluymaekers et al. 2005). In other words, the fact that frequent words tend to be produced with a higher degree of reduction at a static point in time alone does not support the word-specific phonetically detailed lexical representation as such effects can come from on-line factors that affect phonetic implementation of speech sounds (Ernestus 2014; Pierrehumbert 2002).

By the same reasoning, the fact that high frequency words are further along in a diachronic sound change by itself does not necessarily support word-specific phonetic representation either. A general lenition bias which moves all words toward more lenited realization over time, along with synchronic on-line factors that promote further lenition of high frequency words can create...
the appearance of high frequency words being preferentially targeted by reductive sound change. Such a scenario is schematically represented in Figure 1 (a). In the graphs in Figure 1, the x-axis represents time and the y-axis, “vowel duration”, represents a phonetic dimension along which reductive sound change progresses, with a lower value representing a further change. The horizontal line indicates a hypothetical category boundary where the phonological label changes. In Figure 1 (a), high and low frequency words are distributed differently along the phonetic dimension with high frequency words more reduced than low frequency words, an effect attributable to synchronic on-line factors, but high and low frequency words undergo the change at the same rate. When the change has progressed far enough, we may encounter a state where only the high frequency words have advanced in the change far enough to be relabeled or reanalyzed as a different phonological category. It should be noted that the pattern of change in Figure 1 (a) is also consistent with a hybrid exemplar model (Pierrehumbert 2006) where a layer of phonological categories acts to keep the word-specific variation from running rampant. In this model, words can have different phonetic representations but word-specific variation is kept in check and the top down pressure of shared phonological category keeps a sound in high frequency words from leniting exponentially away from the rest of the category.

Figure 1. Schematic representations of interaction between synchronic frequency effect and diachronic frequency effect
The prediction of the exemplar-based model, in its strongest form (Pierrehumbert 2001; Pierrehumbert 2002), on the other hand, can be schematically represented as in Figure 1 (b); high frequency words are overall more reduced but they also undergo the change at a faster rate because each production adds reduced exemplars further precipitating the reductive change. In Figure 1 (a), the frequency effect at the outcome of sound change is a reflection of synchronic frequency effect (Diachrony=Synchrony) while in Figure 1 (b), the synchronic frequency effect is expanded over time and the diachronic frequency effect is expanded compared to the synchronic frequency effect (Diachrony=Synchrony²). So, a study of lexical diffusion in sound change may provide crucial disambiguating evidence as to the status of phonetically rich lexical representation. However, most studies of frequency effects on lexical diffusion of sound change examine the final outcome of a sound change or the frequency effect at a static point in time rather than track the effect of frequency on change over time, and these models are under-differentiated by available evidence.

The first two scenarios sketched in Figure 1 illustrate changes which progress in the same direction as physiologically conditioned lenition and the change affects high frequency words first. However, a distinction between a phonetically motivated reductive sound change and an analogically motivated sound change is not always straightforward and the purported correlation between types of sound change and directions of frequency effect is not always borne out. For example, Phillips (2001) found that a stress shift in English verbs with -ate suffix as in frustrate \( \rightarrow \) frustráte is more prone to affect high frequency words, although this is not a sound change that is phonetically gradient or reductive. Phillips (2001) attributes this frequency effect to the fact that high frequency words are more likely to be analyzed as monomorphemic, rather than as having a stem + suffix structure, allowing the regular stress assignment rule of English to apply. The accent shift in Ancient Greek discussed by Probert (2006) illustrates a more complex interaction of morphological decompositionality and frequency in accentual shift (Probert 2006). Probert (2005) observes that in a group of nouns formed with an adjective-forming affix in Ancient Greek, the accent shifted from a final accent (expected based on the original adjectival affixes) to a recessive accent (a general default accent pattern). As expected, very high frequency words are resistant to this regularizing accent shift and retain their original final accent. Somewhat unexpectedly, very low frequency words also resist this change. Probert
(2006) suggests that very low frequency words are less prone to “demorphologization” and more likely to be analyzed as containing the adjectival suffix and hence more likely to follow the stress pattern expected based on the adjectival suffix. In other words, these two examples of accent shift show that the effect of frequency on morphological compositionality can create low-frequency-last effect even for sound changes that are not physiologically conditioned.

Phillips (2001) also discusses a number of “weakening” sound changes that somewhat unexpectedly affect low frequency words first. For example, the unrounding of /ø(ː)/ to /e(ː)/ in early Middle English affects low frequency words more than high frequency words and also exhibits sensitivity to grammatical category of words. Phillips (2001) analyzes this change as “typologically motivated”; languages without high front rounded vowels tend not to have mid front rounded vowels, and this change affected those dialects of Middle English that recently lost high front rounded vowels. In other words, the change is not a physiologically motivated lenition but a change due to the pressure of the phonological system, i.e., a removal of highly marked structure.

This and other cases of frequency effects led Phillips to propose a refined hypothesis regarding the connection between frequency and sound change that “[s]ound changes which require analysis—whether syntactic, morphological, or phonological—during their implementation affect the least frequent words first (p. 123)”, while “changes which ignore the phonological integrity of segments and the morphological composition of words affect the most frequent words first (p.134).” So, a sound change that may look like a reductive sound change on the surface may in fact affect infrequent words first depending on whether or not the underlying mechanism of the change is analytical.

Assibilation of noun-final coronal plosives in Korean, which turns root-final /tʰ c cʰ/ to /s/ in nouns, is an example where a seemingly phonetically motivated process (Kim 2001) is more properly analyzed as an analogy to a dominant noun alternation pattern (Ito 2010; Ko 1989; Kwak 1984). As expected for an analogically motivated change, high frequency words resist this change (Kang 2003; Kang 2007). Devoicing of stem-final /v/ and /z/ in Dutch presents a case where the process is reductive and gradient at the phonetic level but also shows an analogical
frequency effect; low frequency words are more likely to show devoicing and hypercorrective voicing is also attested, another sign that the change is not purely phonetic (De Schryver et al. 2008).vi

This third possibility of frequency effect on reductive sound change is schematically represented in Figure 1 (c). In this scenario, the sound change involves a category label change, which happens to align with a general direction of synchronic phonetic lenition. So the sound change may look like a phonetic change but the underlying mechanism is in fact analytical and affects the sounds at their category label, not at the phonetic level. As a result, frequent words, despite their generally more lenited phonetic realization within the original category, are more resistant to the change motivated by analogy or markedness pressure from the phonological system.viii

With this background, we now turn to the case of the loss of long vowels in Seoul Korean. The questions we are posing are; 1) is this change actually happening?; 2) are high frequency words more reduced and produced with a shorter vowel duration?; 3) do high and low frequency words undergo change differently and if so how? In the absence of any sound change, due to the synchronic frequency effect on duration (Bell et al. 2009; Pluymaekers et al. 2005), we expect long vowels to be realized with a shorter duration when they occur in high frequency words than in low frequency words. If the sound change is a phonetically gradual reduction, we expect the frequency effects illustrated in Figure 1 (b), while we expect the frequency effect illustrate in Figure 1 (c) if the sound change has an analytical underpinning. Yet another possibility is that we do not find any frequency effect and the sound change affects all words equally as illustrated in Figure 1 (a), modulo the synchronic effect on duration.ix

3. Vowel length contrast in Seoul Korean

In this section, we provide a review of previous studies on Seoul Korean vowel length. Seoul Korean is generally described as having vowel length contrast (Chae 1959; Hur 1960; Lee 1956; Lee 1993; Lee & Ramsey 2011; Lee 1960; Martin 1992) but many also state that the contrast is being lost in younger speakers’ speech and such statements are attested as early as 1951 (Lee 1960; Martin 1951; Martin 1992; Park 1994; Shim et al. 2013; Sohn 1999).x The minimal pairs
in (1) illustrate the contrast.

(1) [nun] ‘eye’ [nu:n] ‘snow’
[pe] ‘pear’ [pe:] ‘double’
[pam] ‘night’ [pa:m] ‘chestnut’
[il] ‘one’ [i:l] ‘work’
[tsʰak-ta] ‘to write’ [tsʰ:k-ta] ‘not plenty’
[cʰan] ‘window’ [cʰa:n] ‘song’

Studies suggest that the long vowels are first weakened and lost in non-initial position and the change further spread to the initial position (Cha 2012; Lee 1960) and in present day Korean, long vowels tend to be limited to the word-initial syllable while underlying long vowels in non-initial position shorten. This alternation is illustrated in (2) (Lee & Ramsey 2011; Lee 1960).

(2) [sa:lam] ‘person’ [nu:n-sal-am] ‘snowman’
[pa:m] ‘chestnut’ [ku:n-pam] ‘roasted chestnut’
[nu:n] ‘snow’ [tsʰan-nun] ‘first snowfall’
[pa:llita] ‘to spread open’ [t’-pallita] ‘to brag’

Underlying long vowels in word-initial position may also shorten in particular morphophonological contexts. In most verbs and adjectives, the underlying long vowels shorten when a vowel-initial suffix is attached as shown in (3a) or when a passive or causative suffix is attached as shown in (3b).

(3) a. [ku:m-t’a] (<ku:lm-ta/) [kulm-’a] ‘to starve’
[na:-tʰa] (<nah-ta/) [na:-ini] (<nahini/) ‘to insert’
[ki:l-ta] [kil-’a] ‘long’
[pu:-t’a] (<pu:s-ta/) [pu-’i] ‘to pour’ (s-irregular)

b. [ka:m-t’a] [ka:m-ki-ta] ‘to be coiled’ (PASSIVE)
[pu:-t’a] (<pu:t-ta/) [pu:l-li-ta] (<put-li-ta/) ‘to soak’ (CAUSATIVE)
[k’o:-ta] [k’o-i-ta] ‘to be entangled’ (PASSIVE)
A study of recordings from the 1930s shows that vowel length contrast is robustly attested in Seoul speakers’ speech from this time period (Cha 2005) but more recent studies are in general agreement that the contrast is being lost in contemporary Seoul Korean. What remains unclear is whether this sound change is a phonetically gradual reduction of long vowels or is an analogical change where the underlying long vowel is reanalyzed/misanalysed as an underlying short vowel, or both.

In the earliest instrumental study we are aware of, Han (1964) examined the vowel duration of 25 minimal pairs with long and short vowel contrast as produced by four Seoul Korean speakers in their 20s and 30s and found a substantial average durational ratio of 2.51 between long and short vowels in citation forms of the words. But, this contrast already exhibited a sign of erosion at the time of Han’s study. Han (1964) observes that there is a group of long vowel words that only some of her informants produced as long. This suggests that some long vowel words were consistently produced with long vowels while other words were not. Han (1964) goes on to speculate that “this is largely due to the lack of distinctive notation in the Korean orthography. […] When dealing with less common words or learned words, only a limited number of people may distinguish them in their conscious speech.” What this statement suggests is that the phonetic contrast of long and short vowels is robust and the change is one of misanalysis of long vowels as short vowels which affects less frequent words first. However, as no production data is provided for these ambiguous vowels, we cannot tell whether these vowels show a phonetic durational reduction or a categorical reanalysis to the short vowel category.

More recent quantitative studies similarly find lexical and individual variation in vowel length contrast realization. Zhi et al. (1990) examined three minimal pairs and three speakers (all males, 19-26 years in age) and found that two of their three speakers produced the contrast consistently with a duration ratio ranging from 1.45 to 2.01 while one speaker did not produce the contrast for two of the pairs and produced a contrast in the reverse direction for one of the pairs. Kahng (1995) found that two of the three speakers (all male and in their 20s) produced a consistent vowel length duration contrast for all or most of the 16 minimal pairs examined, but one speaker showed the correct contrast only for a small subset.¹¹ Kong & Moon (2002), on the other hand,
found that all three of their speakers (one male each in their 20s, 30s, and 40s) produced a substantial and consistent length contrast (a ratio of 1.5 to 2) for the seven minimal pairs examined. What these studies, together, suggest is that while the vowel length contrast is not completely lost in younger speakers’ speech, there is speaker variation (some speakers retain the contrast while others do not) and word-specific effects (the length contrast is retained in some words but not in others). The duration ratio between long and short vowels in more recent studies tend to be smaller, at around 1.5 to 2, than the ratio of 2.51 found in Han (1964)’s study. But, given the difference in the methodology across these studies—Han’s study measured words in isolation while the other studies measured words embedded in a sentence—a direct comparison may not be meaningful.

Studies that directly compare the vowel length production across different age groups also present a similar picture of lexical and speaker variation and find an age-dependent trend of long vowel loss or reduction (Jung & Hwang 2000; Kim 2003; Park 1985). Park (1985) examined the realization of vowel length contrast by 30 Seoul Korean speakers stratified for age. The study examined the speakers’ declarative knowledge of vowel length in 277 commonly used words and their production patterns. Table 1 summarizes the key findings.

<table>
<thead>
<tr>
<th>Speaker’s age</th>
<th>≥ 60s</th>
<th>50s</th>
<th>40s</th>
<th>30s</th>
<th>20s</th>
<th>10s</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Proportion of words produced with a long vowel</td>
<td>62.00</td>
<td>62.75</td>
<td>62.50</td>
<td>49.75</td>
<td>30.75</td>
<td>16.75</td>
</tr>
<tr>
<td>(b) Proportion of words produced with a long vowel out of those the speaker judged to contain a long vowel</td>
<td>94.92</td>
<td>93.29</td>
<td>74.63</td>
<td>69.64</td>
<td>50.39</td>
<td>29.12</td>
</tr>
</tbody>
</table>

In the study, long vowels are defined as those produced 1.5 to 2 times longer than short vowels in an identical context. Based on this criterion, the overall proportion of words produced with a long vowel declines in younger speakers’ speech. While the speakers in their 40s or older produce long vowels for over 60% of the words examined, the rate is lower for younger speakers (Table 1, (a)). Younger speakers also show a large discrepancy between their judgment of vowel length and their actual production; the proportion of long vowel production for words the speakers themselves judged to contain a long vowel is over 90% for speakers in their 50s and 60s
but the rate drops to below 30% for speakers in their teens (Table 1, (b)). Park (1985) interprets this discrepancy between production and judgment as an indication that younger speakers are making a random choice when judging the vowel length of a given word. But, another possibility is that their exposure to older speakers’ speech provides them with implicit knowledge of vowel length pattern even though their own production grammar does not make a consistent-enough distinction to pass the 1.5 ratio test used in the test. Park (1985) also notes that among the vowels that are produced and classified as long, there is a generational difference in the ratio of long vs. short vowels; the ratio is around 2.0 for older speakers but between 1.5 (the lower limit possible by definition) and 1.7 for younger speakers. Park (1985)’s result suggests that the vowel length contrast loss is phonetically gradual but also lexically diffused; in younger speakers’ speech, vowel length contrast is still retained but their long vowels are shorter in duration and the contrast is more robustly preserved in some words than others.

Jung & Hwang (2000) examined the production of 24 minimal pairs (all Sino-Korean disyllabic words) by 12 Seoul Korean speakers stratified for gender and age (from 20s to 70s) and found an effect of both factors. When the duration was averaged across all word pairs, the older speakers showed a statistically significant duration difference between long and short vowels while younger speakers did not. For middle-aged speakers, males patterned with the older speakers while females patterned with the younger speakers. After examining the minimal pair-specific durational contrast for each speaker group, Jung & Hwang (2000) reach the conclusion that the change is one of lexical diffusion where the set of lexical items with long vowels become smaller and smaller in younger speakers’ speech, rather than a phonetically gradual long vowel shortening that affects all long vowels.

A survey by the National Institute of the Korean Language examined the production of 29 long vowel words and 12 short vowel words by 350 speakers of Seoul Metropolitan area, stratified for age, gender, and level of education (Kim 2003). The production was recorded and transcribed but the criteria for determining the vowel length of a particular production are not provided, making it hard to interpret the findings. Figure 2 summarizes the data reported in (Kim 2003). The plot shows the proportion of long vowel realization, aggregated over all words and speakers.
The results show that there is a positive correlation between the age of speakers and the overall percentage of long vowel realization, i.e., older speakers produced more long vowels, as one may expect from the general trend of long vowel reduction in younger speakers. But, interestingly this trend is true of both underlying long and short vowels. In other words, older speakers not only produced more long vowels for underlying long vowels than younger speakers but they also did so for underlying short vowels as well. Kim (2003) interprets this as evidence that even older speakers are losing the contrast, not only mispronouncing long vowels as short but also pronouncing short vowels as long. However, as no information is provided as to the specific criteria for determining the long vs. short vowel realization, we need to be cautious in interpreting the results. One alternative possibility is that the older speakers tend to have a slower speech rate and may have produced all vowels generally longer than younger speakers and this may have affected the categorization of the vowels by the fieldworkers, who were all young Seoul Korean speakers. Regardless of how we may interpret a substantial percentage of long vowel realization of underlying short vowels by older speakers, we do find that the older speakers (50s and 60s)’s production shows a statistically significant difference between long and short vowel realizations, while younger speakers (20s, 30s, and 40s) do not.\textsuperscript{xv}
To summarize, previous studies are generally consistent with the view that younger speakers produce more reduced vowel length contrast. There are suggestions that the change is phonetically gradual and also suggestions of lexical diffusion, with low frequency words particularly susceptible to the change. Also, some evidence for hypercorrective long vowels is found, which is a sign that the sound change is not purely phonetic or that if it were a phonetic sound change, the merger is nearing the end.

4. Data

The data for this study come from The Reading-Style Speech Corpus of Standard Korean (The National Institute of the Korean Language 2005), which contains read speech of 60 male and 60 female speakers of Korean residing in the Seoul metropolitan area. The age of the speakers ranged from 19 to 71 at the time of recording in 2003. For two of the speakers, the sound files were missing or had errors and our analysis is based on data from 118 speakers. The distribution of the speakers by gender and year of birth, inferred from the age and the year of recording, is given in Table 2.

Table 2. Distribution of speakers in the NIKL Corpus by gender and the decade of birth (counts in parentheses include the two speakers with the missing files)

<table>
<thead>
<tr>
<th></th>
<th>1930s</th>
<th>1940s</th>
<th>1950s</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>4</td>
<td>12</td>
<td>4</td>
<td>8</td>
<td>26 (27)</td>
<td>5</td>
<td>60</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
<td>9 (10)</td>
<td>25</td>
<td>3</td>
<td>11</td>
<td>9</td>
<td>60</td>
</tr>
</tbody>
</table>

The speech material consists of well-known short stories and essays, totaling 930 sentences. Of these, 404 sentences were read by all speakers while the rest are read by younger speakers only. More information about the corpus is available in (Kang To appear; Yoon To appear; Yoon & Kang 2012; Yoon & Kang 2014). As the vowel length contrast is mainly retained in the word-initial syllable only, the current study only examined the vowels in the word-initial syllable in the 404 sentences read by all speakers. The analysis only includes monophthongal vowels and vowels preceded by an on-glide or an off-glide (/j/ or /w/) are excluded. In the 404 sentences, a total of 1,260 lemma types, 2,252 wordform types and a total of 3,368 wordform tokens contain a monophthongal vowel in word-initial syllable. Of those 3,368 vowel tokens, 563 (=16.7 %) are
long and the rest are short vowels. The vowel length specification is based on the vowel length marking in *the Great Dictionary of the Korean Language*.\textsuperscript{xvi}

Determining whether a particular speaker’s particular vowel token is realized as phonologically long or short is a difficult problem.\textsuperscript{xvii} Duration of a segment is not only affected by its phonological length but also by a number of contextual factors, such as vowel height, syllable structure, preceding or following segments, word length, position in a prosodic phrase, and speech rate, among others (Crystal & House 1988; Klatt 1976; Lehiste 1970; Maddieson 1985; Turk & Shattuck-Hufnagel 2000; Turk & Shattuck-Hufnagel 2007; Yoon To appear; Yun 2009; Zhi 1993). As a result, there is no absolute value of duration or any other acoustic measure that can determine a vowel as long or short. But, we also expect that if vowel length is contrastive in a speaker’s speech, other things being equal, long vowels should show a longer duration than short vowels. So, in this study, we will use the phonetic duration of a vowel as our measure and examine how the effect of phonological vowel length on phonetic duration changes across speakers of different age group while controlling for other factors that are known to affect segment duration.\textsuperscript{xviii}

The duration measurements were extracted using the forced alignment system for Korean developed by Tae-Jin Yoon (Yoon To appear; Yoon & Kang 2012; Yoon & Kang 2014).\textsuperscript{xix} Files that contained gross errors in alignment, due to incorrect file-text matching in the original corpus or disfluency or reading error by the speakers are discarded. The automatic aligner analyzes the signal in 10 ms frames and assigns each frame a segment label; therefore, the resolution for duration measurements is 10 ms. In the analyses provided below, only vowels with a duration of less than 300 ms were included. Manual inspection of those tokens indicate that such extreme values are often due to alignment errors or are mostly limited to interjections or vowels in utterance or phrase final position. Also excluded are tokens for which no pitch is detected as such voiceless tokens also tend to involve alignment errors as well as completely devoiced and extremely short vowels. After this elimination process, we ended up with a total of 373,733 tokens of vowels in word-initial syllable produced across 118 speakers. In our corpus data, and also likely in the language in general, underlying long vowels are in the minority, comprising around 17.0 % of monophthongal vowels in word-initial syllables. The breakdown of the tokens
by the vowel quality and the underlying vowel length based on the *Great Dictionary of the Korean Language* is provided in Table 3.

<table>
<thead>
<tr>
<th>Vowel Quality</th>
<th>ɑ (ㅏ)</th>
<th>e (ㅐ)</th>
<th>ʌ (ㅓ)</th>
<th>o (ㅗ)</th>
<th>i (ㅣ)</th>
<th>u (ㅜ)</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long</td>
<td>13.2%</td>
<td>36.7%</td>
<td>38.0%</td>
<td>16.8%</td>
<td>24.4%</td>
<td>13.6%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Short</td>
<td>86.8%</td>
<td>63.3%</td>
<td>62.0%</td>
<td>83.2%</td>
<td>75.6%</td>
<td>86.4%</td>
<td>94.7%</td>
</tr>
<tr>
<td>Total</td>
<td>121,210</td>
<td>25,307</td>
<td>4,242</td>
<td>53,634</td>
<td>45,706</td>
<td>25,167</td>
<td>36,946</td>
</tr>
</tbody>
</table>

The frequency counts in our analysis are based on the lemma frequency list published by the National Institute of the Korean Language (조남호 2002).xx This frequency list is based on a corpus of over 1.5 millions words and homophones are disambiguated.xxi For the analysis below, we converted frequency into a categorical variable with three levels (low, mid, and high). The categories are determined by *k*-means clustering of log frequency counts of 3,368 wordforms in the read text.xxii The frequency range of the three categories and the number of tokens for each category is summarized in Table 4.

<table>
<thead>
<tr>
<th>Frequency Category</th>
<th>Low</th>
<th>Mid</th>
<th>High</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency range</td>
<td>0~67</td>
<td>68~1,281</td>
<td>1,330~25,567</td>
<td>0~25,567</td>
</tr>
<tr>
<td>Number of tokens</td>
<td>105,910</td>
<td>162,602</td>
<td>105,221</td>
<td>373,733</td>
</tr>
</tbody>
</table>

5. Analysis

We start with descriptive summaries of the data. Figure 3 (a) shows the mean vowel duration of long and short vowels aggregated over speakers’ decade of birth. As expected, the duration of long vowels is negatively correlated with speakers’ year of birth. But, the duration of the short vowels also show this negative correlation. The general negative correlation between vowel length and year of birth is attributable to the speech rate difference across different age groups; younger speakers generally speak faster than older speakers. This interpretation is supported by other studies that examined the effect of age on speech rate in the same corpus; Kang (2014) found younger speakers produce a shorter vowel duration and this was the case not only for the vowels in the first syllable of a word, which is the focus of the current study, but also for the
vowels in the second syllable. Bang (In progress) calculated the speech rate of each speaker in the same corpus by calculating the number of syllables and phones per second and found that younger speakers show a faster speech rate.

Figure 3. Vowel duration change by speaker’s year of birth

To examine the realization of long vs. short vowel contrast while controlling for the variation in speech rate, the duration of the long vowels are converted to a ratio by dividing the long vowel duration by the mean short vowel duration of each speaker. The mean long/short ratio, aggregated over speakers’ decade of birth, is summarized in Figure 3 (b). The ratio between long and short vowel is over 1.20 for speakers born in 1930s but the ratio falls below 1.00 for speakers born in the 1980s. This is consistent with the general trend of vowel length contrast reduction suggested in the literature reviewed in section 3.

Now we examine the effect of frequency on vowel length realization and its change over time. Figure 3 (c) is the same graph as Figure 3 (b) except that words of high, mid and low frequency are plotted separately. We can make two general observations about this graph. First of all, contrary to the predictions of all three hypotheses in Figure 1, the high frequency words do not have a shorter duration than low frequency words. As for the pattern of change over time, in older generations of speakers, those born in the 1950s or earlier, words of all three frequency types undergo reduction at a comparable rate but in younger generations of speakers, those born in the 1960s or later, mid or low frequency words continue to shorten but high frequency words
resist further reduction. This distinct trajectory of high frequency words seems to suggest that the high frequency words are more resistant to the sound change compared to mid or low frequency words in line with Figure 1 (c).

However, note that the words of different frequency levels consist of words of different phonological shapes occurring in different phrasal contexts and we know that such factors have substantial effect on vowel duration. As a result, we cannot directly compare the duration of different word groups directly and read the frequency effect off of the descriptive statistics. In order to examine the frequency effect properly, we need to control other factors that affect vowel duration. So, we conducted a statistical analysis using a linear mixed-effect model (Baayen et al. 2008), which examines the effect of frequency on vowel duration and the interaction of frequency with speakers’ year of birth while controlling for other factors. The analysis was carried out using the lmer function in the lme4 package (Bates et al. 2011) for R (R Development Core Team 2011).

The data included in the model are 63,365 long vowels in word-initial syllables. In this analysis, we model the effect of speakers’ age and frequency on the realization of long vowel duration. To normalize the speaker-specific durational variation, the duration of each long vowel is divided by the same speaker’s mean short vowel duration. This duration ratio is the dependent variable. Fixed effect predictors included in the model are summarized in Table 5. There are three speaker-level predictors; year of birth (YOB), gender (GENDER), and by-speaker mean short vowel duration (RATE). We expect the duration ratio will be reduced as the speaker’s year of birth increases and therefore expect a negative coefficient for YOB. From the exploratory figure above, we observe that the effect of YOB is not linear and a quadratic term is also included (YOB^2) to model the curved shape of the trajectory. The mean short vowel duration, which we interpret as an indicator of speech rate (RATE), is also included in the model. This predictor is added to take into account the possibility that faster speakers are less likely to retain the durational contrast between long and short vowels. Speaker’s gender (GENDER) is included to test whether male and female speakers produce the contrast differently. Studies of other sound changes in the same corpus—voice onset time merger between lenis and aspirated stops (Kang 2014) and /ɛ/-/e/ merger (Kang To appear)—found that females are ahead of males in these
sound changes in keeping with cross-linguistic tendency for females to lead sound change (Labov 1990). If vowel length merger follows the same pattern, we expect male speakers to show a larger long vowel duration ratio than female speakers.

**Table 5.** The list of fixed effects predictors

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of birth (YOB, YOB^2)</td>
<td>continuous, quadratic</td>
</tr>
<tr>
<td>Frequency (FREQ)</td>
<td>factor (low &lt; mid &lt; high)</td>
</tr>
<tr>
<td>Gender (GENDER)</td>
<td>factor (male, female)</td>
</tr>
<tr>
<td>Speech rate (RATE)</td>
<td>continuous</td>
</tr>
<tr>
<td>Preceding consonant (PREC)</td>
<td>factor (null, m, n, p', t', k', s', c', p, t, k, c, s, pʰ, tʰ, kʰ, cʰ, h)</td>
</tr>
<tr>
<td>Vowel quality (VOWEL)</td>
<td>factor (a, ɛ, e, ʌ, o, i, u, i)</td>
</tr>
<tr>
<td>Syllable structure (SYLL)</td>
<td>factor (closed.obs, closed.son, open.tense, open.C, open.nonC)</td>
</tr>
<tr>
<td>Word length (WORDLENG)</td>
<td>factor (mono, poly)</td>
</tr>
<tr>
<td>Phrase final (IPFINAL)</td>
<td>factor (non-final, final)</td>
</tr>
<tr>
<td>Phrase initial (IPINITIAL)</td>
<td>factor (non-initial, initial)</td>
</tr>
</tbody>
</table>

There are several word-level predictors included in the model. The main predictor of our interest is frequency (**FREQ**). It is a factor with three levels (low, mid, and high) as discussed in Table 4. This factor is Helmert-coded to compare low vs. mid and then low+mid vs. high. Based on the general tendency for high frequency words to show more reduction, we expect a lower duration ratio for high frequency words. To examine how the frequency interacts with the sound change over time, the interaction of **FREQ** with the **YOB** is also included in the model. The three hypotheses sketched in Figure 1 predicts different patterns of interaction; Figure 1 (a) predicts no interaction and Figure 1 (b) and (c) predict an interaction but in opposite directions. According to Figure 1 (b), the **YOB** effect on long vowel duration is stronger in high frequency words (i.e., a steeper slope) than in low frequency words while according to Figure 1 (c), the **YOB** effect on long vowel duration is weakened in high frequency words (i.e., a flatter slope) than in low frequency words.

A number of control predictors that are known to affect vowel duration are also included. These include preceding consonants (**PREC**), vowel quality (**VOWEL**), syllable structure (**SYLL**), word length (**WORDLENG**), and phrasal positions (**IPFINAL** and **IPINITIAL**). For the factor of syllable structure (**SYLL**), based on an exploratory analysis, five levels of syllable structure are defined.
Closed.obs and closed.son represent closed syllable contexts where the coda consonant is an obstruent or a sonorant, respectively. Closed.son includes cases where the coda consonant is underlyingly an obstruent but surfaces as a sonorant due to a regular assimilation process. Open syllables are divided into three levels. Open.tense refers to open syllables followed by a fortis or aspirated consonant. These consonants have long closure duration and are known to shorten the preceding vowel (Yun 2009; Zhi 1993). Open.C refers to open syllables followed by a lenis or sonorant consonant while Open.nonC refers to open syllables followed by /h/, another vowel or word boundary.\textsuperscript{xiii} \textsc{WordLeng} differentiates monosyllabic vs. polysyllabic words. \textsc{IPFinal} and \textsc{IPInitial} distinguish vowels occurring at a phrase boundary from those occurring in phrase medial position. The phrasal boundary is defined by the presence of “silent pause” assigned by the automatic forced aligner (Yoon To appear). We expect the vowels to be longer in absolute phrase-initial or phrase-final position than in phrase-medial position.

To reduce collinearity, numerical variables (\textsc{Yob} and \textsc{Rate}) are centered and all categorical variables (\textsc{Gender}, \textsc{Prec}, \textsc{Vowel}, \textsc{Syll}, \textsc{WordLeng}, \textsc{IPFinal}, and \textsc{IPInitial}) are sum-coded, except for \textsc{Freq}, which is Helmert-coded as explained above. The random effects include \textsc{Word} and \textsc{Speaker}. For \textsc{Speaker}, only a random intercept is included and for \textsc{Word}, a random intercept and a random slope adjustment to \textsc{Yob} (quadratic) are included.

6. Results

Now we turn to the results. All factors included in the model are statistically significant as determined by a Wald chi-square test except that \textsc{Rate} is only marginally significant. \textit{Anova} function of the \textsc{car} package (Fox et al. 2013a) is used for this test. The test statistics are summarized in Table 6.
Table 6. Wald chi-square test of predictors in the linear mixed-effects model

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREC</td>
<td>288.073</td>
<td>15</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>VOWEL</td>
<td>178.752</td>
<td>7</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>SYLL</td>
<td>238.479</td>
<td>4</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>WORDLENG</td>
<td>16.926</td>
<td>1</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>IPFINAL</td>
<td>1927.813</td>
<td>1</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>IPINITIAL</td>
<td>823.720</td>
<td>1</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>GENDER</td>
<td>11.276</td>
<td>1</td>
<td>0.001 **</td>
</tr>
<tr>
<td>RATE</td>
<td>2.890</td>
<td>1</td>
<td>0.089 .</td>
</tr>
<tr>
<td>YOB + YOB $^2$</td>
<td>95.428</td>
<td>2</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>FREQ</td>
<td>13.838</td>
<td>2</td>
<td>0.001 **</td>
</tr>
<tr>
<td>(YOB + YOB $^2$) * FREQ</td>
<td>12.539</td>
<td>4</td>
<td>0.014 *</td>
</tr>
</tbody>
</table>

Table 7 summarizes a coefficient estimate for each predictor and related test statistics. For each fixed effect predictor, a coefficient estimate, a standard error, a t-test statistics and a p-value are provided. The $p$-values are determined using a t-test with a degree of freedom calculated by taking the number of observations (63,365) and subtracting the number of fixed effect parameters (Baayen 2008). The model as a whole explains 49.1% of the variance in the data.
Table 7. A list of fixed effects predictors, coefficient estimates, standard errors, and the p-values.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Estimate</th>
<th>S.E.</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(INTERCEPT)</td>
<td>2.129</td>
<td>0.038</td>
<td>56.338</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /m/)</td>
<td>-0.143</td>
<td>0.039</td>
<td>-3.646</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /n/)</td>
<td>-0.142</td>
<td>0.044</td>
<td>-3.253</td>
<td>0.001 **</td>
</tr>
<tr>
<td>PREC (null vs. /p/)</td>
<td>-0.269</td>
<td>0.112</td>
<td>-2.399</td>
<td>0.016 *</td>
</tr>
<tr>
<td>PREC (null vs. /t/)</td>
<td>0.093</td>
<td>0.112</td>
<td>0.834</td>
<td>0.405</td>
</tr>
<tr>
<td>PREC (null vs. /k/)</td>
<td>-0.278</td>
<td>0.086</td>
<td>-3.24</td>
<td>0.001 **</td>
</tr>
<tr>
<td>PREC (null vs. /c/)</td>
<td>-0.347</td>
<td>0.188</td>
<td>-1.845</td>
<td>0.065</td>
</tr>
<tr>
<td>PREC (null vs. /p/)</td>
<td>-0.194</td>
<td>0.061</td>
<td>-3.179</td>
<td>0.001 **</td>
</tr>
<tr>
<td>PREC (null vs. /t/)</td>
<td>-0.216</td>
<td>0.045</td>
<td>-4.765</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /k/)</td>
<td>-0.396</td>
<td>0.041</td>
<td>-9.627</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /c/)</td>
<td>-0.317</td>
<td>0.041</td>
<td>-7.761</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /s/)</td>
<td>-0.378</td>
<td>0.036</td>
<td>-10.647</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /pʰ/)</td>
<td>-0.364</td>
<td>0.09</td>
<td>-4.027</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /tʰ/)</td>
<td>-0.346</td>
<td>0.188</td>
<td>-1.844</td>
<td>0.065</td>
</tr>
<tr>
<td>PREC (null vs. /cʰ/)</td>
<td>-0.513</td>
<td>0.087</td>
<td>-5.892</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>PREC (null vs. /h/)</td>
<td>-0.512</td>
<td>0.041</td>
<td>-12.354</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>VOWEL (/a/ vs. /e/)</td>
<td>-0.097</td>
<td>0.037</td>
<td>-2.628</td>
<td>0.009 **</td>
</tr>
<tr>
<td>VOWEL (/a/ vs. /æ/)</td>
<td>-0.163</td>
<td>0.07</td>
<td>-2.325</td>
<td>0.02 *</td>
</tr>
<tr>
<td>VOWEL (/a/ vs. /ʌ/)</td>
<td>-0.246</td>
<td>0.036</td>
<td>-6.872</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>VOWEL (/a/ vs. /o/)</td>
<td>-0.154</td>
<td>0.03</td>
<td>-5.205</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>VOWEL (/a/ vs. /i/)</td>
<td>-0.383</td>
<td>0.039</td>
<td>-9.812</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>VOWEL (/a/ vs. /u/)</td>
<td>-0.437</td>
<td>0.043</td>
<td>-10.109</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>VOWEL (/a/ vs. /ɪ/)</td>
<td>-0.483</td>
<td>0.065</td>
<td>-7.433</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>SYLL (closed_obs vs. closed_son)</td>
<td>0.065</td>
<td>0.034</td>
<td>1.898</td>
<td>0.058</td>
</tr>
<tr>
<td>SYLL (closed_obs vs. open_tense)</td>
<td>0.020</td>
<td>0.06</td>
<td>0.33</td>
<td>0.742</td>
</tr>
<tr>
<td>SYLL (closed_obs vs. open_C)</td>
<td>0.260</td>
<td>0.034</td>
<td>7.622</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>SYLL (closed_obs vs. open_nonC)</td>
<td>0.647</td>
<td>0.051</td>
<td>12.645</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>WORDLENG (mono vs. poly)</td>
<td>-0.195</td>
<td>0.047</td>
<td>-4.114</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>IPFINAL (non-final vs. final)</td>
<td>1.266</td>
<td>0.029</td>
<td>43.907</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>IPINITIAL (non-initial vs. initial)</td>
<td>0.409</td>
<td>0.014</td>
<td>28.701</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>GENDER (female vs. male)</td>
<td>0.033</td>
<td>0.01</td>
<td>3.358</td>
<td>0.003 **</td>
</tr>
<tr>
<td>RATE</td>
<td>0.001</td>
<td>0.001</td>
<td>1.7</td>
<td>0.089</td>
</tr>
<tr>
<td>YOB</td>
<td>-15.752</td>
<td>1.706</td>
<td>-9.232</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>YOB ^2</td>
<td>3.469</td>
<td>1.286</td>
<td>2.698</td>
<td>0.007 **</td>
</tr>
<tr>
<td>FREQ (low vs. mid)</td>
<td>-0.015</td>
<td>0.028</td>
<td>-0.54</td>
<td>0.589</td>
</tr>
<tr>
<td>FREQ (low+mid vs. high)</td>
<td>-0.133</td>
<td>0.029</td>
<td>-4.53</td>
<td>&lt;0.001 ***</td>
</tr>
<tr>
<td>YOB * FREQ (low vs. mid)</td>
<td>-3.002</td>
<td>2.202</td>
<td>-1.363</td>
<td>0.173</td>
</tr>
<tr>
<td>YOB^2 * FREQ (low vs. mid)</td>
<td>-0.736</td>
<td>1.353</td>
<td>-0.544</td>
<td>0.587</td>
</tr>
<tr>
<td>YOB * FREQ (low+mid vs. high)</td>
<td>4.091</td>
<td>2.07</td>
<td>1.976</td>
<td>0.048 *</td>
</tr>
<tr>
<td>YOB^2 * FREQ (low+mid vs. high)</td>
<td>2.875</td>
<td>1.21</td>
<td>2.377</td>
<td>0.017 *</td>
</tr>
</tbody>
</table>
We will discuss each predictor in turn along with partial effects plots in Figure 4 and Figure 5, which display the predicted values for each factor with all other factors held constant at their mean. The \textit{effect} function in the \textit{effects} package (Fox et al. 2013b) is used to calculate the partial effect estimates and the 95\% confidence interval. We discuss the control factors first and then discuss the main factors of our interest.

Figure 4. Partial-effect plots of main effects

The consonant that precedes the vowel (\texttt{PREC}) affects the vowel duration. Overall, consonants
that have long aspiration or frication noise tend to shorten the following vowel, as shown in Figure 4 (a). The following vowel quality (VOWEL), more specifically, the following vowel height affects vowel duration; high vowels (/i, i u/) are shorter than mid vowels (/ʌ, o, e, ɛ/) and mid vowels are shorter than the low vowel /a/, as shown in Figure 4 (b). The syllable structure (SYLL) also significantly affects vowel duration; vowels are shorter in closed syllables (Closed.obs or Closed.son) than in open syllables except that open syllables followed by a fortis or aspirated consonant (Open.tense) pattern like a closed syllable in vowel duration. Open syllables followed by a lenis or sonorant consonant (Open.C) is shorter than open syllables followed by /h/, a vowel or a word boundary (Open.nonC). This effect is summarized in Figure 4 (c). WORDLENG, IPFINAL, and IPINITIAL all show a significant effect in the expected direction as shown in Figure 4 (d-f); a vowel is longer in monosyllabic than in polysyllabic words and a vowel is longer when it occurs in phrase-initial or final position than when it occurs in phrase-medial position.

Now we turn to the speaker-level factors. As for GENDER, female speakers overall has a more reduced long vowel duration ratio than male speakers, as shown in Figure 4 (g). This effect is in line with the general trend of female speakers leading various sound changes in this corpus (Kang 2014; Kang To appear), and in sound changes more generally (Labov 1990). Speakers’ speech rate (RATE), as defined by the mean short vowel duration for the speaker, shows that the faster speakers (i.e., speakers with a shorter mean short vowel duration) reduce the vowel length contrast more, as shown in Figure 4 (h) and this effect is marginally significant.

Now we turn to the factors of our main interest; YOB, FREQ, and their interaction. Speakers’ year of birth (YOB) is confirmed to be significant; the younger the speakers, the more reduced the long vs. short vowel contrast, in agreement with previous studies that suggested reduction of this contrast in younger speakers’ speech. The significant quadratic term (YOB^2) indicates that the rate of contrast reduction slows down in the very young speakers’ speech. This is visualized as a curved shape in the partial effect plot in Figure 4 (i). Before we consider the next variable, we want to consider the potential confound of age and speech rate. As discussed above, younger speakers in general produce all vowels with a shorter duration, not only for long vowels but also for short vowels and there is a significant correlation between speakers’ age (YOB) and their
speech rate (\textit{RATE}) \((t = -7.8661; \text{df} = 116; p < 0.001)\). Given this correlation between the two predictors, we need to consider an alternative hypothesis that the age effect is not a reflection of sound change in progress but only an epiphenomenon of a rate effect—i.e., younger speakers tend to speak faster and faster speakers tend to show more reduced durational contrast. To test this alternative hypothesis, we conducted a more stringent test of YOB effect by replacing the YOB with the residual of the YOB against \textit{RATE} as a predictor. In other words, this new model attributes all explanatory power shared by YOB and \textit{RATE} to \textit{RATE} and YOB is given the minimum credit possible. In this new model, the effect of \textit{RATE} is significant \((\chi^2 = 82.828, \text{df} = 1, p < 0.001)\), not surprisingly, but the effect of \text{YOB} + \text{YOB}^2 remains significant as well \((\chi^2 = 86.164, \text{df} = 2, p < 0.001)\) and so is its interaction with \textit{FREQ} \((\chi^2 = 44.374, \text{df} = 4, p < 0.001)\). This confirms that speakers’ age has a robust and independent effect on durational contrast over and above any effect attributable to the speech rate difference across generations.

As for the frequency effect (\textit{FREQ}) on vowel duration, when various structural differences across low, mid, and high frequency words are controlled for, an effect of frequency emerges; long vowels in high frequency words are significantly shorter than those in mid and low frequency words, while there is no significant difference between low and mid frequency words, as shown in Figure 4 (j).

Now that we have established the effects of YOB and \textit{FREQ} in the expected directions, we turn to the interaction of \textit{FREQ} and YOB. Recall from Figure 3 (c), repeated in Figure 5 (a), that vowels in mid and low frequency words show a generally linear trajectory and the shortening trend in the older speakers’ speech is sustained in the speech of younger speakers. For high frequency words, on the other hand, the shortening trend bottoms out for the younger speakers. Our model confirms this interaction of frequency and age; the YOB effect on vowel duration ratio differs significantly between high frequency words and low and mid frequency words. The YOB effect does not differ between low and mid frequency words. This interaction is visually represented in Figure 5 (b).
Figure 5. Interaction of frequency and year of age; (a) actual mean long/short vowel ratio (b) predicted mean long/short vowel ratio

This partial effect plot reveals that long vowels are indeed far more reduced in high frequency words than in low and mid frequency words and they seem to reach the end point much earlier and hence hit the end of the S-curve (Labov 1994) before mid or low frequency words. So, when relevant phonological contexts are controlled for, it does not seem to be the case that the high frequency words are particularly more resistant to this sound change, as one may be tempted to infer based on Figure 5 (a). But high frequency words seem to stop moving further along the change as they already reached the end point and cannot reduce further. So, this trajectory seems inconsistent with an interpretation that this particular sound change is analogical in nature (cf. Figure 1 (c)). To examine if the rate of change is different by frequency during the time when high frequency words are still progressing in change, we fit another model only including the data from speakers who are born before 1960 and found no significant interaction between YOB and FREQ ($\chi^2=1.5806$, df=2, p=0.4537).xxvi In other words, our data show no difference in the rate of change across frequency types and the data support the constant rate hypothesis sketched in Figure 1 (a) over the other two hypotheses sketched in Figure 1 (b) or (c), where high and low frequency words change at different rates.

7. Discussion
Now let us review the questions we presented at the outset of this paper and see how our data can answer them. The first question was if the vowel length contrast merger is actually happening in Seoul Korean and our answer is yes. We found a clear effect of speaker’s year of birth on the durational contrast and confirmed that younger speakers indeed produce a more reduced contrast than older speakers. We also checked to make sure that this effect is not an epiphenomenon of speech rate effect but the age effect is robust even after the effect of faster speech rate of younger speakers is factored out. And we found that the change is almost complete such that the long/short ratio reaches below zero and plateaus out in the youngest speakers’ speech. So, we are observing the end stage of an S-curve in this change.

The others questions we posed are about the effect of frequency. We noted that the synchronic frequency effect on phonetic reduction may give rise to an appearance of diachronic frequency effect, if the frequency effect is examined at a static point in time. To remedy this methodological confound, we tried to disentangle a diachronic effect of frequency on sound change from a synchronic effect on phonetic reduction by examining its dynamic effect (i.e., frequency effect on the rate of change in duration) as well as its static effect (i.e., frequency effect on overall duration). We found a robust effect of frequency on duration; high frequency words have a substantially shorter duration than mid or low frequency words of comparable phonological structure and context. We did find a significant effect of frequency on the trajectory of sound change, but not in the way that we hypothesized. High frequency words differed from mid or low frequency words in that they reached the endpoint earlier and stopped progressing further in the change in younger speakers while mid and low frequency words continue to reduce in younger speakers’ speech. But, when we examine the slope of the change while high frequency words, as well as mid and low frequency words are still progressing in change, words of different frequency types did not differ in their overall rate or slope of change in duration. In short, of the three patterns of frequency sketched in Figure 1, Figure 1 (a) was the closest to the pattern we found.

Now we turn to the question of what this frequency effect or lack thereof suggests about the nature of this sound change and the nature of lexical representation. First of all, we did not find any evidence from frequency effects that the sound change is analogical or analytical in nature—
we did not find high frequency words to be more resistant to the change overall. We also did not find strong support for the of the exemplar-based model of lexical representation, which includes word-specific distribution of phonetically detailed stored exemplars, which get updated with each use of the word. Such a model would predict a faster rate of reduction for high frequency words over low frequency words in a reductive sound change. The observed frequency effect, or lack thereof, is compatible with a model without word-specific phonetic representations where the frequency effect on duration comes from on-line factors that affect phonetic implementation of speech sounds, not through stored tokens of word-specific variants.

References
Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand & Dan Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. Journal of Memory and Language 60.92-111.


Ernestus, Mirjam. 2014. Acoustic reduction and the roles of abstracts and exemplars in speech processing. Lingua 142.27-41.


Fox, John, Sanford Weisberg, Douglas Bates, David Firth, Michael Friendly, Spencer Graves, Richard Heiberger, Rafael Laboissiere, Georges Monette, Henric Nilsson, Brian Ripley, Achim Zeileis & R Core. 2013a. car.

Fox, John, Sanford Weisberg, Jangman Hong, Robert Andersen & Steve Taylor. 2013b. Effects.


Han, Mieko S. . 1964. Studies in the Phonology of Asian Languages II: Duration of Korean Vowels Los Angeles: University of California.


Kim, Ju-Phil. 1990. '표준어 모음'의 심의 경의와 해설. 국어생활 22.190-207.


Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. Language Variation and Change 2.205-54.


—. 2014. 한글 어휘의 발화 발음 및 단모음의 구분. 말소리와음을과학.
Yun, Ilsung. 2009. Vowel duraition and the feature of the following consonant. 말소리와 음성과학 1.41-46.
Acknowledgements to be inserted…

Notable exceptions include Sonderegger & Niyogi (2013) and Hinskens (2014).

Other studies that show high-frequency-words-first effect include Fidelholtz (1975) and Schleef (2013) and studies that show low-frequency-words-first effect include Do et al. (2014), Kang (2003), and Sonderegger & Niyogi (2013), among others.

But, Walker (2012), based on Canadian English data, shows that the frequency effect in t/d deletion is miniscule to non-existent once lexical and phonological effects are factored in. Also, he notes that there is no clear evidence that t/d deletion is in fact a sound change, rather than a synchronic variation.

This is also the prediction of the Constant Rate Hypothesis (Kroch 1989).

We thank Ranjan Sen for bringing this case to our attention.

De Schryver et al. (2008) suggest that orthography likely played a role in this frequency effect; the Dutch orthography reflects the standard form of these words and the speakers would have more exposure to the spelling of high frequency words than low frequency words and this may introduce the frequency effect in the opposite direction from that expected in a reductive sound change.

The picture gets even more complicated as we find cases where within a single sound change frequency effects in opposite directions are found. For example, Sonderegger (2010) examined the English diatonic stress shift in noun-verb pairs, where noun-verb pairs that originally have stress on the second syllable shifts the stress to the initial syllable for nouns (1,2). The overall frequency effect is in the expected direction—i.e., the stress shift is more likely for low frequency than high frequency words—but the frequency effect interacts with the phonological factors such that when the phonological shape of the word strongly prefers stress shifts, the frequency effect is attenuated and even reversed.

Yet another possibility is that the frequency has little to no effect on sound change once all other structural factors are factored in (Dinkin 2008; Walker 2012).

The long vowels of contemporary Seoul Korean developed from the rising tone of Middle Korean (Lee & Ramsey 2001; Martin 1992).

Kahng (1995) also examined the production by three speakers of Munhwae, the standard North Korean and found that they tend to no retain the vowel length contrast.

Each token was categorized as long or short based on the author’s perception supplemented with spectrogram-based measurements. A vowel was considered to be long if it is 1.5-2 times longer than the short vowel in a comparable phonological context.

Such perception-production gap is attested under a sound change in progress (Hay 2006).

Park, Ju-Kyeng. 1985. 현대 한국어의 장단음에 관한 연구. 말소리 11-14.121-31. makes a number of other interesting observations about the generational difference in vowel length production. In older speakers’ speech, the mid back unrounded vowel shows quality differentiation conditioned by the vowel length Lee, Hyun Bok. 1993. Korean. Journal of the International Phonetic Association 23.28-31.; the long /ʌ:/ is realized as more raised and central than the short /ʌ/ while younger speakers produce the long and short /ʌ/ with the same quality. Park (1985)
also note the possibility that that younger speakers lengthen the coda consonant in a syllable with a long vowel, while for older speakers, the vowel itself carries the durational contrast. Our current study did not examine the possibility that younger speakers may retain the vowel length contrast by lengthening the coda consonant. We leave this for future research.

Separate mixed-effects logistic regression models for each age group (dependent variable: the vowel length realization, fixed effect: the underlying vowel length, random effect: word) show that the underlying vowel length is significant for the speakers in their 50s ($z=1.977$, $p=0.0481$) and 60s ($z=2.413$, $p=0.0158$) but not for the speakers in their 20s ($z=0.015$, $p=0.988$), 30s ($z=1.002$, $p=0.316$), and 40s ($z=0.842$, $p=0.4$).

Available online at [http://korean.go.kr/](http://korean.go.kr/). See Kim (1990) for information on how the pronunciation norm, including vowel length specification, on this dictionary has been determined.

Perceptual categorization by native speakers used in Kim (2003)’s survey is a method that is not practical for a corpus study of this scale and the method brings in its own problems as discussed above.

But, this aggregated measure of phonetic duration does not distinguish between two types of changes—all long vowels are reduced in their phonetic duration over time vs. long vowels are not reduced in phonetic duration but some long vowel words are changed to short vowels.


The lemma frequency was chosen partly due to the availability of a reliable frequency list where homophones are carefully disambiguated. The future study will examine the effect of frequency based on wordform frequency counts.

We thank Kevin Tang for suggesting k-means clustering as a way to define the frequency levels. In converting the frequency count to log frequency, the frequency count was raised by 1 to avoid undefined values for zero frequency items.

/h/ systematically deletes in some morpho-phonological contexts and optionally but frequently deletes in intervocalic contexts.

A Tukey’s post-hoc pairwise comparison using the `glht` function of the `multcomp` package (Hothorn et al. 2013) generally confirm the three-way differentiation of high, mid, and low vowels; the differences across different categories are significant and differences within categories are not significant with the following exceptions: no differences are found for /a/ vs. /ɛ/ and /a/ vs. /ɛ/ and /ɛ/ is significantly longer than /ʌ/.

A Tukey’s post-hoc pairwise comparison confirms the following three levels of differentiation: closed.obs, closed.son, open.tense < open.C < open.nonC.

In this model, only a linear term of YOB is used. A model including a quadratic term, and a model with a different age cut-off (born before 1970 or born before 1950) also show no significant interaction of YOB and FREQ.