Lexical diffusion of vowel length merger in Seoul Korean: a corpus-based study

Yoonjung Kang¹, Tae-Jin Yoon², Sungwoo Han³
University of Toronto Scarborough¹,
Cheongju University², Inha University³
LabPhon 14
July 26, 2014
Overview

➢ Usage-based model of phonology:
  - Frequency interacts with sound change reflecting the mechanism behind the change
  - But, studies on the frequency effect on sound change that track the effect over time are rare.

➢ Shortening of long vowels in Seoul Korean, sound change in progress
I. Korean Vowel Length Contrast
Vowel length contrast in Seoul Korean

<table>
<thead>
<tr>
<th>Korean</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>/nun/ 'eye'</td>
<td>/nu:n/ 'snow'</td>
</tr>
<tr>
<td>/pε/ 'pear'</td>
<td>/pε:/ 'double'</td>
</tr>
<tr>
<td>/pɑ:m/ 'night'</td>
<td>/pɑ:m/ 'chestnut'</td>
</tr>
<tr>
<td>/ɪl/ 'one'</td>
<td>/i:l/ 'work'</td>
</tr>
<tr>
<td>/tsʌk-ta/ 'to write'</td>
<td>/tsʌ:k-ta/ 'plenty'</td>
</tr>
<tr>
<td>/mos/ 'nail'</td>
<td>/mo:s/ 'Neg.'</td>
</tr>
</tbody>
</table>
Long vowels are restricted to word-initial position.

- /nᵯːn-s’ʌlmɛ/  
  \[nᵯːns’ʌlmɛ\] ‘toboggan’
- /tsʰʌs-nᵯːn/  
  \[tsʰʌnᵯn\] ‘first snow’
- /soːk-satsʌŋ/  
  \[soːks’adzʌŋ\] ‘behind story’
- /mɑᵯm-soːk/  
  \[mɑᵯm’s’ok\] ‘one’s mind’
- /pʌ:lli-ta/  
  \[pʌ:llida\] ‘open, spread’
- /t’ʌ-pʌ:llita/  
  \[t’ʌbʌllida\] ‘brag’
Some interaction with morphology

- Some verbs/adjectives alternate:
  
<table>
<thead>
<tr>
<th>C-initial suffix</th>
<th>V-initial suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ku:lm-ta] (&gt; [ku:mt’a])</td>
<td>[kulm-ʌ]</td>
</tr>
<tr>
<td>[nʌ:h-ta] (&gt; [nʌ:tʰa])</td>
<td>[nʌ:h-ʌi] (&gt; [nʌi])</td>
</tr>
<tr>
<td>[ki:l-ta]</td>
<td>[kil-ʌ]</td>
</tr>
<tr>
<td>[pu:s-ta] (&gt; [pu:t’a])</td>
<td>[pu-ʌni]</td>
</tr>
</tbody>
</table>

- Not all verbs/adjectives do:
  
  | [ca:k-ta] (> [ca:kt’a]) | [ca:k-ʌ] | ‘small’ |
  | [ʌ:ps-ta] (> [ʌ:pt’a]) | [ʌ:ps-ʌ] (> [ʌps’ʌ]) | ‘absent’ |

- Nouns don’t:
  
<table>
<thead>
<tr>
<th>C-initial suffix</th>
<th>V-initial suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>[sa:lm-to] (&gt; [sa:mdo])</td>
<td>[sa:lm-ɪ]</td>
</tr>
<tr>
<td>[ka:m-kwa] (&gt; [ka:mgwa])</td>
<td>[ka:m-ɪ]</td>
</tr>
</tbody>
</table>
Previous studies

➢ Earlier studies describe vowel length contrast as phonemic, although some individual or lexical variation is noted already.

➢ More recent studies observe generational change and variation in vowel length realization.

➢ Lexical diffusion: certain words are more consistently long than others.
II. Frequency Effect on Sound Change
Frequency and Lexical Diffusion

- Sound change may affect low- and high-frequency words differently.
High frequency words first

- Phonetically motivated changes arising from shortening and lenition affect high-frequency words first.
- If representation of a word gets updated with each use of the word, a reductive change will be more advanced in high frequency words than in low frequency words.
  
  e.g., Schwa deletion in English: every vs. mammary
Low frequency words first

- Analogically motivated changes that involve (re)-analysis affect low-frequency words first.
- High-frequency words form a strong mental representation and resist change motivated by analogy to other forms.
- E.g., Regularization of English irregular past: *weeped* (<wept>) vs. *keeped* (<kept>)
Some changes that affect high-frequency words first are not lenition-based.

- English stress shift in –ate verbs: e.g., frústrate > frustráte

Some (seemingly) reductive changes affect low-frequency words first.

- Middle English mid front vowel unrounding: ö > e
- English yod dropping: tj > t

“Changes which require analysis—whether syntactic, morphological, or phonological—during their implementation affect the least frequent words first; others affect the most frequent words first.”
Synchrony and Diachrony of frequency effect

➢ Synchronic effect of frequency on segment duration: high frequency words are shorter and reduced.

   (Bell et al. 2003, 2009, Ernestus et al. 2006, Gahl 2008…)

➢ But, less clear is how this synchronic frequency effect interact with sound change over time?
A) Diachrony mirrors synchronic frequency effect

*Synchrony=Diachrony*

Diagram showing the relationship between time and duration across different frequency levels (low, mid, high). The diagram illustrates how duration decreases with time, distinguishing between old and young categories.
B) Diachrony **amplifies** synchronic frequency effect

\[ \text{Synchrony}^2 = \text{Diachrony} \]
C) Diachrony introduces an “analytical” frequency effect, opposite of synchrony
D) Mixed

High & Low Last

High & Low First
Which type of change is loss of length contrast in Korean?

- **High** frequency words first
  - Long vowels become short vowels.
  - Shortening/Reduction

- **Low** frequency words first
  - Removal of a marked structure
  - Analogy to a dominant pattern (short vowels)?
    - Long vowels: restricted to word-initial position, estimated to be <20%
III. Data
The Reading-Style Speech Corpus of Standard Korean

- The National Institute of the Korean Language (2002)
- 60 male, 60 female, Seoul-Kyeoungki area
- Age distribution

<table>
<thead>
<tr>
<th>Age</th>
<th>20s</th>
<th>30s</th>
<th>40s</th>
<th>50s</th>
<th>60~70s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>20</td>
<td>20 (19)</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20 (19)</td>
</tr>
</tbody>
</table>

- Apparent time study method (Bailey et al. 1991)
<table>
<thead>
<tr>
<th>Text</th>
<th># of sentences</th>
<th>Duration (mv01)</th>
</tr>
</thead>
<tbody>
<tr>
<td>요람기</td>
<td>51</td>
<td>321.92 sec (5.37 min)</td>
</tr>
<tr>
<td>소나기</td>
<td>87</td>
<td>240.12 sec (4.00 min)</td>
</tr>
<tr>
<td>아름다움에 대하여</td>
<td>69</td>
<td>311.92 sec (5.20 min)</td>
</tr>
<tr>
<td>방망이를 쳐다본 노인</td>
<td>62</td>
<td>275.50 sec (4.59 min)</td>
</tr>
<tr>
<td>문학의 세계와 삶의 세계</td>
<td>47</td>
<td>321.01 sec (5.35 min)</td>
</tr>
<tr>
<td>독서와 인생</td>
<td>54</td>
<td>294.97 sec (4.92 min)</td>
</tr>
<tr>
<td>수난 이대</td>
<td>62</td>
<td>317.31 sec (5.29 min)</td>
</tr>
<tr>
<td>메밀꽃 필 무렵</td>
<td>94</td>
<td>425.98 sec (7.10 min)</td>
</tr>
<tr>
<td>토끼와 자라</td>
<td>60</td>
<td>185.27 sec (3.09 min)</td>
</tr>
<tr>
<td>선녀와 나무꾼</td>
<td>73</td>
<td>295.03 sec (4.92 min)</td>
</tr>
<tr>
<td>호랑이와 꽃감</td>
<td>42</td>
<td>159.92 sec (2.67 min)</td>
</tr>
<tr>
<td>해님 달님</td>
<td>28</td>
<td>124.10 sec (2.07 min)</td>
</tr>
<tr>
<td>그리운 시내가</td>
<td>36</td>
<td>276.49 sec (4.61 min)</td>
</tr>
<tr>
<td>광화문 지하도 아주머니</td>
<td>27</td>
<td>148.40 sec (2.47 min)</td>
</tr>
<tr>
<td>막 지은 밥</td>
<td>17</td>
<td>106.36 sec (1.77 min)</td>
</tr>
<tr>
<td>눈 오던 날</td>
<td>35</td>
<td>147.77 sec (2.46 min)</td>
</tr>
<tr>
<td>숭늉의 지혜</td>
<td>19</td>
<td>95.00 sec (1.58 min)</td>
</tr>
<tr>
<td>깨만 눈동자들 앞에서</td>
<td>27</td>
<td>205.92 sec (3.43 min)</td>
</tr>
<tr>
<td>내 고향 개울 물에서</td>
<td>40</td>
<td>163.34 sec (2.72 min)</td>
</tr>
<tr>
<td>Total</td>
<td>930</td>
<td>4416.40 sec (73.61 min)</td>
</tr>
</tbody>
</table>

From Yoon (2013)
## Seoul Korean vowels

<table>
<thead>
<tr>
<th>i</th>
<th>ɨ</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>Λ</td>
<td>o</td>
</tr>
<tr>
<td>(ɛ)</td>
<td>α</td>
<td></td>
</tr>
</tbody>
</table>
Vowels in word-initial syllables

- 2,250 word types
- 3,366 word tokens
- 1,251 lexemes
- 118 speakers
- <300 ms only, voiced vowels only
- 373,733 tokens of vowels in word-initial syllable
Korean phone-aligner

- Developed by Tae-Jin Yoon (cf. Yoon and Kang 2012, Yoon 2013)
  
  \url{http://korean.utsc.utoronto.ca/kpa/}

- Given a sound file and its Korean transcription, the aligner provides a segmentation of the file into component phones.
IV. Results
Long/Short duration ratio by Year of Birth

By-Speaker Long/Short duration ratio vs Year of Birth:

- 1930s
- 1940s
- 1950s
- 1960s
- 1970s
- 1980s

The ratio decreases over time, indicating a change in speech duration with increasing age.
Linear Mixed Effects Model

- **Dependent variable:** Vowel duration (log)
- **Factors of main interest**
  - YOB + YOB² (Year of Birth, quadratic)
  - Vowel length
  - Frequency (low, mid, high)
  - Vowel Length * [YOB + YOB²]
  - Vowel Length * [YOB + YOB²]* Frequency
- **Control factors**
  - Preceding consonants
  - Vowel height
  - Syllable structure
  - Position in prosodic phrase
  - Word length
  - Gender
- **Random effects**
  - Word: intercept and slope adjustment to [YOB + YOB²]
  - Speaker: intercept
- **59.8% of variance explained by the model**
- **All fixed effects and random effects came out significant.**
Results

- **Vowel Length** * [YOB+YOB²]
  - As YOB increases, vowel length contrast reduces.
  - The reduction rate slows down toward the youngest group.

- **Frequency**
  - Higher the frequency, shorter the duration.
  - Duration: Low freq. > Mid freq. > High freq.

- **Vowel Length** * [YOB+YOB²] * Frequency
  - Reduction slows down/flattens for high frequency words, compared to mid-low frequency words.
  - Difference between mid and low frequency words, not significant
Conclusion

- High frequency words resist the change, opposite of the synchronic frequency effect on duration.
- But, this difference is found mostly in the younger group, where change is already fairly advanced.
- Maybe this “analytical” frequency effect shows up only as change is close to being completed.
- Highlights the need to examine the frequency effect in change as it unfolds over time, rather than the final outcome or the pattern at a static point in time.
Thank you!
Thanks to:

- The Social Sciences and Humanities Research Council of Canada—Partnership Development Grant (# 890-2012-25)
- The National Institute of the Korean Language
- The University of Toronto Scarborough Library Digital Scholarship Unit
- Research Assistants: Yaruna Cooblal, Sohyun Hong, Roobika Karunananthan, Julianna So, Shawna-Kaye Tucker, and Cindy Yee
- LVC group and Phonology/Phonetics group at the University of Toronto