A phonetic comparison of lexical /i/ and epenthetic /i/ in Korean speech corpus

Hyunjin Lee

University of Georgia

Cross-linguistically, an epenthetic vowel can be shorter than its corresponding lexical vowel (Gouskova & Hall 2009). Additionally, it may have different F1 and F2 values compared to its corresponding lexical vowel (Davidson 2006). In Korean, epenthetic /ɨ/ and /i/ are employed in English loanword adaptation to conform English syllable structures to the Korean grammar. In English loanword adaptation in Korean, the vowel /ɨ/ serves as the default epenthetic vowel and is applied to broader contexts compared to the vowel /i/. Specifically, while epenthetic /i/ occurs only after post-alveolar consonants, epenthetic /ɨ/ occurs everywhere except after post-alveolar consonants. For this reason, most previous studies (Kim 2009; Kim & Kochetov 2011) on the phonetic features of Korean epenthetic vowels have primarily concentrated on epenthetic /ɨ/. However, to develop a more fine-grained and comprehensive phonetic model of epenthetic vowels in English loanword adaptation in Korean, there is a need for phonetic investigations on epenthetic /i/. Therefore, this study aims to explore the phonetics of epenthetic /i/ in Korean speech.

I analyzed the corpus 'Korean Broadcast News Speech' (Strassel et al. 2006) from the Linguistic Data Consortium (LDC) as a sample of read broadcast speech. This corpus consists of 18 audio files, amounting to 13 hours of satellite radio news broadcasts delivered by 8 female and 10 male speakers of Korean. To segment all audio files into words and phonemes, the Korean Phonetic Aligner Program Suite (Yoon & Kang 2013) was used, and the alignment results were shown in TextGrid files in Praat (Boersma & Weenink 2021). Once all alignments were corrected, a Praat script was used to extract the phonetic features of lexical /i/ and epenthetic /i/ from Praat TextGrid files. In my Praat script, five formants were detected with the ceiling of 4500 Hz for male speakers and 5500 Hz for female speakers, and the midpoint of vowels were measured for F1 and F2 values. All extracted results were shown in an Excel file. In the data, all epenthetic /i/ examples were loanwords like [mætʃʰi] from English 'match', and all lexical /i/ were native Korean words like [kotʃʰi] 'cocoon'. A total of 1709 vowel tokens were analyzed (epenthetic /i/ = 218, lexical /i/ = 1491).

For the statistical analysis, the linear mixed-effect model was used, and the results show there are no significant differences in F1 (p = 0.2732), F2 (p = 0.5071), and vowel duration (p = 0.0988) between epenthetic /i/ and lexical /i/ in this Korean speech corpus. This suggests that the phonetic features of epenthetic /i/ are not different from those of lexical /i/. This finding may not be surprising given that previous studies (Kim 2009; Kim & Kochetov 2011), comparing the phonetic aspects of Korean /ɨ/ across epenthetic and lexical contexts, also show that there are no formant or durational differences between epenthetic /ɨ/ and lexical /ɨ/.

This study investigated the phonetic features of lexical /i/ and epenthetic /i/ which previous studies did not give much attention to and found that these two types of vowels share the same phonetic features. Therefore, my findings may contribute to filling the gap in understanding the phonetic characteristics of epenthetic vowels in the adaptation of English loanwords in Korean. In future research, assessing phonetic differences between epenthetic /i/ and lexical /i/ across speech styles (e.g., casual speech vs. careful speech) may be of interest, as vowel phonetic features tend to vary across different speech styles. For instance, in Turkish, epenthetic vowels vary more according to the following context than lexical vowels, and this phenomenon strengthens in deaccented casual speech vs. careful speech (Bellik 2019).

References

Bellik, J. (2019). Vowel Intrusion in Turkish Onset Clusters. UC Santa Cruz Ph.D. dissertation.

Boersma, P. & Weenink, D. (2021). Praat: Doing phonetics by computer [Computer program], Version 6.2.01. https://www.fon.hum.uva.nl/praat/.

Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34(1), 104-137.

Gouskova, M., & Hall, N. (2009). Acoustics of epenthetic vowels in Lebanese Arabic. In Parker, S. (Eds.), *Phonological argumentation: Essays on evidence and motivation*, 203-225.

Kadenge, M, & Mudzingwa, C. (2012). Comparing ChiShona loanwords of monolingual and bilingual speakers: An Optimality Theory analysis. *South African Journal of African Languages, 32*(2), 141-151.

Kim, J. (2009). Insertion Preferred to Deletion in Learner Speech: A Study of Korean English. *Studies in English Language & Literature* 51(3), 101-124.

Kim, K., & Kochetov, A. (2011). Phonology and phonetics of epenthetic vowels in loanwords: Experimental evidence from Korean. *Lingua* 121(3), 511-532.

Strassel, S., Martey, N., & Graff, D. (2006). Korean Broadcast News Speech. https://catalog.ldc.upenn.edu/LDC2006S42 (24 August, 2006)

Yoon, T. & Kang, Y. (2013). The Korean Phonetic Aligner Program Suite. https://korean.utsc.utoronto.ca/kpa/