

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

Hyunjin Lee
(University of Georgia)

Lee, Hyunjin. 2025. A phonetic comparison of epenthetic /i/ and lexical /i/ in Korean spontaneous speech. *Linguistic Research* 42(1): 203-225. This study investigates the acoustic features (i.e., vowel formants and durations) of epenthetic /i/ and lexical /i/ in Korean spontaneous speech. While epenthetic vowels often differ phonetically from lexical vowels in terms of formants and/or duration across many languages (Davidson 2006; Gouskova and Hall 2009), previous studies on Korean epenthetic /i/ and lexical /i/ in read speech have shown no significant acoustic differences between these two vowels (Kim 2009; Kim and Kochetov 2011). Given that speech styles influence the acoustic features of vowels (Koopmans-van Beinum 1980; Harmegnies and Poch-Olivé 1992), it is worth examining the acoustic features of epenthetic /i/ and lexical /i/ across different speech styles in Korean. Thus, in this study, political debates from the Korean Broadcast News Speech (Strassel et al. 2006) were analyzed as Korean spontaneous speech. The results showed no significant differences in the first and second formants between epenthetic /i/ and lexical /i/; however, epenthetic /i/ had a shorter duration than lexical /i/ in spontaneous speech. This shorter duration is possibly a result of the less prominent nature of epenthetic /i/ in terms of duration compared to lexical /i/, and this difference becomes more noticeable in spontaneous speech where phonetic reduction is more common than in read speech where articulation is more precise. (University of Georgia)

Keywords Korean epenthetic vowels, Korean lexical vowels, speech styles, spontaneous speech, speech corpus

1. Introduction

Vowel epenthesis is a common phonological change in loanword adaptation in various languages. Many languages take advantage of vowel epenthesis as a repair strategy for adapting illicit non-native inputs (Kang 2003; Uffmann 2006; Hall 2011). Korean

also uses /i/ epenthesis to avoid illicit inputs in loanword adaptations. Specifically, in Korean, vowel epenthesis occurs: (a) to break up illicit consonant clusters, and (b) to avoid illicit simple codas.

In many languages, epenthetic vowels and their corresponding lexical vowels differ in the first formant (F1), the second formant (F2), and/or durations. For example, in American English, an epenthetic schwa has lower F2 compared to a lexical schwa (Kondon 1994; Flemming 2004; Davidson 2006). In addition, Miner (1979) demonstrates that epenthetic vowels in Winnebago are shorter than lexical vowels. Generally, two factors have been considered as reasons for the acoustic differences between epenthetic and lexical vowels. First, the acoustic differences may be related to 'undershoot', a phenomenon in which a speaker does not fully reach the target articulation of a sound (Van Son and Pols 1992; Flemming 2002; Davidson 2006). For example, Davidson (2006) argues that native English speakers attempt to produce foreign consonant clusters with the English epenthetic schwa as closely as possible to those produced by native speakers of the target language, and this effort leads to a shrinking of the vowel space, causing 'undershoot'. As a result, the English epenthetic schwa exhibits different acoustic features from its lexical schwa. Second, the underspecification of epenthetic vowels may account for the acoustic differences between epenthetic and lexical vowels. According to previous studies (Browman and Goldstein 1990; Smorodinsky 2002), epenthetic vowels may be inherently underspecified for target articulation. That is, epenthetic vowels lack specific articulatory targets and are shaped by the surrounding linguistic environment, much like transitional vowels. As a result, the formants of epenthetic vowels vary more based on neighboring segments than those of lexical vowels. Korean epenthetic /i/ aligns with this view. According to Oh (1992), the epenthetic /i/ is specified only as [+high] in the underlying form, and it lacks a specified [backness] which is relevant to F2. Thus, it is possible that F2 of epenthetic /i/ may vary more significantly than that of lexical /i/ depending on the adjacent segments. When vowel backness is unspecified, its contextual variability can extend to specified height due to the interdependence of articulatory dimensions. Earlier studies (Fant 1960; Keating 1988) argue that F1 and F2 are inherently linked by the physical constraints of the vocal tract, as tongue positioning for backness and height is interdependent. Consequently, unspecified backness can influence height, even when height is fully specified, causing both F1 and F2 to vary dynamically. In addition, the lack of a fully specified

articulatory target can also result in short vowel durations. If a vowel does not have a fully specified articulatory gesture, the articulators transition quickly between the surrounding consonantal gestures instead of fully forming a vowel gesture. This can result in a shorter vowel duration compared to a fully specified vowel (Hall 2006).

Previous studies (Kim 2009; Kim and Kochetov 2011) on the acoustic differences between epenthetic /i/ and lexical /i/ in Korean show that they have the same F1, F2 and duration. According to Kim and Kochetov (2011), vowel formants and durations of epenthetic /i/ and lexical /i/ are the same in read speech produced by native standard Korean speakers. In addition, epenthetic /i/ shows the same contextual variations as lexical /i/ does. Specifically, the place of articulation of the preceding consonant has significant effects on F2 of both epenthetic /i/ and lexical /i/ in Korean. F2 is lowered when the preceding consonant is labial and raised when the preceding consonant is coronal due to the coarticulatory effects of consonant place on nearby vowels. Also, Kim (2009) investigates the acoustic differences between epenthetic and lexical vowels in word internal positions in a Kyungsang dialect in Korean and found no acoustic differences between epenthetic /i/ and lexical /i/ in this dialect.

It is well known that speech styles can affect the acoustic features of vowels (Koopmans-van Beinum 1980; Van Bergem and Koopmans-van Beinum 1989; Harmegnies and Poch-Olivé 1992). For example, Harmegnies and Poch-Olivé (1992) show that Spanish vowels exhibit a stronger schwa tendency in spontaneous speech. Van Bergem and Koopmans-van Beinum (1989) also argue that Dutch vowels become more reduced in spontaneous speech than in formal speech. Vowel durations can also change depending on speech style (Smiljanic and Bradlow 2008; DiCanio et al. 2015; Bellik 2019). For instance, Smiljanic and Bradlow (2008) found that both tense and lax vowels are shorter in conversational speech than in clear speech in English and Croatian. More importantly, lexical and non-lexical vowels can exhibit distinct behaviors depending on speech style. Bellik (2019) argues that as speech transitions from careful to casual, Turkish lexical vowels are shortened, whereas non-lexical vowels remain unchanged.

This study explores whether meaningful acoustic differences exist between epenthetic /i/ and lexical /i/ in Korean spontaneous speech, as previous studies on this subject have primarily focused on Korean read speech. Spontaneous speech such as conversational, casual, fast, and natural is different from read speech (Tucker and Mukai 2023: 3). In spontaneous speech, phonetic reduction such as segment deletion,

shortening, or partial articulation may occur more frequently (Greenberg 1999; Warner and Tucker 2011; Tucker and Mukai 2023). This makes it worthwhile to investigate the acoustic differences between epenthetic /i/ and lexical /i/ in Korean spontaneous speech, as phonetic reduction may influence both. Accordingly, this study hypothesizes that (a) epenthetic /i/ has a shorter duration than lexical /i/, and (b) epenthetic /i/ exhibits greater contextual variation in vowel formants than lexical /i/ depending on the influence of surrounding consonants.

2. Methods

2.1 Stimuli

This study focuses on comparing F1, F2, and vowel duration of epenthetic /i/ and lexical /i/ in Korean spontaneous speech. To this end, political debates which have been used as spontaneous speech in many studies (Bruce and Touati 1992; Llisterri 1992; Kim et al. 2012; Dufour et al. 2014) were analyzed. The data came from the corpus ‘Korean Broadcast News Speech’ (Strassel et al. 2006). Four political debates include two and a half hours of satellite radio recordings by three male and three female Korean speakers. A total of 834 stimuli were used, as shown in Table 1.

Table 1. Total number of stimuli collected from three male and three female speakers

	Male	Female	Total
Lexical /i/	405	327	732
Epenthetic /i/	59	43	102
Total	464	370	834

All English loanwords and Korean lexical stimuli are listed in either the Standard Korean Language Dictionary or the Naver Dictionary. Stimuli appear in Appendix A.¹

¹ Due to the large number of stimuli, a representative sample set of 35 for each epenthetic /i/ and lexical /i/ is included in Appendix A.

2.2 Data labeling and extraction

To segment all audio files into words and phonemes, the Korean Phonetic Aligner Program Suite (Yoon and Kang 2013) was used in the study. The alignment results were shown in TextGrid files in Praat (Boersma and Weenink 2024), as shown in Figure 1.

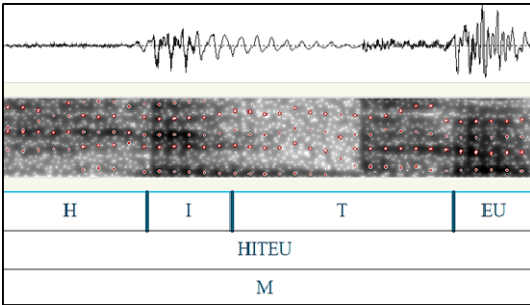


Figure 1. English loanword *hit*

Figure 1 shows epenthetic /i/ in English loanword *hit*. In Figure 1, ‘EU’ represents the target /i/, and ‘M’ stands for a male speaker (and ‘F’ stands for a female speaker). All stimulus alignments were checked manually based on the following measurement criteria: (a) the starting point of target vowel is the onset of F2 and (b) the end point of target vowel is the offset of F2 or the point where the spectrogram and waveform show a sudden change resulting from the following consonant or word-final silence (Renwick 2012). Once all alignments were manually corrected, a Praat script was used to extract the acoustic features of epenthetic /i/ and lexical /i/ from Praat TextGrid files. In the Praat script, five formants were detected with the ceiling of 5000 Hz for male speakers and 5500 Hz for female speakers, and the midpoints of vowels were measured for F1, F2, and F3.² Devoiced epenthetic /i/ and lexical /i/ sometimes occurred after a voiceless sound or between voiceless sounds, and they did not exhibit the typical vowel formants. Thus, all devoiced epenthetic /i/ and lexical /i/ were excluded from the analysis.

2 Since F1 and F2 values are considered the most significant for identifying vowels (Raina et al. 2014), F3 was not analyzed as an acoustic feature in this study.

2.3 Data analysis

The extracted acoustic data were organized in Excel, as shown in Tables 2.1 and 2.2.

Table 2.1. Sample dataset of epenthetic /i/ and lexical /i/³

Subject	Gender	Word	Origin	PrecedingC -Place	Preceding C-Aspiration
1	F	WEOKEUSJOPDO	Epenthetic	Dorsal	Aspirated
1	F	GEUREOM	Lexical	Dorsal	Unaspirated
⋮	⋮	⋮	⋮	⋮	⋮

Table 2.2. Sample dataset of epenthetic /i/ and lexical /i/

Following C-Place	Syllable Structure	Word Position	Prosodic Position	F1	F2	Duration
None	Open	Medial	Internal	437.08	1579.82	0.0432
None	Open	Initial	Initial	384.27	1684.19	0.0411
⋮	⋮	⋮	⋮	⋮	⋮	⋮

In Tables 2.1 and 2.2, the data include subject number, gender, word, origin (epenthetic or lexical), place of articulation of the preceding and following consonants (coronal, dorsal, or labial), aspiration of the preceding consonant (aspirated or unaspirated), syllable structure in which epenthetic /i/ or lexical /i/ occurs (closed or open), word position (initial, medial, or final), and prosodic position (initial or internal), F1, F2, and duration.

Generally, the place of articulation of the preceding consonant has a greater effect on the vowel formant values than that of the following consonant, and this influences F2 more than F1 (Stevens and House 1963; Hillenbrand et al. 2001). In this study, the places of articulation for both the preceding and following consonants were included in the data. Also, the aspiration of the preceding consonant can influence the acoustic features of vowels; a vowel tends to be pronounced shorter after an aspirated consonant than after an unaspirated one (Chung et al. 1999). Therefore, the aspiration of the preceding consonant was included in the data for statistical analysis. Next, syllable structures may affect the acoustic features of vowels. Specifically, vowels generally have longer durations in open syllables than in closed syllables (Rositzke 1939; Monsen 1974; Choi and Jun 1998; Curtis 2002). Syllable structures

³ Note that due to space constraints, the original data table has been divided into two parts: Tables 2.1 and 2.2.

can also affect the vowel formants. Storme (2017) examined the acoustic features of close-mid and open-mid vowels in a French dialect where mid vowels follow the loi de position and found that mid vowels are consistently lower, and peripheral vowels are more centralized in closed syllables compared to open syllables. For this reason, the syllable structure was also included in the data. In addition, word positions can influence the acoustic characteristics of vowels. For example, vowels in word-final syllables tend to have longer durations compared to those in non-final syllables (O'Shaughnessy 1980). Thus, in this study, epenthetic and lexical vowels were categorized based on their positions within words: initial, medial, and final. Lastly, the acoustic characteristics of vowels are affected by their position within an utterance. According to Chung et al. (1999), Korean vowels are longest in clause-final positions and are longer in phrase-initial positions than in phrase-internal positions. However, in the data, there were no instances of epenthetic /i/ and lexical /i/ occurring in clause-final positions, as speakers tended to use formal clause-ending suffixes such as [ta] and [jo] in political debates. As a result, epenthetic /i/ and lexical /i/ were categorized as either phrase-initial or phrase-internal based on their prosodic position.

3. Results

3.1 F1 and F2

In this section, F1 and F2 of epenthetic /i/ and lexical /i/ were examined. First, the following plots show F1 and F2 of epenthetic /i/ and lexical /i/ in male and female spontaneous speech.

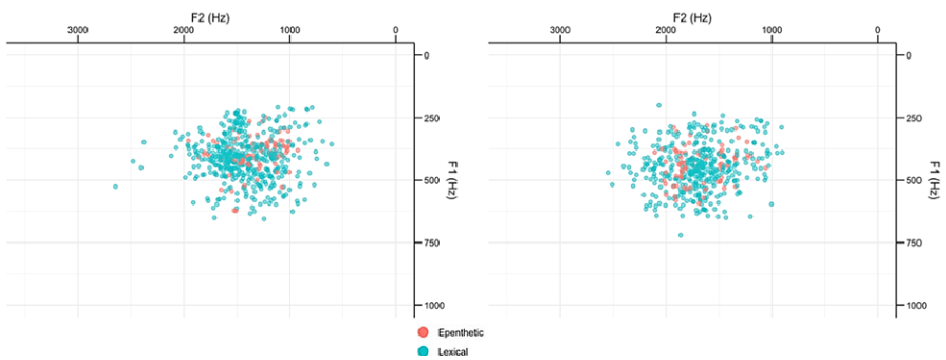


Figure 2. F1 and F2 (Hz) of epenthetic /i/ and lexical /i/ for male speakers (left) and female speakers (right)

Table 3. Mean and standard deviation (SD) of F1 and F2 (Hz) in male and female speakers

	Male		Female	
	F1	F2	F1	F2
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Epenthetic /i/	391.34 (25.71)	1381.01 (135.22)	439.28 (31.65)	1688.72 (137.28)
Lexical/i/	411.38 (34.09)	1423.68 (128.96)	451.54 (27.59)	1696.57 (116.41)

Figures 2.1 and 2.2 illustrate the distribution of F1 and F2 for epenthetic /i/ and lexical /i/. In male speech, F1 of epenthetic /i/ ranges from 250 Hz to 622 Hz, and its F2 ranges from 926 Hz to 1862 Hz. F1 of lexical /i/ spans from 208 Hz to 655 Hz, and its F2 ranges from 652 Hz to 2645 Hz. In female speech, F1 of epenthetic /i/ ranges from 281 Hz to 592 Hz, and its F2 ranges from 1060 Hz to 2106 Hz. Also, F1 of lexical /i/ ranges from 202 Hz to 719 Hz, and its F2 ranges from 890 Hz to 2600 Hz.

Table 3 presents the mean formant values and SD of F1 and F2. The mean F1 and F2 values of lexical /i/ are higher than those of epenthetic /i/ in both male and female speech. To determine if the numerical differences between epenthetic /i/ and lexical /i/ were statistically significant, a statistical analysis was performed using the lmer() function from the lme4 package (Bates et al. 2024) in R (R Core Team 2024). Several mixed-effects models with different fixed and random effects were formulated based on data, and ANOVA was used to determine which model provided the best

fit. The best mixed-effects model had four independent variables as fixed effects. Fixed effects included gender (male or female), place of articulation of the preceding consonant (coronal, dorsal, or labial), aspiration of the preceding consonant (aspirated or unaspirated), and origin (epenthetic or lexical). Also, this model included random effects 'subject' and 'word'.

Table 4.1. Mixed-effects model results for F1

	Estimate	Std.Error	t-value	Pr(> t)
(Intercept)	442.037	22.594	19.559	<0.0001 ***
Gender Male	-42.394	15.502	-2.708	0.0069 **
Place Coronal	-34.549	28.432	-1.214	0.2252
Place Labial	-25.891	24.786	-1.045	0.2971
Origin Lexical	11.239	20.567	0.547	0.5449
Aspiration Yes	-10.508	25.345	-0.415	0.6785

Table 4.2. Mixed-effects model results for F2

	Estimate	Std.Error	t-value	Pr(> t)
(Intercept)	1579.039	50.124	31.505	<0.0001 ***
Gender Male	-284.711	40.579	-7.017	<0.0001 ***
Place Coronal	197.436	31.049	6.529	<0.0001 ***
Place Labial	-26.933	17.741	-4.743	0.0403 *
Origin Lexical	21.985	49.992	1.001	0.6620
Aspiration Yes	32.149	58.938	0.534	0.5855

As presented in Table 4.1, only gender has a significant effect on F1 ($p = 0.0069$), indicating that male speakers exhibit lower F1 than female speakers. On the other hand, origin ($p = 0.5449$), preceding consonant (coronal: $p = 0.2252$, labial: $p = 0.2971$), and aspiration ($p = 0.6785$) did not show significant effects on F1. In other words, the F1 values of epenthetic /i/ and lexical /i/ are not significantly different from each other and do not vary based on the place of articulation or the aspiration of the preceding consonant.

Next, Table 4.2 shows the statistical results for F2. In Table 4.2, origin and aspiration of the preceding consonant are not significant factors (origin: $p = 0.6620$, aspiration: $p = 0.5855$), indicating that F2 of epenthetic /i/ and lexical /i/ are statistically similar and do not vary under these conditions. On the other hand, gender and place of articulation of the preceding consonant are significant factors (gender: $p < 0.0001$, coronal: $p < 0.0001$, labial: $p = 0.0403$). These results indicate that (a) male speakers

produce lower F2 than female speakers, (b) F2 is higher when the place of articulation of the preceding consonant is coronal, and (c) F2 is lower when it is labial.

Figure 3 and Tables 5.1 and 5.2 show the mean and SD of epenthetic /i/ and lexical /i/ based on the place of articulation of the preceding consonant.

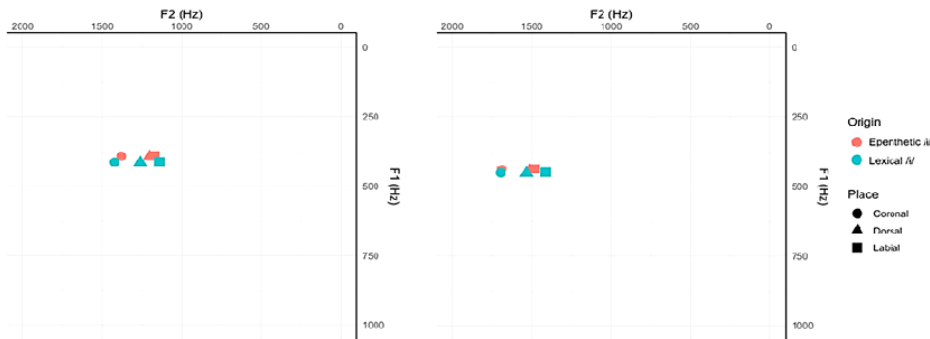


Figure 3. Mean F1 and F2 (Hz) of epenthetic /i/ and lexical /i/ for male speakers (left) and female speakers (right)

Table 5.1. Mean F1 and F2 (Hz) with SD for epenthetic /i/ and lexical /i/ in male speakers by place of articulation of the preceding consonant

Epenthetic /i/				Lexical /i/			
	Coronal	Dorsal	Labial		Coronal	Dorsal	Labial
F1	391.12 (28.58)	390.83 (29.02)	388.95 (29.35)	F1	411.38 (43.23)	413.02 (39.49)	410.19 (33.81)
F2	1380.04 (122.04)	1207.52 (135.39)	1168.71 (145.85)	F2	1423.68 (124.49)	1261.58 (133.39)	1138.38 (147.62)

Table 5.2. Mean F1 and F2 (Hz) with SD for epenthetic /i/ and lexical /i/ in female speakers by place of articulation of the preceding consonant

Epenthetic /i/				Lexical /i/			
	Coronal	Dorsal	Labial		Coronal	Dorsal	Labial
F1	434.38 (35.25)	445.96 (31.67)	432.24 (33.48)	F1	451.54 (37.32)	453.20 (36.14)	449.71 (34.87)
F2	1692.14 (165.31)	1512.39 (145.87)	1479.89 (159.05)	F2	1696.57 (147.43)	1534.47 (155.28)	1411.27 (150.94)

In Tables 5.1 and 5.2, the mean F2 is highest after coronal consonants and lowest after labial consonants for both epenthetic /i/ and lexical /i/ across genders. To further explore the differences in F2 values among coronal, dorsal, and labial places of articulation in epenthetic /i/ and lexical /i/, Tukey's pairwise comparisons from the

emmeans package (Lenth 2024) were conducted in R (R Core Team 2024).

Table 6. Tukey's pairwise comparisons of F2 by place of articulation of the preceding consonant

	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Epenthetic/i/	Coronal – Labial	212.252	19.457	10.911	0.0021 **
	Coronal – Dorsal	179.741	17.824	10.087	0.0034 **
	Labial – Dorsal	-32.510	11.362	-2.860	0.0213 *
	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Lexical/i/	Coronal – Labial	285.289	24.387	11.697	0.0015 **
	Coronal – Dorsal	162.109	14.456	11.213	0.0027 **
	Labial – Dorsal	-123.201	13.092	-9.410	0.0168 *

As Table 6 shows, when the preceding consonant is coronal, F2 is higher than when it is either labial or dorsal for both epenthetic /i/ (labial: $p = 0.0021$, dorsal: $p = 0.0034$) and lexical /i/ (labial: $p = 0.0015$, dorsal: $p = 0.0027$).

When the preceding consonant is labial, F2 is lower than when it is dorsal for both epenthetic /i/ ($p = 0.0213$) and lexical /i/ ($p = 0.0168$). That is, the F2 values of epenthetic /i/ and lexical /i/ are highest after the coronal consonant, followed by the dorsal consonant, and lowest after the labial consonant. To explore whether significant F2 differences exist between epenthetic /i/ and lexical /i/ due to the place of articulation, Tukey's pairwise comparisons were conducted.

Table 7. Tukey's pairwise comparisons of F2 by place of articulation between epenthetic /i/ and lexical /i/

Place	Origin	Estimate	Std.Error	t.ratio	Pr(> z)
Coronal	Epenthetic – Lexical	-4.428	16.230	-0.272	0.7853
Dorsal	Epenthetic – Lexical	-22.079	15.873	-1.390	0.1643
Labial	Epenthetic – Lexical	68.621	23.459	2.926	0.0852

In Table 7, the results indicate that the numerical differences in F2 between epenthetic /i/ and lexical /i/ are not statistically significant within each place condition (coronal: $p = 0.7853$, dorsal: $p = 0.1643$, labial: $p = 0.0852$). This means that the F2 formant values for epenthetic /i/ and lexical /i/ are similar for each specific place of articulation.

To summarize, F1 and F2 of epenthetic /i/ and lexical /i/ were analyzed in this section, and the results confirmed that these acoustic features are not meaningfully different from each other, and the F2 values of both vowel types vary similarly

depending on the place of articulation of the preceding consonant.

3.2 Vowel duration

The durations of epenthetic /i/ and lexical /i/ were measured to compare their temporal properties. Figure 4 shows the mean durations of epenthetic /i/ and lexical /i/. The vowel durations were transformed into logarithmic form to manage skewed data and reduce the effect of outliers.



Figure 4. Mean vowel durations (logarithmic scale) of epenthetic /i/ and lexical /i/

Table 8. Mean (sec) and SD of epenthetic /i/ and lexical /i/

Epenthetic /i/		Lexical /i/	
Mean	SD	Mean	SD
0.0415	0.0171	0.0506	0.0158

As shown in Figure 4 and Table 8, lexical /i/ has a longer duration than epenthetic /i/. To investigate whether this durational difference is statistically significant, the mixed-effects model was applied to vowel duration using the lmer() function in R (R Core Team 2024). ANOVA was used to select the best-fit model among several mixed-effects models with varying fixed and random effects. The best mixed-effects model included five independent variables as fixed effects: gender (male or female), place of articulation of the preceding consonant (coronal, dorsal, or labial), aspiration of the preceding consonant (aspirated or unaspirated), origin (epenthetic or lexical), and prosodic position (phrase-initial or phrase-internal). Additionally, the model

incorporated random effects for 'subject' and 'word'.

Table 9. Mixed-effects model results for vowel duration

	Estimate	Std.Error	t-value	Pr(> t)
(Intercept)	0.0565	0.0032	7.6562	<0.0001 ***
Gender Male	-0.0053	0.0024	-2.2083	0.0412 *
Place Coronal	-0.0018	0.0015	-1.2009	0.2318
Place Labial	0.0022	0.0017	1.2941	0.1978
Origin Lexical	0.0035	0.0021	1.7499	0.0423 *
Aspiration Yes	-0.0076	0.0029	-2.6207	0.0087 **
Phrase Internal	-0.0030	0.0008	-3.8214	0.0002 **

In Table 9, the statistical results show that the place of articulation of the preceding consonant does not significantly influence vowel duration (coronal: $p = 0.2318$, labial: $p = 0.1978$). On the other hand, the results confirm that (a) male speakers produce shorter vowel durations than female speakers ($p = 0.0412$), (b) lexical /i/ has a longer duration than epenthetic /i/ ($p = 0.0423$), (c) the vowel duration is longer when the preceding consonant is unaspirated compared to when it is aspirated ($p = 0.0087$), and (d) vowels in phrase-internal positions have shorter durations compared to those in phrase-initial positions ($p = 0.0002$). To investigate whether these tendencies apply equally to both epenthetic /i/ and lexical /i/, Tukey's pairwise comparisons were conducted.

First, as Table 10 shows, vowel durations are shorter after aspirated consonants than after unaspirated consonants in both epenthetic /i/ and lexical /i/ (epenthetic /i/: $p = 0.0042$, lexical /i/: $p = 0.0368$).

Table 10. Tukey's pairwise comparisons of vowel durations by aspiration of the preceding consonant

	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Epenthetic /i/	Yes – No	-0.0061	0.0026	-2.3461	0.0042 **
	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Lexical /i/	Yes – No	-0.0042	0.0023	-2.0870	0.0368 *

Next, both epenthetic /i/ and lexical /i/ in the phrase-internal position are shorter (epenthetic: $p = 0.0163$, lexical: $p = 0.0093$) as shown in Table 11.

Table 11. Tukey's pairwise comparisons of vowel durations by prosodic position

	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Epenthetic /i/	Phrase Internal – Phrase Initial	−0.0032	0.0013	−2.4614	0.0163 *
	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Lexical /i/	Phrase Internal – Phrase Initial	−0.0028	0.0009	−3.1101	0.0093 **

Gender is also statistically significant for vowel durations ($p = 0.0412$) in Table 9, indicating that vowel durations produced by male speakers are shorter than those produced by female speakers. The following table presents the mean durations and SD of epenthetic /i/ and lexical /i/ by gender.

Table 12. Mean (sec) and SD by gender

Epenthetic /i/				Lexical /i/			
Male		Female		Male		Female	
Mean	SD	Mean	SD	Mean	SD	Mean	SD
0.0399	0.0121	0.0431	0.0142	0.0485	0.0112	0.0525	0.0136

In Table 12, females produce longer vowel durations than males in both epenthetic /i/ and lexical /i/. Tukey's pairwise comparisons were used to check whether these differences between male and female speakers were statistically significant for both epenthetic /i/ and lexical /i/.

Table 13. Tukey's pairwise comparisons of vowel durations by gender

	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Epenthetic /i/	Female – Male	0.0032	0.0008	3.7490	0.0019 **
	Contrast	Estimate	Std.Error	t.ratio	Pr(> z)
Lexical /i/	Female – Male	0.0041	0.0011	3.8011	0.0005 **

In Table 13, females produce both epenthetic /i/ and lexical /i/ with longer durations than males (epenthetic /i/: $p = 0.0019$, lexical /i/: $p = 0.0005$).

In summary, this section examined the vowel durations of epenthetic /i/ and lexical /i/, and the results showed that epenthetic /i/ is shorter than lexical /i/. On the other hand, epenthetic /i/ and lexical /i/ exhibited similar behavior depending on linguistic factors. Both epenthetic /i/ and lexical /i/ are longer after unaspirated consonants than after aspirated ones. In addition, the vowel durations of both epenthetic /i/ and

lexical /i/ were influenced by prosodic contexts: phrase-initial position resulted in longer durations than phrase-internal position. Finally, female speakers produced both epenthetic /i/ and lexical /i/ with longer durations than male speakers.

4. General discussions and limitations

This study aimed to investigate the acoustic differences between epenthetic /i/ and lexical /i/ in Korean spontaneous speech and hypothesized that epenthetic /i/ would show greater contextual variation in formants and a shorter duration compared to lexical /i/. First, regarding the greater contextual variation in formants, the results showed that the preceding consonant did not affect the F1 values of either epenthetic /i/ or lexical /i/. In contrast, the preceding consonant influenced F2 of both epenthetic /i/ and lexical /i/. Specifically, F2 was highest when the preceding consonant was coronal, followed by dorsal, and lowest when it was labial, a tendency also found in Korean read speech (Kim and Kochetov 2011). Many studies show that the place of articulation of the preceding consonant affects the following vowels (Cooper et al. 1952; Kondo 1994; Koopmans-van Beinum 1994). This may be due to the influence of coarticulation, a phenomenon in which the shape of the vocal tract for one sound affects the production of the preceding and/or following sound (Marchal 2009; Goldstein 2010). For example, earlier studies (Cooper et al. 1952; Liberman et al. 1954; Kishon-Rabin et al. 2003; Kerdpol 2012; Jachova et al. 2021) show that consonants with different places of articulation may have different F2 values. Specifically, a coronal consonant has a higher F2 compared to a labial consonant. These variations in F2, depending on the place of articulation, may affect the following vowel.

With regard to the expected greater variation in epenthetic /i/ compared to lexical /i/, the results indicated that the F2 of epenthetic /i/ did not exhibit greater variation than that of lexical /i/. Both vowels showed similar levels of variation depending on the place of articulation of the preceding consonant. Consequently, the findings in this study suggest that F1 and F2 of epenthetic /i/ and lexical /i/ are identical in Korean spontaneous speech.

Importantly, this study found the durational difference between epenthetic /i/ and lexical /i/ in Korean spontaneous speech. Specifically, epenthetic /i/ has a shorter

duration than lexical /i/. This durational difference was not observed in studies based on read speech. The durational difference between epenthetic /i/ and lexical /i/ may arise from the different characteristics of spontaneous and read speech. According to previous studies (Trouvain et al. 2001; Schwab and Avanzi 2015; Tucker and Mukai 2023), spontaneous speech is characterized by a faster articulation rate and greater variability compared to read speech. For example, phonetic reduction such as shortening and deletion may occur in spontaneous speech (Greenberg 1999; Warner and Tucker 2011; Tucker and Mukai 2023). In contrast, in read speech, speakers tend to articulate sounds more consciously with minimal phonetic reduction which may result in diminished durational differentiation between vowels. Therefore, it is possible that the durational difference between epenthetic /i/ and lexical /i/ is minimized in read speech.

The difference in durations between epenthetic /i/ and lexical /i/ may be attributed to their distinct roles in Korean. In many languages including Korean, epenthetic vowels are generally used to satisfy phonotactic constraints rather than to convey lexical meaning (Kang 2003; Uffmann 2006; Hall 2011). As a result, epenthetic vowels often exhibit less phonetic prominence such as shorter durations and/or greater centralization in formants, compared to lexical vowels (Miner 1979; Kondo 1994; Flemming 2004; Davidson 2006). In Korean, while both epenthetic /i/ and lexical /i/ share the same F1 and F2 values, epenthetic /i/ may be produced more quickly due to its non-lexical nature. This tendency is further amplified in spontaneous speech where phonetic reduction is more pronounced. As a result, epenthetic /i/ may have a shorter duration than lexical /i/ in this study.

Prosodic contexts influenced the durations of both epenthetic /i/ and lexical /i/: longer durations were observed in phrase-initial positions. This prosodic effect has been found cross-linguistically and known as domain-initial strengthening (Byrd and Saltzman 2003; Tabain 2003; Cho and Keating 2009). Segments in domain-initial positions are enhanced, exhibiting features such as longer durations. These enhancements make the boundaries more salient, effectively signaling the start of a new phrase. In other words, this kind of effect can be understood as increasing the salience of the boundary for the listener and aiding in the recognition and processing of prosodic structure (Byrd and Saltzman 2003).

Also, this study demonstrates that the vowel durations of female speakers are longer than those of male speakers in both epenthetic /i/ and lexical /i/. This finding is

consistent with previous studies (Simpson and Ericsson 1998; Ericsson and Ericsson 2001) on gender-specific durational differences. For example, Simpson and Ericsson (1998) found that vowel durations of female speakers of German and American English are longer than those of male speakers. Simpson and Ericsson (1998) argue that gender-specific durational differences may result from differences in the vocal tracts of males and females. Specifically, biomechanical differences in vocal tracts of males and females may cause different articulatory distances when achieving the same phonetic targets. The larger vocal tract dimensions in male speakers require greater articulatory speed which leads to faster speech rates than in female speakers. As a result, males tend to speak more quickly and produce shorter vowel durations.

In this study, a total of 102 epenthetic /i/ were analyzed, which is significantly fewer than the 732 lexical /i/ instances, so it is necessary to acknowledge the limitations of the number of analyzed epenthetic /i/ instances. Additionally, to make the results more reliable, the same speakers should participate in both the read speech and spontaneous speech tasks, and the results from these tasks should be compared to reduce variability between speakers. Also, the English loanword stimuli used in this study have been frequently used in Korean, which increases the likelihood that they have already been nativized. As a result, observing significant phonetic differences between epenthetic /i/ and lexical /i/ may be challenging. Although eliciting non-word stimuli in spontaneous speech is difficult, investigating epenthetic vowels through non-word English stimuli remains a valuable avenue for research.

5. Conclusion

Speech styles affect the acoustic features of vowels in many languages (Koopmans-van Beinum 1980; Harmegnies and Poch-Olivé 1992) and can also influence durations of lexical and non-lexical vowels differently (Bellik 2019). Most studies on the acoustic differences between epenthetic /i/ and lexical /i/ in Korean focused on read speech; therefore, this study analyzed epenthetic /i/ and lexical /i/ in Korean spontaneous speech. The results show that while there is no difference between the F1 values of epenthetic /i/ and lexical /i/, nor between their F2 values, epenthetic /i/ is produced with a shorter duration than lexical /i/ in spontaneous speech. This durational

difference is possibly due to the less prominent nature of epenthetic /i/. This becomes more apparent in spontaneous speech where phonetic reduction is more pronounced than it is in read speech where articulation is clearer.

References

- Bates, Douglas, Martin Mächler, Ben Bolker, and Steven Walker. 2024. Fitting linear mixed-effects models using lme4, R package version 1.1-36. <https://cran.r-project.org/package=lme4>.
- Bellik, Jennifer A. 2019. *Vowel intrusion in Turkish onset clusters*. Phd Dissertation. University of California, Santa Cruz.
- Boersma, Paul and David Weenink. 2024. Praat: Doing phonetics by computer [Computer program], Version 6.4.19. <https://www.fon.hum.uva.nl/praat/>.
- Browman, Catherine and Louis Goldstein. 1990. Representation and reality: Physical systems and phonological structure. *Journal of Phonetics* 18(3): 411-424.
- Bruce, Gösta and Paul Touati. 1992. On the analysis of prosody in spontaneous speech with exemplification from Swedish and French. *Speech Communication* 11(4-5): 453-458.
- Byrd, Dani and Elliot Saltzman. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31(2): 149-180. [https://doi.org/10.1016/S0095-4470\(02\)00085-2](https://doi.org/10.1016/S0095-4470(02)00085-2).
- Cho, Taehong and Patricia Keating. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37(4): 466-485. <https://doi.org/10.1016/j.wocn.2009.08.001>.
- Choi, Seong-Won and Jong-Ho Jun. 1998. Are Korean fortis and aspirated consonants geminates? *Linguistic Research* 34(3): 521-546.
- Chung, Hyun-Song, Kyong-Sok Kim, and Mark Huckvale. 1999. Consonantal and prosodic influences on Korean vowel duration. *Proceedings of Eurospeech 1999*, 1-4. Grenoble: ISCA.
- Cooper, Franklin S., Pierre C. Delattre, Alvin M. Liberman, John M. Borst, and Louis J. Gerstman. 1952. Some experiments on the perception of synthetic speech sounds. *The Journal of the Acoustical Society of America* 24: 597-606.
- Curtis, Emily. 2002. Moraic structure and segment duration in Korean. *Japanese/Korean Linguistics* 10: 378-391.
- Davidson, Lisa. 2006. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34(1): 104-137. <https://doi.org/10.1016/j.wocn.2005.03.004>.
- DiCanio, Christian, Ho-Sung Nam, Jonathan D. Amith, Rey Castillo García, and Douglas Whalen. 2015. Vowel variability in elicited versus spontaneous speech: Evidence from Mixtec. *Journal of Phonetics* 48: 45-59. <https://doi.org/10.1016/j.wocn.2014.10.003>.
- Dufour, Richard, Yannick Estève, and Paul Deléglise. 2014. Characterizing and detecting sponta-

- neous speech: Application to speaker role recognition. *Speech Communication* 56: 1-18.
<https://doi.org/10.1016/j.specom.2013.07.007>.
- Ericsson, Christine and Anna M. Ericsson. 2001. Gender differences in vowel duration in read Swedish: Preliminary results. *Working Papers* 49: 34-37.
- Fant, Gunnar. 1960. *Acoustic theory of speech production: With calculations based on X-Ray studies of Russian articulations*. Hague: Mouton.
- Flemming, Edward. 2002. *Auditory representations in phonology*. New York, NY: Routledge.
- Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.), *Phonetically-based phonology*, 232-276. Cambridge: Cambridge University Press.
- Goldstein, Bruce. 2010. *Sensation and perception*. Belmont: Wadsworth, Cengage Learning.
- Gouskova, Maria and Nancy Hall. 2009. Acoustics of epenthetic vowels in Lebanese Arabic. In Steve Parker (eds.), *Phonological argumentation: Essays on evidence and motivation*, 203-225. London: Equinox.
- Greenberg, Steven. 1999. Speaking in shorthand—A syllable-centric perspective for understanding pronunciation variation. *Speech Communication* 29(2): 159-176.
- Hall, Nancy. 2006. Cross-linguistic patterns of vowel intrusion. *Phonology* 23(3): 387-429.
<https://doi.org/10.1017/S0952675706000996>.
- Hall, Nancy. 2011. Vowel epenthesis. *The Blackwell Companion to Phonology*, 1-21.
- Hall, Nancy. 2013. Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics* 41(2): 133-143. <https://doi.org/10.1016/j.wocn.2012.12.001>.
- Harmegnies, Bernard and Dolors Poch-Olivé. 1992. A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. *Speech Communication* 11(4): 429-437.
- Hillenbrand, James M., Michael J. Clark, and Terrance M. Nearey. 2001. Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America* 109(2): 748-763. <https://psycnet.apa.org/doi/10.1121/1.1337959>.
- Keating, Patricia. 1988. Underspecification in phonetics. *Phonology* 5(2): 275-292.
- Kerdpol, Karnthida. 2012. Formant transitions as effective cues to differentiate the places of articulation of Ban Pa La-u Sgaw Karen Nasals. *Journal of Humanities Regular* 15(2): 21-38.
- Kim, Hyun-Ju. 2009. The phonology and phonetics of unaccentable vowels in Kyungsang Korean. *The Japanese/Korean Linguistics* 18: 343-354.
- Kim, Kyung-Min. 2008. Coronals in epenthesis in loanwords with comparison to velars. *Toronto Working Papers in Linguistics* 30: 53-68.
- Kim, Kyung-Min and Kochetov Alexei. 2011. Phonology and phonetics of epenthetic vowels in loanwords: Experimental evidence from Korean. *Lingua* 121(3): 511-532.
- Kim, Samuel, Fabio Valente, and Alessandro Vinciarelli. 2012. Automatic detection of conflicts in spoken conversations: Ratings and analysis of broadcast political debates. *Proceedings of International Conference on Acoustics, Speech and Signal Processing 2012*, 5089-5092. Piscataway: IEEE.

- Kishon-Rabin, Liat, Meytal Dayan, and Orit Michaeli. 2003. Effect of second-formant transitions on the perception of Hebrew voiced stop consonants. *Journal of Basic and Clinical Physiology and Pharmacology* 14(2): 151-164.
- Kondo, Yuko. 1994. Targetless schwa: Is that how we get the impression of stress-timing in English? *Proceedings of the Edinburgh Linguistics Department Conference 94*, 63-76. Edinburgh: Edinburgh University Press.
- Koopmans-van Beinum, Florian. 1980. *Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions*. Phd Dissertation. University of Amsterdam.
- Koopmans-van Beinum, Florian. 1994. What's in a schwa?: Durational and spectral analysis of natural continuous speech and diphones in Dutch. *Phonetica* 51(1-3): 68-79. <https://doi.org/10.1159/000261959>.
- Lehmann, Winfred and Raymond M. Heffner. 1943. Notes on the length of vowels. *American Speech* 18: 208-215.
- Lenth, Russell V. 2024. Emmeans: Estimated marginal means. R package version 1.6.3. <https://CRAN.R-project.org/package=emmeans>.
- Lieberman, Alvin M., Pierre C. Delattre, Franklin S. Cooper, and Louis J. Gerstman. 1954. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied* 68(8): 1-13.
- Llisterri, Joaquim. 1992. Speaking styles in speech research. *Proceedings of Workshop on Integrating Speech and Natural Language*, 1-28. Stroudsburg: ACL.
- Marchal, Alain. 2009. *From speech physiology to linguistic phonetics*. London: ISTE.
- Miner, Kenneth. 1979. Dorsey's law in Winnebago-Chiwere and Winnebago accent. *International Journal of American Linguistics* 45(1): 25-33.
- Monsen, Randall B. 1974. Durational aspects of vowel production in the speech of deaf children. *Journal of Speech and Hearing Research* 17(3): 386-398.
- Oh, Mi-Ra. 1992. Palatal consonants, labial consonants, and vowel epenthesis in Korean loanword. *Linguistics Journal of Korean* 17(1): 141-161.
- O'Shaughnessy, Douglas. 1981. A study of French vowel and consonant durations. *Journal of Phonetics* 9(4): 385-406.
- Raina, Diksha, Sangita Chakraborty, and Makarand R. Velankar. 2014. Automatic classification of instrumental music and human voice using formant analysis. *International Journal of Advanced Research in Computer Science & Technology* 2(2): 242-245.
- Renwick, Margaret E. L. 2012. *Vowels of Romanian. Historical, phonological and phonetic studies*. Phd Dissertation. Cornell University.
- Rositzke, Harry A. 1938. Vowel length in general American speech. *Language* 15(2): 99-109.
- Schwab, Sandra and Mathieu Avanzi. 2015. Regional variation and articulation rate in French. *Journal of Phonetics* 48: 96-105. <https://doi.org/10.1016/j.wocn.2014.10.009>.
- Shin, Dong-Jin and Paul Iverson. 2014. Phonetic investigation of epenthetic vowels produced

- by Korean learners of English. *Journal of the Korean Society of Speech Sciences* 6(4): 17-26.
- Simpson, Adrian P. and Christine Ericsdotter. 1998. Sex-specific durational differences in English and Swedish. *Proceedings of 15th international congress of phonetic sciences*, 1113-1116. Barcelona: Universitat Autònoma de Barcelona.
- Smiljanic, Rajka and Ann R. Bradlow. 2008. Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics* 36(1): 91-113.
- Smorodinsky, Iris. 2002. *Schwas with and without active gestural control*. Phd Dissertation. Yale University.
- Stevens, Kenneth N. and Arthur S. House. 1963. Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research* 6(2): 111-128.
- Storme, Benjamin. 2017. The loi de position and the acoustics of French mid vowels. *Glossa: A Journal of General Linguistics* 2(1): 1-25. <https://doi.org/10.5334/gjgl.300>.
- Strassel, Stephanie, Nii Martey, and David Graff. 2006. Korean Broadcast News Speech. <https://catalog.ldc.upenn.edu/LDC2006S42>. (24 August, 2006)
- Summers, W. Van. 1987. Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *Journal of the Acoustical Society of America* 82: 847-863.
- Susman, Amelia L. 1943. *The accentual system of Winnebago*. Phd Dissertation. Columbia University.
- Tabain, Marija. 2003. Effects of prosodic boundary on /aC/ sequences: Articulatory results. *The Journal of the Acoustical Society of America* 113(5): 2834-2849. <https://doi.org/10.1121/1.1564013>.
- Trouvain, Jürgen, Jacques Koreman, Attilio Erriquez, and Bettina Braun. 2001. Articulation rate measures and their relation to phone classification in spontaneous and read German speech. *Proceedings of the 16th International Congress of the Phonetic Sciences*, 155-158. Saarbrücken: Saarland University Press.
- Tucker, Benjamin V. and Yoichi Mukai. 2023. *Spontaneous speech*. Cambridge: Cambridge University Press.
- Uffmann, Christian. 2006. Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua* 116(7): 1079-1111.
- Uffmann, Christian. 2007. *Vowel epenthesis in loanword adaptation*. Berlin: Max Niemeyer Verlag.
- Van Bergem, Dick and Florian Koopmans-van Beinum. 1989. Vowel reduction in natural speech. *Proceedings of Eurospeech 1989*, 2285-2288. Grenoble: ISCA.
- Van Son, Rob and Louis C. W. Pols. 1992. Formant movements of Dutch vowels in a text, read at normal and fast rate. *The Journal of the Acoustical Society of America* 92(1): 121-127. <https://doi.org/10.1121/1.404277>.
- Warner, Natasha and Benjamin V. Tucker. 2011. Phonetic variability of stops and flaps in spontaneous and careful speech. *The Journal of the Acoustical Society of America* 130(3): 1606-1617.

Yoon, Tae-Jin and Yoon-Jung Kang. 2013. The Korean phonetic aligner program suite.
<https://korean.utsc.utoronto.ca/kpa/>.

Yun, Il-Sung. 2009. Vowel duration and the feature of the following consonant. *Phonetics and Speech Sciences* 1(1): 41-46.

Appendix A: Stimuli for epenthetic /i/ and lexical /i/

Epenthetic /i/	Lexical /i/
APATEUGA 아파트가	ADEUGBADEUG 아득바득
BEURAENDEU 브랜드	APEUGEDO 아프게도
CEKEUREUL 체크를	BOGEUBSOREAL 보급소를
GEURAEPEUDO 그래프도	BUDEUREOUN 부드러운
GOLPEUREUL 골프를	CHUSEUREUGO 추스르고
HAEPEUNEENGEURO 해프닝으로	DEULRINEUN 들리는
EENSENTEEBEUGA 인센티브가	DEUREEN 드린
EENPEULAGA 인프라가	DEUDKONEUN 들고는
KAEMPEUREUL 캠프를	DEUGSEELDO 득실도
KADEU 카드	DEUNOPIN 드높인
KEUREINEE 크레인이	EEDEUKDO 이득도
RUTEU 루트	GADEUK 가득
MAIKEUGA 마이크가	GEUPGEUPHAESEO 급급해서
NETEUWEOKEUDO 네트워크도	GEUREOM 그림
NEGEOTEEBEU 네거티브	GEUMJU 금주
NOTEUGA 노트	GEUGSOSU 극소수
PAIPEUREUL 파이프를	HEUTEUREOJEEM 호트러짐
PATEUNEODO 파트너도	JEENGEUPDO 진급도
PEURAIBOSEE 프라이버시	JAGEUGJEOGEN 자극적인
PEURO 프로	KEUGEEGA 크기가
PEUROSESEUGA 프로세스가	KEULSUROG 클수록
POINTEU 포인트	KEUPNEEDA 겁니다
SELPEU 셀프	MAEDEUBDO 매듭도
SEULEOMPEU 슬럼프	MAEDEUJISNEUN 매듭짓는
SEUKEURAEBDO 스크램도	MODEUN 모든
SEUKEULIBTEUGA 스크립트가	NEUREEN 느린
SEUPEGTEUREOMEE 스펙트럼이	NEULSANG 늘상
SEUTEURAIKEU 스트라이크	SEUBDOGA 습도가
SEUTEURESEUGA 스트레스가	SEUSEUREOM 스스럼
SOPEUTEUWEEO 소프트웨어	SEULPEUBNEEDA 슬픔니다
TEIBEUL 테이블	SODEUGSUJUN 소득수준
TEUREOGEE 트럭이	SUGEUBDO 수급도
TEURAEGEUL 트랙을	TEUJEEBEE 트집이

TEUROPEENEUN 트로피는	TEUGBYEOLHEE 특별히
WEOKEUSJOPDO 워크숍도	TEUGSUHAN 특수한

Hyunjin Lee
Ph.D. Student
Department of Linguistics
University of Georgia
142 Gilbert Hall,
Athens, GA 30602, USA
E-mail: hyunjin.lee25@uga.edu

Received: 2024. 09. 24.
Revised: 2025. 02. 13.
Accepted: 2025. 02. 16.