Korean Journal of English Language and Linguistics, **Vol 26, February 2026, pp. 329-362**
DOI: 10.15738/kjell.26..202602.329

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Speech Rate Effects on the Phonetic Realization of Epenthetic Vowels in English Loanwords in Korean*

Hyunjin Lee (Pusan National University)

Received: September 11, 2025
Revised: January 12, 2026
Accepted: January 12, 2026

Lee, Hyunjin
Lecturer, Department of Language and Information
Pusan National University
2 Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan, Korea
Tel: +82-51-512-0311
Email: hjlee66904@gmail.com

* An earlier version of this work was presented at the Summer 2025 Meeting of the Korean Association of Language Sciences (KALS).

ABSTRACT

Lee, Hyunjin. 2026. Speech rate effects on the phonetic realization of epenthetic vowels in English loanwords in Korean. *Korean Journal of English Language and Linguistics 26*, 329-362.

This study investigates the phonetic realization of epenthetic vowels in English loanwords in Korean under different speech rate conditions, based on read speech. Thirty native speakers of Seoul Korean participated in a production experiment in which they read sentences at normal and fast rates. Both epenthetic /ɨ, i/ and their lexical counterparts were analyzed for F1, F2, and duration values. The results revealed that epenthetic and lexical vowels did not differ significantly in their formant values or durations at either speech rate. While vowel durations shortened overall in fast speech, no evidence of centralization was observed for epenthetic and lexical /ɨ/, suggesting that it maintains stable phonetic properties across rates. Gender, syllable position and syllable structure exerted significant effects on vowel duration, but these applied equally to epenthetic and lexical vowels. These findings indicate that epenthetic vowels are fully integrated into the phonological system and function as phonological segments with defined phonetic targets, rather than as transitional vowels in read speech.

KEYWORDS

English loanwords in Korean, vowel epenthesis, speech rate

1. Introduction

Across languages, lexical items borrowed from a donor language often undergo phonological repair strategies so that they conform to the sound system of the borrowing language. One of the most widespread repair strategies observed in loanword phonology is vowel epenthesis, where a vowel is inserted to resolve phonotactically illicit structures such as consonant clusters (Hall 2011, Kang 2003, Uffmann 2006). In Korean, vowel epenthesis either resolves illicit consonant clusters or prevents the occurrence of disallowed coda consonants in the adaptation of English loanwords. Specifically, since consonant clusters are impermissible in Korean, speakers produce epenthetic /ɨ/ to break up consonant clusters in English loanwords. In addition, given that Korean codas are limited to [p], [t], [k], [n], [m], [ŋ], and [l], epenthetic /ɨ/ is also inserted to block other consonants from surfacing in coda position. In certain contexts, epenthetic /i/ is also employed, specifically following the English post-alveolar consonants /ʒ/, /ʃ/, /ʤ/, and /ʧ/. In a wide range of languages, epenthetic vowels often differ phonetically from their lexical counterparts (Bellik 2019, Davidson 2006, Gouskova and Hall 2009, Hall 2013, Miner 1979). Davidson (2006), for example, found that English speakers frequently use an epenthetic schwa to resolve illicit consonant clusters, and that this epenthetic vowel is typically shorter in duration with lower F1 and F2 than a lexical schwa. Similarly, Gouskova and Hall (2009) demonstrate that in Lebanese Arabic, epenthetic /i/ is markedly shorter than lexical /i/. In contrast, Korean does not show such distinctions. For example, studies by Kim (2009) and Kim and Kochetov (2011) indicate that epenthetic and lexical vowels in Korean display comparable F1, F2 and vowel durations in sentence-level read speech.

Speech rate is known to affect the phonetic realization of vowels. Numerous studies have documented vowel centralization at faster speaking rates in languages such as English (Agwuele et al. 2008, Miller 1981, Turner et al. 1995), Spanish (Nadeu 2014), and Korean (Igeta et al. 2017, Lee et al. 2003). Within a single language, however, vowels may respond differently to changes in rate. Hirata and Tsukada (2003), for example, investigated F1 and F2 of five short Japanese vowels /i, e, a, o, u/ and their long counterparts /iː, eː, aː, oː, uː/ across slow and fast rates. Their results revealed greater effects of speech rate on some vowels than on others. In particular, speech rate has a greater impact on the short mid vowels /e/ and /o/ than on other vowels; under fast speech conditions, their F1 formant values showed more centralization. Speech rate also affects vowel duration (Gay 1978, Magen and Blumstein 1993, Smith 2002), with vowels generally becoming shorter at faster rates than at slower or normal ones. For instance, studies on Korean have shown that its eight lexical vowels /i, e, ɛ, ə, u, ɨ, o, ɑ/ are consistently shorter in fast speech than in slow speech (Magen and Blumstein 1993).

Research on the phonetic characteristics of Korean lexical vowels across different speech rates remains relatively limited. For instance, Lee et al. (2003) showed that while vowel durations decrease as speaking rate increases, the F1 and F2 values of Korean monophthongs showed no systematic differences across slow, normal, and fast speech-rate conditions. By contrast, Son (2017) investigated Korean lexical vowels /ɑ, i, u, æ, ɛ, o, ʌ, ɨ/ and found that the formant values of /ɑ, i, u, ɛ, o/ changed as speech rate increased, with the magnitude of change varying across vowels. Igeta et al. (2017) investigated Korean back vowels /u, o/ and reported increased centralization with faster speech rates, particularly for female speakers. However, previous studies have been conducted under a range of methodological conditions that may limit the generalizability of the findings. In particular, earlier research has often relied on relatively small speaker samples, included speakers of a single gender, or involved speakers with heterogeneous regional dialect backgrounds. As a result, it remains unclear to what extent the reported effects reflect general properties of Korean speech or are influenced by participant-related factors such as sample size or dialectal variation.

With respect to epenthetic vowels, more importantly, there have been almost no studies examining the effects of speech rate on them. Yet, speech rate may influence these two types of vowels in distinct ways. One possibility is that, in contrast to lexical vowels, epenthetic vowels vary in their occurrence across different speech rates. For instance, Tajima et al. (2002) showed that epenthetic vowels may be omitted at faster speech rates in Japanese-accented English. Another possibility is that speech rate can either reduce or amplify the phonetic differences between epenthetic and lexical vowels. Gouskova and Hall (2009) reported that in Lebanese Arabic, lexical vowels are longer than epenthetic vowels in slow speech, but this difference vanishes at faster rates (Hall 2013). On the other hand, speech rate may also create new differences: Bellik (2019) found that while Turkish lexical vowels shorten in fast, casual speech relative to careful speech, non-lexical vowels remain constant because they are already reduced to the minimal duration set by Turkish gestural timing.

Proficiency in the source language significantly shapes how loanwords are produced and perceived, often leading to variation in the realization of the same borrowed forms (Best and Tyler 2007, Paradis and LaCharité 2012). For example, differences between bilingual and monolingual speakers in loanword production have been widely documented (LaCharité and Paradis 2005, Kadenge and Mudzingwa 2012). Kadenge and Mudzingwa (2012) report that in the adaptation of English loanwords into chiShona, bilingual speakers may produce segments not found in the chiShona sound system. Since chiShona lacks /l/, monolingual speakers generally replace it with /r/, whereas bilingual speakers may preserve /l/, reflecting the influence of English phonology. Accordingly, although the present study investigates native Korean speakers, it also examines whether patterns vary as a function of English proficiency.

Therefore, the present study aims to provide a more systematic examination of the phonetic characteristics of epenthetic and lexical vowels across normal and fast speech rates. Slow speech was not included because it can elicit artificial hyper-articulation, such as exaggerated separation of syllable boundaries or unnaturally prolonged vowel durations, which could obscure naturally occurring speech-rate variation. Accordingly, the present study investigates whether epenthetic vowels exhibit similar or different phonetic realizations within each speech-rate condition (normal vs. fast). If epenthetic vowels are fully phonological segments, they are expected to exhibit formant and durational patterns comparable to those of lexical vowels across changes in speech rate.

2. Methods

2.1 Participants

All thirty participants were native speakers of Seoul Korean, born and raised in Seoul, South Korea; 15 were women and 15 were men (M = 31.6 years, SD = 2.51, range = 27–35). This selection reduces potential dialectal variation that might otherwise influence the production data. Of these participants, twenty-five were residing in Seoul at the time of the study and had never lived in an English-speaking country, while the remaining five had spent time in Georgia, USA (mean duration = 3.4 months). According to responses on the English Proficiency Self-Rating Questionnaire, however, even those with overseas experience continued to use Seoul Korean regularly in academic, professional, and social settings.

Thus, the five participants with overseas experience were also included in the analysis as part of the same participant group. The questionnaire was adapted from Park and Ziegler (2014), which is based on the Common European Framework of Reference for Languages (CEFR). It is available through the Instruments and Data for Research in Language Studies (IRIS) database (https://iris-database.org/) for research use. Participants also reported their current English proficiency across four skills—reading, listening, speaking, and writing—using five-point rating scales via a questionnaire. Scores from these four domains were summed to obtain an overall proficiency score. Based on the resulting composite scores, participants were classified into four CEFR-aligned proficiency groups: Group A (17–20; n = 8), Group B (13–16; n = 8), Group C (9–12; n = 7), and Group D (4–8; n = 7).

2.2 Procedures

The experiment began with the normal speech rate condition. In the normal rate condition, participants were asked to read each sentence twice at their natural speaking pace in Korean, as presented on a computer screen. The experiment consisted of separate blocks at the normal speech rate, presented in the following order: epenthetic /ɨ/, epenthetic /i/, lexical /ɨ/, and lexical /i/, with an optional five-minute break between blocks. The fast rate condition employed the same sequence, except that participants were instructed to read each sentence as rapidly as possible within a five-second window. In each trial, a beep sounded one second after pressing the Enter key, prompting participants to read the displayed sentence until a second beep occurred five seconds later.

Recordings for twenty-five participants were conducted in quiet environments, including a library study room and a seminar room in Seoul, Korea, while the remaining five participants were recorded in a sound-attenuated booth at the Linguistics Laboratory of the University of Georgia, USA. All recordings in the production experiments were made using a Marantz digital recorder and a Shure headworn dynamic microphone.

2.3 Stimuli

Stimuli were real words listed in the Standard Korean Language Dictionary (2024) or Naver Dictionary. Epenthetic /ɨ/ was controlled for preceding consonant place (alveolar vs. bilabial vs. velar), syllable structure (closed vs. open), and vowel position within the word (final syllable vs. non-final syllable).

In many languages, the place of articulation of a consonant influences F2 of the following vowel (Cooper et al. 1952, Kerdpol 2012, Kim and Kochetov 2011). In particular, vowels following labial consonants typically show lower F2, whereas those after coronal consonants display higher F2, reflecting coarticulatory effects of consonantal place on adjacent vowels (Kim and Kochetov 2011). Accordingly, in this study, epenthetic /ɨ/ was examined following alveolar (/t, t^h/), bilabial (/p, p^h/), and velar (/k, k^h/) consonants in English loanwords. The stimuli for epenthetic /ɨ/ were evenly distributed across three places of articulation, with 12 tokens each (e.g., six following /t/ and six following /t^h/ in the alveolar category), across alveolar, bilabial, and velar consonants. The same distribution applies to lexical /ɨ/.

Vowels are generally longer in open than closed syllables (Choi and Jun 1998, Curtis 2002, Monsen 1974). Hence, the stimuli included epenthetic /ɨ/ in both contexts. In closed syllables, /ɨ/ mainly occurred between a consonant and /l/ (e.g., /pl, tl, kl, bl, dl, gl/ → [p^hɨl], [t^hɨl], [k^hɨl], [pɨl], [tɨl], [kɨl]). Epenthetic /ɨ/ also occurs within the 'sm' cluster, but this cluster appears in only a few Korean loanwords. Thus, epenthetic /ɨ/ in closed syllables primarily surfaced between a consonant and /l/ in this study. In addition, since vowel position affects duration, with word-final vowels generally lengthened (van Santen 1992, Windmann et al. 2015, White et al. 2020), epenthetic /ɨ/ was categorized in this study as occurring in either non-final or final syllables. Table 1 shows the distribution of epenthetic /ɨ/ stimuli by syllable structure and vowel position, and the same distribution applies to lexical /ɨ/.

Table 1. Distribution of Stimuli by Syllable Structure and Vowel Position

Syllable structure	Vowel position	Number of stimuli
CV	Non-word-final	12
CV	Word-final	12
CVC	Word-final	12

In the case of epenthetic /i/, it only occurs after English post-alveolar consonants and appears in word-final open syllables. Accordingly, 10 stimuli following /tɕ/ and 10 stimuli following /tɕʰ/ were used, and the same design applies to lexical /i/. All detailed stimuli are described in Appendices A and B, and they were embedded in the following Korean carrier sentence.

1) 우리는 각각 __X__ 발음을 하고 있습니다.
[ulinɨn kakkak __X__ palɨmɨl hako isɨmnita]
'We are individually pronouncing X.'

As illustrated in carrier sentence (1), each stimulus X was framed by the stop consonants /k/ and /p/, which respectively preceded and followed the target item to facilitate extraction (Magen and Blumstein 1993). To minimize potential bias from Korean orthography, all English loanword stimuli were provided in English.

All speech data were manually transcribed. Based on these transcriptions, audio files were segmented into words and phonemes using a Korean forced alignment tool (Yoon 2024). The resulting alignments were visualized as TextGrid files in Praat (Boersma and Weenink 2024). To identify and correct potential errors arising from the automatic alignment process, all alignments were manually inspected according to the following measurement criteria: (a) the onset of the target vowel was defined as the point at which F2 began, and (b) the offset of the target vowel was defined as the point at which F2 ended or the point at which a sudden change was observed in the spectrogram and waveform due to the following consonant. Formant values were then extracted at the midpoint of the vowel.

3. Results

3.1 Speech Rate

In this study, speech rate was defined as the number of syllables produced in the carrier sentence divided by the total sentence duration, measured from sentence onset to offset and including intervening pauses, yielding a measure of syllables per second (Laver 1994, Son 2017). At the group level, the mean sentence duration in the normal speech-rate condition was approximately 2.89s, whereas it was substantially shorter in the fast speech-rate condition, at approximately 1.86s. Accordingly, speech rate increased from approximately 5.71 syllables per second in the normal-rate condition to approximately 8.87 syllables per second in the fast-rate condition. This result is consistent with previous research (Kim 2009, Lee et al. 2003, Lee and Ko 2004, Son 2017). For example, Lee and Ko (2004) showed that the speech rate of normal Korean speech ranges from approximately 5.60 to 6.29 syllables per second, whereas fast speech is produced at a rate of approximately 7.04 to 8.14 syllables per second.

To determine whether the observed differences reflected inter-speaker variability in baseline speaking rate, the normal-rate and fast-rate conditions were compared at both the group and individual participant levels. Results showed that, for all participants, sentence durations were shorter and syllables-per-second values were higher in the fast-rate condition than in the normal-rate condition. This pattern was confirmed by a paired t-test on participant-level mean syllables per second, which revealed a significantly higher speech rate in the fast-rate condition. Together, these results indicate that the contrast between normal and fast speech rates was consistently maintained across participants, confirming the effectiveness of the speech-rate manipulation at both the group and individual levels.

3.2 Normal Speech Rate

3.2.1 Epenthetic /ɨ/ and lexical /ɨ/

For epenthetic /ɨ/, a total of 2,160 tokens were collected: (a) by place of articulation, the stimuli were evenly distributed across three categories, yielding 720 tokens per place of articulation and 2,160 tokens in total; and (b) by syllable structure and vowel position, 12 CV non-word-final stimuli yielded 720 tokens, 12 CV word-final stimuli yielded 720 tokens, and 12 CVC word-final stimuli yielded 720 tokens, resulting in a total of 2,160 tokens. However, 1 alveolar CV non-word-final token and 1 velar CVC word-final token produced by female speakers were phonetically unclear and therefore excluded from the analysis. In addition, 2 alveolar CV non-word-final tokens, 1 bilabial CV word-final token, and 1 bilabial CVC word-final token produced by male speakers were phonetically unclear and excluded. As a result, a total of 2,154 epenthetic /ɨ/ tokens were analyzed.

Next, for lexical /ɨ/, a total of 2,160 tokens were collected using the same design as for epenthetic /ɨ/. However, 1 bilabial CV non-word-final token, 1 bilabial CV word-final token, and 1 velar token produced by female speakers, as well as 1 alveolar CVC word-final token and 1 bilabial CV word-final token produced by male speakers, were not clear and therefore excluded from the analysis. Consequently, 2,155 lexical /ɨ/ tokens were included in the analysis. Figure 1a–b and Table 2 show the mean F1 and F2 values at the normal speech rate, with triangles and circles representing the mean values for individual speakers.

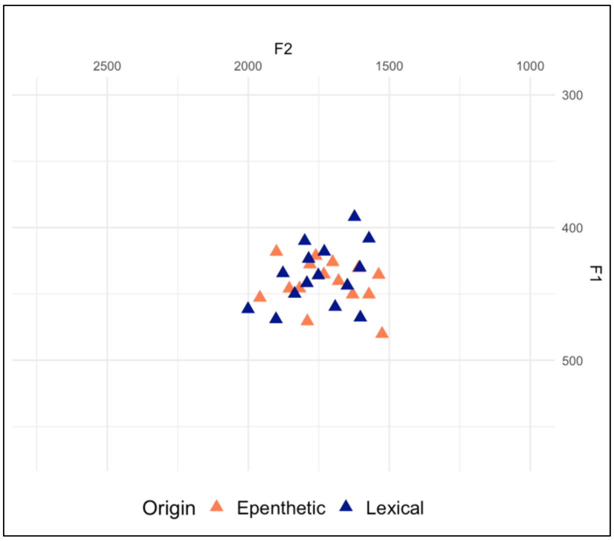

Figure 1a. Mean F1 and F2 (Hz) of Epenthetic and Lexical /ɨ/ for 15 Female Speakers at NSR (Normal Speech Rate)

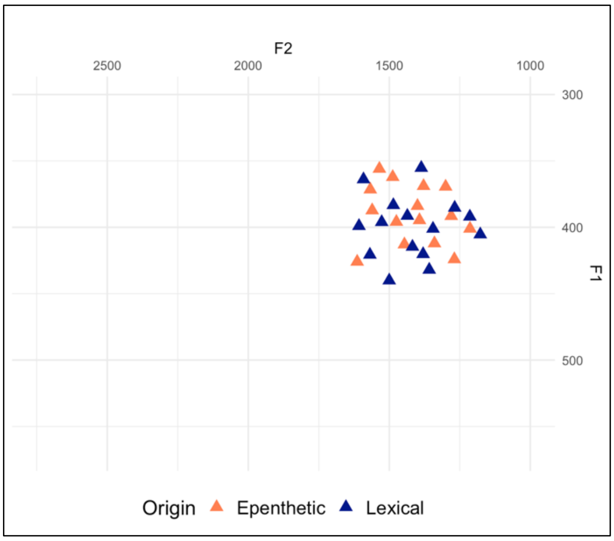

Figure 1b. Mean F1 and F2 (Hz) of Epenthetic and Lexical /ɨ/ for 15 Male Speakers at NSR

Table 2. Mean F1 and F2 (Hz) with Standard Deviation (SD) of Epenthetic and Lexical /ɨ/ at NSR

	Epenthetic /ɨ/		Lexical /ɨ/	
	Female	Male	Female	Male
F1	445.01 (35.05)	378.74 (30.98)	441.79 (33.59)	392.07 (30.28)
F2	1722.85 (124.82)	1418.74 (115.63)	1749.42 (118.14)	1405.32 (116.35)

For F1 and F2, mixed-effects linear regression models were fitted using the lmer() function in R (R Core Team 2025) to compare epenthetic and lexical /ɨ/. The fixed-effects structure was determined via likelihood ratio tests comparing nested models (anova), starting from an initial model that included gender (female vs. male)[1], place of articulation (alveolar vs. bilabial vs. velar), speech rate (normal vs. fast), vowel origin (epenthetic vs. lexical), and English proficiency group with random intercepts for Speaker and Word. The best-fitting model is reported in Table 3. Categorical predictors were treatment-coded with prespecified reference levels (female for gender, epenthetic vowels for vowel origin, and velar context for place of articulation); accordingly, the intercept represents the reference condition and fixed-effect estimates represent deviations from that reference. Random intercepts were included for Speaker and Word, with no random slopes.

[1] As the analyses in this study relied on non-normalized formants, gender was included as an independent factor to account for the fact that female speakers typically exhibit higher formant frequencies than male speakers due to shorter vocal tract length (Fant 1966, Kent and Read 2002, Peterson and Barney 1952).

Table 3. Results for F1 and F2 of Epenthetic and Lexical /ɨ/ at NSR

F1	Estimate	Std. Error	*t*-value	Pr (>| *t* |)
(Intercept)	451.0063	61.1072	7.3805	<0.0001 ***
Gender Male	−61.1987	43.4875	−1.4072	0.0359 *
Place Alveolar	23.3241	19.4996	1.1961	0.0688
Place Bilabial	−16.1649	16. 9483	−0.9537	0.3402
Origin Lexical	2.0734	11.3419	0.1820	0.8524
F2				
(Intercept)	1649.4215	106.2247	15.5276	<0.0001 ***
Gender Male	−344.1029	37.0853	−9.2786	<0.0001 ***
Place Alveolar	142.7627	59.2928	2.4077	0.0167 *
Place Bilabial	−37.4918	17.8692	−2.0981	0.0362 *
Origin Lexical	6.5812	29.5644	0.2226	0.7043

Table 3 shows no significant differences in F1 or F2 between epenthetic and lexical /ɨ/ (F1: p = 0.8524; F2: p = 0.7043). Gender effects were observed, with males producing lower F1 and F2 (F1: p = 0.0359; F2: p < 0.0001). Place of articulation influenced F2, raising values after alveolars (p = 0.0167) and lowering them after bilabials (p = 0.0362). Tukey's pairwise tests assessed F2 differences between epenthetic and lexical /ɨ/ across places of articulation.

Table 4. Tukey's Pairwise Comparisons of F2 across Places of Articulation at NSR

Female	Place	Comparison	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
	Alveolar	Epenthetic – Lexical	–12.0091	5.5546	–2.1621	0.1604
	Velar	Epenthetic – Lexical	1.1844	0.9908	1.1954	0.8527
	Bilabial	Epenthetic – Lexical	–19.8713	8.7460	–2.2720	0.3841
Male						
	Alveolar	Epenthetic – Lexical	9.2517	5.9454	1.5561	0.0925
	Velar	Epenthetic – Lexical	10.4934	6.4378	1.6299	0.1064
	Bilabial	Epenthetic – Lexical	5.5352	4.4963	1.2311	0.2186

In Table 4, no significant F2 differences were found between epenthetic and lexical /ɨ/ across places of articulation for either females (alveolar: p = 0.1604; velar: p = 0.8527; bilabial: p = 0.3841) or males (alveolar: p = 0.0925; velar: p =0.1064; bilabial: p = 0.2186).

Taken together, origin itself (epenthetic vs. lexical) did not have a meaningful effect on the F1 or F2 of /ɨ/ at the normal speech rate. The observed differences can be summarized as follows: across both epenthetic and lexical vowels, female speakers produced higher F1 and F2 values than male speakers, and F2 increased following alveolars and decreased following bilabials as a function of place of articulation. Also, the magnitude of these place-of-articulation-related F2 changes did not differ between epenthetic and lexical /ɨ/.

Figure 2 illustrates the log-transformed durations of epenthetic and lexical /ɨ/ for male and female speakers at a normal speech rate. Table 5 presents the mean durations of epenthetic and lexical /ɨ/, and Table 6 reports the corresponding statistical results based on the same log-transformed values.

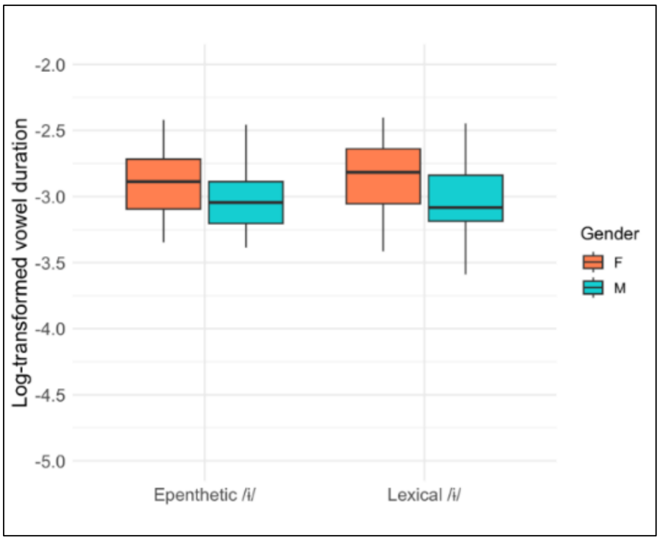

Figure 2. Vowel Durations (Log-transformed) of Epenthetic and Lexical /ɨ/ at NSR

Table 5. Mean Durations (s) with SD of Epenthetic and Lexical /ɨ/ at NSR

Normal rate	Epenthetic /ɨ/		Lexical /ɨ/	
	Female	Male	Female	Male
	0.0579 (0.0146)	0.0509 (0.0139)	0.0591 (0.0158)	0.0513 (0.0163)

Table 6. Results for Durations of Epenthetic and Lexical /ɨ/ at NSR

	Estimate	Std. Error	*t*-value	Pr (>| *t* |)
(Intercept)	0.0573	0.0041	13.9756	<0.0001 ***
Gender Male	−0.0072	0.0032	−2.2499	0.0476 *
Origin Lexical	0.0014	0.0011	1.2726	0.3276
Position Final	0.0071	0.0029	2.4483	0.0198 *
Structure Open	0.0016	0.0015	1.0667	0.3192
Rate Fast	−0.0315	0.0197	−1.5990	0.0058 **

(Note that this model was fit to the full dataset (normal + fast), with speech rate included as a fixed effect. Rate-specific results were obtained via Tukey post-hoc comparisons based on the best-fitting model reported in Table 6.)

Vowel duration was analyzed using mixed-effects linear regression in R (lmer; R Core Team 2025). Model selection was performed by comparing nested models using likelihood ratio tests (anova), beginning with an initial model that included gender[2] , origin, vowel position, syllable structure, English proficiency group, and speech rate as fixed effects and random intercepts for Speaker and Word. Table 6 presents the best-fitting model. Categorical predictors were specified using treatment coding with female, epenthetic /ɨ/, non-final position, closed syllables, and the normal speech rate set as the reference levels. Under this parameterization, the intercept corresponds to the predicted vowel duration for the reference condition, and each fixed-effect coefficient reflects

[2] Given that non-normalized duration measures were used, vowel duration was expected to vary by gender, as male speakers have been shown to produce shorter vowels than female speakers (Holt et al. 2015, Jacewicz and Fox 2010, Simpson and Ericsdotter 1998).

the deviation from that reference. Random intercepts were included for Speaker and Word, and no random slopes were specified. This model captured the combined influence of these variables across both normal and fast speech conditions, rather than isolating the normal rate. To examine effects specific to the normal rate, Tukey's pairwise comparisons were conducted using the best-fitting model. The analysis revealed no significant durational difference between epenthetic and lexical /ɨ/ at the normal speech rate (p = 0.9301). Also, Tukey's pairwise tests at the normal rate assessed whether epenthetic and lexical /ɨ/ showed similar patterns across factors.

Table 7. Tukey's Pairwise Comparisons of Durations by Gender, Vowel Position, and Syllable Structure at NSR

Origin	Comparison (gender)	Estimate	Std. Error	t-value	Pr (>|t|)
Epenthetic /ɨ/	Male – Female	−0.0064	0.0028	−2.2857	0.0301 *
Lexical /ɨ/	Male – Female	−0.0076	0.0031	−2.4516	0.0225 *
Origin	Comparison (vowel position)	Estimate	Std. Error	t-value	Pr (>|t|)
Epenthetic /ɨ/	Non final – Final	−0.0119	0.0044	−2.7045	0.0124 *
Lexical /ɨ/	Non final – Final	−0.0091	0.0043	−2.1163	0.0388 *
Origin	Comparison (syllable structure)	Estimate	Std. Error	t-value	Pr (>|t|)
Epenthetic /ɨ/	Open – Closed	0.0092	0.0043	2.1395	0.0383 *
Lexical /ɨ/	Open – Closed	0.0088	0.0038	2.3158	0.0169 *

As shown in Table 7, both epenthetic and lexical /ɨ/ at the normal speech rate were significantly shorter when (a) produced by male speakers (epenthetic /ɨ/: p = 0.0301; lexical /ɨ/: p = 0.0225), (b) occurring in non-final positions (epenthetic /ɨ/: p = 0.0124; lexical /ɨ/: p = 0.0388), and (c) occurring in closed syllables (epenthetic /ɨ/: p = 0.0383; lexical /ɨ/: p = 0.0169).

3.2.2 Epenthetic /i/ and lexical /i/

A total of 1,200 epenthetic /i/ tokens were elicited (600 from female speakers and 600 from male speakers). Of these, 5 tokens from female speakers (4 following [tɕʰ] and 1 following [tɕ]) and 4 tokens from male speakers (2 following [tɕʰ] and 2 following [tɕ]) were excluded due to unclear productions. Similarly, 1,200 lexical /i/ tokens were collected, of which 1 token following [tɕʰ] from a female speaker and 5 tokens from male speakers (3 following [tɕʰ] and 2 following [tɕ]) were excluded due to unclear pronunciation. Thus, in total, 2,385 vowels (1,191 epenthetic and 1,194 lexical) remained for analysis. Figure 3a-b and Table 8 present the mean F1 and F2 values at the normal speech rate.

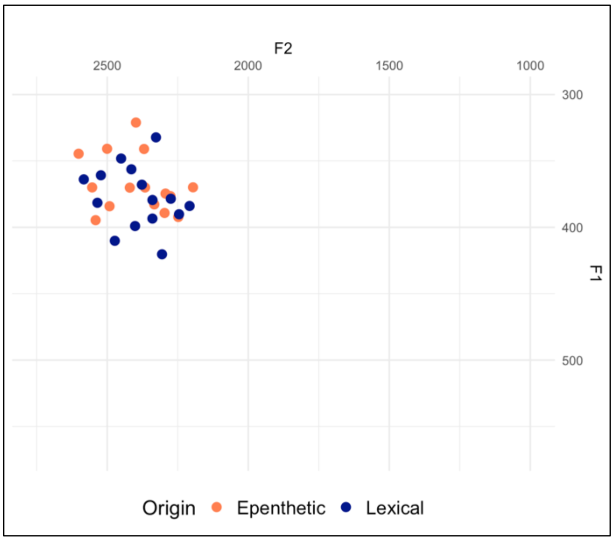

Figure 3a. Mean F1 and F2 (Hz) of Epenthetic and Lexical /i/ for 15 Female Speakers at NSR

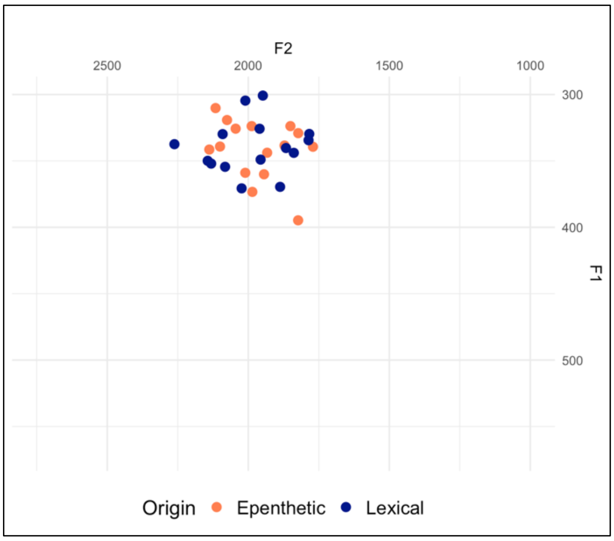

Figure 3b. Mean F1 and F2 (Hz) of Epenthetic and Lexical /i/ for 15 Male Speakers at NSR

Table 8. Mean F1 and F2 (Hz) with SD of Epenthetic and Lexical /i/ at NSR

	Epenthetic /i/		Lexical /i/	
	Female	Male	Female	Male
F1	368.01 (34.23)	328.76 (36.34)	379.59 (35.81)	339.51 (28.67)
F2	2394.43 (159.17)	1975.82 (139.61)	2382.64 (159.75)	1991.38 (135.85)

Mixed-effects linear regression models fitted with the lmer() function were used to examine differences in F1 and F2 between epenthetic and lexical /i/. The best-fitting model was selected by comparing nested models using likelihood ratio tests (anova) and included gender (female vs. male), origin (epenthetic vs. lexical), and speech rate (fast vs. normal) as fixed effects, with random intercepts for Speaker and Word. Categorical variables were coded using a treatment scheme, with the intercept corresponding to the predicted F1/F2 values for epenthetic /i/ produced by female speakers under the normal speech rate condition.

Table 9. Results for F1 and F2 of Epenthetic and Lexical /i/ at NSR

F1	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	378.6012	8.1051	46.7115	<0.0001 ***
Gender Male	−40.0083	10.3137	−3.8791	<0.0001 ***
Origin Lexical	11.1726	17.6256	0.6339	0.5262
Rate Fast	47.6917	16.4609	2.8973	0.0038 *
F2				
(Intercept)	2182.6143	35.7192	61.1048	<0.0001 ***
Gender Male	−391.2632	46.8915	−8.3440	<0.0001 ***
Origin Lexical	9.8967	33.3093	0.2971	0.7671
Rate Fast	−96.3413	29.7254	−3.2411	0.0012 **

According to Table 9, the best-fitting mixed-effects model incorporated speech rate as a fixed factor, reflecting effects across both normal and fast conditions rather than isolating the normal rate. To assess outcomes at the normal rate, Tukey's pairwise comparisons were conducted. Results indicated that epenthetic and lexical /i/ did not differ significantly in F1 or F2 for females (F1: $p = 0.8181$; F2: $p = 0.6159$) or for males (F1: $p = 0.6199$; F2: $p = 0.9361$) speakers.

Overall, vowel origin (epenthetic vs. lexical) did not significantly influence the F1 or F2 values of /i/ at the normal speech rate. With respect to gender, female speakers exhibited higher F1 and F2 values than male speakers across both epenthetic and lexical vowels.

Next, Figure 4 and Table 10 present the mean duration of epenthetic and lexical /i/, and Table 11 shows statistical results for their durations.

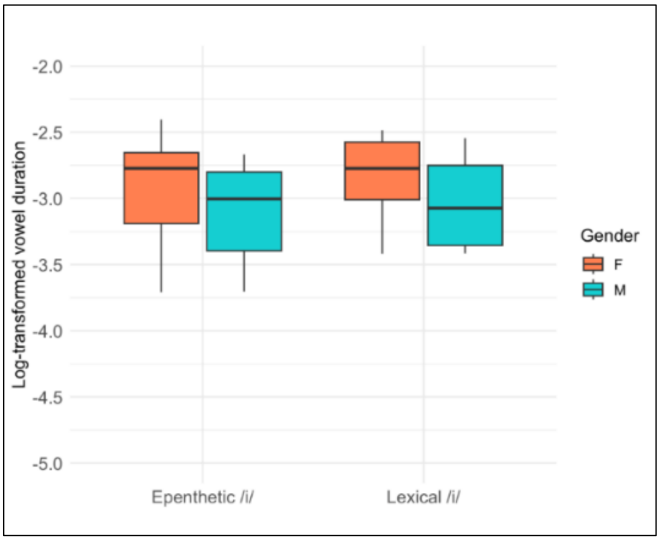

Figure 4. Vowel Durations (Log-transformed) of Epenthetic and Lexical /i/ at NSR

Table 10. Mean Durations (s) with SD of Epenthetic and Lexical /i/ at NSR

Normal rate	Epenthetic /i/		Lexical /i/	
	Female	Male	Female	Male
	0.0585 (0.0192)	0.0517 (0.0174)	0.0599 (0.0185)	0.0507 (0.0166)

Table 11. Results for Durations of Epenthetic and Lexical /i/ at NSR

	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	0.0463	0.0157	2.9490	0.0020 **
Gender Male	−0.0079	0.0032	−2.4688	0.0372 *
Origin Lexical	0.0051	0.0201	0.2537	0.8029
Rate Fast	−0.0174	0.0078	−2.2308	0.0021 **

Table 11 presents results reflecting the combined effects of variables across both normal and fast speech rates. To examine outcomes specific to the normal rate, Tukey's pairwise comparisons were carried out. The analysis showed no significant durational differences between epenthetic and lexical /i/ at the normal rate (p = 0.9640). Moreover, the comparisons revealed that both epenthetic and lexical /i/ were significantly shorter when produced by male speakers (epenthetic /i/: p = 0.0399; lexical /i/: p = 0.0401).

3.3 Fast Speech Rate

3.3.1 Epenthetic /ɨ/ and lexical /ɨ/

For epenthetic /ɨ/, female speakers produced 1,398 tokens, while male speakers produced 1,620 tokens. Of these, 12 tokens from female speech were excluded due to insufficient acoustic clarity (1 alveolar CVC word-final, 4 alveolar CV word-final, and 7 velar CVC word-final). Similarly, 25 tokens from male speech were excluded (2 alveolar CV non-word-final, 4 alveolar CV word-final, 1 bilabial CV non-word-final, 10 bilabial CVC word-final, 1 velar CVC word-final, 1 velar CV non-word-final, and 6 velar CVC word-final). After these exclusions, 2,981 epenthetic /ɨ/ tokens remained for analysis. For lexical /ɨ/, 1,391 tokens were produced by female speakers and 1,618 tokens by male speakers. Among these, 16 tokens from female speech were excluded because they were acoustically unclear (1 alveolar CVC non-word-final, 1 alveolar CV word-final, 3 bilabial CV non-word-final, 6 bilabial CV word-final, 2 velar CVC word-final, and 3 velar CV word-final). In addition, 19 tokens from male speech were excluded (1 alveolar CV word-final, 2 alveolar CV non-word-final, 6 alveolar CVC non-word-final, 3 bilabial CV non-word-final, 2 velar CVC word-final, and 5 velar CV word-final). Consequently, 1,375 tokens from female speakers and 1,599 tokens from male speakers were retained, yielding a total of 2,974 lexical /ɨ/ tokens for the final analysis. The mean F1 and F2 values at the normal speech rate are presented in Figures 5a-b and Table 12.

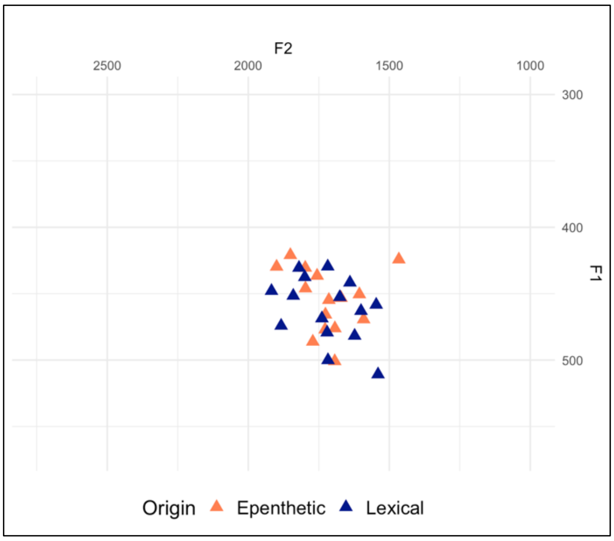

Figure 5a. Mean F1 and F2 (Hz) of Epenthetic and Lexical /ɨ/ for 15 Female Speakers at FSR (Fast Speech Rate)

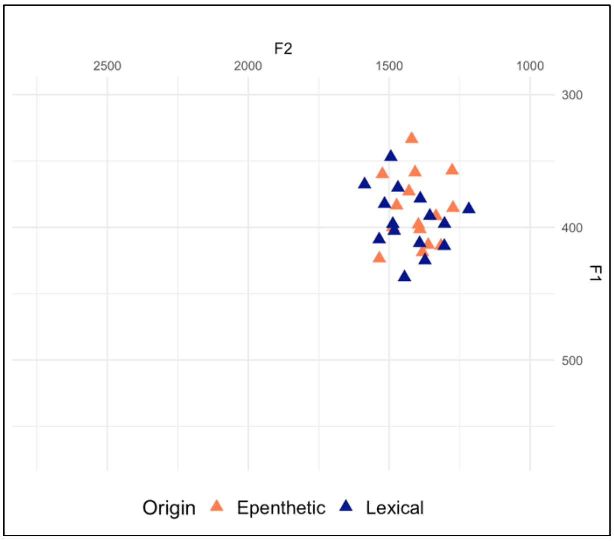

Figure 5b. Mean F1 and F2 (Hz) of Epenthetic and Lexical /ɨ/ for 15 Male Speakers at FSR

Table 12. Mean F1 and F2 (Hz) with SD of Epenthetic and Lexical /ɨ/ at FSR

	Epenthetic /ɨ/		Lexical /ɨ/	
	Female	Male	Female	Male
F1	459.89 (23.49)	386.74 (20.38)	462.91 (25.85)	397.73 (23.41)
F2	1725.25 (118.38)	1395.27 (102.29)	1710.54 (120.02)	1426.28 (115.82)

Mixed-effects models (lmer) were used to test differences in F1 and F2 between epenthetic and lexical /ɨ/. The anova-selected model, which was identical to the normal-rate model, included gender (female vs. male), place of articulation (alveolar vs. bilabial vs. velar), and origin (epenthetic vs. lexical) as fixed effects, with subject and word as random factors. Categorical variables were treatment-coded, such that the intercept represents the predicted F1/F2 values for epenthetic /ɨ/ in a velar context, produced by female speakers at the fast speech rate.

Table 13. Results for F1 and F2 of Epenthetic and Lexical /ɨ/ at FSR

F1	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	425.8025	15.6942	27.1312	<0.0001 ***
Gender Male	−50.3942	7.6024	−6.6287	<0.0001 ***
Place Alveolar	10.1396	14.5396	0.6974	0.4856
Place Bilabial	15.0911	12.0315	1.2543	0.2097
Origin Lexical	1.5329	4.6021	0.3331	0.7392
F2				
(Intercept)	1432.4623	49.6229	28.8670	<0.0001 ***
Gender Male	−363.7855	26.6742	−13.6381	<0.0001 ***
Place Alveolar	34.3417	30.2114	1.1367	0.2556
Place Bilabial	−20.5241	59.7426	−0.3435	0.7312
Origin Lexical	13.6098	41.3652	0.3290	0.7421

As in Table 13, gender significantly affected both F1 and F2, with male speakers producing lower values than females (F1: $p < 0.0001$; F2: $p < 0.0001$). No significant differences were observed between epenthetic and lexical /ɨ/ (F1: $p = 0.7392$; F2: $p = 0.7421$), and place of articulation did not significantly influence F2 at the fast speech rate (alveolar: $p = 0.2556$; bilabial: $p = 0.7312$). To further determine whether place of articulation consistently failed to affect the F2 of epenthetic and lexical /ɨ/, additional statistical tests were conducted.

Table 14. Tukey's Pairwise Comparisons of F2 across Places of Articulation at FSR

Female	Place	Comparison	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
	Alveolar	Epenthetic – Lexical	7.4531	5.5126	1.3521	0.1771
	Velar	Epenthetic – Lexical	–9.7128	7.0027	–1.3869	0.1655
	Bilabial	Epenthetic – Lexical	–2.5354	3.9678	–0.6390	0.5224
Male						
	Alveolar	Epenthetic – Lexical	–5.0214	4.1810	–1.2009	0.3078
	Velar	Epenthetic – Lexical	7.3487	6.2971	1.1672	0.2431
	Bilabial	Epenthetic – Lexical	–3.4925	4.7196	–0.7401	0.4569

In Table 14, F2 did not significantly differ between epenthetic and lexical /ɨ/ by place of articulation at the fast speech rate for either females (alveolar: $p = 0.1771$, velar: $p = 0.1655$, bilabial: $p = 0.5224$) or for males (alveolar: $p = 0.3078$, velar: $p = 0.2431$, bilabial: $p = 0.4569$).

Overall, vowel origin (epenthetic vs. lexical) did not exert a meaningful influence on the F1 or F2 values of /ɨ/ at the fast speech rate. Across both epenthetic and lexical vowels, female speakers exhibited higher F1 and F2 values than male speakers. However, F2 did not vary according to place of articulation, and the extent of any place-of-articulation-related F2 variation did not differ between epenthetic and lexical /ɨ/.

Figure 6 and Table 15 show the mean durations of epenthetic and lexical /ɨ/, with statistical results reported in Table 16. The best-fitting lmer model, selected via nested model comparisons using likelihood ratio tests (anova), included gender (female vs. male), origin (epenthetic vs. lexical), word position (final vs. non-final), syllable structure (closed vs. open), and speech rate (fast vs. normal) as fixed effects, with subject and word as random factors. A treatment-coding approach was applied to categorical predictors, such that the intercept reflects the expected vowel duration for epenthetic /ɨ/ produced by female speakers in closed syllables in non-final position at the normal speech rate.

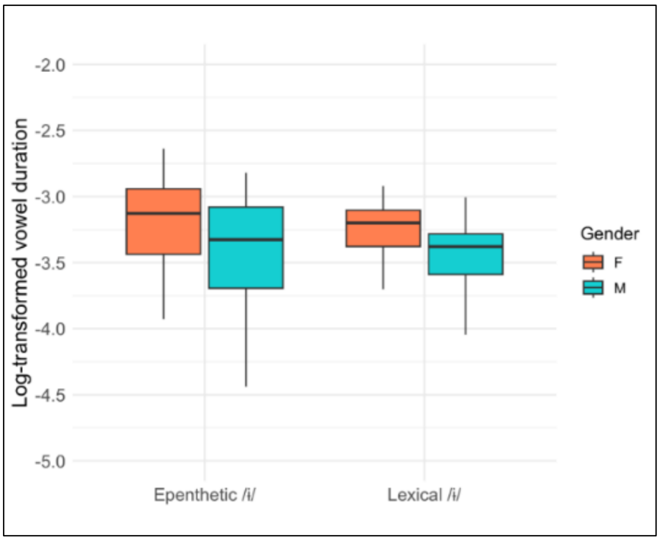

Figure 6. Vowel Durations (Log-transformed) of Epenthetic and Lexical /ɨ/ at FSR

Table 15. Mean Durations (s) with SD of Epenthetic and Lexical /ɨ/ at FSR

Fast rate	Epenthetic /ɨ/		Lexical /ɨ/	
	Female	Male	Female	Male
	0.0431 (0.0133)	0.0352 (0.0128)	0.0420 (0.0127)	0.0343 (0.0115)

Table 16. Results for Durations of Epenthetic and Lexical /ɨ/ at FSR

	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	0.0573	0.0041	13.9756	<0.0001 ***
Gender Male	−0.0072	0.0032	−2.2499	0.0476 *
Origin Lexical	0.0014	0.0011	1.2726	0.3276
Position Final	0.0071	0.0029	2.4483	0.0198 *
Structure Open	0.0016	0.0015	1.0667	0.3192
Rate Fast	−0.0315	0.0197	−1.5990	0.0058 **

Table 16 showed speech rate as a fixed effect; thus, Tukey’s tests were used to isolate the fast rate. No significant durational difference was found between epenthetic and lexical /ɨ/ (p = 0.9011). Further tests examined gender, vowel position, and syllable structure effects at the fast rate in Table 17.

Table 17. Tukey's Pairwise Comparisons of Durations by Gender, Vowel Position, Syllable Structure at FSR

Origin	Comparison (gender)	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
Epenthetic /ɨ/	Female – Male	0.0079	0.0032	2.4688	0.0143 *
Lexical /ɨ/	Female – Male	0.0064	0.0031	2.0645	0.0375 *
Origin	Comparison (vowel position)	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
Epenthetic /ɨ/	Non final – Final	−0.0013	0.0007	−1.8571	0.0712
Lexical /ɨ/	Non final – Final	0.0016	0.0009	1.7778	0.0699
Origin	Comparison (syllable structure)	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
Epenthetic /ɨ/	Open – Closed	0.0012	0.0007	1.7143	0.0931
Lexical /ɨ/	Open – Closed	−0.0015	0.0010	−1.5000	0.1302

As Table 17 presents, both epenthetic and lexical /ɨ/ were shorter when produced by male speakers (epenthetic /ɨ/: $p = 0.0143$; lexical /ɨ/: $p = 0.0375$). In contrast, vowel position had no effect (epenthetic /ɨ/: $p = 0.0712$; lexical /ɨ/: $p = 0.0699$) and syllable structure likewise did not influence vowel duration (epenthetic /ɨ/: $p = 0.0931$; lexical /ɨ/: $p = 0.1302$).

3.3.2 Epenthetic /i/ and lexical /i/

A total of 1,648 epenthetic /i/ tokens were obtained, including 766 tokens from female speakers and 882 tokens from male speakers. Among these, 2 tokens from female speakers (1 following [tɕʰ] and 1 following [tɕ]) and 6 tokens from male speakers (2 following [tɕʰ] and 4 following [tɕ]) were excluded because their productions lacked sufficient clarity. Additionally, 1,651 lexical /i/ tokens were collected (759 from female speakers and 892 from male speakers). Of these, three tokens from female speakers (two following [tɕʰ] and one following [tɕ]) and two tokens from male speakers (one following [tɕʰ] and one following [tɕ]) were excluded for the same reason. Following these exclusions, a total of 3,286 vowels—1,640 epenthetic and 1,646 lexical—were retained for the final analysis. Figures 7a-b and Table 18 present the mean F1 and F2 values at the fast speech rate.

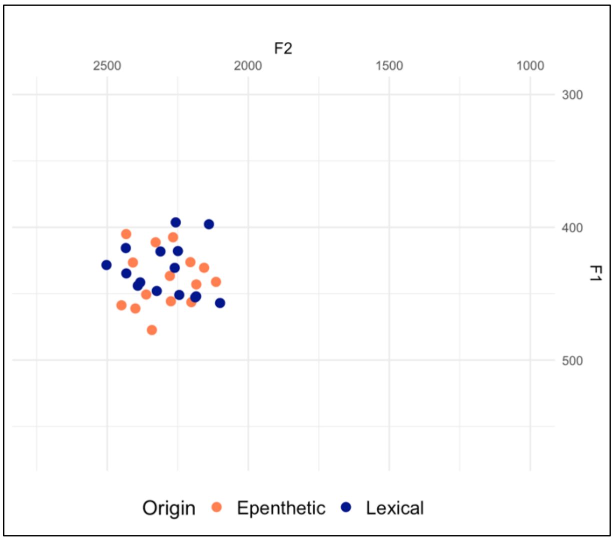

Figure 7a. Mean F1 and F2 (Hz) of Epenthetic and Lexical /i/ for 15 Female Speakers at FSR

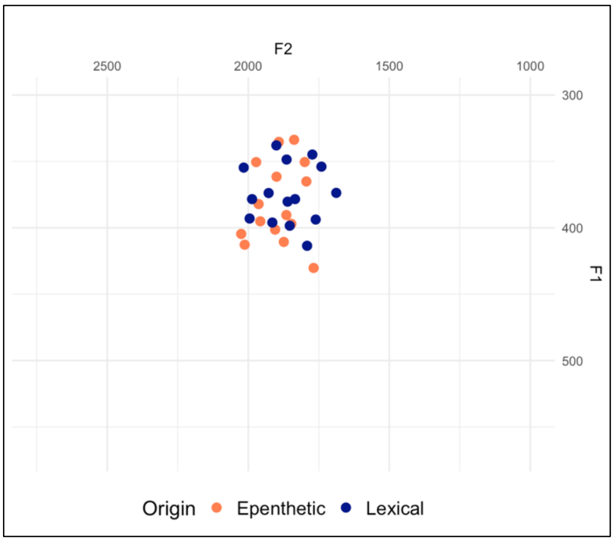

Figure 7b. Mean F1 and F2 (Hz) of Epenthetic and Lexical /i/ for 15 Male Speakers at FSR

Table 18. Mean F1 and F2 (Hz) with SD of Epenthetic and Lexical /i/ at FSR

	Epenthetic /i/		Lexical /i/	
	Female	Male	Female	Male
F1	435.87 (31.88)	369.01 (33.04)	427.25 (35.02)	373.52 (27.48)
F2	2309.25 (123.33)	1901.63 (104.92)	2292.32 (134.49)	1872.71 (121.05)

Mixed-effects models (lmer) tested F1, F2, and duration differences between epenthetic and lexical /i/. The anova-selected model, consistent with the normal rate analysis, included gender (female vs. male), origin (epenthetic vs. lexical), and speech rate (fast vs. normal) as fixed effects, and subject and word as random factors.

Table 19. Results for F1 and F2 of Epenthetic and Lexical /i/ at FSR

F1	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	378.6012	8.1051	46.7115	<0.0001 ***
Gender Male	−40.0083	10.3137	−3.8791	<0.0001 ***
Origin Lexical	11.1726	17.6256	0.6339	0.5262
Rate Fast	47.6917	16.4609	2.8973	0.0038 *
F2				
(Intercept)	2182.6143	35.7192	61.1048	<0.0001 ***
Gender Male	−391.2632	46.8915	−8.3440	<0.0001 ***
Origin Lexical	9.8967	33.3093	0.2971	0.7671
Rate Fast	−96.3413	29.7254	−3.2411	0.0012 **

Table 19 shows that the final model included speech rate as a fixed effect. Tukey's tests at the fast rate examined the effects of gender and origin.

Table 20. Tukey's Pairwise Comparisons of F1 and F2 of Epenthetic and Lexical /i/ by Origin at FSR

		Comparison (origin)	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
Female	F1	Epenthetic /i/ – Lexical /i/	8.6213	2.9783	2.8947	0.5661
	F2	Epenthetic /i/ – Lexical /i/	16.9338	5.6103	3.0183	0.7359
Male	F1	Epenthetic /i/ – Lexical /i/	−3.5827	1.4881	−2.4075	0.8152
	F2	Epenthetic /i/ – Lexical /i/	9.2916	4.0783	2.2783	0.5638

As presented in Table 20, the F1 and F2 of epenthetic /ɨ/ did not differ from those of lexical /ɨ/ for either female (F1: $p = 0.5661$; F2: $p = 0.7359$) or male (F1: $p = 0.8152$; F2: $p = 0.5638$) speakers.

In sum, vowel origin (epenthetic vs. lexical) did not have a significant effect on the F1 or F2 values of /i/, and female speakers produced higher F1 and F2 values than male speakers across both epenthetic and lexical vowels at the fast speech rate.

Next, Figure 8 and Table 21 present the mean duration of epenthetic and lexical /ɨ/, and Table 22 provides the statistical results for their durations.

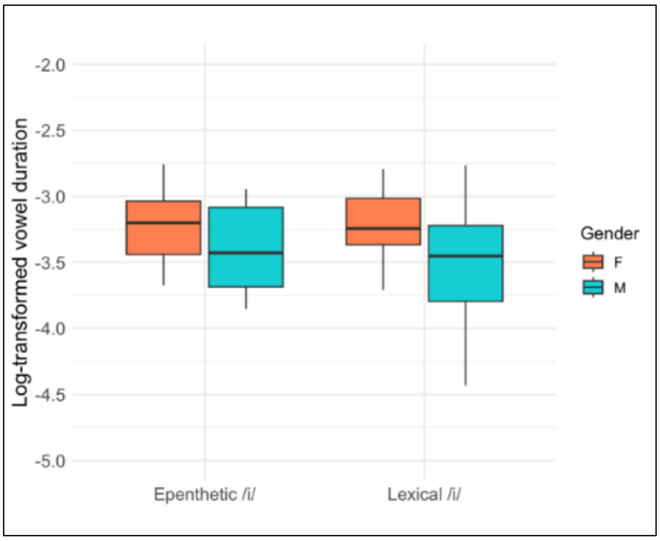

Figure 8. Vowel Durations (Log-transformed) of Epenthetic and Lexical /i/ at FSR

Table 21. Mean Durations (s) with SD of Epenthetic and Lexical /i/ at FSR

Fast rate	Epenthetic /i/		Lexical /i/	
	Female	Male	Female	Male
	0.0422 (0.0117)	0.0347 (0.0111)	0.0425 (0.0106)	0.0339 (0.0129)

Table 22. Results for Durations of Epenthetic and Lexical /i/ at FSR

	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	0.0463	0.0157	2.9490	0.0020 **
Gender Male	−0.0079	0.0032	−2.4688	0.0372 *
Origin Lexical	0.0051	0.0201	0.2537	0.8029
Rate Fast	−0.0174	0.0078	−2.2308	0.0021 **

As shown in Table 22, the model captured the combined influence of variables across both normal and fast speech rates; therefore, Tukey's pairwise comparisons were conducted specifically for the fast rate. The results indicated no significant durational difference between epenthetic and lexical /i/ at the fast rate (p = 0.0719). However, both vowels were significantly longer when produced by female speakers (epenthetic /ɨ/: p = 0.0357; lexical /i/: p = 0.0124).

3.4 Overview of the Shift from Normal Speech to Fast Speech

Figure 9 presents the mean F1 and F2 for both rates.

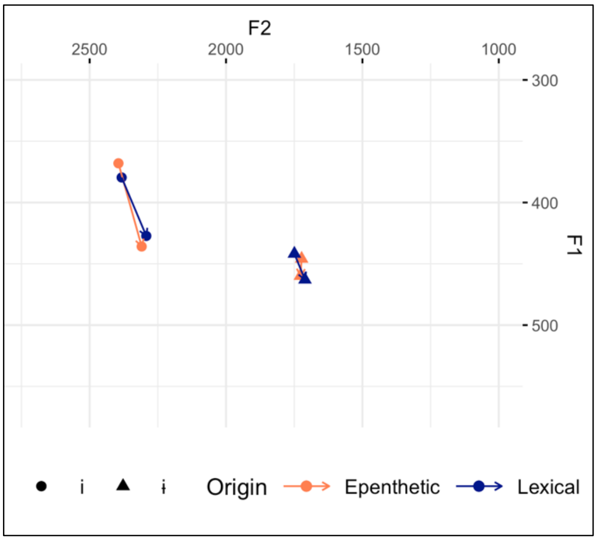

Figure 9a. F1 and F2 Shifts (Hz) in Epenthetic and Lexical Vowels for Female Speakers, with Arrows Indicating the Transition from NSR to FSR

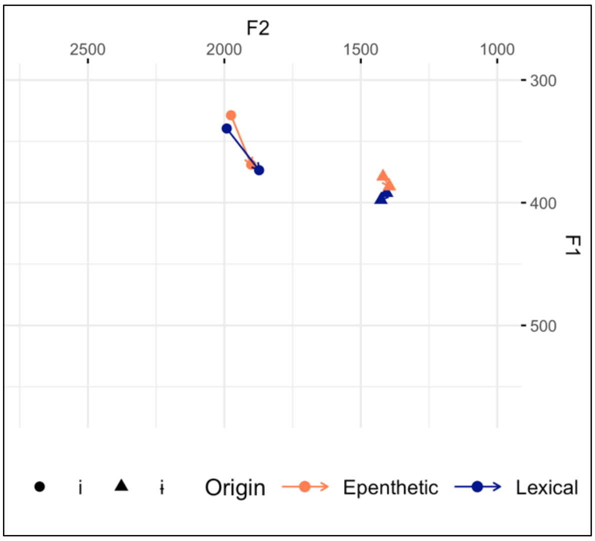

Figure 9b. F1 and F2 Shifts (Hz) in Epenthetic and Lexical Vowels for Male speakers, with Arrows Indicating the Transition from NSR to FSR

Figures 9a and 9b shows that both female and male speakers exhibited greater variation in F1 and F2 across speech rates for epenthetic and lexical /i/ than for epenthetic and lexical /ɨ/. Next, Figure 10 shows durational differences by speech rate.

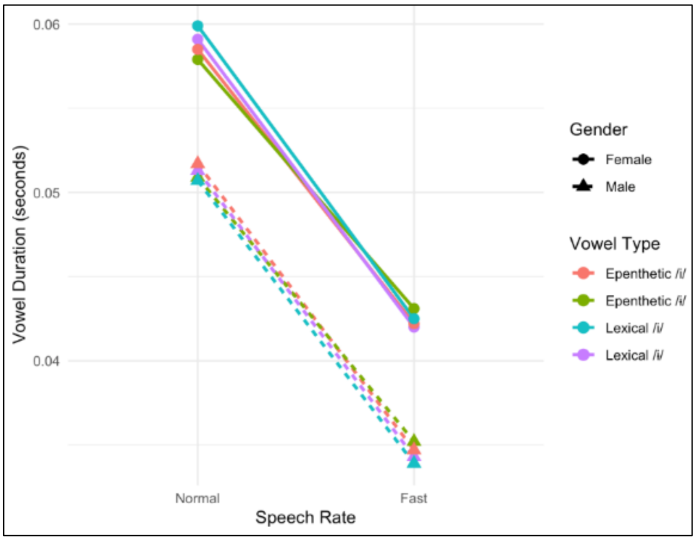

Figure 10. Vowel Durational Shifts (s) in Epenthetic and Lexical vowels

As shown in Figure 10, vowel durations were consistently shorter at the fast speech rate, regardless of origin (epenthetic vs. lexical), gender (female vs. male), or vowel type (/ɨ/ vs. /i/).

In sum, whereas the F1 and F2 of epenthetic and lexical /ɨ/ showed relatively little change, those of epenthetic and lexical /i/ varied noticeably with speech rate. In particular, speakers produced epenthetic and lexical /i/ with significantly higher F1 and lower F2 at the fast rate than at the normal rate, a pattern consistent with vowel centralization in fast speech. With respect to duration, all vowels were shorter at the fast rate than at the normal rate. The following section presents statistical analyses that substantiate these observations.

3.4.1 Epenthetic /ɨ/ and lexical /ɨ/

In §3.2.1 and §3.3.1, the best-fitting models for the F1 and F2 of epenthetic and lexical /ɨ/ did not retain speech rate. To directly test its influence, speech rate was later incorporated as a fixed effect, as reported in Table 23.

Table 23. Results for F1 and F2 of Epenthetic and Lexical /ɨ/

F1	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	450.1804	4.5945	97.9825	<0.0001 ***
Gender Male	−55.6561	5.0649	−10.9886	0.0112 *
Place Alveolar	19.0419	6.0117	3.1675	0.3537
Place Bilabial	−10.2936	7.0212	−1.4661	0.4549
Origin Lexical	4.5232	4.0703	1.1113	0.2661
Rate Fast	7.3828	3.4795	2.1218	0.0702

F2				
(Intercept)	1715.0857	22.9238	74.8168	<0.0001 ***
Gender Male	−312.0502	26.4776	−11.7854	0.0141 *
Place Alveolar	293.5315	59.2924	4.9506	0.0312 *
Place Bilabial	−51.3314	17.8697	−2.8725	0.0497 *
Origin Lexical	12.1769	18.7062	0.6510	0.5159
Rate Fast	−11.4873	15.5399	−0.7392	0.5372

Table 23 shows results consistent with the best-fitting model for gender, place of articulation, and origin. Notably, speech rate did not significantly affect the F1 or F2 values of epenthetic and lexical /ɨ/ (F1: p = 0.0702; F2: p = 0.5372). Tukey's pairwise comparisons were then carried out to examine whether this pattern held for both epenthetic /ɨ/ and lexical /ɨ/.

Table 24. Tukey's Pairwise Comparisons of F1 and F2 of Epenthetic /ɨ/ and Lexical /ɨ/ by Speech Rate

			Comparison	Estimate	Std. Error	t-value	Pr (>|t|)
Female	F1	Epenthetic /ɨ/	Fast – Normal	14.5349	20.2438	0.7179	0.4726
		Lexical /ɨ/	Fast – Normal	19.1472	20.4345	0.9371	0.3488
	F2	Epenthetic /ɨ/	Fast – Normal	3.3052	15.8143	0.2093	0.8342
		Lexical /ɨ/	Fast – Normal	–37.9455	42.3972	–0.8956	0.3706
Male	F1	Epenthetic /ɨ/	Fast – Normal	9.4918	15.4087	0.6161	0.5378
		Lexical /ɨ/	Fast – Normal	5.9684	13.5337	0.4415	0.6591
	F2	Epenthetic /ɨ/	Fast – Normal	–19.6426	39.6020	–0.4969	0.1905
		Lexical /ɨ/	Fast – Normal	21.4031	46.4275	0.4618	0.3924

As in Table 24, the F1 and F2 of epenthetic and lexical /ɨ/ were not influenced by speech rate.

Both epenthetic and lexical /ɨ/ were produced with shorter durations at the fast speech rate across genders. Since the best-fitting mixed-effects model already incorporated speech rate as a fixed factor, these findings replicate the results reported in §3.2.1 and §3.3.1.

Table 25. Results for Durations of Epenthetic and Lexical /ɨ/

	Estimate	Std. Error	t-value	Pr (>|t|)
(Intercept)	0.0573	0.0041	13.9756	<0.0001 ***
Gender Male	−0.0072	0.0032	−2.2499	0.0476 *
Origin Lexical	0.0014	0.0011	1.2726	0.3276
Position Final	0.0071	0.0029	2.4483	0.0198 *
Structure Open	0.0016	0.0015	1.0667	0.3192
Rate Fast	−0.0315	0.0197	−1.5990	0.0058 **

Table 25 indicates that vowels were significantly shorter in duration at the fast rate (p = 0.0058).

3.4.2 Epenthetic /i/ and lexical /i/

Since speech rate was already modeled as a fixed effect, the findings in §3.2.2 and §3.3.2 are reconsidered here.

Table 26. Results for F1 and F2 of Epenthetic and Lexical /i/

F1	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	378.6012	8.1051	46.7115	<0.0001 ***
Gender Male	−40.0083	10.3137	−3.8791	<0.0001 ***
Origin Lexical	11.1726	17.6256	0.6339	0.5262
Rate Fast	47.6917	16.4609	2.8973	0.0038 *
F2				
(Intercept)	2182.6143	35.7192	61.1048	<0.0001 ***
Gender Male	−391.2632	46.8915	−8.3440	<0.0001 ***
Origin Lexical	9.8967	33.3093	0.2971	0.7671
Rate Fast	−96.3413	29.7254	−3.2411	0.0012 **

Table 26 shows significant effects of speech rate on F1 ($p = 0.0038$) and F2 ($p = 0.0012$). Tukey's tests evaluated whether this held for both epenthetic and lexical /i/.

Table 27. Tukey's Pairwise Comparisons of F1 and F2 of Epenthetic and Lexical /i/ by Speech Rate

	F1		Comparison	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
Female	F1	Epenthetic /i/	Fast – Normal	70.5233	22.3464	3.1559	0.0043 **
		Lexical /i/	Fast – Normal	51.3840	21.2682	2.4162	0.0156 *
	F2	Epenthetic /i/	Fast – Normal	−89.5918	32.9623	−2.7181	0.0067 **
		Lexical /i/	Fast – Normal	−92.4056	36.1807	−2.5544	0.0106 *
Male	F1	Epenthetic /i/	Fast – Normal	44.9682	14.4825	3.1058	0.0018 **
		Lexical /i/	Fast – Normal	35.3475	15.7871	2.2396	0.0251 *
	F2	Epenthetic /i/	Fast – Normal	−61.2155	24.0154	−2.5499	0.0108 *
		Lexical /i/	Fast – Normal	−56.3893	20.2257	−2.7884	0.0053 **

Table 27 demonstrates that speech rate significantly affected the F1 and F2 of both epenthetic and lexical /i/ for male and female speakers. Across genders, F1 increased while F2 decreased, indicating vowel centralization at the fast speech rate for both epenthetic and lexical /i/.

With respect to vowel duration, both epenthetic and lexical /i/ were shorter at the fast speech rate across gender. Since the best-fitting mixed-effects model had already incorporated speech rate as a fixed factor, these findings replicate the results presented in §3.2.2 and §3.3.2.

Table 28. Results for Durations of Epenthetic and Lexical /i/

	Estimate	Std. Error	*t*-value	Pr (>|*t*|)
(Intercept)	0.0463	0.0157	2.9490	0.0020 **
Gender Male	−0.0079	0.0032	−2.4688	0.0372 *
Origin Lexical	0.0051	0.0201	0.2537	0.8029
Rate Fast	−0.0174	0.0078	−2.2308	0.0021 **

In Table 28, vowels were shorter at the fast rate than at the normal rate ($p = 0.0021$).

4. General Discussions and Limitations

F1 and F2 of epenthetic and lexical /ɨ/ were comparable at both the normal and fast speech rates. At the normal rate, the F2 values of both vowels increased when preceded by alveolars and decreased when preceded by bilabials. This pattern aligns with previous research showing that vowel F2 varies with the place of articulation of the preceding consonant, likely due to coarticulatory influence (Cooper et al. 1952, Kerdpol 2012, Liberman et al. 1954). Labial consonants tend to have lower F2 values than coronal consonants, and this tendency may influence the F2 formant values of following vowels due to coarticulation. By contrast, at the fast speech rate, the F2 values of both epenthetic and lexical /ɨ/ did not show systematic variation across places of articulation in this study. This pattern may be explained by F2 lowering in coronal and dorsal consonants under fast speech conditions. Agwuele et al. (2008) argue that as speech rate increases, F2 values in coronal and dorsal decrease to a greater extent than in labial stops, and they attribute this pattern to greater coarticulatory compression at these places of articulation under fast speech conditions. A similar mechanism may account for the present findings, in which F2 values at the fast speech rate did not differ as a function of place of articulation.

The durations of epenthetic and lexical /ɨ/ were similar across speech rates, showing parallel patterns. At the normal rate, both vowels were lengthened in final word position, open syllables, and productions by female speakers. First, vowel lengthening in final word positions is well documented (van Santen 1992, Windmann et al. 2015, White et al. 2020). Windmann et al. (2015) conducted a corpus-based analysis of vowel duration using the Aix-MARSEC corpus of British English broadcast speech. By explicitly controlling for prominence and within-word position, they showed that vowels are consistently longer in word-final than in non-final positions across stress categories, providing evidence for word-final vowel lengthening. White et al. (2020) argue that English listeners exploit lengthened vowels in word-final position as a cue for predicting upcoming word boundaries, facilitating word segmentation during speech processing. Also, previous studies (Duanmu 1996, Klatt 1976, O'Shaughnessy 1981) argue that word-final lengthening may be attributed to pre-pausal lengthening, whereby vowels lengthen before a pause. O'Shaughnessy (1981), for instance, observed longer durations for French vowels and consonants in word-final syllables, attributing the effect to an immediately following pause. This mechanism may also explain the longer durations of Korean vowels in word-final position in the present study, given that the target words were consistently followed by a brief pause in the reading task. However, at the fast speech rate, word-final lengthening was no longer observed in this study. Similar findings were reported by Albano (1999) for Brazilian Portuguese, where pre-pausal lengthening (i.e., word-final lengthening) diminished at faster rates, resulting in vowels in final and non-final positions having comparable durations. Albano attributed this pattern to speech-rate-induced compression, which reduces pre-pausal vowel length to approximate that of other vowels.

Next, in many languages, including Korean, vowels in open syllables are typically longer than those in closed syllables (Curtis 2002, Oh 2016, Rositzke 1939). For example, Oh (2016) examined V, CV, VC, and CVC syllable types in Korean and found that vowels were longest in V syllables, followed by CV, then VC, and shortest in CVC syllables. Similarly, Rositzke (1939) reported that both monophthongs and diphthongs were longer in open than in closed syllables in General American English. Previous studies (Fowler 1981, Klatt 1976, Lindblom and Rapp 1973) suggest that this pattern is linked to compensatory shortening, whereby vowels reduce in length relative to other intrasyllabic segments while retaining their phonetic identity. As a result, vowels are shortened when additional consonants are present in the syllable coda, allowing speakers to maintain stable syllable duration. In the present study, however, compensatory shortening of epenthetic and lexical /ɨ/ was not observed at the fast speech rate. This finding aligns with prior work showing that speech rate can reshape temporal patterns across languages (Smith 2000, 2002). Smith (2002) argues that temporal patterns arise from two distinct sources: prosodic and phonetic effects. Prosodic effects are linked to a segment's position in an utterance, such as utterance final-syllable lengthening, while phonetic effects are shaped by adjacent segments, as in vowel lengthening before voiced stops. Smith (2002) further argues that prosodic effects tend to remain stable across different speech rates, whereas phonetic effects vary with speech rate. For example, vowel lengthening before voiced stops decreases as speech rate increases, resulting in shorter segments relative to the normal rate. In contrast, utterance final-syllable lengthening did not diminish at faster rates. These findings suggest that prosodic temporal patterns are more resistant to changes in speech rate because speakers have internalized them as production strategies and perceptual cues (de Jong and Zawaydeh 1999, Smith 2002). In the present study, vowel durations were shorter in closed syllables than in open syllables. This compensatory vowel shortening is likely attributable to the influence of adjacent segments rather than to prosodic position. Consequently, the duration difference between vowels in closed and open syllables may diminish under fast speech due to temporal compression.

Gender also influenced the duration of both epenthetic and lexical vowels at the two speech rates, with male speakers producing shorter vowels than female speakers. This tendency has been observed cross-linguistically (Holt et al. 2015, Jacewicz and Fox 2015, Simpson and Ericsdotter 1998). Simpson and Ericsdotter (1998) suggest that such gender-based durational differences stem from anatomical differences in the vocal tract. Since male speakers typically have larger vocal tracts, they must move their articulators across greater distances to reach the same phonetic targets. To compensate, they increase their articulatory speed, thereby producing shorter vowel durations overall. In addition, sociolinguistic factors may also account for this pattern. Previous studies have shown that female speakers tend to produce clearer speech than male speakers (Albuquerque et al. 2021, Bradlow et al. 2003). Clear speech is generally understood as an articulatory strategy aimed at realizing phonetic targets more distinctly, and it is typically accompanied by an expansion of the vowel formant space and an increase in segmental duration (Picheny et al. 1986, 1986). From this perspective, the longer vowel durations observed in female speakers may be a byproduct of clearer speech.

English proficiency was not retained as a fixed effect in the final best-fitting model. This suggests that it did not substantially improve model fit or account for a meaningful portion of the variance in the production of epenthetic /i/. However, to further examine whether English proficiency influenced epenthetic vowel realization, additional analyses were conducted that included proficiency as a predictor. The overall pattern of results for the other predictors remained unchanged compared to the model without proficiency, and English proficiency was not statistically significant in any of the models ($p > 0.05$). Although previous studies (LaCharité and Paradis 2005, Kadenge and Mudzingwa 2012) showed that loanword adaptation patterns may vary as a function of proficiency, such findings have typically been derived from comparisons between native and bilingual speakers. In contrast, the participants in the present study were all native Korean speakers, which may account for the absence of an English proficiency effect.

This study demonstrated that epenthetic and lexical /ɨ/ showed only vowel shortening at the fast speech rate, whereas epenthetic and lexical /i/ exhibited both centralization and shortening. Such centralization and vowel shortening are widely observed under increased speech rate. Reduced vowels tend to approximate the schwa /ə/ (Moon and Lindblom 1994, Stetson 1951), a process referred to as "undershoot," or phonetic vowel reduction. Undershoot occurs when shortened vowel duration prevents speakers from fully reaching the articulatory target, leading to centralization of the vowel space (Guenther 1995, Koopmans-van Beinum 1980, Lindblom 1963). Although undershoot is often associated with unstressed or weakly stressed vowels in stress-timed languages, it is not exclusive to them, since vowel durations are shaped by additional factors. One such factor is speech rate, which can itself trigger undershoot (Fourakis 1991, Jaworski 2009, Nadeu 2014). At faster rates, the reduced time for articulation limits movement toward vowel targets, causing vowel formants to centralize as durations shorten. For instance, Nadeu (2014) showed that in Catalan, speech rate exerts a stronger influence on undershoot than stress, because it reduces vowel duration more than the absence of stress does. In Korean, undershoot has been observed as an effect of fast speech (Igeta et al. 2017, Son 2017). Son (2017) compared the formants of /ɑ, i, u, æ, ɛ, o, ʌ, ɨ/ at normal and fast rates and found that F1 for /ɑ/ decreased, F2 for /i/ and /ɛ/ decreased, and F2 for /u/ and /o/ increased at the fast rate. Although not all vowels displayed undershoot, several were centralized under fast speech conditions.

In the present study, both epenthetic and lexical /i/ showed increased F1, decreased F2, and shorter durations at the fast rate, suggesting centralization due to undershoot. In contrast, the F1 and F2 of both epenthetic and lexical /ɨ/ remained stable despite shorter durations, indicating that undershoot did not fully occur. Jaworski (2009) similarly noted that not all vowels are equally subject to undershoot. His investigation of Polish, Spanish, and Russian revealed that high vowels (/i, u/) are more vulnerable to undershoot during fast speech, whereas central vowels (/a, ɨ/) are least affected. Jaworski (2009) attributes these findings to vocal tract inertia, the tendency of articulators to maintain their current motion or position during speech. This inertia limits the ability of the articulators to reach target configurations, particularly in fast speech. The resulting process, termed inertial lenition, includes undershoot as one of its manifestations. Jaworski (2009) further argues that because vocal tract inertia affects all speech production, inertial lenition is universal. He proposes a cross-linguistic hierarchy of vowel susceptibility to inertia: high vowels /i, u/ > mid vowels /e, o/ > central vowels /a, ɨ/ (2009:126). High vowels /i/ and /u/ occupy the most peripheral positions in the vowel space, making them more difficult to fully articulate under inertia at faster speech rates. Consequently, these vowels are more prone to undershoot. Central vowels, by contrast, may demand the least articulatory effort because they are already positioned near the center of the vowel space. As a result, they are less affected by inertia and exhibit less undershoot than high vowels. Jaworski's (2009) hierarchy of susceptibility to inertia accounts for the current findings: both epenthetic and lexical /i/ underwent undershoot, whereas epenthetic and lexical /ɨ/ did not. Within the Korean vowel system, /i/ is the most peripheral vowel, making it more vulnerable to undershoot in fast speech due to vocal tract inertia. Korean /ɨ/, however, presents a more complex case. Phonologically, it is often categorized as a back vowel, but several studies further argue that Korean /ɨ/ is phonetically central with slight tongue retraction (Lee and Ramsey 2011, Umeda 2022). Given this classification, Korean /ɨ/ is less affected by inertia than the high vowel /i/.

The present study demonstrated that epenthetic vowels in Korean exhibit F1, F2, and duration values comparable to those of lexical vowels in read speech. Across both normal and fast speech rates, epenthetic vowels showed no evidence of acoustic reduction relative to lexical vowels. These findings suggest that, in a controlled read-speech context, epenthetic vowels are not realized as underspecified elements or merely transitional vowels arising from articulatory movement. Rather, they appear to be produced with stable and clearly defined phonetic targets, similar to lexical vowels. In this respect, the findings support the view that epenthetic vowels may function as full phonological segments in Korean, diverging from accounts that analyze them as gradient, schwa-like transitional elements. However, this interpretation should be treated with caution, particularly in the case of /ɨ/. As /ɨ/ is inherently centralized within the vowel inventory, even if some degree of reduction occurs at faster speech rates, such changes may not be clearly observable in the F1–F2 space. Accordingly, while the absence of centralization may provide evidence supporting phonological stability of epenthetic /ɨ/, it should not be considered a decisive indicator in isolation, but rather interpreted in conjunction with the vowel's intrinsic acoustic properties. Future research should incorporate additional phonetic measures beyond formant values to more comprehensively evaluate whether epenthetic vowels maintain stable phonetic targets across different speech conditions.

As a final note, it should be acknowledged that the participants in this study displayed some degree of bilingualism. Strictly speaking, monolinguals are defined as individuals who speak only one language (Kemp

2009), but the definition can be broadened. Escudero et al. (2014) characterize monolinguals as individuals who rely on their L1 for daily communication, have not resided in a different language environment for more than one month, and have received only minimal, classroom-based L2 instruction, typically from teachers with an L1 accent and with a focus on reading and grammar rather than communicative practice. To account for such contexts, the term "functional monolinguals" is also used. By this definition, 25 of the 30 participants in this study can be regarded as functional monolinguals. However, five participants had resided in the United States for longer than one month, although they primarily used Korean in daily life at school, work, and in social interactions. Thus, the present study acknowledges that the participants may exhibit a limited degree of bilingualism and recognizes as a limitation the possibility that differences in English exposure may have influenced speech rate strategies or familiarity with loanwords.

Also, this study employed nativized English loanwords in Korean as stimuli. Since these items are listed in Korean dictionaries, they can be assumed to have been in use for a considerable period and are fully integrated into the language. Consequently, Korean speakers may produce epenthetic vowels in such loanwords in the same way as lexical vowels, yielding comparable formant values and durations. To further examine this issue, it would be useful to investigate epenthetic vowel production in nonce words. For instance, Korean speakers could be asked to read random English consonant clusters such as /mkd/. Given that /ɨ/ is the default epenthetic vowel in Korean, speakers would likely insert /ɨ/ to break up these clusters, and such epenthetic vowels could then be systematically compared with lexical vowels.

Next, in the present study, the experiment was conducted in a fixed order, with the normal speech rate condition preceding the fast speech rate condition. Therefore, the possibility of an order effect cannot be entirely ruled out. The results observed in the fast-rate condition may have been influenced, to some extent, by task familiarization or strategic adjustments. Future research should counterbalance the order of speech-rate conditions to examine this possibility more rigorously.

Lastly, speaker familiarity with loanwords constitutes another limitation of this study. Some loanwords are widely used in Korean, whereas others are less common. In such cases, the phonetic realization of epenthetic vowels may vary depending on the speaker's familiarity with a given word. Along the same lines, some of the Korean words used as stimuli are used less frequently than others. Differences in word usage frequency or speaker familiarity may have influenced variation in pronunciation. Specifically, one stimulus [sukʰɨl] is not strictly considered a standard form and may therefore have lower familiarity. However, additional statistical analyses conducted excluding this item showed that the overall pattern of results remained unchanged. Thus, future research would benefit from controlling for the frequency of both English loanwords and Korean words, as well as speaker familiarity, in order to more fully assess their potential influence.

5. Conclusion

This paper examined the phonetic characteristics of epenthetic and lexical vowels in Korean across different speech rates. The findings indicate that epenthetic vowels and their lexical counterparts did not differ in their phonetic properties at either normal or fast rates in read speech. Although the durations of epenthetic and lexical /ɨ/ shortened under fast speech, they did not undergo centralization, a result that may be explained by vocal tract inertia

At the normal rate, epenthetic and lexical /ɨ/ were lengthened in open syllables and in word-final position, but these effects disappeared at the fast rate. Such patterns appear to result from temporal compression associated with fast speech (Albano 1999, de Jong and Zawaydeh 1999, Smith 2002). Additionally, gender was found to influence vowel duration. Male speakers produced shorter vowel durations than female speakers for both epenthetic and lexical vowels across speech rates. Such gender-based differences are well documented in a range of languages, including English, Swedish, and Korean (Holt et al. 2015, Lee and Jin 2016, Simpson and Ericsdotter 1998). Since male speakers typically have larger vocal tracts, they compensate by increasing articulatory speed, which results in generally shorter vowel durations. Also, female speakers tend to produce clearer speech than male speakers (Albuquerque et al. 2021, Bradlow et al. 2003), and it may contribute to their longer vowel durations.

Overall, this study demonstrated that epenthetic vowels in Korean exhibit F1, F2 and durations comparable to those of lexical vowels in read speech. Thus, this finding suggests that epenthetic vowels may function as full phonological segments with defined phonetic targets, rather than as transitional schwas arising from articulatory movement (Davidson 2006, Davidson and Stone 2003) in read speech. However, given that /ɨ/ is inherently centralized, the absence of centralization at a faster speech rate may be attributable not only to the phonological stability of epenthetic vowels but also to the acoustic properties of /ɨ/ itself. Accordingly, the lack of centralization may not be taken as a decisive indicator of epenthetic /ɨ/ stability, but should instead be interpreted in light of the vowel's acoustic characteristics.

References

Agwuele, A., H. Sussman and B. Lindblom. 2008. The effect of speaking rate on consonant vowel coarticulation. *Phonetica* 65(4), 194-209.

Albano, E. 1999. A gestural solution for some glide epenthesis problems. In *Proceedings of the 14th International Congress of Phonetic Sciences*, 1785-1788.

Albuquerque, L., R. Valente, A. Teixeira, C. Oliveira and D. Figueiredo. 2021. Age and gender effects in European Portuguese spontaneous speech. *Loquens* 8(1-2), 1-9.

Bellik, J. A. 2019. The effect of speech style and deaccentuation on vowel intrusion in Turkish complex onsets. In *Proceedings of the 19th International Congress of Phonetic Sciences*, 3498-3502.

Best, C. and M. Tyler. 2007. Nonnative and second language speech perception. *Language Learning & Language Teaching* 17, 13-34.

Boersma, P. and D. Weenink. 2024. Praat. Available online at https://www.fon.hum.uva.nl/praat/

Bradlow, A. R., N. Kraus and E. Hayes. 2003. Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research* 46(1), 80-97

Cangi, M. E., A. Isildar, A. Tekin and A. B. Sarac. 2020. A preliminary study of normative speech rate values of Turkish speaking adults. *ENT Updates* 10, 381-389.

Choi, S. and J. Jun. 1998. Are Korean fortis and aspirated consonants geminates? *Language Research* 34(3), 521-546.

Cooper, F. S., P. C. Delattre, A. M. Liberman, J. M. Borst and L. J. Gerstman. 1952. Some experiments on the perception of synthetic speech sounds. *The Journal of the Acoustical Society of America* 24, 597-606.

Curtis, E. 2002. Moraic structure and segment duration in Korean. *Japanese/Korean Linguistics* 10, 378-391.

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

Davidson, L. and M. Stone. 2003. Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. In *Proceedings of the West Coast Conference on Formal Linguistics* 22, 165-178.

de Jong, K. and B. A. Zawaydeh. 1999. Stress, duration, and intonation in Arabic word-level prosody. *Journal of Phonetics* 27(3), 3-22.

Duanmu, S. 1996. Pre-juncture lengthening and foot binarity. *Studies in the Linguistic Sciences* 26(1-2), 95-115.

Escudero, P., B. Sisinni and M. Grimaldi. 2014. The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *Journal of the Acoustical Society of America* 135(3), 1577-1584.

Fant, G. 1966. A note on vocal tract size factors and non-uniform F-pattern scaling. *Speech Transmission Laboratory-Quarterly Progress and Status Report* 4, 22-30.

Fourakis, M. 1991. Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical Society of America* 90(4), 1816-1827.

Fowler, C. A. 1981. A relationship between coarticulation and compensatory shortening. *Phonetica* 38(1-3), 35-50.

Gay, T. 1978. Effect of speaking rate on vowel formant movements. *The Journal of the Acoustical Society of America* 63(1), 223-230.

Gouskova, M. and N. Hall. 2009. Acoustics of epenthetic vowels in Lebanese Arabic. In S. Parker, ed., *Phonological Argumentation: Essays on Evidence and Motivation*, 203-225. London: Equinox.

Guenther, F. H. 1995. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review* 102(3), 594-621.

Hall, N. 2011. Vowel epenthesis. *The Blackwell Companion to Phonology* 3, 1-21.

Hall, N. 2013. Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic. *Journal of Phonetics* 41(2), 133-143.

Hirata, Y. and K. Tsukada. 2003. The effects of speaking rates and vowel length on formant movements in Japanese. In *Proceedings of the 2003 Texas Linguistics Society Conference: Coarticulation in Speech Production and Perception*, 73-85.

Holt, Y. F., E. Jacewicz and R. A. Fox. 2015. Variation in vowel duration among Southern African American English speakers. *American Journal of Speech-Language Pathology* 24(3), 460-469.

Igeta, T., S. Hiroya and T. Arai. 2017. Overlapping of /o/ and /u/ in modern Seoul Korean: Focusing on speech rate in read speech. *Phonetics and Speech Sciences* 9(1), 1-7.

Jacewicz, E. and R. A. Fox. 2015. Eliciting sociophonetic variation in vowel duration. In *Proceedings of the 18th International Congress of Phonetic Sciences*, 1-5.

Jaworski, S. 2009. Inertial and non-inertial lenition processes. *Poznań Studies in Contemporary Linguistics* 45(1), 103-129.

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

Kang, Y. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20, 219-273.

Kemp, C. 2009. *The Exploration of Multilingualism*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kent, R. D. and C. Read. 2002. *The Acoustic Analysis of Speech*. San Diego: Singular Publishing Group.

Kerdpol, K. 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, H. 2009. The phonology and phonetics of unaccentable vowels in Kyungsang Korean. *Paper presented in the Japanese/Korean Linguistics* 18.

Kim, K. and A. Kochetov. 2011. Phonology and phonetics of epenthetic vowels in loanwords: Experimental evidence from Korean. *Lingua* 121(3), 511-532.
Kim, S. 2009. A preliminary study on the relationship between speech rate and prosodic unit generation in Korean read speech. *EONEOHAG* 53, 225-254.
Klatt, D. H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America* 59(5), 1208-1221.
Koopmans-van Beinum, F. 1980. *Vowel Contrast Reduction: An Acoustic and Perceptual Study of Dutch Vowels in Various Speech Conditions*. Doctoral dissertation, University of Amsterdam, Amsterdam, Netherlands.
LaCharité, D. and C. Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36(2), 223-258.
Laver, J. 1994. *Principles of phonetics.* Cambridge: Cambridge University Press.
Lee, K. and I. Jin. 2016. A study of durations for Korean vowels at isolated utterance condition. *Audiology and Speech Research* 12(3), 150-156.
Lee, K. M. and S. Ramsey. 2011. *A History of the Korean Language*. Cambridge: Cambridge University Press.
Lee, S. H. and H. Go. 2004. Speech rate and the acoustic features of Korean segments. *The Journal of the Acoustical Society of Korea* 23(2), 162-172.
Lee, S. H. Go, Y. Han and J. Kim. 2003. Effects of speaking rate on Korean vowels. *The Journal of the Acoustical Society of Korea* 22(1), 14-22.
Liberman, A., P. C. Delattre, F. S. Cooper and L. 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.
Lindblom, B. 1963. Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America* 35(11), 1773-1781.
Lindblom, B and K. Rapp. 1973. Some temporal regularities of spoken Swedish. *Papers in Linguistics from the University of Stockholm* 21, 1-59.
Magen, H. and S. Blumstein. 1993. Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics* 21(4), 387-409.
Miller, J. L. 1981. Effects of speaking rate on segmental distinctions. In P. D. Eimas, and J. L. Miller, eds., *Perspectives on the study of speech*, 39-74. Hillsdale: Erlbaum.
Miner, K. 1979. Dorsey's Law in Winnebago-Chiwere and Winnebago accent. *International Journal of American Linguistics* 45(1), 25-33.
Monsen, R. B. 1974. Durational aspects of vowel production in the speech of deaf children. *Journal of Speech and Hearing Research* 17(3), 386-398.
Moon, S. and B. Lindblom. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America* 96(1), 40–55.
Nadeu, M. 2014. Stress- and speech rate-induced vowel quality variation in Catalan and Spanish. *Journal of Phonetics* 46(1), 1-22.
Naver. 2024. Naver Dictionary. Available online at https://dict.naver.com/
National Institute of Korean Language. 2024. Standard Korean Language Dictionary. Available online at https://stdict.korean.go.kr/main/main.do/
Oh, J. 2016. The aspect of duration in Korean vowels according to the phonological conditions. *The Society of Korean National Language and Literature* 57, 185-207.
O'Shaughnessy, D. 1981. A study of French vowel and consonant durations. *Journal of Phonetics* 9(4), 385-406.

Paradis, C. and D. LaCharité. 2012. The Influence of attitude on the treatment of interdentals in loanwords: Ill-performed importation. *Catalan Journal of Linguistics* 11, 97-126.
Park, H. and N. Ziegler. 2014. Cognitive shift in the bilingual mind: Spatial concepts in Korean-English bilinguals. *Bilingualism: Language and Cognition* 17(2), 410-430.
Peterson, G. E. and H. L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24(2), 175-184.
Picheny, M. A., N. I. Durlach and L. D. Braida. 1985. Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research* 28(1), 96-103.
Picheny, M. A., N. I. Durlach and L. D. Braida. 1986. Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research* 29(4), 434-446.
R Core Team. 2025. R: A language and environment for statistical computing. Available online at https://www.r-project.org/
Rositzke, H. A. 1939. Vowel length in general American speech. *Language* 15(2), 99-109.
Simpson, A and C. Ericsdotter. 1998. Sex-specific durational differences in English and Swedish. *15th ICPhS Barcelona*, 1113-1116.
Smith, B. L. 2000. Variations in temporal patterns of speech production among speakers of English. *The Journal of the Acoustical Society of America* 108(1), 2438-2442.
Smith, B. L. 2002. Effects of speaking rate on temporal patterns of English. *Phonetica* 59(4), 232-244.
Son, I. 2017. Comparative study on formant displacement in reading rate between Korean and English vowels of Korean speakers. *Studies in English Language and Literature* 59(3), 239-261.
Stetson, R. H. 1951. *Motor phonetics: A study of speech movements in action*. Amsterdam: North-Holland Publishing Company.
Tajima, K., D. Erickson and K. Nagao. 2002. Production of syllable structure in a second language: Factors affecting vowel epenthesis in Japanese-accented English. *The Indiana University Linguistics Club Working Papers* 2(2), 77-91.
Turner, G. S., K. Tjaden and G. Weismer. 1995. The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech and Hearing Research* 38(5), 1001-1013.
Uffmann, C. 2006. Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua* 116(7), 1079-1111.
Umeda, H. 2022. Korean vowels. *Gengo Kenkyu Anthology* 2, 25-41.
Umeda, N. 1975. Vowel duration in American English. *The Journal of the Acoustical Society of America* 58(2), 434-445.
van Santen, J. P. H. 1992. Contextual effects on vowel duration. *Speech Communication* 11(6), 513-546.
White, L., S. Benavides-Varela and K. Mády. 2020. Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues? *Journal of Phonetics* 81, 1-14.
Windmann, A., J. Šimko and P. Wagner. 2015. Polysyllabic shortening and word-final lengthening in English. In *Proceedings of Interspeech* 2015, 36-40.
Yoon, T.-J. 2024. Korean Forced Aligner. Available online at https://tyoon.net/
Zellou, G. 2011. Pharyngealization of French loanwords in dialectal Moroccan Arabic: An acoustic analysis of bilingual speakers. *Dialectologia* 6, 95-108.

Appendix A

Epenthetic /ɨ/	
12 epenthetic /ɨ/ in non-final CV syllable structure	
그램 [kɨlæm] ‘gram’	크랩 [kʰɨlæp] ‘crab’
그룹 [kɨlup] ‘group’	크림 [kʰɨlim] ‘cream’
드림 [tɨlim] ‘dream’	트리 [tʰɨli] ‘tree’
드럼 [tɨlʌm] ‘drum’	트랩 [tʰɨlæp] ‘trap’
브롬 [pɨlom] ‘brom(ine)’	프로 [pʰɨlo] ‘pro(gram)’
브로 [pɨlo] ‘bro(ther)’	프리 [pʰɨli] ‘free’
12 epenthetic /ɨ/ in final CV syllable structure	
도그 [tokɨ] ‘dog’	보트 [potʰɨ] ‘boat’
머그 [mʌkɨ] ‘mug’	미트 [mitʰɨ] ‘meat’
노크 [nokʰɨ] ‘knock’	커브 [kʰʌpɨ] ‘curve’
다크 [tɑkʰɨ] ‘dark’	허브 [hʌpɨ] ‘hurb’
모드 [motɨ] ‘mode’	로프 [lopʰɨ] ‘rope’
카드 [kʰɑtɨ] ‘card’	터프 [tʰʌpʰɨ] ‘tough’
12 epenthetic /ɨ/ in final CVC syllable structure	
고글 [kokɨl] ‘goggles’	리틀 [litʰɨl] ‘little’
이글 [ikɨl] ‘eagle’	셔틀 [sjʌtʰɨl] ‘shuttle’
서클 [sʌkʰɨl] ‘circle’	더블 [tʌpɨl] ‘double’
태클 [tʰækʰɨl] ‘tackle’	마블 [mɑpɨl] ‘marble’
누들 [nutɨl] ‘noodle’	애플 [æpʰɨl] ‘apple’
미들 [mitɨl] ‘middle’	와플 [wɑpʰɨl] ‘waffle’
Epenthetic /i/	
10 epenthetic /i/ after final /tɕ/	
라지 [lɑtɕi] ‘large’	에지 [ɛtɕi] ‘edge’
루지 [lutɕi] ‘luge’	조지 [tɕotɕi] ‘George’
비지 [pitɕi] ‘busy’	저지 [tɕʌtɕi] ‘judge’
배지 [pætɕi] ‘badge’	피지 [pʰitɕi] ‘Fiji’

이지 [itɕi] ‘easy’	퍼지 [pʰʌtɕi] ‘fudge’
10 epenthetic /i/ after final /tɕʰ/	
마치 [mɑtɕʰi] ‘march’	토치 [tʰotɕʰi] ‘torch’
매치 [mætɕʰi] ‘match’	피치 [pʰitɕʰi] ‘pitch’
워치 [wʌtɕʰi] ‘watch’	패치 [pʰætɕʰi] ‘patch’
아치 [ɑtɕʰi] ‘arch’	히치 [hitɕʰi] ‘hitch’
터치 [tʰʌtɕʰi] ‘touch’	해치 [hætɕʰi] ‘hatch’

Appendix B

Lexical /ɨ/	
12 lexical /ɨ/ in non-final CV syllable structure	
그림 [kɨlim] ‘picture’	크기 [kʰɨki] ‘size’
그네 [kɨnɛ] ‘swing’	크다 [kʰɨtɑ] ‘big’
드글 [tɨkɨl] ‘swarming with’	트림 [tʰɨlim] ‘burp’
드문 [tɨmun] ‘rare’	트집 [tʰɨtɕip] ‘nitpick’
널브리다 [nʌlpɨliɑ] ‘scatter’	배고프다 [pægopʰɨtɑ] ‘hungry’
시나브로 [sinɑpɨlo] ‘slowly but steadily’	애달프다[ætɑlpʰɨtɑ] ‘heartrending’
12 lexical /ɨ/ in final CVC syllable structure	
서글 [sʌkɨl] ‘warm and friendly demeanor’	비틀 [itʰɨl] ‘staggering’
지글 [tɕikɨl] ‘sound of something frying’	베틀 [pɛtʰɨl] ‘loom for weaving’
뭉클 [muŋkʰɨl] ‘emotional’	입을 [ipʰɨl] ‘wear (future)’
수클 [sukʰɨl] ‘writing applied well after learning’	깁을 [kipʰɨl] ‘sew (future)’
구들 [kutɨl] ‘Korean underfloor heating system’	헤플 [hɛpʰɨl] ‘wasteful (future)’
버들 [pʌtɨl] ‘willow tree’	슬플 [sɨlpʰɨl] ‘sad (future)’
12 lexical /ɨ/ in non-final CVC syllable structure	
글자 [kɨltɕɑ] ‘letter’	틀다 [tʰɨltɑ] ‘turn’
글씨 [kɨls’i] ‘handwriting’	틀니 [tʰɨlli] ‘denture’
무클하다 [mukʰɨlhɑtɑ]	씹을수록 [s’ipɨls’ulok]

‘rotting and becoming mushy’	‘the more one chews’
클클 [kʰɨlkʰɨl] ‘chuckling’	잡을수록 [tɕɑpɨls’ulok] ‘the more one holds on’
들판 [tɨlpʰɑn] ‘field’	고플수록 [kopʰɨls’ɛla] ‘the hungrier one is’
들깨 [tɨlk’æ] ‘perilla seed’	아플수록 [ɑpʰɨls’ɛla] ‘the sicker one is’
Lexical /i/	
10 lexical /i/ after final /tɕ/	
가지 [kɑtɕi] ‘branch’	소지 [sotɕi] ‘possession’
거지 [kʌtɕi] ‘beggar’	오지 [otɕi] ‘outback’
무지 [mutɕi] ‘ignorance’	유지 [jutɕi] ‘maintenance’
미지 [mitɕi] ‘unknown’	차지 [tɕʰɑtɕi] ‘occupation’
사지 [sɑtɕi] ‘limbs’	처지 [tɕʰʌtɕi] ‘position’
10 lexical /i/ after final /tɕʰ/	
가치 [kɑtɕʰi] ‘value’	수치 [sutɕʰi] ‘shame’
구치 [kutɕʰi] ‘custody’	이치 [itɕʰi] ‘reason’
비치 [pitɕʰi] ‘equipping’	어치 [ʌtɕʰi] ‘worth of’
배치 [pætɕʰi] ‘arrangement’	차치 [tɕʰatɕʰi] ‘leaving aside’
사치 [sɑtɕʰi] ‘luxury’	처치 [tɕʰʌtɕʰi] ‘handling’

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary