Journal of Pan-Pacific Association of Applied Linguistics, 27(1), 43–60.
https://doi.org/10.25256/PAAL.27.1.3

Covert Contrasts in Korean EFL Learners' Interlanguage Phonology*

Jayeon Lim and Misun Seo**

Lim, J., & Seo, M. (2023). Covert contrasts in Korean EFL learners' interlanguage phonology. *Journal of Pan-Pacific Association of Applied Linguistics, 27*(1), 43-60.

This study examined how Korean EFL learners in different proficiency levels attest contrasts of L2 phonology. To this end, twenty Korean EFL learners in high and low levels of English proficiency produced words with word-final contrasts of /ʃ/ vs. /ʃi/, /tʃ/ vs. /tʃi/, and /dʒ/ vs. /dʒi/. Acoustic analyses of lexical vs. epenthetic vowels and lexically vs. derivationally final sibilants were conducted. The results of the acoustic analyses revealed that Korean learners differentiated them at the acoustic level by employing different vowel ratios or F1 for the produced forms with a final lexical or epenthetic vowel which did not differ at the segmental level. On the other hand, the low group did not acoustically differentiate the two types of vowels after /dʒ/ at all. With the production of a final sibilant either lexical or derivational, Korean learners also made distinctions in terms of different frication ratios. The results of the study found L2 learners distinguish segmental contrasts acoustically, which is different from L1 speakers, and the proficiency level partially mediates this. The results implicate the existence of interlanguage phonology with its independent system.

Keywords: L2 production, covert contrasts, acoustic distinctions, interlanguage phonology

1 Introduction

The term of "interlanguage" was first introduced by Selinker (1972) who claimed that L2 learners' language, however deviant it may be from the target language, has a system of its own. Since then, a plethora of research has supported the interlanguage hypothesis. For instance, studies on L2 grammar have shown that there are patterns of errors from which the universal developmental stages of language learning can be construed; studies of interlanguage variability have recognized that learners' interlanguage systems

* This work was supported by the 2022 Research Fund of the University of Seoul.

** First Author: **Jayeon Lim**, Professor, Department of English Language and Literature, University of Seoul
Corresponding Author: **Misun Seo**, Professor, Department of English Language and Literature, Hannam University

exhibit variability due to factors such as L1, learning environment, and/or individual factors such as strategies, motivation, etc.; interlanguage pragmatics found that generalizations can be made regarding learners' pragmatic competence. To summarize, these studies evidenced the presence of the independent and complete system of interlanguage as proposed by the Interlanguage Hypothesis in that learners' language system conforms within the boundary of language universals.

Despite abundant research in domains of L2 such as grammar and pragmatics, the inquiry into interlanguage phonology has received limited attention. Barrientos Contreras (2018) has pointed out that it may be due to a lack of interaction between the field of phonology and second language acquisition (SLA). Additionally, it could have been due to the widely accepted Critical Period Hypothesis, where phonological development of reaching the native speaker level or even showing developmental progress as adults is robustly refuted.

Nonetheless, an increasing body of research has been employing the term 'interlanguage phonology' (Major, 1998; Tarone, 1979) capturing the intermediate nature of the categories produced by L2 learners. That is, Selinker's notion of interlanguage stands in phonology as well, supporting the presence of independent categories of interlanguage sounds. The present study sets out to investigate the nature of the L2 contrasts by examining Korean EFL learners. Specifically, the study investigates interlanguage phonology by examining word-final contrasts of /ʃ/ vs. /ʃi/, /tʃ/ vs. /tʃi/, and /dʒ/ vs. /dʒi/ produced. The participants included Korean EFL learners in different proficiency levels in order to examine developmental patterns. Thus, the study sets out to examine whether L2 learners will exhibit acoustic distinctions even when segmental distinctions are absent in their production of the above-mentioned segmental contrasts.

2 Literature Review

Early studies supporting the interlanguage hypothesis have mostly investigated L2 grammar and pragmatics, yet interlanguage research has also focused on L2 phonology in some later studies. Some key areas of L2 interlanguage phonology have focused on how learners develop their accents and pronunciation in the target language in terms of segmental and suprasegmental features as well as phonological processes such as assimilation, deletion, substitution, etc. in their interlanguage productions. In doing so, studies have focused on the role of L1 transfer as well as language universals. In this vein, studies in interlanguage phonology are in line with those of interlanguage in other domains such as grammar or pragmatics in that both L1 transfer and/or language universals seem to be operating. Thus, the results have provided some informative insights to language teachers and the development of effective strategies for improving learners' phonological proficiency.

More recently, what is more specifically relevant to interlanguage phonology is "covert contrasts", where L2 learners fail to recognize the phonemic contrasts that exist in the target language in their pronunciation but may still undergo making acoustical differences. This means that learners may be unaware of or unable to produce specific phonemic distinctions that exist in the L2 they are learning. (Song & Eckman, 2019). Yet, according to Li et al. (2009), these nonnatives do produce contrastive sounds that exhibit statistically significant acoustic differences between the two phonemes. Thus, the results are in line with Selinker in that learners do exhibit a system of their own.

What needs to be noted is that L2 learners attempt to realize meaning differences in their interlanguage phonology. Although such efforts may not always turn out to be successful since native language speakers may fail to notice the differences between the attempted sounds and the target ones, it is significant in that the attempt results in significant acoustic differences. By and large, L2 learners who make these significant acoustic differences seem to be aware of what is crucial for meaningful communication.

The implication of covert contrasts was demonstrated in studies of L1 children and those with disordered speech. In Macken and Barton (1980), it was shown that children learning voicing contrasts would go through the stage of making significantly different VOTs even when the acoustic differences were not detected phonemically by adults. Additionally, according to Li et al. (2009), speech-disordered children who exhibit covert contrasts turn out to make a better prognosis than those who do not. Thus, producing significant acoustic differences will lead to better development.

However, research on interlanguage development, more specifically on interlanguage phonology, has not focused on examining what kind of significant acoustic differences L2 learners create, but rather on the fact that L2 phonology fails to make phonemic differences of L1 phonology. Although there is a vast number of studies that have reported the findings from L2 perception and production with results on the sound systems of interlanguage phonology, Song and Eckman (2019) rightfully mentioned that focusing on acoustic differences has been the focus in only a few recent studies.

Eckman et al. (2014) investigated the acquisition of the English /s/-/z/ by Spanish native speakers where [s] and [z] are allophones of the phoneme /s/. Results indicated that only seven of the 14 participants were able to make phonemic distinctions between [s] and [z] in English. However, the acoustic analysis revealed that four of the other seven, who were unable to make phonemic distinctions, were able to administer a statistically significant acoustic difference measured by frication. In a further study by Eckman et al. (2015), L1 speakers of Arabic where [p] and [b] are allophones of the phoneme /b/ were investigated. It was found that five participants out of nine did not distinguish between /p/ and /b/, but two out of those five still could show a statistically significant difference in VOT between the two phonemes. In both studies, it was found that covert contrasts were realized among at least some of the L2 learners.

In the previous studies by Lim and Seo (2016), Korean EFL learners participated in the study to answer the question of whether they will exhibit acoustic differences when producing sound contrasts of /ʃ/ and /ʃi/. The results indicated that the participants exhibited differences according to their proficiency levels. Specifically, when producing word-final /ʃ/ and /ʃi/, both high- and low-level learners exhibited error patterns of deleting vowels or producing epenthetic vowels in the final positions. However, unlike the low group who did not distinguish epenthetic vowels (i.e., vowels produced incorrectly) from lexical vowels (i.e., vowels produced correctly), the high group exhibited significant acoustic differences between epenthetic and lexical vowels. For instance, the high group's epenthetic vowels exhibited shorter vowel ratio, higher F1 and lower F2 than lexical vowels. While the low group did not distinguish epenthetic vowels from lexical vowels, the high group distinguished the two with a possibility of treating epenthetic vowels as transitional ones between target-like and incorrect realizations.

In a further study investigating L2 production of English affricates and fricatives in word-final position across three different recordings, Seo and Lim (2023) attested that the high group's production showed better accuracy as the recordings progressed. The study also mentioned that word frequency may have mediated the results since words ending with consonants are more frequent than those ending with vowels. Overall, the reviewed studies have made similar claims in that L2 learners make acoustically significant distinctions even when no distinctions are made at the segmental level.

In continuation of Lim and Seo (2016) and Seo and Lim (2023), the present study focuses on the production of fricatives and affricates, i.e., /ʃ/ vs. /ʃi/, /tʃ/ vs. /tʃi/, and /dʒ/ vs. /dʒi/. In their production of the above contrasts, would Korean EFL learners accurately manifest the consonants and vowels in word-final position? There are four possibilities: (1) accurately produced forms with a lexically final sibilant (i.e., *fish* as [fɪʃ]) (2) accurately produced forms with a lexical vowel (i.e., *fishy* as [fɪʃi]) (3) inaccurately produced forms with a derivationally final sibilant by deleting a final vowel (i.e., *fishy* as [fɪʃ]) and (4) inaccurately produced forms with an epenthetic vowel (i.e., *fish* as [fɪʃV]). Out of the above four, (1) and (2) would indicate correctly pronounced contrasts and thus be referred to as overt contrasts. On the other hand, (3) and (4) would indicate misrepresentations and may include covert contrasts. Our goal would be to examine whether Korean EFL learners would exhibit covert contrasts when pronouncing the contrasts mentioned above by acoustically comparing (1) and (3) as well as (2) and (4). Also, the study would examine whether the presence of covert contrasts is different depending on the consonant types. Finally, the study sets out to investigate if there is a developmental trend as L2 proficiency increases. Thus, the following research questions were postulated:

(1) Do Korean EFL learners show covert contrasts in the production of the L2 word-final contrasts of /ʃ/ vs. /ʃi/, /tʃ/ vs. /tʃi/, and /dʒ/ vs. /dʒi/?
(2) Would their production be influenced by different consonants?
(3) What is the developmental trend as L2 proficiency increases?

3 Research Methodology

3.1 Participants

Twenty Korean learners of English who were university students conducted the production experiment. Ten (5 males, 5 females) of them had TOEIC scores of 600s or below and they were considered the low group. The other ten (5 males, 5 females) with TOEIC scores of 900s were regarded as the high group. In addition, four native speakers of English (2 males, 2 females) performed the production experiment to provide native English speakers' baseline data. All the native speakers of English were students at a university in the U.S.A.

3.2 Stimuli

The stimuli of the production experiment were 30 pairs of English words with the contrasts of /ʃ/ vs. /ʃi/, /tʃ/ vs. /tʃi/, or /dʒ/ vs. /dʒi/ in word-final position. A full list of the words is provided in Table 1.

Table 1. Words for the Production Experiment

Contrasts	Words
/ʃ/ vs. /ʃi/	fish-fishy, mush-mushy, dash-dashy, mesh-meshy, bush-bushy ash-ashy, push-pushy, wash-washy, marsh-marshy, dish-dishy
/tʃ/ vs. /tʃi/	munch-munchy, catch-catchy, witch-witchy, peach-peachy botch-botchy, bunch-bunchy, beach-beachy, itch-itchy, touch-touchy, punch-punchy
/dʒ/ vs. /dʒi/	hedge-hedgy, wedge-wedgy, ridge-ridgy, budge-budgie, podge-podgy, cage-cagey, edge-edgy, dodge-dodgy, pudge-pudgy, ledge-ledgy

3.3 Procedure

Each word in Table 1 was put in a carrier sentence "please say __________" to elicit natural speech. For each participant, three lists of sentences including target words were made by three different randomization processes. Therefore, each participant made recordings three times for each word embedded in a carrier sentence. Participants were asked to read the sentences as naturally as possible. Korean learners of English made recordings in a quiet room while native English speakers did in a sound-attenuated room. All recordings were made at a 44,100 Hz sampling rate and 16 bit quantization. Through the production experiment, 3600 tokens (20 speakers × 60 words × 3 repetitions) were collected from Korean learners of English and 720 tokens (4 speakers × 60 words × 3 repetitions) from native speakers of English.

3.4 Analysis

The 4320 tokens collected were analyzed by using the *Praat* program. Based on the spectrogram and waveform of each token, it was first examined whether or not the final segment was correctly pronounced. For the tokens ending in /ʃi/, /tʃi/, or /dʒi/, Korean learners of English exhibited two pronunciation patterns: correct pronunciation (i.e., *fishy* as [fɪʃi]) and incorrect pronunciation by deleting the final vowel /i/ (i.e., *fishy* as [fɪʃ]). In the case of the tokens ending in /ʃ/, /tʃ/, or /dʒ/, Korean learners of English pronounced them correctly (i.e., *fish* as [fɪʃ]) or incorrectly by inserting a vowel in final position (i.e., *fish* as [fɪʃV]). Of the tokens produced by native speakers of English, there was one token mispronounced. One native English speaker incorrectly pronounced 'hedgy' as 'hedge' in one recording and it was not included in the acoustic and statistical analyses.

The acoustic analysis involved measurements of word duration. Because some of the stimuli used in the experiment did not contain a word-initial consonant, word duration was measured by excluding a word-initial consonant. For the produced forms with a final vowel (i.e., both a lexical vowel when pronouncing *fishy* as [fɪʃ**i**] and an epenthetic vowel when pronouncing *fish* as [fɪʃ**V**]), measurements of vowel duration were made. The first and last full glottal pulses displayed in the waveform and spectrogram were considered the indicators of the onset and offset of a vowel. The ratio of vowel duration within word duration was calculated to investigate the difference between lexical and epenthetic vowels. In addition, at the midpoint of a lexical or epenthetic vowel, the first and second formant (i.e., F1 and F2) frequencies were measured. For the produced forms with a final sibilant (i.e., both a lexically final sibilant when pronouncing *fish* as [fɪʃ] and a derivationally final sibilant when pronouncing *fishy* as [fɪʃ]), frication duration was measured by referring to the onset and offset of aperiodic noise in the waveform and spectrogram. Then, the ratio of frication duration within word duration was calculated.

4 Results

Statistical analyses were made separately for lexical vs. epenthetic vowels and for lexically vs. derivationally final sibilants.

4.1 Lexical vs. epenthetic vowels

The number of tokens with lexical or epenthetic vowels within each subject group is illustrated in Table 2 below.

Table 2. Number of Tokens with Lexical or Epenthetic Vowels

	Low	High	Native
Lexical Vowel	417	669	359
Epenthetic Vowel	303	57	

Of the 900 tokens ending in /ʃi/, /tʃi/, or /dʒi/ produced by the low group, 417 tokens (i.e., 46.3%) were correctly pronounced with a lexical vowel. In the case of the high group, 669 tokens out of 900 (i.e., 74.3%) were correctly pronounced with a lexical vowel. As for the words ending in /ʃ/, /tʃ/, or /dʒ/, the low group exhibited incorrect pronunciation in 303 tokens out of 900 (i.e., 33.7%) by inserting a vowel in final position while the high group did so in 57 tokens out of 900 (i.e., 6.3%). Detailed analysis of the accuracy data is provided in Seo and Lim (2023).

For the 1805 tokens illustrated in Table 1, a multivariate analysis of variance (MANOVA) was conducted with the ratio of a final vowel within a word, F1, and F2 as dependent variables. In the analysis, independent variables were subjects' group (low vs. high vs. native), vowel type (lexical vs. epenthetic), and consonant type before a final vowel (/ʃ/ vs. /tʃ/ vs. /dʒ/). There was a significant interaction of subjects' group * vowel type with the ratio of the final vowel ($df = 1$, $F = 19.919$, $p < .05$) and F1 ($df = 1$, $F = 212.067$, $p < .05$). Figure 1 illustrates the interaction of subjects' group * vowel type with vowel ratio.

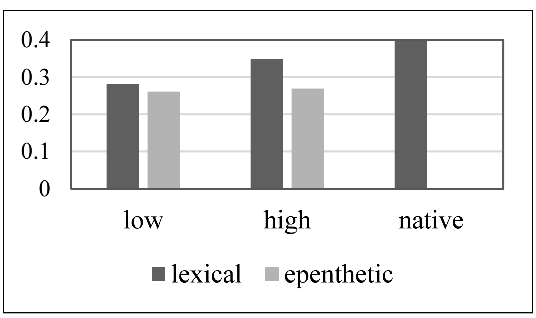


Figure 1. Mean final vowel ratio by group * vowel type

The mean vowel ratio of a lexical vowel was the highest for the native at .396 followed by the high at .349 and the low at .282. According to a post-hoc test of LSD, there were significant differences for native vs. high, high vs. low, and native vs. low ($p < .05$). Thus, Korean learners of English showed a tendency to pronounce a final lexical vowel with a shorter duration than native English speakers and this pattern was more evident for the low group. As shown in Figure 1, both the low and high groups pronounced an epenthetic vowel significantly shorter than a lexical vowel (.282 vs. .261 for the low; .349 vs. .269 for the high) ($p < .05$) while the vowel ratio difference between the two was greater within the high group.

Figure 2 illustrates the interaction of subjects' group * vowel type with F1.

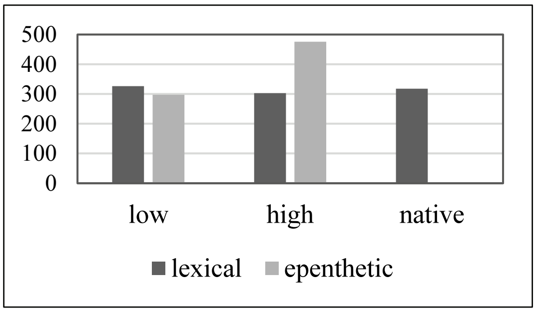
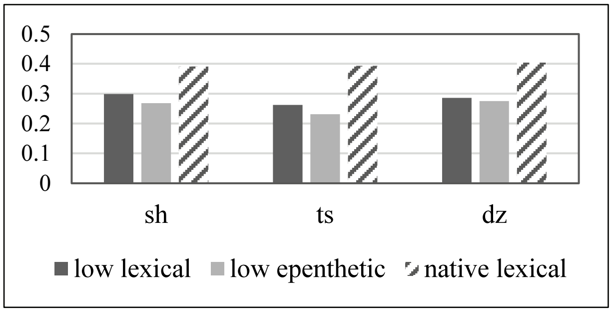


Figure 2. Mean F1 by group * vowel type

The mean F1 of a lexical vowel was the highest for the low at 326.4 Hz followed by the native at 317.7 Hz and the high at 302.7 Hz. A post-hoc test of LSD revealed significant differences for low vs. high and high vs. native ($p < .05$) and no significant difference for low vs. native ($p = .276$). Thus, the high group's mean F1 of a lexical vowel was significantly lower than the low or native group. The low group's mean F1 of an epenthetic vowel was significantly lower than that of a lexical vowel (297 Hz vs. 326.4 Hz) ($p < .05$). The high group showed the opposite pattern: the mean F1 of an epenthetic vowel was significantly higher than that of a lexical vowel (475.4 Hz vs. 302.7 Hz) (p < .05). In addition, the mean F1 difference between the two types of vowels was much greater within the high group.

As for the interaction of subjects' group * consonant type * vowel type, a marginally significant interaction was attested with vowel ratio ($df = 2$, $F = 2.722$, $p = .066$) and a significant interaction with F1 ($df = 2$, $F = 56.784$, $p < .05$). Figure 3 shows the interaction of consonant type * vowel type with vowel ratio within the low and native groups.



Note: sh = ʃ, ts = tʃ, dz = dʒ

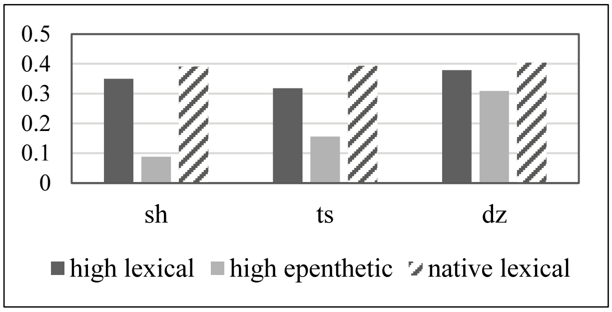
Figure 3. Low's and native's mean vowel ratio by consonant type * vowel type

In the native English group, the mean vowel ratio of a lexical vowel was the highest for /dʒ/ at .404 followed by /tʃ/ at .393 and /ʃ/ at .391. In the low group, it was the highest for /ʃ/ at .299 followed by /dʒ/ at .286 and /tʃ/ at .262. Within the native group, the mean vowel ratio of a lexical vowel was not significantly different according to consonant type ($df = 2$, $F = 1.509$, $p = .223$). On the other hand, the low group's mean vowel ratio of a lexical vowel was significantly different according to consonant type ($df = 2$, $F = 4.46$, $p < .05$). A post-hoc test of LSD revealed significant differences for /ʃ/ vs. /tʃ/ and /dʒ/ vs. /tʃ/ ($p < .05$). Thus, the mean vowel ratio was significantly higher for /ʃ/ and / dʒ/ than for / tʃ/.

In the case of an epenthetic vowel, the low group's vowel ratio was highest after /dʒ/ at .275 followed by after /ʃ/ at .268 and after /tʃ/ at .231. As for the low group's vowel ratio of an epenthetic vowel, there was a marginally significant difference according to consonant type ($df = 2$, $F = 2.957$, $p = .054$). According to a post-hoc test of LSD, a significant difference was attested for /tʃ/ vs. /dʒ/ ($p < .05$), illustrating that the mean vowel ratio of an epenthetic vowel was significantly lower after /tʃ/ than after /dʒ/.

The low group's vowel ratio of a lexical and an epenthetic vowel was significantly lower than the native group's vowel ratio of a lexical vowel ($p < .05$). Within the low group, the mean vowel ratio difference between a lexical and an epenthetic vowel was marginally significant after /ʃ/ ($t = 3.666$, $p = .507$) and after /tʃ/ ($t = 3.385$, p = .067) while it was not significant after /dʒ/ ($t = .653$, $p = .42$). Therefore, the low group showed a tendency to differentiate an epenthetic vowel from a lexical vowel in terms of the vowel ratio after /ʃ/ and /tʃ/, but the two types of vowels were not distinguished by the vowel ratio after /dʒ/.

Figure 4 illustrates the interaction of consonant type * vowel type with vowel ratio within the high and native groups.



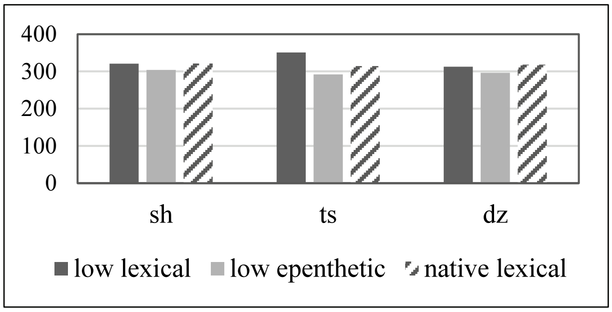
Note: sh = ʃ, ts = tʃ, dz = dʒ

Figure 4. High's and native's mean vowel ratio by consonant type * vowel type

In the high group, the mean vowel ratio of a lexical vowel was the highest for /dʒ/ at .379 followed by /ʃ/ at .35and /tʃ/ at .318. There was a significant difference regarding the vowel ratio of a lexical vowel according to consonant type (*df* = 2, *F* = 5.821, *p* < .05). According to a post-hoc test of LSD, a significant difference was found for /dʒ/ vs. /tʃ/ (*p* < .05). The high group's mean vowel ratio of an epenthetic vowel was the highest for /dʒ/ at .309 followed by /tʃ/ at .156 and /ʃ/ at .088. The vowel ratio of an epenthetic vowel was found to be significantly different according to consonant type (*df* = 2, *F* = 14.183, *p* < .05). A post-hoc test of LSD showed significant differences for /dʒ/ vs. /tʃ/ and /dʒ/ vs. /ʃ/, illustrating that the vowel ratio of an epenthetic vowel was significantly higher when the vowel occurs after /dʒ/.

The high group's vowel ratio of a lexical vowel was significantly different from the native group's after /ʃ/ (*t* = 26.079, *p* < .05) and after /tʃ/ (*t* = 102.885, *p* < .05) while it was not after /dʒ/ (*t* = .755, *p* = .386). In addition, as for all consonant types, the high group's vowel ratio of an epenthetic vowel was significantly different from the native group's vowel ratio of a lexical vowel ((*t* = 75.165, *p* < .05 for /ʃ/; *t* = 1145.749, *p* < .05 for /tʃ/; *t* = 45.963, *p* < .05 for /dʒ/). The vowel ratio difference between a lexical and an epenthetic vowel was significantly different for /ʃ/ (*t* = 50.071, *p* < .05) and /tʃ/ (*t* = 442.689, *p* < .05) while it was not for /dʒ/ (*t* = 2.17, *p* = .142). Within the high group, compared with the low group, there was a more evident pattern differentiating the two types of vowels by realizing an epenthetic vowel with a lower vowel ratio after /ʃ/ and /tʃ/. However, after /dʒ/, there was no vowel ratio difference between a lexical and an epenthetic vowel as in the low group.

Figure 5 shows the interaction of consonant type * vowel type with F1 within the low and native groups.



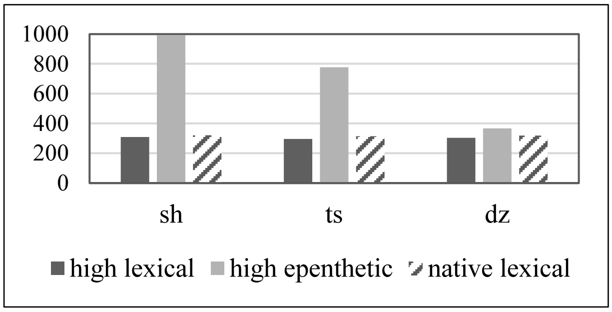
Note: sh = ʃ, ts = tʃ, dz = dʒ

Figure 5. Low's and native's mean F1 by consonant type * vowel type

Within the native group, F1 of a lexical vowel was the highest after /ʃ/ at 320.9 Hz followed by after /dʒ/ at 318.1 Hz and /tʃ/ at 314.1 Hz. According to an analysis of variance (ANOVA), the native group's F1 of a lexical vowel was not significantly different according to consonant type (df = 2, F = .625, p = .536). Within the low group, F1 of a lexical vowel was the highest after /tʃ/ at 350.6 Hz followed by after /ʃ/ at 320.4 Hz and after /dʒ/ at 312.6 Hz. In the case of the low group's lexical vowel, there was no significant difference of F1 according to consonant type (df = 2, F = 1.958, p = .142). In addition, with respect to F1 of a lexical vowel, the low and native groups were not significantly different for all consonant types (t = .003, p = .957 for /ʃ/; t = 2.443, p = .119 for /tʃ/; t = .23, p = .632 for /dʒ/).

In the low group, F1 of an epenthetic vowel was the highest after /ʃ/ at 304 Hz followed by after /dʒ/ at 295.9 Hz and after /tʃ/ at 292 Hz. An ANOVA revealed no significant difference according to consonant type (df = 2, F = .202, p = .817). The F1 difference between the low's epenthetic vowel and the native's lexical vowel was significant for /dʒ/ (t = 4.362, p = .038) and marginally significant for /tʃ/ (t = 3.341, p = .069) while it was not significant for /ʃ/ (t = 1.314, p = .253). As for the low group's F1 difference between a lexical and an epenthetic vowel, a significant difference was found only after /tʃ/ (t = 3.876, p = .05). After /ʃ/ or /dʒ/, there was no significant difference.

Figure 6 exhibits the interaction of consonant type * vowel type with F1 within the high and native groups.



Note: sh = ʃ, ts = tʃ, dz = dʒ

Figure 6. High's and native's mean F1 by consonant type * vowel type

In the high group, F1 of a lexical vowel was the highest after /ʃ/ at 308.2 Hz followed by after /dʒ/ at 303.6 Hz and after /tʃ/ at 296.1 Hz. An ANOVA revealed no significant difference of F1 according to consonant type ($df = 2$, $F = 1.105$, $p = .332$). Regarding F1 of a lexical vowel, the high and native groups did not show any significant differences after /ʃ/ ($t = 2.813$, $p = .094$) and after /dʒ/ ($t = 1.772$, $p = .184$). On the other hand, the two groups showed a significantly different F1 after /tʃ/ ($t = 7.156$, $p < .05$).

As for an epenthetic vowel, the high group's F1 was the highest after /ʃ/ at 992.5 Hz followed by after /tʃ/ at 776.8 Hz and after /dʒ/ at 366.8 Hz. According to an ANOVA, F1 of an epenthetic vowel was significantly different by consonant type ($df = 2$, $F = 6.197$, $p < .05$). A post-hoc test of LSD showed significant differences for /ʃ/ vs. /dʒ/ and /tʃ/ vs. /dʒ/ ($p < .05$), illustrating that F1 of an epenthetic vowel was significantly higher after /ʃ/ or /tʃ/ than after /dʒ/. In the case of the F1 difference between the high's epenthetic vowel and the native's lexical vowel, significant differences were observed for /ʃ/ ($t = 88.739$, $p < .05$) and /tʃ/ ($t = 50.628$, $p < .05$) while there was no significant difference for /dʒ/ ($t = 2.791$, $p = .097$). In addition, within the high group, F1 between a lexical and an epenthetic vowel was significantly different for all consonant types ($t = 129.407$, $p < .05$ for /ʃ/; $t = 87.902$, $p < .05$ for /tʃ/; $t = 5.424$, $p < .05$ for /dʒ/). F1 of an epenthetic vowel was significantly higher than that of a lexical vowel for all consonant types.

4.2 Lexically vs. derivationally final sibilants

Table 3 below illustrates the number of tokens with lexically or derivationally final sibilants within each subject group.

Table 3. Number of Tokens with Lexically or Derivationally Final Sibilants

	Low	High	Native
Lexically	597	844	360
Derivationally	483	230	

Of the 900 tokens with word-final /ʃ/, /tʃ/, or /dʒ/ produced by the low group, 597 tokens (i.e., 66.3%) were correctly pronounced by correctly realizing a lexically final sibilant without vowel insertion. In the high group, 844 tokens out of 900 (i.e., 93.8%) were correctly pronounced with a lexically final sibilant. In the case of the words ending in /ʃi/, /tʃi/, or /dʒi/, the low group showed incorrect pronunciation in 483 tokens out of 900 (i.e., 53.7%) with a derivationally final sibilant created by deleting a final vowel. On the other hand, the high group did so in 230 tokens out of 900 (i.e., 25.6%). Detailed analysis of the accuracy data is provided in Seo and Lim (2023).

For the 2514 tokens illustrated in Table 1, an ANOVA was performed with the ratio of frication within a word as a dependent variable. In the analysis, independent variables were subjects' group (low vs. high vs. native), final sibilant type (lexical vs. derivational), and consonant type involved (/ʃ/ vs. /tʃ/ vs. /dʒ/). A significant interaction of subjects' group * final sibilant type was attested ($df = 1$, $F = 3.996$, $p < .05$). Figure 7 illustrates the finding.

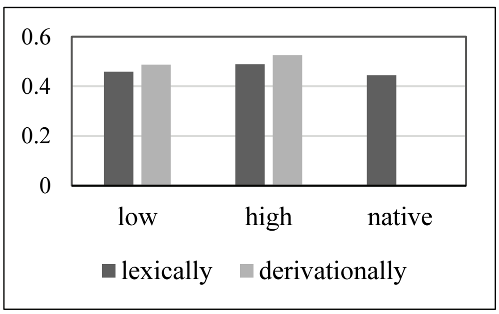


Figure 7. Mean frication ratio by group * final sibilant type

The frication ratio of a lexically final sibilant was the highest for the high at .489 followed by the low at .459 and the native at .445. According to an ANOVA, the frication ratio of a lexically final sibilant was significantly different by subjects' group ($df = 2$, $F = 18.933$, $p < .05$). In terms of a post-hoc test of LSD, significant differences were found for high vs. low and high vs. native ($p < .05$) while there was no significant difference for low vs. native ($p = .089$). Therefore, the high group's frication ratio of a lexically final sibilant was significantly higher than the other two groups.

As for a derivationally final sibilant, the high group illustrated a higher frication ratio than the low group (.526 vs. .487) and the difference was significant ($t = 11.637$, $p < .05$). Thus, the high group tended to pronounce both types of final sibilants with longer duration. An ANOVA was performed to see if Korean learners' frication ratio of a derivationally final sibilant was significantly different from the native's frication ratio of a lexically final sibilant. According to the analysis, there was a significant main effect of subjects' group (*df* = 2, *F* = 28.345, *p* < .05). A post-hoc test of LSD showed significant differences for low vs. high, high vs. native, and low vs. native (*p* < .05). Thus, Korean learners showed a tendency to pronounce a derivationally final sibilant with longer duration compared with the native's lexically final sibilant. Noteworthy is that the high group's frication ratio of a derivationally final sibilant was higher than the low group's. As for the frication ratio difference between the two types of sibilants, significant differences were attested in the low (*t* = 10.559, *p* < .05) and high (*t* = 15.273, *p* < .05) groups. Thus, both groups of Korean learners differentiated the two types of sibilants by realizing a derivationally final sibilant with a higher frication ratio than a lexically final sibilant.

5 Discussion

The results of the present study showed that both low and high group learners of Korean differentiated lexical vowels from epenthetic vowels in terms of different acoustic realizations. The vowel ratio of an epenthetic vowel was significantly lower than that of a lexical vowel within both groups of Korean learners. Thus, Korean learners pronounced an epenthetic vowel with shorter duration than a lexical vowel and this trend was more evident in the high group. In addition, Korean learners also differentiated the two types of vowels in terms of F1. The low group showed significantly lower F1 for an epenthetic vowel than for a lexical one while the high group's F1 was significantly higher for an epenthetic vowel than for a lexical one. F1 is an acoustic property inversely related to vowel height (Ladefoged & Johnson, 2015). Therefore, the results showed that the low group pronounced an epenthetic vowel as a higher vowel than a lexical vowel while the high group pronounced an epenthetic vowel as a lower vowel than a lexical vowel. The F1 difference between the two types of vowels was greater in the high group.

These results illustrate that Korean learners made covert contrasts in their production of the word-final contrasts of /ʃ/ vs. /ʃi/, /tʃ/ vs. /tʃi/, and /dʒ/ vs. /dʒi/. The contrasts should be made at the segmental level, that is, in terms of the absence or presence of a final vowel. However, when Korean learners pronounced the contrasts with the same produced forms having a final vowel (i.e., a lexical or epenthetic vowel), they made contrasts by realizing the two types of vowels differently at the acoustic level. Thus, although their covert contrasts were not readily identified, Korean learners were aware of the phonological contrasts and tried to differentiate the contrasts by employing different acoustic values. The greater vowel ratio or F1 difference between the two types of vowels observed in the high group suggests that the high group made more efforts to differentiate the contrasts in their production.

Korean learners' acoustic properties of a lexical or an epenthetic vowel were affected by the type of a consonant positioned before a vowel. Within the low group, the vowel ratio of a lexical or an epenthetic vowel was higher after /ʃ/ or /dʒ/. This result might have been influenced by L1 and a language universal factor. In Korean, [ʃ] occurs as an allophone of /s/ before /i/ or /j/ (Shin, 2000). Considering this L1 phonology, Korean learners' pronunciation of [ʃ] will be easier when the vowel [i] occurs after the consonant, which may have induced the relatively higher vowel ratio of a final vowel after [ʃ]. In addition, according to Kang (2003), an epenthetic vowel is more likely to occur after a voiced stop than after a voiceless stop. This is because of the perceptual similarity between a vowel and the release/voicing of a voiced stop. Taking this into account, an epenthetic vowel will be more likely to be inserted after the voiced /dʒ/ than after the voiceless /ʃ/ or /tʃ/ and this may have caused the higher vowel ratio after /dʒ/, too. Within the high group, the vowel ratio of a lexical vowel was significantly higher after /dʒ/ than after /tʃ/. However, as for an epenthetic vowel, the vowel ratio was significantly lower after /ʃ/ (as well as /tʃ/) than after /dʒ/ and it is not clear why the high group's vowel ratio of an epenthetic vowel was low after /ʃ/. For a detailed explanation of such a pattern, more research will be needed.

There was the influence of a prevocalic consonant on the acoustic difference between a lexical and an epenthetic vowel. The low group differentiated the two types of vowels in terms of the higher vowel ratio for lexical vowels than for epenthetic vowels after /ʃ/ or /tʃ/. However, after /dʒ/, there was no vowel ratio difference between the two types of vowels. Regarding F1, the low group distinguished the two types of vowels only after /tʃ/. That is, a lexical vowel after /tʃ/ exhibited higher F1 than an epenthetic vowel. After /ʃ/ or /dʒ/, a lexical and an epenthetic vowel were not significantly different with respect to F1. Thus, the low group did not differentiate the two types of vowels after /dʒ/ even at the acoustic level, not to mention at the segmental level.

In the case of the high group, like the low group, the two types of vowels were distinguished in terms of the vowel ratio after /ʃ/ or /tʃ/. The vowel ratio of a lexical vowel was significantly higher than that of an epenthetic vowel after /ʃ/ or /tʃ/. On the other hand, the vowel ratio of the two types of vowels was not significantly different after /dʒ/. With respect to F1, regardless of a prevocalic consonant type, the high group differentiated the vowels in terms of significantly higher F1 for an epenthetic vowel than for a lexical vowel. Thus, regarding the contrasts not differentiated at the segmental level, the high group made covert contrasts at the acoustic level by employing either the vowel ratio or F1. Overall, the study revealed a significant difference between the high and low groups.

In the case of lexically final sibilants (i.e., [fɪʃ] from *fish*) and derivationally final sibilants (i.e., [fɪʃ] from *fishy*), both groups of Korean learners differentiated the two types of sibilants by realizing derivationally final sibilants with higher frication ratio than lexically final sibilants. This result shows that Korean learners differentiated the word-final contrasts of /ʃ/

vs. /ʃi/, /tʃ/ vs. /tʃi/, and /dʒ/ vs. /dʒi/ in their own way at the acoustic level when they mispronounced /ʃi/, /tʃi/, and /dʒi/ as [ʃ], [tʃ], and [dʒ], respectively, resulting in no contrast at the segmental level. This tendency was more evident in the high group, as indicated by the greater frication ratio difference between the two types of fricatives.

6 Conclusion

This study examined the word-final contrasts of /ʃ/ vs. /ʃi/, /tʃ/ vs. /tʃi/, and /dʒ/ vs. /dʒi/ produced by Korean learners of English at two different proficiency levels (i.e., high and low), focusing on the acoustic analyses of lexical vs. epenthetic vowels and lexically vs. derivationally final sibilants. As for the produced forms not differentiated at the segmental level by containing a final lexical or epenthetic vowel, Korean learners differentiated them at the acoustic level by employing different vowel ratios or F1 although the low group did not acoustically differentiate the two types of vowels after /dʒ/ at all. In the case of the produced forms with a final sibilant (i.e., lexically or derivationally), Korean learners also made covert contrasts in terms of different frication ratios. The results of the present study showed that segmental contrasts can be realized at the acoustic level within L2 learners' production, but not at the segmental level.

To summarize, this study showed that the presence of significant acoustic differences between sound contrasts by Korean EFL learners. The results substantiated the claim made by the interlanguage hypothesis mentioned earlier, in that L2 learners showed a system of their own different from native speakers at the segmental level but significant at the acoustic level. In that sense, we can make a claim that interlanguage phonology substantiates a pattern of its own evidencing the cognitive and mental representations of distinctions. Additionally, a significant difference between the high and low groups indicates that development has occurred as the proficiency level increased. Whether the high group will continue to develop to show segmental level distinctions is beyond the scope of this study. A further study involving near-native speakers or those who began learning as young children will be able to shed light on this question.

References

Barrientos Contreras, F. A. (2018). *Perceptual representations in interlanguage phonology: Subcategorial learning in late-learners with a smaller vowel inventory* [Unpublished doctoral dissertation, The University of Manchester].

Eckman, F., Iverson, G., & Song, J. Y. (2014). Covert contrast in the acquisition of second language phonology. In F. T., Ashley & J. Barlow (Eds.), *Perspectives on phonological theory and acquisition: Papers in honor of Daniel A. Dinnsen* (pp. 25-48). John Benjamins Publishers.

Eckman, F., Iverson, G., & Song, J. Y. (2015). Overt and covert contrast in L2 phonology. *Journal of Second Language Pronunciation 1*, 254-278.

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

Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th ed.). Cengage Learning.

Li, F., Edwards, J., & Beckman, M. E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics, 37,* 111-124.

Lim, J., & Seo, M. (2016). Korean Learners' Production of English Sound Contrast: Focusing on Word-Final /ʃ/ and /ʃi/. *Journal of Asia TEFL, 13*(1), 16-30.

Macken. M. A., & Barton, D. (1980). A longitudinal study of the acquisition of the voicing contrast in American-English word-initial stops, as measured by voice onset time. *Journal of Child Language, 7*, 41-74.

Major, R. (1998). Interlanguage phonology: An introduction. *Studies in Second Language Acquisition, 20*(2), 131-137.

Selinker, L. (1972). Interlanguage. *IRAL: International Review of Applied Linguistics in Language Teaching, 10*, 209-241.

Seo, M., & Lim, J. (2023). Factors affecting L2 production of English affricates and fricatives in word-final position. *Modern Studies in English Language and Literature, 67*(1), 265-85.

Shin, J. (2000). *Malsoriui ihae* [*Understanding speech sounds*]. Hanguk Munhwasa.

Song, J., & Eckman, (2019). Covert contrasts in the acquisition of English high front vowels by native speakers of Korean, Portuguese, and Spanish. *Language Acquisition, 26*(4), 436-456.

Tarone, E. (1979). Interlanguage as chameleon. *Language Learning, 29*, 181-191.

Jayeon Lim, Professor
Department of English Language and Literature
University of Seoul
163 Seoulsiripdaero, Dongdaemun-gu
Seoul, 02504, Korea
E-mail: limjy@uos.ac.kr

Misun Seo, Professor
Department of English Language and Literature
Hannam University
70 Hannam-ro Daedeok-gu
Daejeon, 34430, Korea
E-mail: misunseo@hnu.kr

Received: June 10, 2023
Revised: July 1, 2023
Accepted: July 5, 2023