Korean Journal of English Language and Linguistics, Vol 25, April 2025, pp. 472-492
DOI: 10.15738/kjell.25..202504.472

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

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

http://journal.kasell.or.kr

Investigating the Relationship Between Production and Perception of Word-Final Stop Voicing Contrast in Korean High School EFL Learners*

Sowon Kang · Hyunkee Ahn (Seoul National University)

Received: February 10, 2025
Revised: March 8, 2025
Accepted: March 29, 2025

Kang, Sowon (first author)
Ph. D. student, Department of English Education, College of Education
Seoul National University
Email: thdnjs1009@snu.ac.kr

Ahn, Hyunkee
(corresponding author)
Professor, Department of English Language Education & Learning Sciences Research Institute
College of Education, Seoul National University
Email: ahnhk@snu.ac.kr

* An earlier version of this work was presented at the 2023 KASELL International Conference held in Seoul. This study expands upon the findings of my master's thesis, "Korean High School EFL Learners' Production and Perception of Word-final Voicing Contrast," completed at Seoul National University in 2019. The thesis has been revised and condensed for publication in this journal.

ABSTRACT

Kang, Sowon and Hyunkee Ahn. 2025. Investigating the relationship between production and perception of word-final stop voicing in Korean high school EFL learners. *Korean Journal of English Language and Linguistics* 25, 472-492.

The current study analyzed Korean high school EFL learners' production and perception of English word-final stop voicing contrasts, exploring the relationship between these skills. Results revealed that lower-proficiency learners had significantly lower production intelligibility and perceptual identification ability compared to native speakers. Higher-proficiency learners showed native-like perception but lower production intelligibility. No significant correlation was found between production intelligibility and perceptual identification ability among Korean EFL learners, suggesting these are distinct skills. Analysis of cue sensitivity showed that higher-proficiency learners' production and perception cue hierarchy resembled native speakers', while lower-proficiency learners differed critically, particularly in their insensitivity to vowel cues in both production and perception. These results suggest that Korean EFL learners' difficulties with word-final stop voicing vary with proficiency. While perception may improve naturally with English learning, production seems to require dedicated training, even for higher-proficiency learners.

KEYWORDS

word final stop, cue hierarchy, voicing contrast, production, perception

1. Introduction

A person's linguistic background significantly influences their language production and perception, particularly when phonological rules differ between languages (Munro et al. 2006). Korean and English have distinct phonological systems, especially regarding syllable coda positions. Korean allows only seven consonants in the coda, including three lax oral stops /p t k/, but not tense stops /p* t* k*/ or aspirated stops /p^h t^h k^h/, due to the coda neutralization rule (Cho 2000). Unlike Korean, where neutralization limits coda consonants, English allows both voiced and voiceless stops in syllable-final positions, maintaining their contrast. Release part can play an important role in distinguishing them in English. Though closure releasing does not have any contrastive function in both languages, English word-final stops are not always but often released (Bent and Bradlow 2003, Tsukada et al. 2004), while Korean word-final stops are always unreleased (Kim 1998, Kim-Renaud 1974, Tsukada et al. 2004).

These phonological and phonetic differences may lead to difficulties for Korean EFL learners, particularly in accurately distinguishing and producing word-final voiced and voiceless stops in English. Such challenges can result in communication breakdowns, underscoring the importance of addressing these issues in language education. This is particularly critical for high-frequency contrasts like word-final /t/ vs. /d/, which play a significant role in English communication due to their high functional load—the extent to which they differentiate meaning in numerous minimal pairs (Catford 1987, King 1967).

Thus, the current study aims to provide a detailed analysis of Korean high school EFL learners' production and perception of word-final stop voicing contrasts, shedding light on their cue hierarchy and how it impacts both skills. It explores the relationship between production and perception skills in this context, examining how these two skills interact and whether they develop in parallel. The study considers the possibility that one skill may lag behind the other, emphasizing the importance of addressing imbalances to support learners' overall proficiency. Therefore, this study established the following research questions and conducted experiments to find answers to them:

1) What is the relationship between Korean high school EFL learners' ability to produce intelligible word-final stop voicing contrasts and their skill in perceptually identifying these contrasts?
2) How do the English proficiency levels and the place of articulation (POA) of the word-final stop affect Korean EFL learners' production intelligibility and perceptual identification ability for word-final stop voicing contrasts?
3) How can the answers to the first and second research questions be acoustically explained through cue-sensitivity?

2. Literature Review

2.1 Acoustic Cues for Distinguishing Stop Coda Voicing Contrasts and Their Hierarchy

Voice Onset Time (VOT) and F0, typically used to distinguish stop sound voicing (Holt et al. 2001, Ohde 1984), are commonly employed to distinguish voicing in stop consonants (Holt et al. 2001, Ohde 1984). However, these cues are less applicable for word-final stops in English, as voiced stops in this position frequently lack vocal fold vibration (Kang 2014). Consequently, researchers have identified alternative acoustic cues that are crucial for distinguishing word-final voiced and voiceless stops.

One prominent cue is preceding vowel duration, which is significantly longer before voiced stops (Chen 1970, Flege et al. 1992, Hayes-Harb et al. 2008, Hillenbrand et al. 1984, Klatt 1976, Peterson and Lehiste 1960). Additionally, stop closure duration is longer for voiceless stops (Chen 1970, Denes 1955, Flege et al. 1992, Hayes-Harb et al. 2008), while closure voicing duration tends to be longer for voiced stops (Flege et al. 1992, Hayes-Harb et al. 2008, Hillenbrand et al. 1984). The release burst duration also serves as a potential cue, though its importance varies (Hayes-Harb et al. 2008, Kang 2014). Furthermore, frequency-based cues, such as a lower F1 offset frequency for voiced stops, have been observed (Flege et al. 1992, Hillenbrand et al. 1984).

Building on these established acoustic cues, Kang (2014) investigated the hierarchy of cues used by native English speakers to produce and perceive word-final stop voicing contrasts. In the production experiment, participants were asked to read experimental stimuli, and measurements of pre-consonantal vowel duration, closure duration, and release burst duration were statistically analyzed. For the perception experiment, Kang (2014) manipulated sound files recorded by American English speakers. The original files were segmented into three parts: the preceding vowel, the closure, and the release burst. These segments were recombined to create stimuli (e.g., voiced vowel + voiced closure + voiced release). Participants listened to the manipulated stimuli and identified the words as having either word-final voiced or voiceless stops. Their responses were analyzed to determine the robustness of each cue.

As shown in Table 1, findings revealed that vowel duration was the primary cue for distinguishing word-final voicing in both production and perception. However, the release burst was prominent only in perception, while closure voicing was significant in production but less so in perception for unreleased word-final stops. These results highlight an asymmetrical relationship between production and perception in word-final stop voicing.

Table 1. Native Speakers' Cue Hierarchy (Kang 2014)

Environment	Cue Types	Production	Perception
Word-final Released	Main Cues	Vowel Duration	Vowel Duration **Release**
	Subsidiary Cues	Closure **Release**	Closure
Word-final Unreleased	Main Cues	Vowel Duration **Voicing during the Closure**	Vowel Duration
	Subsidiary Cues		**Voicing During the Closure**

Note: Cues in bold=asymmetrical cue

This hierarchy of acoustic cues plays a crucial role in distinguishing and producing word-final stop voicing contrasts. Furthermore, it may explain why L2 speakers differ from native speakers in their ability to produce and perceptually distinguish voicing contrasts.

2.2 EFL Learners' Production and Perception of Word-final Stop Voicing

Unlike native English speakers, EFL learners often struggle with producing and perceiving the voicing contrast of word-final stops due to varying coda restrictions across languages. Numerous studies have examined these differences, highlighting the challenges faced by EFL learners in comparison to native English speakers.

Flege et al. (1992) explored whether Mandarin and Spanish speakers—whose native languages do not allow voiced or voiceless stops in coda positions—could effectively distinguish English word-final /t/ and /d/. In a

listening test, both groups demonstrated lower proficiency in distinguishing word-final voiced and voiceless stops compared to native English speakers. Furthermore, L2 learners' proficiency levels did not significantly influence their ability to differentiate these sounds. In the production test, Mandarin and Spanish speakers exhibited smaller acoustic differences between /d/ and /t/ than native English speakers.

Similarly, Smith et al. (2009) examined German speakers, whose language devoices word-final consonants, thereby neutralizing voicing contrasts in speech. Their study revealed that German speakers could produce voicing contrasts in English word-final stops using vowel duration, closure duration, and voicing duration. However, these acoustic cues were less robust than those produced by native English speakers. A perception test further showed that German speakers had better intelligibility for native-accented voicing contrasts than for German-accented ones.

Crowther and Mann (1992) investigated Arabic speakers, focusing on two acoustic cues: vocalic duration and F1 offset frequency. The results showed that early learners were more sensitive to vocalic duration than late learners, with early learners achieving near-native sensitivity in perception tests. For F1 offset frequency, Arabic speakers produced smaller distinctions between voiced and voiceless stops compared to native speakers but demonstrated near-native sensitivity in perception.

Similar to other EFL learners, Korean speakers also face difficulties in producing and perceiving voicing contrasts in word-final stops. Son (2009) conducted an experiment comparing 10 Korean learners and 10 native speakers. The study analyzed the ratio of voiceless stop duration and vowel duration to the whole rhyme duration. Korean learners exhibited a higher proportion of voiceless coda in voiceless stop rhymes and a lower proportion of vowels in voiced stop rhymes than native speakers. Additionally, the correlation between vowel duration and word-final consonant voicing was weaker for Korean learners. Cho and Shin (2013) analyzed Korean learners' use of four acoustic cues—vowel duration, F1 offset frequency, closure voicing, and closure duration. Korean learners displayed significantly lower sensitivity to vowel duration, F1 offset frequency, and closure voicing compared to native speakers, while closure duration did not differ significantly. Kim (2011) further observed that the effect of voicing on vowel length was less pronounced in Korean speakers than in native speakers. Korean speakers also released word-final stops less frequently than native speakers.

Furthermore, Kang (2007) investigated the impact of overseas experience on Korean learners' production and perception of word-final stop voicing. Four groups were tested: native English speakers, Koreans with American residence experience, Koreans without such experience, and Korean elementary students. Results showed that Koreans with overseas experience performed better in both production and perception tasks. However, even these learners were not fully native-like. Temporal features produced by Koreans without overseas experience were more similar to native speakers than those of Koreans with overseas experience. In perception, both groups of Korean speakers performed similarly in word-final release environments, but those with overseas experience outperformed their counterparts in non-release environments.

These findings underscore the challenges faced by EFL learners, including Koreans, in mastering word-final stop voicing contrasts. Variations in acoustic cue sensitivity, influenced by L1 background, age of acquisition, and L2 experience, play a significant role in shaping learners' abilities to produce and perceive these contrasts. Although some previous studies have explored L2 speakers' production and perception of English word-final stop voicing separately, the relationship between these two aspects remains underexplored. While studies suggest a significant correlation between general production and perception skills (Beach et al. 2001, Lopez-Soto and Kewley-Port 2009, Tsushima and Hamada 2005), as well as sensitivity to acoustic cues (Flege 1999, Nishi and Rogers 2002), some studies have found no significant link (Lee 2006, Peperkamp and Bouchon 2011). The findings vary depending on the specific phonetic and phonological aspects examined. For instance, Cheng and

Zhang (2009) found a significant correlation for consonants but not for vowels, while Flege (1999) showed significant results for vowels.

Overall, there is no agreement on whether there is a significant correlation between speech production and perception among L2 learners. It is still uncertain if those with strong English perception skills are also proficient in production, or if the opposite is true. Similarly, it is unclear whether an EFL learner's accurate perception of acoustic cues corresponds with their ability to produce those cues. This ambiguity is why the current study seeks to explore this area further. Specifically, investigating Korean EFL learners' cue hierarchy could provide valuable insights. This research revealed intriguing production-perception asymmetries among native speakers, with vowel duration serving as the primary cue for both production and perception, while other cues showed varying importance. Comparing and analyzing how this cue hierarchy manifests in Korean EFL learners would allow for a closer examination of potential asymmetries and provide a deeper insights into Korean learners' acquisition process for these challenging phonetic features.

3. Method

3.1 Experiment 1: Production Test

3.1.1 Participants

The experiment was conducted on 8 male Native GA (General American) English speakers (average age: 39 years; age range: 27–55 years) and 16 Korean male first- and second-year high school students who speak the Seoul standard dialect and attend school in the Bundang or Siheung region. The L1 diversity of the high school participants was restricted, as it could be a factor influencing the experimental results. Additionally, none of the Korean participants had any experience living abroad in English-speaking countries. The gender of the participants was controlled to prevent any potential gender effects on formant frequency.

The Korean participants were split into two equal groups: higher-proficiency and lower-proficiency English learners. To categorize the Korean students, two assessment methods were used: (i) a pronunciation test where students read a passage aloud, which native English speakers then rated for comprehensibility and accent, (ii) a listening comprehension test based on TOEIC questions. Combining scores from both assessments, the top-performing half of the Korean students were identified as higher-proficiency learners, with the rest classified as lower-proficiency learners.

To evaluate the intelligibility of the participants' productions, ten native GA speakers (six males, four females) served as listeners and assessed the recorded speech. These listeners were independent from the native GA speakers who participated in the production test, ensuring that the evaluation was conducted by a separate group of raters.

3.1.2 Materials

The study utilized 12 monosyllabic words from Table 2 as test stimuli. These words formed six minimal pairs, with each pair differing only in the voicing of the final stop consonant. The chosen words were carefully balanced to represent different POAs for the final stop sound. All words in the set used the vowel /æ/ before the final consonant. Due to limitations in finding suitable minimal pairs within the specified phonological context, the nonce word *pag* was included as part of the stimuli.

Table 2. Production Test Stimuli

Bilabial	Alveolar	Velar
tap/tab	bat/bad	back/bag
cap/cab	pat/pad	pack/pag

3.1.3 Procedure

The production experiment took place in a quiet room. The test participants read five randomized lists. Each list contained 12 monosyllabic words in Table 1 and some fillers. In the list, each word was embedded in the carrier sentence, "Please say ____." The experimenter assisted the participants to read the sentences in right speed and flow and recorded their speech.

The carrier sentence was intentionally kept simple to minimize any potential influence on participants' pronunciation. Since some Korean high school learners have very limited English proficiency, a long or complex carrier sentence could itself affect their production. To avoid this, a short and straightforward structure was used, ensuring that the focus remained on the target words rather than the difficulty of the sentence.

10 native GA English speakers evaluated the intelligibility of the word-final stop voicing contrasts produced by the participants. Using a computer and a headphone, they listened to audio recordings from the 24 participants and responded to intelligibility assessment questions. For each question, listeners chose between two options: a word ending with a voiceless consonant or one ending with a voiced consonant. They were instructed to select the option that most closely matched the sound in the provided audio file. The assessment excluded the first and last randomized recordings, resulting in three evaluated instances per word.

3.2 Experiment 2: Perception Test

3.2.1 Original stimuli

The original stimuli were produced by a female GA speaker temporarily staying in Korea who has lived in Massachusetts, U.S.A for most of her life. She read four randomized lists. Each list consisted of the words presented in Table 2 along with some fillers embedded in a carrier sentence adapted from Hayes-Harb et al. (2008): "I like you to say ____ some of the time, but now I'm going to say ____." Two words that form minimal pairs were placed in the same sentence, and two versions for each minimal pair were recorded: "I like you to say element1 some of the time, but now I'm going to say element2," and "I like you to say element2 some of the time, but now I'm going to say element1." The speaker was asked to release the final stop for the first two trials and not to release for the last two trials.

Unlike the production experiment, which used a simple carrier sentence to minimize cognitive load for Korean high school learners, the perception test employed a longer sentence to provide a more natural linguistic environment for the target words. This choice ensured that the target words were naturally embedded in a phrasal context, allowing for more native-like prosody and facilitating better acoustic control over final stop realizations. Additionally, the minimal pair contrast within the same sentence helped reinforce participants' awareness of subtle phonetic differences.

As with the production experiment, the first and the last trials for each sentence were removed. Therefore, 24 sentence tokens (12 sentences * 2 repetitions) and 48 word tokens (12 words * 2 sentences * 2 repetitions) excluding the fillers were used as the experiment stimuli.

Table 3. Cue Manipulation (adapted from Kang 2014)

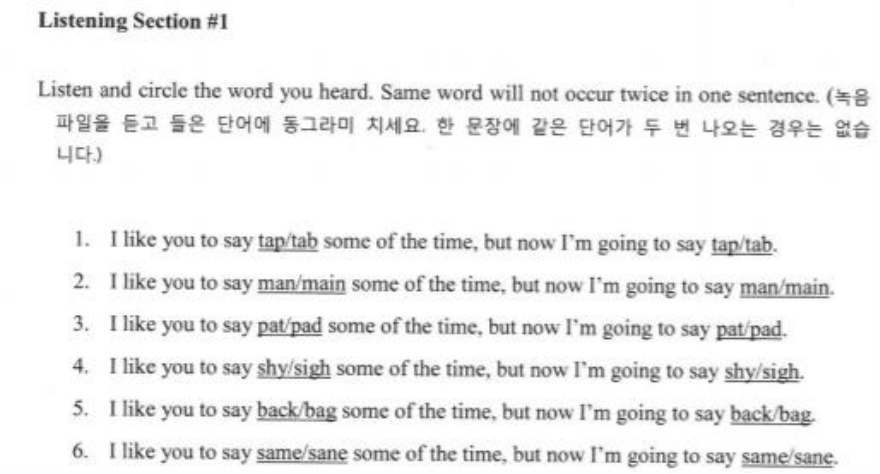
Release Burst	Cue Manipulation	Numbers
Word-final released	(1) voiced V+voiced closure+voiced release	8 tokens * 8 versions * 6 pairs = 384
	(2) voiced V+voiceless closure+voiced release	
	(3) voiced V+voiced closure+voiceless release	
	(4) voiced V+voiceless closure+voiceless release	
	(5) voiceless V+voiced closure+voiced release	
	(6) voiceless V+voiced closure+voiceless release	
	(7) voiceless V+voiceless closure+voiced release	
	(8) voiceless V+voiceless closure+voiceless release	
Word-final unreleased	(1) voiced V+voiced closure	4 tokens * 4 versions * 6 pairs = 96
	(2) voiced V+voiceless closure	
	(3) voiceless V+voiced closure	
	(4) voiceless V+voiceless closure	

3.2.2 Manipulated stimuli

The released stimuli were segmented into the vowel part, the closure part, and the release part, and the unreleased segments were spliced into the vowel part and the closure part using *Praat* 6.0.43. Following Kang (2014), the spliced segments were combined and manipulated as in Table 3. There were 480 stimuli (384 word-final released and 96 word-final unreleased; 144 same-spliced and 336 cross-spliced) in total. The same-spliced ones were used to check the reliability of the manipulated stimuli.

3.2.3 Participants and procedure

The perception test involved the same 16 Korean high school EFL learners and 8 native English speakers from the production experiment. The test comprised two sections. The first, an identification test, presented 24 trials of original stimuli sentences (12 word-final released and 12 unreleased) plus 20 fillers. This section lasted 3 to 4 minutes, with randomly ordered questions requiring participants to circle answers that best matches the words they heard for each blank, noting that the two blanks in a sentence needed different answers, as shown in Figure 1. The second section, a cue match test, consisted of 480 trials of manipulated stimuli presented as isolated words. After each audio file, participants selected the word from given minimal pairs that most closely resembled the stimulus. Both sections used a 2-second interstimulus interval, and participants were not allowed to listen to audio files more than once.



Listening Section #1

Listen and circle the word you heard. Same word will not occur twice in one sentence. (녹음 파일을 듣고 들은 단어에 동그라미 치세요. 한 문장에 같은 단어가 두 번 나오는 경우는 없습니다.)

1. I like you to say tap/tab some of the time, but now I'm going to say tap/tab.
2. I like you to say man/main some of the time, but now I'm going to say man/main.
3. I like you to say pat/pad some of the time, but now I'm going to say pat/pad.
4. I like you to say shy/sigh some of the time, but now I'm going to say shy/sigh.
5. I like you to say back/bag some of the time, but now I'm going to say back/bag.
6. I like you to say same/sane some of the time, but now I'm going to say same/sane.

Figure 1. Identification Test

3.3 Analysis

3.3.1 Production intelligibility and perceptual identification ability analysis

The intelligibility of each participant's production was analyzed based on the responses of 10 listeners. An intelligibility score for each speaker was calculated by averaging the ratio of correct answers from the 10 listeners for each production. These scores were subjected to a repeated measure analysis of variance (ANOVA) with group and POA as independent variables. For the perception analysis, each participant's ratio of correct answers in the identification test was calculated. These scores were then analyzed using a repeated-measure three-way ANOVA, with the identification test scores as the dependent variable and group, POAs, and releasing as independent variables. To examine the correlation between production and perception abilities of the word-final stop voicing contrast, intelligibility assessment scores and identification test scores were subjected to Pearson R correlation analysis.

3.3.2 Cue sensitivity analysis

The acoustic analysis of word-final released segments involved measuring vowel duration, F1 offset frequency, closure duration, closure voicing duration, and release duration using *Praat* 6.0.43. For unreleased segments, only vowel duration, F1 offset frequency, and closure voicing duration were measured, following standards from previous studies. These acoustic cues were measured based on the measurement standard Cho and Shin (2013), Fledge et al. (1992), and Hayes-Harb et al. (2008) used (Table 4). Additionally, the proportion of trials in which the word-final stop was released was calculated for each group.

Table 4. Acoustic Cue Measurements (adapted from Cho and Shin 2013, Fledge et al. 1992, Hayes-Harb et al. 2009)

Parts	Acoustic Cues	Measurement
preceding vowel	vowel duration	measure from the first positive peak in the periodic portion of each waveform to constriction of the word-final stop
	F1 offset frequency	measure at the end of a vowel, at 20 ms, and at 40 ms before the end of the vowel, and average the measurements over the three points
closure	closure voicing duration	measure from the point of constriction to the last positive peak in any low-amplitude, sinusoidal voicing in the stop closure interval
	closure duration	measure from constriction to the beginning of the release burst
release	duration of release	measure from the release burst to the end of the word

A repeated-measures three-way ANOVA (group × voicing × POA) was conducted on the released-trial proportion as the dependent variable. Similarly, the same analysis was applied to the values of each acoustic cue to examine their variations across the factors. Furthermore, to test the simple main effects, repeated measures one-way ANOVA in which the independent variable is voicing was conducted separately for the native English speakers (NE), for the higher-proficiency Korean EFL learners (HK), and for the for the lower-proficiency Korean EFL learners (LK). However, considering that the size of voicing effects on vowel duration and F1 frequency can each be affected by individual speakers' speech rate and original tongue position for /æ/, ratios were additionally used to minimize such confounding effects as suggested by Flege et al. (1992). For these two acoustic cues, the ratio between the values of the cue for voiced and voiceless stimuli was calculated. The ratio was submitted to repeated measures two-way ANOVA (group × POA).

In the perception analysis, as in Kang (2014), the match ratio of each acoustic cue was measured by counting the score for target cue when it led to the whole voicing response. For example, if a participant selected voiceless option for the "voiced V + voiceless closure + voiced release" stimuli, closure scored 1 point. Then, cue robustness was calculated by dividing the match ratio by the number of whole trial. The results were statistically analyzed, again separately for the word-final released and unreleased tokens. The repeated measures ANOVA (part × voicing × POA) was conducted for each group to check how the gaps among the cue-robustness of vowel, closure, and release part differ across the three groups, and how POA affects them.

Based on the results of these analyses, the production and perception cue hierarchies were systematically compared and contrasted. For both production and perception, the vowel cue, closure cue, and release cue were categorized as either main cues or subsidiary cues. This categorization was determined based on statistical significance and the mean values of cue robustness.

4. Results and Discussion

4.1 Production Intelligibility and Perceptual Identification Ability Analysis

Table 5 displays the average intelligibility and identification test scores for each speaker group. A two-way repeated-measures ANOVA (POA × group) for intelligibility scores revealed a significant group effect [$F(2,69)=55.788$, $p=.000$] and a significant group × POA interaction effect [$F(4,63)=9.520$, $p=.000$]. The significant group effect indicates that speaker groups differed in their ability to produce intelligible voicing

contrasts for word-final stops.

Table 5. Mean Intelligibility and Identification Test Score (%)

Group	Intelligibility	Identification Test
NE	99.27 (0.83)	98.96 (1.93)
HK	74.79 (12.67)	96.88 (3.69)
LK	56.77 (5.88)	86.98 (11.67)

Note: numbers in parenthesis=standard deviation

Post-hoc Bonferroni tests showed that native English speakers produced significantly more intelligible stop voicing contrasts compared to both Korean EFL learner groups [NE-HK=24.4793, p=.000; NE-LK=42.5, p=.000]. Additionally, higher-proficiency Korean EFL learners' production was significantly more intelligible than that of lower-proficiency learners [HK-LK=18.0208, p=.001]. This suggests that while Korean learners may not produce perfectly intelligible word-final contrasts like native English speakers, their intelligibility may improve as their English proficiency increases.

The significant group × POA interaction effect suggests that POA impacts intelligibility differently across speaker groups. Post-hoc tests revealed that Korean EFL learners' productions were notably less intelligible for bilabial sounds [HK: alveolar-bilabial=16.563, p=.000; velar-bilabial=24.063, p=.002; LK: alveolar-bilabial=20.938, p=.036]. In contrast, POA did not significantly affect the intelligibility of native English speakers' production [F(2,21)=.089, p=.916].

Moreover, a three-way repeated-measures ANOVA (group × POA × release) for perceptual identification test scores reported a non-significant effect of group × POA [F(4,135)=1.367, p=.262] and group × POA × release interaction [F(4,126)=1.491, p=.222], whereas the effects of group [F(2,141)=6.406, p=.007] and group × release interaction [F(2,128)=5.471, p=.012] were both significant. The significant group effect indicates that participant groups differed in their ability to correctly perceive word-final stop voicing contrasts. Post-hoc Bonferroni tests showed no significant difference between higher-proficiency Korean EFL learners and native English speakers [NE-HK=2.08, p=1.000]. However, lower-proficiency learners scored significantly lower than both native English speakers [NE-LK=11.98, p=.009] and higher-proficiency Korean learners [HK-LK=9.90, p=.035]. The significant group × release interaction implies that English proficiency impacts the ability to identify voicing in unreleased word-final stops more than in released ones. Lower-proficiency learners perceived significantly fewer unreleased stimuli correctly compared to native English speakers and higher-proficiency learners [NE-LK=17.707, p=.001; HK-LK=13.542, p=.007]. For released stimuli, the difference across the three groups was not significant [F(2,69)=2.333, p=.122].

The analysis reveals an asymmetrical relationship between EFL learners' production intelligibility and perceptual identification ability for word-final stop voicing contrasts. Higher-proficiency EFL learners demonstrated near-native perceptual identification abilities, yet their production intelligibility for these contrasts was significantly lower than that of the native English speakers. Furthermore, a Pearson correlation analysis showed no significant relationship between Korean EFL learners' production intelligibility scores and their identification test scores [r=.332, p=.210]. This discrepancy between the ability to produce intelligible word-final voicing contrasts and accurately perceive them suggests a gap between EFL learners' production and perception skills. To further investigate this discrepancy, the next section will present an analysis of the acoustic cues in participants' productions and their perceptual cue-robustness.

4.2 Cue Sensitivity Analysis

4.2.1 Production: acoustic cue analysis

Before delving into the acoustic analysis for each cue, it would be worthwhile to examine the ratio of released tokens for each group. There was a notable difference in the number of word-final released tokens across groups. As shown in Table 6, native English speakers released word-final stops significantly more often than Korean EFL learners. A repeated-measures three-way ANOVA (group × voicing × POA) demonstrated significant effects of group [F(2,141)=16.668, p=.000], voicing [F(1,142)=6.047, p=.023], and POA [F(2,141)=17.703, p=.000]. However, interaction effects were not significant, except for group × POA [F(4,135)=4.198, p=.006]. Additionally, pairwise comparisons indicated that voiced stops were released more frequently than voiceless stops overall [VD-VL=9.954, p=0.023], but the effect of voicing did not differ significantly across groups (group × voicing [F(2,138) =1.972, p=0.164]).

Table 6. Mean and Standard Deviation of Word-final Released Tokens Ratio

Group	Mean (%)	SD
NE	87.20	26.46
HK	48.30	36.10
LK	30.20	29.71

The post-hoc Bonferroni test reported that native English speakers released significantly more tokens than the Korean EFL learners [NE-HK=38.885, p=.003; NE-LK=59.940, p=.000], but the higher-proficiency learners did not produce significantly more word-final released tokens than the lower-proficiency learners [HK-LK=18.054, p=.263]. The Korean L1 speakers' tendency to unrelease the English word-final stop could be explained by the phonological behavior of Korean: Korean word-final stops are always unreleased (Kim 1998, Kim-Renaud 1974, Tsukada et al. 2004). It can be predicted that phonological characteristics of the participants' L1 affected their production of L2. Furthermore, as indicated in Figure 2, this tendency was particularly pronounced in the bilabial position, where the Korean EFL learners produced a noticeably smaller number of released tokens compared to other positions. The relatively low number of word-final released tokens in the bilabial position might be one of the cause for the low intelligibility of the Korean EFL learners' productions in bilabial position, taking into account that English word final stops can be less intelligible if they are produced without any release burst (Malécot 1958, Wang 1959).

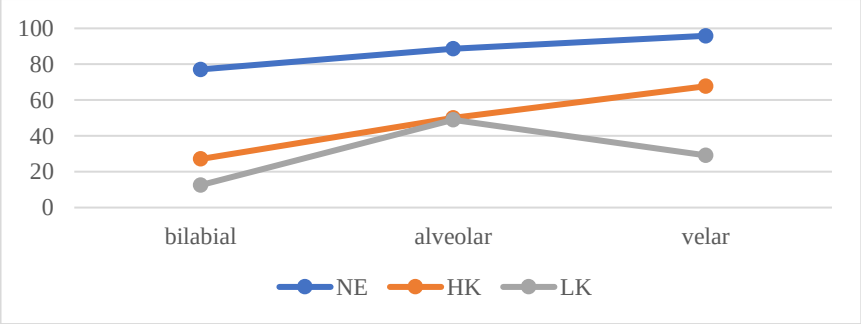
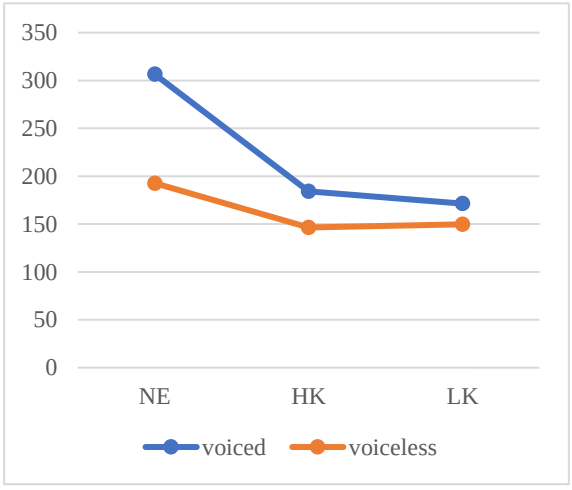
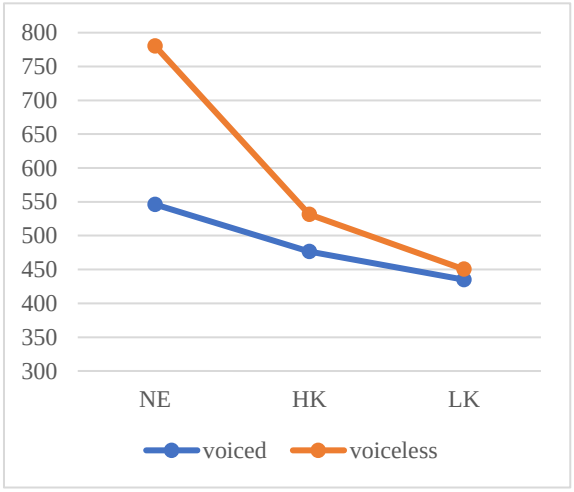


Figure 2. Mean Ratio of Word-final Released Tokens for Each POA

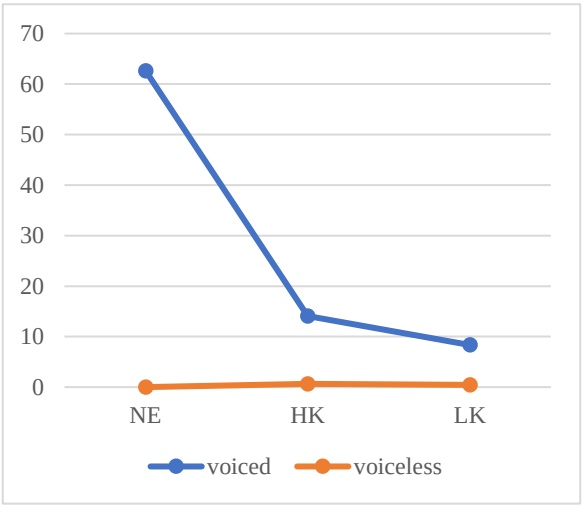
To further explore the acoustic properties related to the production differences, Figure 3 presents the average acoustic cue values for each group, contrasting voiced and voiceless counterparts. These cues include preceding vowel duration, preceding vowel F1 offset frequency, closure voicing duration, and closure duration.



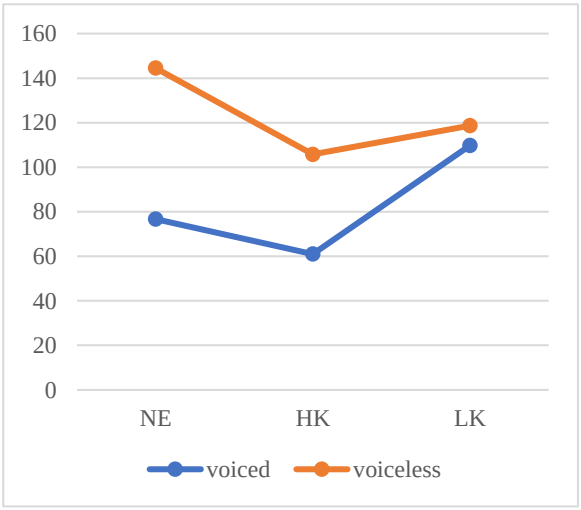
(a) preceding vowel duration



(b) preceding vowel F1 offset frequency



(c) closure voicing duration



(d) closure duration

Figure 3. Mean Acoustic Cue Values of Each Group

For preceding vowel duration, a repeated measures three-way ANOVA showed significant group × voicing interaction [$F(2,138)=39.708$, $p=.000$], but non-significant group × voicing × POA interaction [$F(4,126)=1.518$, $p=.214$]. Simple main effect tests revealed that all three groups produced significant differences in preceding vowel length between word-final voiced and voiceless stops [NE: $F(1,14)=130.520$, $p=.000$; HK: $F(1,14)=23.397$,

p=.002; LK: *F*(1,14)=19.293, *p*=.003]. To account for speech rate variations, a repeated measures two-way ANOVA was conducted on the ratio of vowel durations preceding voiced to voiceless stops. This analysis showed a significant group effect [*F*(1,70)=18.045, *p*=.000] but insignificant group × POA interaction [*F*(4,63)=.823, *p*=.518], aligning with the three-way ANOVA results. Post-hoc Bonferroni tests revealed that native speakers produced significantly larger durational differences between vowels preceding voiced and voiceless stops compared to the Korean participants [NE-HK=.368, *p*=.001; NE-LK=.472, *p*=.000]. The two groups of Korean EFL learners, however, produced similar durational differences [HK-LK=.104, *p*=.663]. Therefore, it could be concluded that Korean EFL learners produced less robust difference for word-final voicing contrast in terms of preceding vowel duration and that their English proficiency had little effects on the robustness of this difference.

Regarding F1 offset frequency, a repeated measures three-way ANOVA revealed significant effects for group × voicing interaction [*F*(2,138)=44.610, *p*=.000] and group × voicing × POA interaction [*F*(4,126)=7.673, *p*=.000]. Simple main effects tests showed that voicing significantly affected F1 offset frequency of preceding vowels for native English speakers and higher-proficiency Korean EFL learners [NE: *F*(1,13)=122.058, *p*=.000, HK: *F*(1,13)=7.251, *p*=.031], but not for lower-proficiency learners [*F*(1,13)=4.507, *p*=.071]. To control the confounding effects, a repeated measure two-way ANOVA was conducted using the ratio of vowel F1 offset frequency preceding voiceless to voiced stops. This analysis showed significant effects for group [*F*(2,69)=29.108, *p*=.000] and group × POA interaction [*F*(4,63)=7.792, *p*=.000]. Post-hoc Bonferroni tests revealed that native English speakers produced significantly more robust differences in vowel F1 offset frequencies between word-final voiced and voiceless stops compared to Korean EFL learners [NE-HK=.352, *p*=.000; NE-LK=.430, *p*=.000]. However, higher-proficiency Korean learners did not differ significantly from lower-proficiency learners [NE-LK=.078, *p*=.620]. The significant group × voicing × POA interaction indicates that POA is also important. Post-hoc test showed that native English speakers exhibited larger differences between voiced and voiceless tokens for word-final velar stops compared to other POAs [velar-bilabial=.516, *p*=.014; velar-alveolar=.401, *p*=.032]. Higher-proficiency Korean learners also showed the largest gap for velar stops [velar-bilabial=.200, *p*=.038; velar-alveolar=.205, *p*=.027], though smaller than native speakers. However, lower-proficiency Korean learners showed similar F1 differences across all three POAs [*F*(2,21)=.083, *p*=.921].

Next, for closure voicing duration, an analysis of closure voicing duration for word-final voiceless stops revealed minimal voicing across all three groups, as depicted in Figure 3. A repeated measures two-way ANOVA, examining the effects of group and POA on closure voicing duration of word-final voiced stops, yielded significant results for both the group factor [*F*(2,69)=29.260, *p*<.001] and the group × POA interaction [*F*(4,63)=4.334, *p*=.005]. Further investigation through Bonferroni post-hoc tests demonstrated that native English speakers produced substantially longer closure voicing durations for word-final voiced stops compared to both groups of Korean speakers learning English [NE-HK=48.55, *p*=.000; NE-LK=54.30, *p*=.000]. No significant difference was found between the higher-proficiency and lower-proficiency learners [HK-LK=5.75, *p*=1.000]. Considering that the Korean phonological system lacks a voiced stop, the Korean EFL learners might had some difficulties producing enough amount of voicing during the closure because they are not used to producing voicing during the closure. The significant group × POA interaction indicates that the degree of differences of closure voicing duration for the three groups varies across the POAs. The native English speakers had the longest closure voicing duration for the bilabials, while the Korean learners had the shortest closure voicing duration for them.

Furthermore, the repeated measures three-way ANOVA on closure duration revealed a significant interaction between group and voicing [*F*(2,138)=4.878, *p*=.030] but insignificant interaction between group, voicing, and POA [*F*(4,126)=2.002, *p*=.129]. Further analysis of simple main effects showed that both native English speakers and higher-proficiency Korean EFL learners produced significant differences in closure duration between word-

final voiced and voiceless stops [NE: $F(1,13)=81.401$, $p<.001$; HK: $F(1,13)=26.866$, $p<.001$]. However, the lower-proficiency learners did not show this distinction [$F(1,13)=4.354$, $p=.056$]. The previous studies by Cho and Shin (2013) and Flege et al. (1992) presented an insignificant group × voicing interaction for Korean EFL learners' closure duration, but these studies did not differentiate the learners' proficiency levels. The results of this experiment are basically similar to those of the previous studies, but they suggest that the pattern may change as proficiency increases.

Figure 4 illustrates the release burst duration of word-final voiced and voiceless stops for each group. Notably, higher-proficiency Korean EFL learners demonstrated a considerable durational difference between the releases of word-final voiced and voiceless stops, while native English speakers showed only a small gap. A repeated measures one-way ANOVA, with voicing as the independent variable, revealed a significant effect for higher-proficiency Korean learners [$F(1,13)=9.202$, $p=.007$], but not for native English speakers [$F(1,13)=1.729$, $p=.202$] or lower-proficiency learners [$F(1,13)=.110$, $p=.746$]. This finding aligns with Kang's (2014) study, which also found no statistical difference in release burst duration between word-final voiced and voiceless stops in native English speakers' productions. However, the significant difference exhibited by higher-proficiency Korean learners remains puzzling. It is possible that these learners, in their attempt to distinguish between voiced and voiceless stops, may have exaggerated the voiceless release. This could be because they find it challenging to differentiate these sounds through vowel and closure characteristics, which are less evident to non-native speakers.

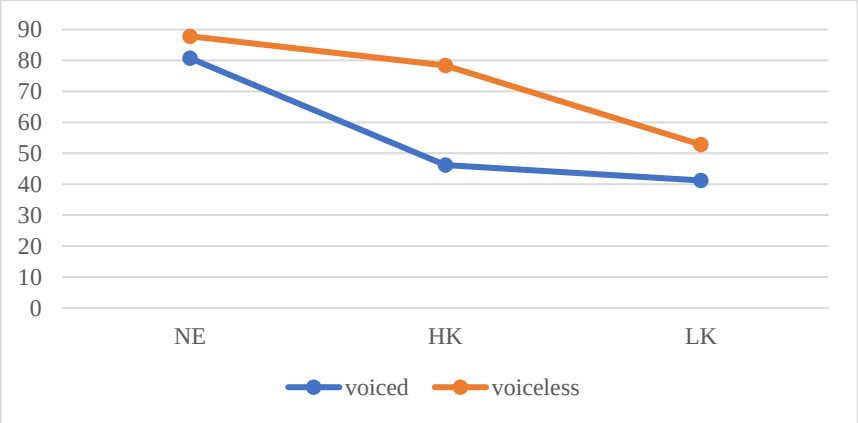


Figure 4. Mean Release Burst Duration

4.2.2 Perception: cue-robustness analysis

Table 7 displays the results for the three experimental groups across four same-spliced and eight cross-spliced stimulus types. For the same-spliced stimuli, all three groups demonstrated accuracy rates similar to those observed in the identification test, particularly for the released tokens. There was a slight decrease in accuracy for all groups when responding to the unreleased stimuli, but it could still be assumed that the participants' responses to the cue-match test were fairly reliable.

The results for the cross-spliced stimuli in Table 7 illustrate how each acoustic cue—vowel, closure, and release—contributes to the overall voicing perception. Each stimulus consists of a specific combination of these cues, allowing for an analysis of which element most strongly influenced participants' judgments. For example, the 101 stimulus (voiced vowel + voiceless closure + voiced release) isolates the closure as the primary voiceless cue. If participants categorized this stimulus as voiceless, it indicates that they relied on the closure cue despite the

presence of a voiced vowel and release. Similarly, the 011 stimulus (voiceless vowel + voiced closure + voiced release) assesses the influence of vowel voicing. A high match ratio for voiceless responses in this condition would suggest that participants primarily relied on the vowel cue.

Table 7. Result of Cue-match Test (%)

Stimulus type		NE		HK		LK	
		Response Type		Response Type		Response Type	
		VD	VL	VD	VL	VD	VL
Same-spliced	111	**99.22**	0.78	**97.66**	2.34	**94.53**	5.47
	000	1.82	**98.18**	8.07	**91.93**	8.33	**91.67**
	11	**95.83**	4.17	**75.52**	24.48	**63.02**	36.98
	00	23.44	**76.56**	31.25	**68.75**	23.44	**76.56**
Cross-spliced	101	85.25	**14.58**	91.41	**8.59**	82.55	**17.45**
	110	84.37	**15.63**	80.47	**19.53**	53.12	**46.88**
	100	**55.99**	44.01	**57.55**	42.45	**33.59**	66.41
	011	90.10	**9.90**	84.9	**15.10**	78.65	**21.35**
	010	**26.30**	73.70	**22.66**	77.34	**19.79**	80.21
	001	**47.40**	52.60	**66.93**	33.07	**51.56**	48.44
	10	**80.21**	**19.79**	**70.83**	29.17	**42.71**	**57.29**
	01	**42.71**	**57.29**	37.50	**62.50**	**45.31**	**54.69**

Note: 1=voiced; 0=voiceless; Numbers in bold=match ratio

Table 8. Cue-robustness for Released Stimuli (%)

Part	Voicing	NE	HK	LK
vowel	VD	55.99	57.55	33.59
	VL	9.90	15.10	21.35
	MEAN	**32.94**	**36.33**	**27.47**
closure	VD	26.30	22.66	19.79
	VL	14.58	8.59	17.45
	MEAN	**20.44**	**15.63**	**18.62**
release	VD	47.40	66.93	51.56
	VL	15.63	19.53	46.88
	MEAN	**31.51**	**43.23**	**48.88**

Note: Numbers in bold=cue-robustness

Table 8 presents the cue-robustness for released stimuli, calculated using the cue-match ratio of the cross-spliced stimuli. Cue-robustness refers to the extent to which a particular cue influences listeners' voicing judgments, with higher values indicating a stronger reliance on that cue. Cue-match ratio is stimulus-specific and reflects how often a cue aligns with the perceived voicing, while cue-robustness is a cumulative measure that captures the overall influence of a cue across different conditions. The cue-robustness table aggregates these cue-match ratios to highlight which cues were most influential in shaping voicing judgments. The table presents the match ratios for each cue type in both voiced and voiceless conditions, along with the mean robustness for each cue across the

different groups. A repeated measures three-way ANOVA (part × voicing × POA) on native English speakers' cue-robustness for released stimuli revealed a significant effect of part [$F(2,141)=13.421$, $p=.001$]. Bonferroni tests showed significant differences between vowel and closure parts [vowel-closure=12.500, $p=.001$] and between release and closure parts [release-closure=11.068, $p=.003$], but not between vowel and release parts [vowel-release=1.1432, $p=.645$]. This indicates that native English speakers perceive word-final stops more sensitively based on vowel and release parts than the closure part. Significant interactions were also found for part × POA [$F(4,135)=17.232$, $p<.001$], part × voicing [$F(2,138)=25.287$, $p<.001$], and part × voicing × POA [$F(4,126)=23.274$, $p<.001$]. The match ratio differences between parts varied across POA and voicing. For alveolars, the release part's match ratio was significantly larger than the other two parts [ANOVA: $F(2,45)=3.749$, $p<.001$; Bonferroni: release-vowel=23.05, $p=.007$]. Conversely, for velars, the vowel part had a higher match ratio [ANOVA: $F(2,45)=10.165$, $p=.002$; Bonferroni: vowel-release=25.3937, $p=.003$]. Bilabials showed no significant differences between parts [ANOVA: $F(2,45)=.218$, $p=.807$]. Notably, for velar stimuli, vowel match ratios differed significantly between voiced and voiceless stimuli [$F(2,13)=1701.000$, $p<.001$], while closure match ratios did not [$F(2,13)=.024$, $p=.882$]. This pattern was not observed in other POAs.

A repeated measures three-way ANOVA (part × voicing × POA) on higher-proficiency learners' cue-robustness for released stimuli showed a significant effect of part [$F(2,141)=47.522$, $p<.001$]. Bonferroni tests revealed significant differences between vowel and closure match ratios [vowel-closure=20.703, $p<.001$] and between release and closure [release-closure=27.604, $p<.001$], but not between vowel and release [vowel-release=6.703, $p=.174$]. This suggests that higher-proficiency learners' perceptual cue-hierarchy for word-final released stimuli is similar to that of native English speakers, with vowel and release parts being more salient than closure. Furthermore, significant interaction effects were found for part × voicing [$F(2,138)=28.733$, $p<.001$], part × POA [$F(2,138)=4.197$, $p=.009$], and part × voicing × POA [$F(4,126)=7.748$, $p<.001$]. In alveolar and velar positions, the difference between match ratios of vowel and release parts were statistically insignificant [alveolar: vowel-release=14.844, $p=.337$; velar: vowel-release=11.719, $p=.069$]. However, in the bilabial position, the release part was more perceptually salient than the vowel part [release-vowel=17.578, $p=.022$]. This pattern is the opposite of native English speakers, where no significant differences were observed for bilabials, and vowels were more salient for velars. Moreover, the closure match ratio for voiced was larger than that for the voiceless in bilabial and alveolar position [bilabial: voiced-voiceless=24.219, $p=.003$; alveolar: voiced-voiceless=25.781, $p=.009$], while it was the opposite in the velar position [voiceless-voiced=7.813, $p=.028$].

A repeated measures three-way ANOVA (part × voicing × POA) on lower-proficiency learners' cue-robustness for released stimuli showed a significant effect of part [$F(2,141)=16.475$, $p=.005$]. Bonferroni tests revealed that the release part had significantly higher cue robustness than both vowel and closure parts [release-vowel=21.745, $p=.014$; release-closure=30.599, $p=.001$], while vowel and closure parts were statistically similar [vowel-closure=8.854, $p=.135$]. This suggests that lower-proficiency Korean EFL learners have a different perceptual cue hierarchy compared to native English speakers and higher-proficiency EFL learners. While the latter two groups showed similar cue robustness for vowel and release parts, lower-proficiency learners demonstrated significantly lower cue-robustness for vowel parts than release parts. The ANOVA also revealed a significant part × POA interaction [$F(4,135)=46.430$, $p<.001$], but insignificant interactions for part × voicing [$F(4,135)=.269$, $p=.620$] and part × voicing × POA [$F(4,126)=3.846$, $p=.091$]. In bilabial and alveolar positions, the match ratio of release part was significantly higher than that of the vowel [bilabial: release-vowel=37.109, $p=.008$; alveolar: release-vowel=37.5000, $p=.012$]. However, in the velar position, there was no significant difference between vowel and release parts [vowel-release=9.375, $p=.167$]. As with the higher-proficiency learners, lower-proficiency learners exhibited a pattern opposite to that of native English speakers who showed no significant differences for bilabials

and a preference for the vowel part in velars.

Table 9 presents the cue-robustness for unreleased stimuli. For the native English speaker group, the repeated measure three-way ANOVA revealed the significant effect of part [$F(1,94)=33.851$, $p=.001$], which indicates that vowel had a considerably higher cue-robustness than the closure in unreleased condition. It also revealed significant interaction effects of part and POA [$F(4,87)=22.108$, $p=.000$]. The difference between the match ratios of vowel and closure was insignificant in bilabial [$F(1,30)=.020$, $p=.890$], while they were significant in alveolar and velar [alveolar: $F(1,30)=65.258$, $p=.000$; velar: $F(1,30)=67.014$, $p=.000$].

A repeated measures three-way ANOVA conducted on higher-proficiency Korean EFL learners' cue-robustness for unreleased stimuli revealed a significant effect of part [$F(1,94)=89.600$, $p<.001$], indicating that these learners demonstrated substantially higher cue-robustness for vowel compared to closure in unreleased word-final stops. This finding aligns with the perceptual cue hierarchy observed in native English speakers for the same type of stimuli. The ANOVA also revealed a significant interaction effect between part and POA [$F(4,87)=15.608$, $p<.001$] for higher-proficiency Korean EFL learners. This interaction indicates that the relationship between vowel and closure cue-robustness varied depending on the POA. Specifically, for bilabial sounds, there was no significant difference between the match ratios of vowel and closure parts [$F(1,30)=2.749$, $p=.142$]. However, significant differences were found for other POAs [alveolar: $F(1,30)=23.253$, $p=.002$; velar: $F(1,30)=54.50$, $p<.001$]. This pattern of results is similar to the findings observed in the native English speakers' response.

Table 9. Cue-robustness for Unreleased Stimuli (%)

Part	Voicing	NE	HK	LK
vowel	VD	80.21	70.83	42.71
	VL	57.29	62.50	54.69
	MEAN	**68.75**	**66.67**	**48.70**
closure	VD	42.71	29.17	57.29
	VL	19.79	37.50	45.31
	MEAN	**31.25**	**33.33**	**51.30**

Note: Numbers in bold = cue-robustness

For lower-proficiency Korean EFL learners, the analysis of cue-robustness in word-final unreleased stops yielded different results compared to native English speakers and higher-proficiency learners. A repeated measures three-way ANOVA revealed insignificant effect of part [$F(1,94)=.119$, $p=.740$], indicating that lower-proficiency learners did not show a significant difference between the cue-robustness of vowel and closure parts. This contrasts with the native speakers and higher-proficiency learners, who demonstrated significantly higher cue robustness for vowel parts than closure parts. The ANOVA also showed a significant part × POA interaction [$F(4,87)=10.134$, $p=.002$], suggesting that the relative importance of vowel and closure cues varied across POAs. For bilabial sounds, the cue-robustness of closure was significantly higher than that of vowel [$F(1,30)=8.191$, $p=.024$]. However, no significant differences were found between vowel and closure parts for alveolar [$F(1,30)=.845$, $p=.388$] or velar positions [$F(1,30)=5.395$, $p=.053$]. This indicates that the results are completely opposite for the native English speaker group and the higher-proficiency Korean EFL learner group.

4.2.3 Comparison between production and perception cue hierarchy

Table 10 presents the production and perception cue-hierarchy for each group investigated. Native English speakers demonstrated an asymmetrical relationship between production and perception cues. While they did not produce significant differences in release burst duration between voiced and voiceless stimuli, they were highly sensitive to the release part in perception. Conversely, they showed the opposite pattern for the closure part.

Higher-proficiency EFL learners exhibited a cue-hierarchy similar to native speakers, with one notable difference: they produced a significant gap in release duration between voiced and voiceless stimuli, unlike native speakers. For this group, the release part served as a main cue in both production and perception.

Lower-proficiency learners showed more significant discrepancies from native speakers in their cue hierarchy. While they resembled native speakers in using the release part as a main perceptual cue but not a production cue, they differed in their treatment of vowel cues. Lower-proficiency learners produced a significant vowel duration difference between voiced and voiceless stimuli but did not show a significant difference in F1 offset frequency. Still, the ability of lower-proficiency learners to produce meaningful vowel duration differences remains questionable, as the average difference was only about 20ms, despite being statistically significant. Furthermore, they were not particularly sensitive to vowel cues in terms of perception.

Table 10. Production and Perception Cue-Hierarchy for Each Group

Group	Cue types	Production	Perception
NE	Main cues	vowel duration F1 offset frequency closure voicing duration (closure duration)	vowel (release)
	Subsidiary cues	(release duration)	closure
HK	Main cues	vowel duration F1 offset frequency closure voicing duration (closure duration) (release duration)	vowel (release)
	Subsidiary cues	-	closure
LK	Main cues	vowel duration closure voicing duration	(release)
	Subsidiary cues	F1 offset frequency (closure duration) (release duration)	vowel closure

Note: Cues in parenthesis=only for the word-final released stimuli

5. Conclusion

This study revealed that Korean EFL learners' productions of word-final stops were less intelligible than those of native English speakers due to smaller acoustic differences between voiced and voiceless stimuli, as well as a tendency to produce unreleased stops. Higher-proficiency learners demonstrated near-native perceptual identification abilities, while lower-proficiency learners struggled with vowel cues and showed lower

identification skills. These results were explained by the cue-hierarchy analysis. The higher-proficiency EFL learners' production and perception cue hierarchy resembled that of native English speakers, though there were differences in the specific patterns for each POA. While the cue hierarchy in production was similar, the difference between voiced and voiceless sounds was less robust compared to the native English speakers. However, the lower-proficiency learners showed critical differences from native English speakers in that were insensitive to the vowel cues in both production (F1 offset frequency) and perception.

The findings suggest that English proficiency significantly affects perception of word-final stop voicing but has less impact on production. The lack of correlation between production intelligibility and perceptual identification ability among Korean EFL learners indicates that production and perception skills develop independently. Consequently, to address this gap, the study recommends implementing separate speaking and pronunciation instructions, which is currently being overlooked in in Korean high school EFL classrooms.

The current study has several limitations that should be addressed in future research. Firstly, the number of participants was too small due to the limited availability of native English speakers in Korea. While most previous studies included more than 10 participants per group, this study only had 8 per group. Consequently, the results may not be generalizable, and future research should involve a larger sample size. Additionally, the results of the cue match test might not be entirely reliable, as some participants reported difficulty maintaining attention towards the end. Listening to 480 stimuli and answering questions was very tiring. Future studies should consider reducing the number of stimuli to improve reliability. Lastly, participants seemed to exaggerate some acoustic cues during their recorded utterances, which may not represent their natural speech. This means the results may not fully reflect participants' natural production of word-final stops. Future research should consider this and examine how participants produce stimuli at various speech rates to analyze more natural productions.

References

Beach, E. F., D. Burnham and C. Kitamura. 2001. Bilingualism and the relationship between perception and production: Greek-English bilinguals and Thai bilabial stops. *International Journal of Bilingualism* 5(2), 221-235.

Bent, T. and A. Bradlow. 2003. The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America* 114(3), 1600-1610.

Catford, J. C. 1987. Phonetics and the teaching of pronunciation: A systematic description of English phonology. In J. Morley, ed., *Current Perspectives on Pronunciation: Practices Anchored in Theory*. TESOL.

Chen, M. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phoentica* 22, 129-159.

Cheng, B. and Y. Zhang. 2009. The relationship between speech perception and production in second language learners. *The Journal of the Acoustical Society of America* 125(4). 2754.

Cho, H and J. A. Shin. 2013. Korean learners' production of the voicing contrast in English word-final stops. *Korean Journal of English Language and Linguistics* 13(3), 695-716.

Cho, S. 2000. A study on the structure of the syllable in Korean. *Journal of Korean Language Education* 11(1), 131-151.

Crowther, C and V. Mann. 1992. Use of vocalic cues to final consonant voicing in the perception and production of English CVC's by native speakers of Arabic. *The Journal of the Acoustical Society of America* 91, 2472.

Denes, P. 1955. Effect of duration on the perception of voicing. *The Journal of the Acoustical Society of America*

27(4), 761-764.
Flege, J. E., M. J. Munro and L. Skelton. 1992. Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America* 92, 128-143.
Flege, J. E. 1999. The relation between L2 production and perception. *Proceedings of the XIVth International Congress of Phonetics Sciences*, 1273-1276.
Hayes-Harb, R., Smith, B. L., Bent, T., and A. R. Bradlow. 2008. The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics* 36, 664-679.
Hillenbrand, J., D. Ingrisano, B. Smith and J. E. Flege. 1984. Perception of the voiced-voiceless contrast in syllable-final stops. *The Journal of the Acoustical Society of America* 76, 18-26.
Holt, L. L., Lotto, A. J., and K. R. Kluender. 2001. Influence of fundamental frequency on stop-consonant voicing perception: A case of learned covariation or auditory enhancement?. *The Journal of the Acoustical Society of America* 109(2), 764-773.
Kang, S. 2007. A study on Korean students' production and perception of English word-final stop voicing. *Speech Sciences* 14(1), 105-119.
Kang, S. 2014. A study of asymmetry relations between perception and production in the word-final stop voicing. *Studies in Phonetics, Phonology and Morphology* 20(1), 3-22.
Kang, S. 2019. *Korean High School EFL Learners' Production and Perception of English Word-final Stop Voicing Contrast*. Master's thesis. Seoul National University.
Kim, H. 1998. A phonetic characterization of release and nonrelease: The case of Korean and English. *Linguistic Research* 34. 347-367.
Kim, J. 2011. A study on the correlation between English word-final stop and vowel duration produced by speakers of Korean. *Phonetics and Speech Sciences* 3(1), 15-22.
Kim-Renaud, Y. K. 1974. *Korean Consonantal Phonology*. Ph.D dissertation, University of Hawaii, Manoa.
King, R. D. 1967. Functional load and sound change. *Language* 43, 831-852.
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.
Lee, S. 2006. The relationship between perception and production in L2 English stress acquisition. *Studies in Phonetics, Phonology and Morphology* 12(3), 661-685.
Lopez-Soto, T. and D. Kewley-Port. 2009. Relation of perception training to production of codas in English as a Second Language. *The Journal of the Acoustical Society of America* 125(4). 2756.
Malécot, A. 1958. The role of releases in the identification of released final stops: A series of tape-cutting experiments. *Language: Journal of the Linguistic Society of America* 34(3), 370-380.
Munro, M. J., T. M. Derwing and S. L. Morton. 2006. The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition* 28(1), 111-131.
Nishi, K. and C. L. Rogers. 2002. On the relationship between perception and production of American English vowels by native speakers of Japanese: A pair of case studies. *The Journal of the Acoustical Society of America* 112(5), 2250.
Ohde, R. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America* 75(1), 224-230.
Peperkamp, S. and C. Bouchon. 2011. The relation between perception and production in L2 phonological processing. *Proceedings of Interspeech* 12, 161-164.
Peterson, G. E. and I. Lehiste. 1960. Duration of syllable nuclei in English. *The Journal of Acoustical Society*

America 32, 693-703.

Smith, B.L., R. Hayes-Harb, M. Bruss and A. Harker. 2009. Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics* 37, 257-275.

Son, I. 2009. A duration comparison between native speakers of Korean and native speakers of English in the rhyme with a stop coda. *Studies in Modern Grammar* 60, 137-155.

Tsukada, K., D. Birdsong, M. Mack, H. Sung, E. Bialystok and J. Fledge. 2004. Release bursts in English word-final voiceless stops produced by native English and Korean adults and children. *Phonetica* 61(2-3), 67-84.

Tsushima, T. and M. Hamada. 2005. Relation between perception and production ability during a speech training course. *The Journal of the Acoustical Society of America* 117(4). 2427.

Wang, W. 1959. Transition and release as perceptual cues for final plosives. *Journal of Speech, Language, and Hearing Research* 2(1), 66-73.

Examples in: English
Applicable Languages: English
Applicable Level: All

