Analysis of Korean Word-Initial Stops:
A Phonetic Study on Non-Native Speakers of Korean

Corey Fraser

Supervisor: Dr. Chandan Narayan

Table of Contents

1 Introduction and Background

Research in second language (L2) acquisition consistently demonstrates that some L2 sounds present greater challenges in both perception and production compared to others for learners. The difficulty in acquiring specific L2 sounds is often linked to the influence of first language (L1) phonological knowledge. It is generally assumed that learning an L2 sound is easier when it closely resembles an L1 sound and more challenging when it significantly differs. However, L2 speech perception research indicates that the process is more complex than merely comparing L2 sounds to L1 sounds. A well-documented instance of this challenge is observed in L1 Japanese speakers learning English as an L2. These learners often struggle with the /l/-/r/ contrast in English (Goto, 1971; Miyawaki et al., 1975). The issue arises because the closest related phoneme in Japanese is perceptually intermediate between /l/ and /r/, consequently interfering with the learners' attempts to acquire this novel distinction in English (Lotto et al., 2004). This example underscores the interplay between L1 phonological influence and the acquisition of L2 sounds, highlighting the complexities of the underlying processes that learners undergo during L2 acquisition.

Second Language Acquisition (SLA) research encompasses a broad spectrum of topics related to how individuals learn languages beyond their native tongue. Within this extensive field, significant attention has been devoted to understanding the mechanisms and outcomes of learners' encounters with novel phonetic distinctions in L2 learning. As noted by Chang (2009), a substantial body of research in second language acquisition has explored how L2 learners acquire laryngeal categories that differ from those in their L1. However, most of this research has focused on languages with two laryngeal categories that differ between L1 and L2 in terms of the same primary acoustic cue: voice onset time (VOT). This focus is likely due to the prominence

of languages that have two-way distinctions within their phonemic inventory in comparison to languages that exhibit three-way or four-way distinctions (Lisker & Abramson, 1964).

The present study examines non-native speakers' production of novel phonemes in Korean, which has a three-way laryngeal contrast. While Korean is not an understudied language in the field of SLA research, it was selected for this study for two key reasons. First, Korean is increasingly becoming a prominent choice for language learners, and this rising popularity is expected to yield a diverse range of speakers with varying levels of proficiency. Second, Korean features a typologically unique laryngeal distinction within its phonemic inventory that is not present in many other languages (Lisker & Abramson, 1964). These two reasons will be explored in more detail in the following sections.

1.1 Rising Popularity of Korean in Language Learning

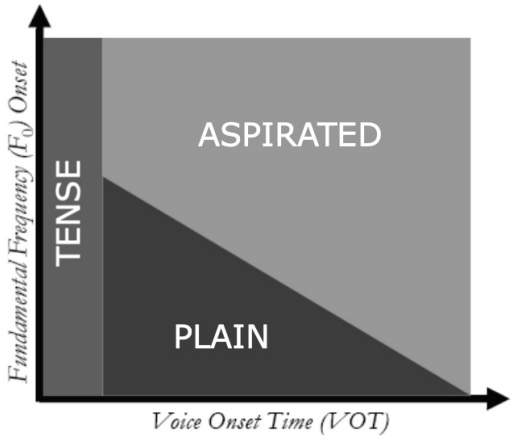
According to the 2021 MLA Language Enrollment Survey, despite a 16.6% decline in language enrollment at U.S. universities and colleges between 2016 and 2021, Korean was one of the three languages that saw an increase in enrollment during this period, with a 38.3% rise, making it one of the top 15 most studied languages (Lusin et al., 2023). As outlined by Wang and Pyun (2020), several factors have contributed to the growing popularity of the Korean language. These include Korea's economic success and elevated status on the global stage, the Korean government's strategic support of Korean language studies overseas through agencies such as the Korea Foundation and the Academy of Korean Studies, and the immense popularity of Korean popular culture, also known as Hallyu or the "Korean Wave" (Wang & Pyun, 2020). Hallyu emerged in the early 1990s with the success of Korean television dramas and films in Asian countries, particularly China and Japan. It later spread to other Asian regions and was eventually embraced by audiences worldwide, evolving into a global phenomenon. Wang and Pyun (2020)

found that two-thirds of participants enrolled in Korean language courses cited an interest in Korean pop culture as a significant catalyst for their interest in learning Korean and a key motivational factor. The increasing prominence of Korean as a second language was a primary factor in the selection of Korean for this study as a primary focus of this study was to investigate the realization of novel Korean phonemes by non-native speakers across various proficiency levels. Due to the increasing interest in Korean as a foreign language over the past few decades, it was anticipated that participants with a broad spectrum of L2 proficiency levels would be available for this research.

1.2 Korean Three-way Laryngeal Contrast

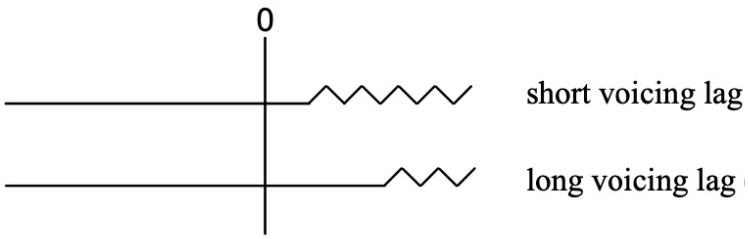
Similar to the difficulties faced by L1 Japanese speakers learning English as an L2 and encountering the novel /l/-/r/ contrast, L2 Korean learners face significant challenges when acquiring the phonemic inventory of Korean. Korean presents L2 learners with a distinct phonological contrast in plosive articulation, featuring three categories of voiceless plosives: plain, tense, and aspirated. These plosives, as realized by native Korean speakers, are distinguished by the use of both voice onset time (VOT) and fundamental frequency (F0) as acoustic cues. Tense plosives are characterized by short-lag VOT and high F0, plain plosives by long-lag VOT and low F0, and aspirated plosives by long-lag VOT and high F0 (see Figure 1) (Chang & Mandock, 2019; Kang & Nagy, 2016; Lee & Jongman, 2012; Ryu, 2017; Seo et al., 2022).

Figure 1: Schematic of the Korean laryngeal contrast in terms of voice onset time and fundamental frequency onset (based on Kim 2004, altered from Chang, 2009).



This contrast is novel for many L2 speakers because it does not exist in most other languages, with the exceptions of Thai and Eastern Armenian (Lisker & Abramson, 1964).. Most languages use a two-way distinction with VOT as the primary acoustic cue (Lisker & Abramson, 1964). For instance, in English, plosives are categorized into voiced and voiceless contrasts (/b/-/p/, /d/-/t/, /g/-/k/). The primary acoustic cue used in English to differentiate these contrasts is VOT, with voiced plosives typically having a short-lag VOT value (voicing begins shortly after the release of the stop) and voiceless plosives being realized with a long-lag VOT value (voicing begins considerably later than the release of the stop) (Figure 2) (Lisker & Abramson, 1964).

Figure 2: Schematic illustration of negative VOT, short lag VOT and long lag VOT (adapted from Simon, 2010)



Before outlining the specific research goals and hypotheses of this study, the forthcoming section will provide a comprehensive review of previous SLA research and theoretical

frameworks pertinent to phonemic perception and production, as well as an examination of research focusing on the acquisition of the Korean three-way laryngeal contrast. Following this, the study's hypotheses and research objectives will be presented. Subsequently, Section 3 will outline the methodological approach employed in this study will be detailed including participants, data collection, and analysis. Section 4 consists of description and analysis of the results. Section 5 will offer an in-depth discussion of the findings, address the study's limitations, and propose avenues for future research building upon this work.

2 Literature Review

The study of L2 phoneme acquisition has garnered significant attention within the field of SLA research, particularly concerning how learners navigate the challenges posed by novel phonetic distinctions. Central to understanding L2 phoneme acquisition are two prominent theoretical models: Flege & Bohn's (2021) Revised Speech Learning Model (SLM-r) and Best's (1994) Perceptual Assimilation Model (PAM). These models offer valuable insights into how L2 learners perceive and produce sounds that are not present in their L1. Additionally, this review will examine previous research on the acquisition of Korean phonemes, particularly the challenges faced by non-native speakers of varying L1 backgrounds, including longitudinal studies and research on proficiency effects.

2.1 Theoretical Frameworks

2.1.1 Revised Speech Learning Model: SLM-R

The Speech Learning Model (SLM), originally posited by Flege (1995), is a theoretical framework that outlines that the most difficult L2 sounds to acquire accurately are not novel sounds, but rather sounds that are similar to those of the L1. This is because similar sounds

become perceptually linked with sounds in the L1, which the SLM posits interferes with the formation of distinct representations of the L2 sounds (Flege, 1995). As a result, L2 learners might exhibit varying degrees of difficulty depending on whether the L2 sounds fit into L1 categories or represent novel distinctions (Flege, 1995). For example, if an L2 learner encounters a sound in the target language that closely resembles a sound from their native language, they are likely to perceive and produce the L2 sound using the existing phonetic framework of the L1. This can lead to inaccuracies and a failure to develop a separate phonetic category for the L2 sound. In contrast, truly novel sounds that have no counterpart in the L1 are more likely to be perceived and produced as distinct phonetic entities, albeit with initial difficulty.

The SLM was revisited and expanded in 2021 by Flege and Bohn, leading to the revised model known as SLM-r (Flege & Bohn, 2021). The revised model accounts for how phonetic systems reorganize over the lifespan of a learner in response to the phonetic input received during naturalistic L2 learning. This updated version of the SLM places greater emphasis on the dynamic nature of phonetic learning and the continual interaction between L1 and L2 phonetic systems (Flege & Bohn, 2021). SLM-r suggests that the process of acquiring L2 sounds involves not only the initial stage of learning, but also ongoing adjustments and refinements influenced by continued exposure to the L2 (Flege & Bohn, 2021). These adjustments are shaped by the learner's perceptual experiences and the specific phonetic characteristics of the L2 input they receive (Flege & Bohn, 2021). As learners are exposed to more varied and naturalistic L2 input, their phonetic categories may undergo reorganization, leading to more accurate production and perception of L2 sounds over time (Flege & Bohn, 2021). As articulated by Flege (2021), “L2 speech learning is profoundly shaped by perceptual biases induced by the L1 phonetic system” (p. 24). This statement emphasizes that L2 learners often interpret new L2 sounds through the

lens of their existing L1 phonetic categories, which can either facilitate or hinder the acquisition process depending on the degree of similarity between L1 and L2 sounds.

2.1.2 Perceptual Assimilation Model: PAM

In line with the SLM-r, the Perceptual Assimilation Model (PAM), proposed by Best (1994), offers a complementary perspective by focusing on how L2 sounds are perceived through the lens of the L1 phonetic system. Both models recognize the influence of L1 on L2 acquisition but differ in their primary focus: SLM-r is concerned with the long-term adaptation of the phonetic system, whereas PAM is centered on the immediate perceptual categorization of L2 sounds based on L1 categories.

The PAM provides a framework for understanding how listeners perceive non-native phonetic segments. According to PAM, L2 sounds are assimilated into existing L1 phonetic categories based on their perceived similarity (Best, 1994). PAM's focus on perceptual assimilation underscores the role of L1 phonetic categories in shaping how L2 sounds are initially processed. This model suggests that the ease or difficulty of acquiring L2 contrasts is largely determined by the degree to which L2 sounds can be mapped onto existing L1 categories (Best, 1994). For instance, if an English speaker learning Korean perceives the Korean tense and aspirated plosives as variants of a single English plosive, they may struggle to accurately produce and distinguish these sounds.

As previously discussed, while the SLM-r emphasizes the gradual reorganization of the phonetic system, PAM highlights the immediate perceptual challenges faced by L2 learners. This difference in focus makes PAM particularly useful for predicting and explaining early-stage L2 perception difficulties, whereas SLM-r provides a broader perspective on long-term phonetic

development. Despite these differences, both models acknowledge the profound impact of L1 on L2 phonetic acquisition. By examining how L1-induced perceptual biases shape L2 learning, both PAM and SLM-r offer valuable insights into the intricacies of L2 sound acquisition. Together, these models provide a comprehensive understanding of the challenges and processes involved in L2 phonetic acquisition, offering a dual perspective that encompasses both immediate and evolving aspects of L2 learning.

2.2 Prior Research on Korean Three-Way Laryngeal Distinction

Much research has been conducted on the acquisition and perception of Korean laryngeal plosive distinctions by non-native speakers, particularly those whose first languages have a two-way distinction with similar phonemes. Studies have focused on L1 backgrounds such as English (Chang, 2009; Seo et al., 2020), French (Lee, 2022), and Mandarin (Ryu, 2017). These studies have consistently shown that L2 Korean learners face challenges in accurately utilizing both fundamental frequency and voice onset time to produce and perceive Korean's three-way plosive distinctions—plain, tense, and aspirated. The research indicates that L2 speakers from L1 backgrounds with a two-way distinction predominantly rely on VOT to differentiate plosives, while F0, which is a crucial cue for native Korean speakers, is used less frequently by L2 learners. The findings across these studies suggest that non-native speakers often impose their L1 phonetic categories on L2 sounds, a phenomenon consistent with the SLM-r and PAM. According to these frameworks, learners may struggle with the novel three-way distinction in Korean because they initially categorize L2 sounds based on the phonetic distinctions of their L1. This early categorization can lead to difficulties in adapting to new phonemic distinctions that do not align with their existing L1 categories, such as the use of fundamental frequency (F0) as an acoustic cue.

Moreover, proficiency effects have been observed in the acquisition of these plosives. Ryu (2017) examined L1 Mandarin speakers learning Korean and found that higher proficiency levels were associated with improved acoustic perception of the three distinct plosive sounds. This finding lends support to the SLM-r framework by indicating that increased exposure to the L2, which typically accompanies higher proficiency, facilitates more accurate categorization and production of L2 phonemes. As learners progress in their proficiency, their ability to perceive and produce the phonemic distinctions of the L2 becomes more refined, suggesting a gradual reorganization of phonetic categories within their mental lexicon. The results of these prior studies highlight the relationship between L1 phonological influences and L2 acquisition. The evidence supports the theoretical frameworks of SLM-r and PAM by illustrating how L2 learners' reliance on L1 categories affects their acquisition of L2 phonemes and how proficiency impacts their ability to overcome these challenges.

This review of the theoretical framework and prior research on L2 phoneme acquisition and production provides a thorough foundation for the research goals and hypotheses of the current study.

2.3 CURRENT STUDY

The present study is focused on analyzing the non-native production of the Korean three-way laryngeal contrast, particularly within the context of Korean plosives. The central aim is to investigate how non-native speakers produce this novel three-way distinction, in line with established studies such as those conducted by Chang (2009) and Seo et al., (2020). These studies have identified notable variation in how L2 learners realize this phonemic distinction, revealing that non-native speakers often exhibit diverse phonetic realizations. This research seeks

to expand upon these findings by providing a more nuanced understanding of these variations, particularly in relation to the phonetic cues used by non-native speakers.

Furthermore, the study aims to explore the effects of proficiency on the production of the three-way laryngeal contrast. Building on the work of Ryu (2017), which demonstrated that L2 learners with higher proficiency levels in Korean are better able to perceive distinct phonemes compared to less proficient learners, this study will further investigate how proficiency influences production accuracy and the ability to achieve the novel three-way distinction. By examining these proficiency effects, the research intends to contribute valuable insights into how varying levels of language proficiency impact the acquisition of complex phonemic contrasts.

The main research question of this study is as follows:

1. Do non-native speakers utilize F0 and/or VOT to distinguish between the three-way laryngeal contrast in Korean?

Additionally, a subordinate research question was proposed:

2. Does increased L2 proficiency affect the use of F0 and VOT in the production of the three-way laryngeal contrast in Korean?

By addressing these research questions, the study aims to provide a comprehensive understanding of how non-native speakers produce the Korean three-way laryngeal distinction. Insights gained from this research may contribute to the broader field of SLA by highlighting the intricate interplay between L1 phonetic influence and L2 phonemic acquisition. The findings may also shed light on the role of proficiency in the successful adoption of novel phonetic distinctions.

With these objectives in mind, the subsequent section will detail the methodology employed in this study, including participant selection, data collection procedures, and analytical techniques, to systematically investigate the outlined research questions.

3 Methodology

3.1 Participants

Two participants groups were involved in this study. The baseline group comprised five native Korean speakers who currently reside within the Greater Toronto Area in Ontario, Canada. The ages of the baseline participants ranged from 28 years old (two participants), 30 years old (two participants), and 31 years old (one participant). All participants were born in Korea, across various cities (see Table 1). Due to the variation within the places in Korea that these participants were born in, dialectal variation was expected to occur between the baseline group. However, it is important to note that Korean dialects do not exhibit differences in onset consonants and mainly differ in mid-syllable consonants, vowel tone and duration, and grammatical alterations (Tranter, 2012). In the background questionnaire that the baseline group received, participants were asked about their place of birth, their age, the first language of their parents, the language most used at home, and if they spoke any languages other than English or Korean and their proficiency level in said languages on a scale of beginner, intermediate, advanced, near native (see Appendix B for the full list of questions included in the questionnaire).

Table 1: Summary of questionnaire responses for Native Korean speakers

Native Korean speakers	Parent's L1	Language most often used at home	P.O.B	Dialect[1]
Speaker 1	Korean	Korean	Chungju, South Korea	Cheongchung
Speaker 2	Korean	Korean	South Korea[2]	n/a
Speaker 3	Korean	Korean	Daejon, South Korea	Cheongchung
Speaker 4	Korean	Korean	Gwangmyeong, South Korea	Seoul
Speaker 5	Korean	Korean	Gwangju, South Korea	Jeolla

The target group consisted of 18 non-native Korean speakers. The ages of these participants ranged from 22 to 53, with a median age of 27. The duration of Korean learning varied between participants, ranging from 0.5 years to 32 years, with a median of 5 years. Additionally, the duration of Korean learning that occurred within a formal classroom setting varied as well, with the duration ranging between 1 to 8 years, with a median of 3 years.

Similar to the baseline group, the non-native speakers were also asked to complete a questionnaire to gather additional information about their language background and their Korean learning experience. This included the duration of learning, methods of learning, weekly Korean usage, and self-rated proficiency in reading, writing, speaking, and listening. See **Table 2** for a summary of the non-native speakers' language background and Korean learning experience.

Participants were recruited through various channels, including email recruitment via York University's Korean language department, specifically targeting students in beginner,

[1] Based on participant's P.O.B.
[2] Specific P.O.B was not provided.

intermediate, and advanced Korean language courses, as well as through the linguistics departments at York University, the University of Toronto, and the University of Hawaii at Manoa. Additional recruitment efforts included social media postings on Facebook and Instagram, along with word-of-mouth referrals from previous participants and the primary researcher. All recruitment was consistent with ethical practices outlined by the Office of Research Ethics, York University.

Table 2: Summary of Non-Native Korean-Speaking Participants' Native Languages, Self-Reported Korean Proficiency, and Additional Non-Native Languages

Non-Native Korean Speakers	*Native Language*	*Korean Proficiency: Reading*	*Korean Proficiency: Speaking*	*Korean Proficiency: Listening*	*Korean Proficiency: Writing*	*Additional non-native languages? (Y/N)*
NNS1	English	Intermediate	Intermediate	Intermediate	Intermediate	Y
NNS2	English	Intermediate	Intermediate	Intermediate	Intermediate	Y
NNS3	English	Advanced	Intermediate	Intermediate	Intermediate	Y
NNS4	English	Beginner	Beginner	Beginner	Beginner	N
NNS5	English	Intermediate	Beginner	Beginner	Intermediate	Y
NNS6	English	Intermediate	Advanced	Advanced	Intermediate	N
NNS7	English	Advanced	Intermediate	Intermediate	Intermediate	Y
NNS8	English	Intermediate	Intermediate	Intermediate	Intermediate	Y
NNS9	English	Intermediate	Intermediate	Intermediate	Intermediate	Y
NNS10	English	Intermediate	Advanced	Advanced	Intermediate	Y
NNS11	English	Intermediate	Intermediate	Intermediate	Intermediate	Y
NNS12	English	Beginner	Beginner	Beginner	Beginner	Y
NNS13	English	Intermediate	Intermediate	Advanced	Intermediate	Y
NNS14	English	Intermediate	Beginner	Intermediate	Intermediate	N
NNS15	French	Intermediate	Intermediate	Intermediate	Intermediate	Y
NNS16	French	Advanced	Intermediate	Advanced	Intermediate	Y
NNS17	French	Advanced	Advanced	Intermediate	Intermediate	Y
NNS18	Mongolian	Beginner	Beginner	Beginner	Beginner	Y

3.2 Test Materials

The stimuli for this study consisted of nine plosive consonants which are detailed in Table 3 with the Korean orthography and the phonemic transcription that will be adhered to for this study which follows the Yale Romanization of Korean[3].

Table 3: Target Phonemes of this study

Place of Articulation	Plain	Tense	Aspirated
Bilabial	ㅂ /p/	ㅃ /pp/	ㅍ /ph/
Alveolar	ㄷ /t/	ㄸ /tt/	ㅌ /th/
Velar	ㄱ /k/	ㄲ /kk/	ㅋ /kh/

Each target consonant was represented by six stimuli within the wordlist created and served as the onset consonant in a CVC-structured word, with the nucleus being a monophthong vowel (see Table 4 for examples from wordlist).

Table 4: Example stimuli words with Hangul and Romanization representations

Place of Articulation	Plain	Tense	Aspirated
Bilabial	방 /pang/	빵 /ppang/	팔 /phal/
Alveolar	달 /tal/	딸 /ttal/	팀 /thim/
Velar	금 /kum/	꿈 /kkwum/	큰 /khun/

Among these six stimuli, three consisted of elementary Korean words that participants were likely to have encountered previously, either on their own or within larger compound words. The remaining three stimuli were either more intermediate Korean words, which were unlikely to

[3] https://en.wikipedia.org/wiki/Yale_romanization_of_Korean

appear in general vocabulary learning, or pseudowords that adhered to the phonotactic rules of Korean but did not have any actual meaning in the language. The common words were selected through introspection by the primary researcher, who is a non-native Korean speaker, in addition to consulting a Korean-English online dictionary to find advanced words and to ensure that the pseudowords did not correlate to any existing Korean words[4]. The inclusion of elementary, intermediate, and pseudowords was intended to ensure that participants encountered unfamiliar words, helping to distinguish between genuine phonetic acquisition and potentially the mere memorization of common word pronunciations.

The words were initially embedded in the fixed carrier phrase: *igeun* [CVC]*ieyo* ("This is a [CVC_{noun}]/This is [CVC_{adj}]"), following the methodology of Chang (2019). However, the presence of the nasal phoneme /n/ preceding the target phoneme introduced challenges in accurately measuring the Voice Onset Time (VOT) of the target plosives. Consequently, the phrasal structure was revised to: [CVC]*ieyo* ("It is a [CVC_{noun}]/It is [CVC_{adj}]"). This modification, which removed the preceding nasal sound, allowed for more precise VOT measurements to be recorded.

Incorporating a fixed carrier phrase aimed to minimize potential intonation variations, specifically list intonation, which is the tendency for the intonation or pitch of each word in a list to rise until the final word, which typically falls in pitch. By having participants produce the target utterances within a phrase, as opposed to producing them in isolation, the expectation was that this would prevent list intonation from influencing their pitch, thereby allowing for a clearer assessment of their phonetic acquisition and production.

[4] The words were classified as elementary, intermediate, or advanced as determined by consulting the Naver Korean-English dictionary and the TOPIK classification scale

In addition to the target words, 16 filler words were included in the stimuli which brought the total number of utterances per participant to 70. The wordlist was manually randomized to ensure that no two utterances of the same target phoneme occurred consecutively, and each participant received the wordlist in the same order. Each target sentence was written in Korean orthography and provided to the participants in PDF format. Overall, 984 target tokens were collected.

For the complete wordlist, see Appendix A.

3.3 Procedure

The survey platform Google Forms was used to administer the task to participants remotely. Given the constraints of this study, participants were unable to be recorded with high-quality audio recording equipment, such as a sound booth or specialized microphones. Instead, participants were instructed to record themselves using a laptop or smartphone with an external microphone, such as headphones with a built-in microphone or a USB microphone. It has been previously demonstrated that self-administered recordings using headset microphones are sufficient for certain types of acoustic analysis, including VOT measurements (see Narayan, 2023). To minimize environmental noise, participants were directed to record themselves in a quiet location, away from external disturbances, including other people, fans, traffic noise, and other ambient sounds. Additionally, the audio recordings were processed using Audacity's Noise Removal feature to eliminate as much background noise as possible and were saved in .wav format.

3.4 Measurements And Analysis

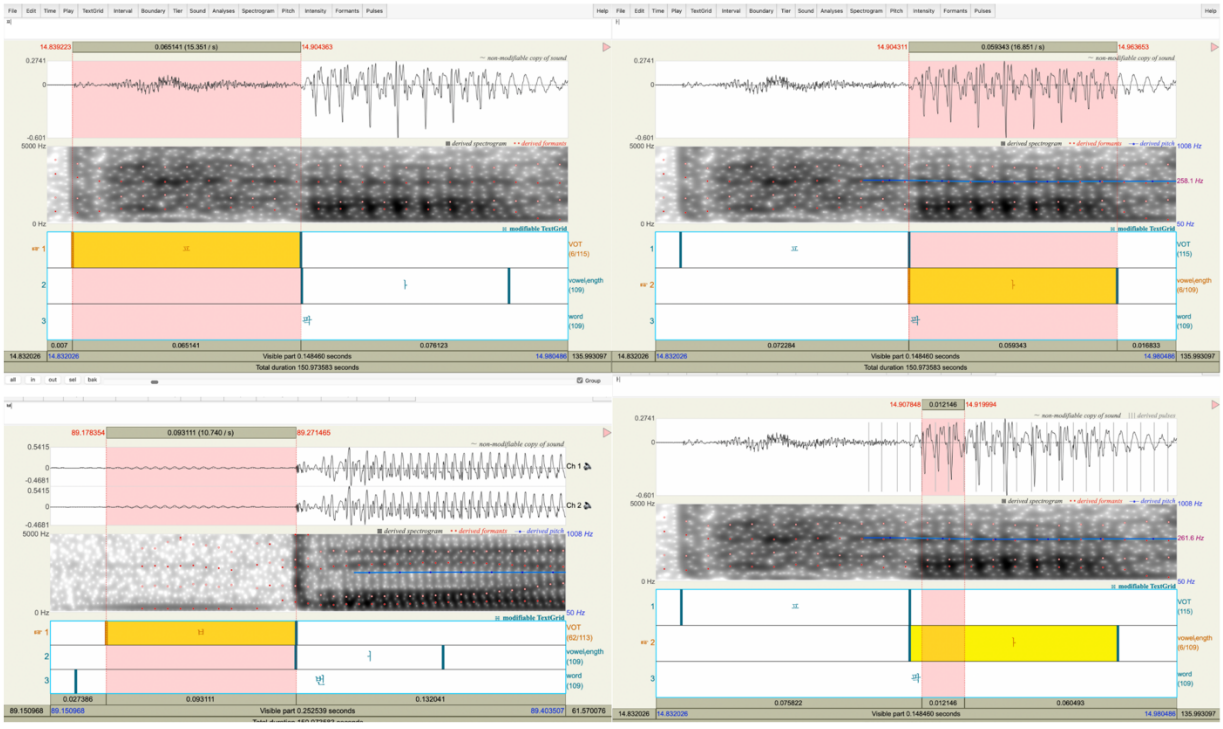
The audio recordings were manually annotated in a Praat (Broersma & Weenink, 2024) TextGrid by consulting a wide-band spectrogram and the sound waveform. Boundaries for VOT and vowel duration were delineated for each utterance. For VOT measurements, following the methodology of Narayan (2023), the starting boundary for positive VOT values was established at the initial release of the plosive, as indicated by a high concentration of energy on the spectrogram and waveform, while the ending boundary was marked at the onset of periodic voicing. For negative VOT values, the starting boundary was identified at the presence of a voice bar, and the ending boundary was marked at the release of the plosive. The duration of each plosive's VOT was collected with a Praat script designed to measure the duration (in seconds) between the beginning and end boundaries for each utterance, converting these measurements into milliseconds by multiplying by one thousand.

For vowel duration measurement, again following Narayan (2023), the vowel duration measurement began at the onset of periodic voicing (corresponding with the ending boundary of VOT) and the end of the vowel was "marked by three co-occurring events: (1) a dramatic change in amplitude in the waveform, (2) a change in the energy in the formants accompanied by a change in complexity in the waveform indicating a loss of energy in F2 and F3, and (3) the onset of aperiodicity" (Narayan, 2023).

To measure the fundamental frequency (F0) of the following vowel, the average pitch across a steady state of three glottal periods within the onset of the vowel duration was manually measured and recorded following the methodology of Chang (2019). To put all speakers on the same F0 scale, the F0 measurements were standardized to z-scores by learner. The z-scores were calculated using R by subtracting the speaker's mean F0 and dividing by the square root of the

speaker's variance in F0 for each F0 value within each plosive laryngeal type (plain, tense, and aspirated) (see Figure 3 for examples of measurements).

Figure 3: Top left image: positive VOT measurement; Top right image: vowel duration measurement; Bottom left image: negative VOT measurement; Bottom right image: F0 measurement



4 Results

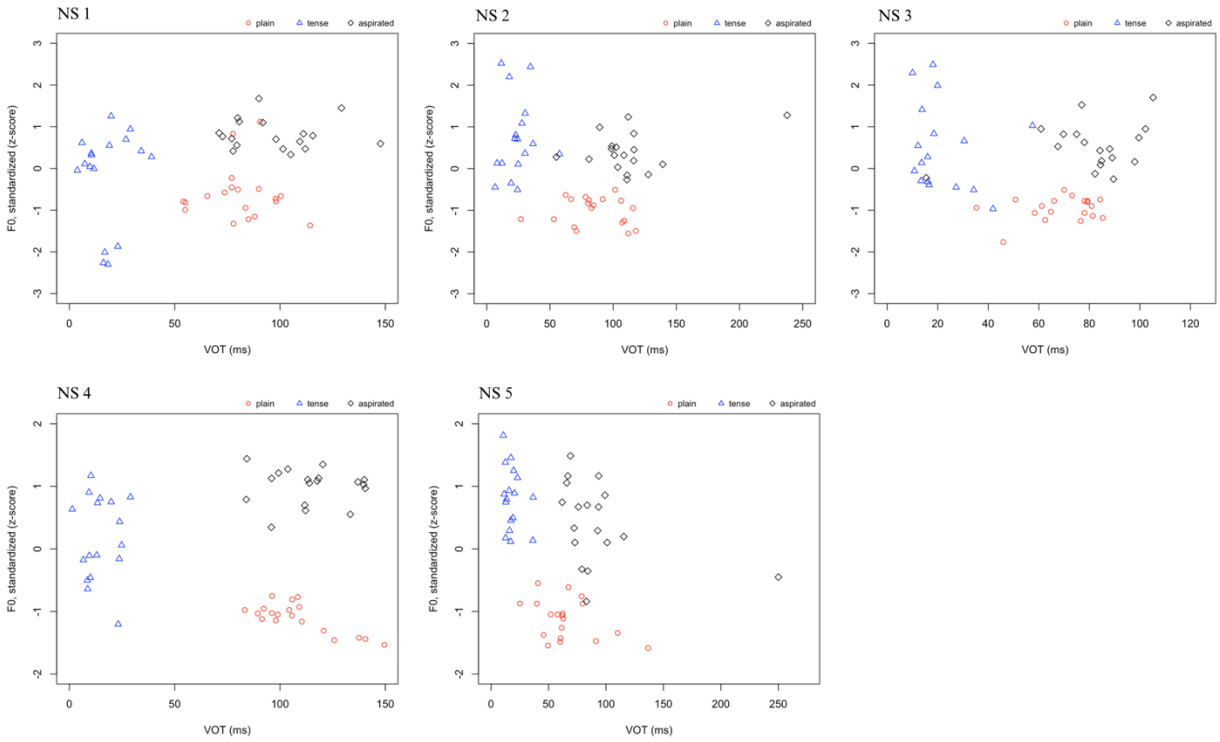
The results of this study comprise both quantitative data from the production task and qualitative data from the language questionnaire. To begin, below I will first examine the quantitative results of the production task performed by the native Korean and non-native Korean speakers.

4.1 Quantitative Results

The following section presents an examination of the native Korean speakers' performance on the production task. In all plots, plain productions are represented by red circles, tense productions by blue triangles, and aspirated productions by black diamonds.

4.1.1 Native Korean Speakers

Figure 4: Representative scatter plots of productions of native Korean speakers (top: NS1, NS2, NS3, bottom: NS4, NS5)



The phonetic spacing of the native Korean speakers (Figure 4) is generally consistent with previous research (Chang, 2009; Cho et al., 2002; Kang, 2006; Kim, 2004) regarding the realization of Korean laryngeal categories in terms of VOT and F0 onset. Tense plosives are characterized by short VOT and high F0 values. Plain plosives are produced with long VOT and lower F0 values. Aspirated plosives exhibit long VOT and high F0 values. While there is some overlap between the categories—specifically, overlap in F0 values between aspirated and tense

plosives, and in VOT values between aspirated and plain plosives—there is no overlap within both VOT and F0 for these Korean laryngeal categories. The results of the native Korean speakers' production indicate that VOT and F0 both function as acoustic cues for distinguishing among the three laryngeal types.

Additionally, consistent with previous research on the tonogenesis-like sound change in Seoul Korean (Choi, 2020; Kang, 2013), where the primary phonetic feature distinguishing plain and aspirated plosives is shifting from VOT to F0, this pattern is also evident in the present data. This observation aligns with the fact that all native speaker participants were born in the 1990s, and research on the topic indicates that this change has been in the process since the 1960s (Silva, 2006). This phenomenon is most clearly exemplified by NS4, who was born in the Seoul area and presumably speaks a Seoul dialect. The phonetic spacing for NS4 reveals a complete overlap in VOT for plain and aspirated plosives, with F0 values serving as the sole distinguishing feature between these two plosive categories (Figure 4, NS4).

4.1.2 Non-Native Korean Speakers

In line with prior research, there is significant variability in the phonetic spacing of plain, tense, and aspirated plosives among the non-native speaker group. To analyze these results comprehensively, the non-native speakers were categorized into seven groups based on their production patterns and acoustic cues used for differentiating these laryngeal contrasts. Linear Discriminant Analysis (LDA), a machine learning classification technique, was utilized to statistically categorize and organize the speakers based on their acoustic feature distributions. LDA enabled precise classification by identifying patterns and overlaps in plain, tense, and aspirated productions, allowing for grouping of speakers with similar phonetic characteristics. The LDA output for each speaker is presented as a confusion matrix, illustrating the accuracy

and misclassifications in the predictions. Within each matrix, the columns represent the predicted plosive types, while the rows represent the actual plosive types. This method provided a systematic approach to understanding the variations in how non-native speakers produce these laryngeal contrasts. Each group will be examined in detail below.

Figure 5: Representative scatter plot of productions of Group 1 speaker (NNS11)

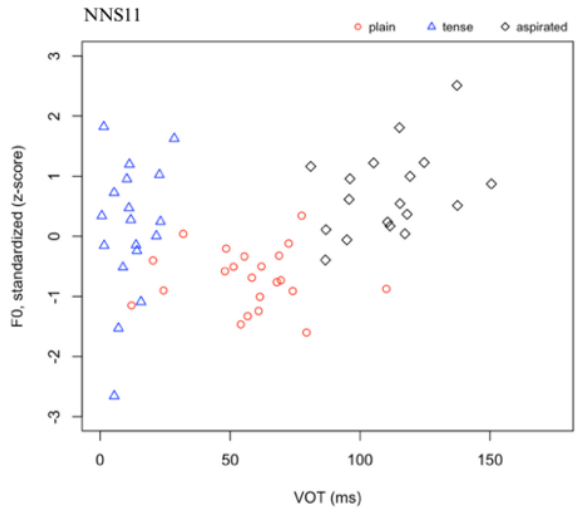


Table 5: LDA Confusion Matrix of NNS11

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	17	1	0
Plain	1	17	0
Tense	0	4	18

Group 1 (Figure 5) comprises a single non-native speaker, NNS11, whose phonetic spacing closely mirrors that of the baseline group (Figure 4). This speaker effectively employs VOT and F0 as acoustic cues for distinguishing between plain, tense, and aspirated plosives. Tense plosives are characterized by short VOT and higher pitch. Plain plosives display mid-range VOT, while aspirated plosives exhibit the longest VOT. There is an overlap in pitch values

between plain and tense plosives, with plain plosives having a lower pitch range. Additionally, aspirated and tense plosives also overlap in pitch values, with aspirated plosives having a higher pitch range. The confusion matrix for NNS11 reflects the minimal prediction overlap between the plosive types, with only a slight overlap observed within the plain and tense plosives (Table 5).

Figure 6: Representative scatter plot of productions of Group 2 speakers (NNS8, NNS2)

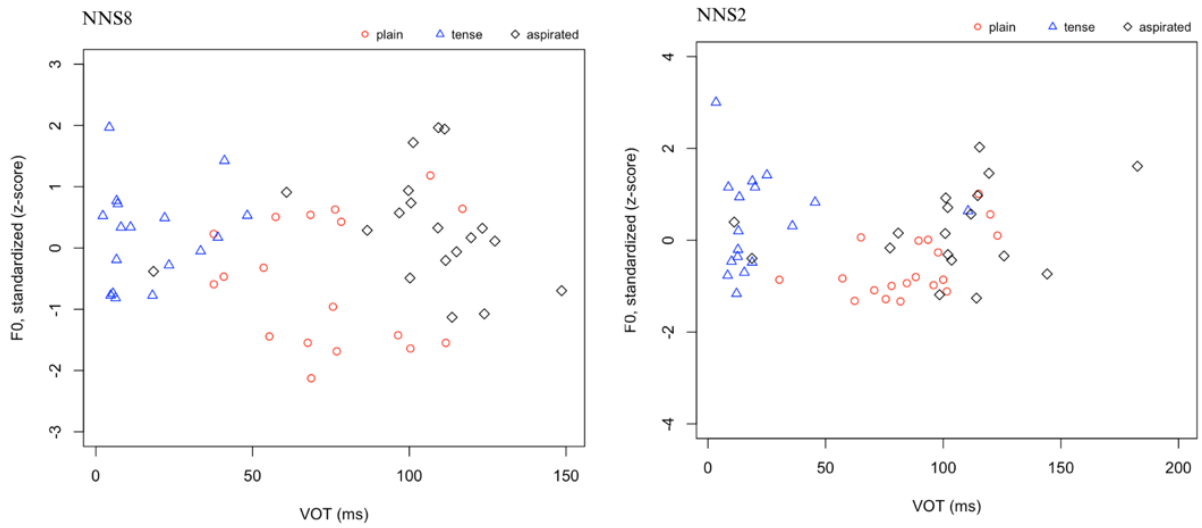


Table 6: LDA Confusion Matrix of NNS8

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	17	4	0
Plain	1	11	0
Tense	1	3	17

Table 7: LDA Confusion Matrix of NNS2

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	10	5	1
Plain	6	13	0
Tense	2	1	16

The phonetic spacings in Group 2 (Figure 6) indicate that these speakers (NNS8 and NNS2) primarily use VOT as the acoustic cue for distinguishing between plain, tense, and aspirated plosives. Tense plosives are characterized by short VOT values. Plain plosives display mid-range VOT values, while aspirated plosives are characterized by the longest VOT values. Compared to Group 1, there is less reliance on F0 as an acoustic cue by Group 2's speakers. Plain and aspirated plosives show slight variation in F0 with most values overlapping, and the F0 values for tense plosives overlap with those of both plain and aspirated productions. The confusion matrices for NNS8 and NNS4 reflect the overlap of plosive types by demonstrating overlap among predictions within the three plosive categories (Table 6 and Table 7, respectively).

Figure 7: Representative scatter plot of productions of Group 3 speaker (NNS5)

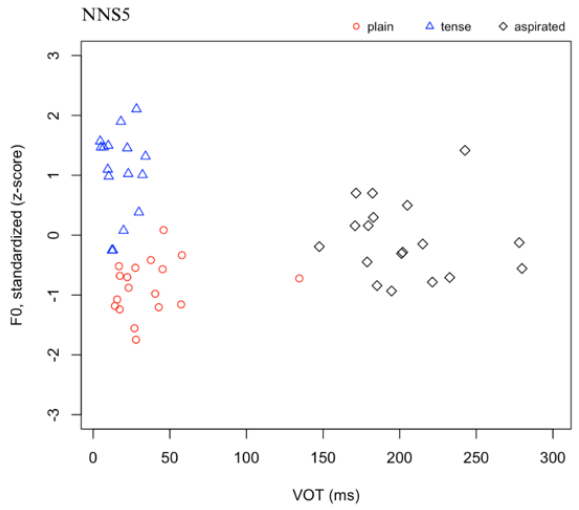


Table 8: LDA Confusion Matrix of NNS5

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	18	1	0
Plain	0	17	3
Tense	0	0	13

Group 3 (Figure 7) comprises of only one speaker, NNS5. The phonetic spacing indicates two two-way distinctions. First, both tense and plain plosives have short VOT values, while aspirated plosives have long VOT values. Second, plain and tense plosives are distinguished by their F0 values, with plain plosives having lower F0 values and tense plosives having higher F0 values. This F0 distinction does not extend to aspirated plosives, as there is overlap in the F0 values of both plain and tense plosives. The confusion matrices for NNS5 show no overlap in the prediction of tense plosives and minimal overlap in the prediction of aspirated and plain plosives (Table 8)

Figure 8: Representative scatter plot of productions of Group 4 speakers (NNS4, NNS7)

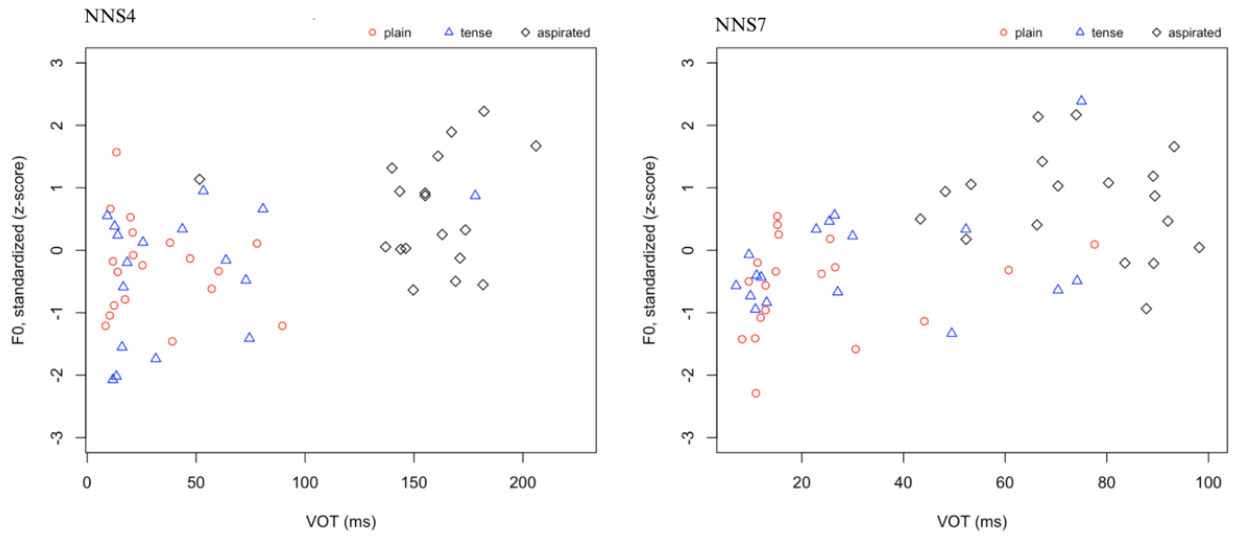


Table 9: LDA Confusion Matrix of NNS4

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	17	0	1
Plain	1	14	11
Tense	0	5	5

Table 10: LDA Confusion Matrix of NNS7

LDA Prediction	*Aspirated*	*Plain*	*Tense*

Aspirated	16	2	4
Plain	0	13	9
Tense	2	4	4

The phonetic spacings of the speakers in Group 4 (Figure 8; NNS4 and NNS7) also indicate the presence of two two-way distinctions similar to Group 3. Plain and tense plosives both exhibit short VOT values, while aspirated plosives have long VOT values, consistent with the results from Group 3. However, regarding the use of F0 as a distinguishing acoustic cue, the phonetic spacing of Group 4's speakers shows that plain and tense plosives overlap in F0 values, whereas aspirated plosives have higher F0 values. The predictions outlined in the confusion matrices for NNS4 (Table 9) and NNS7 (Table 10) further highlight this overlap, showing considerable overlap between tense and plain plosives.

Figure 9: Representative scatter plot of productions of Group 5 speakers (top: NNS12, NNS17, bottom: NNS9)

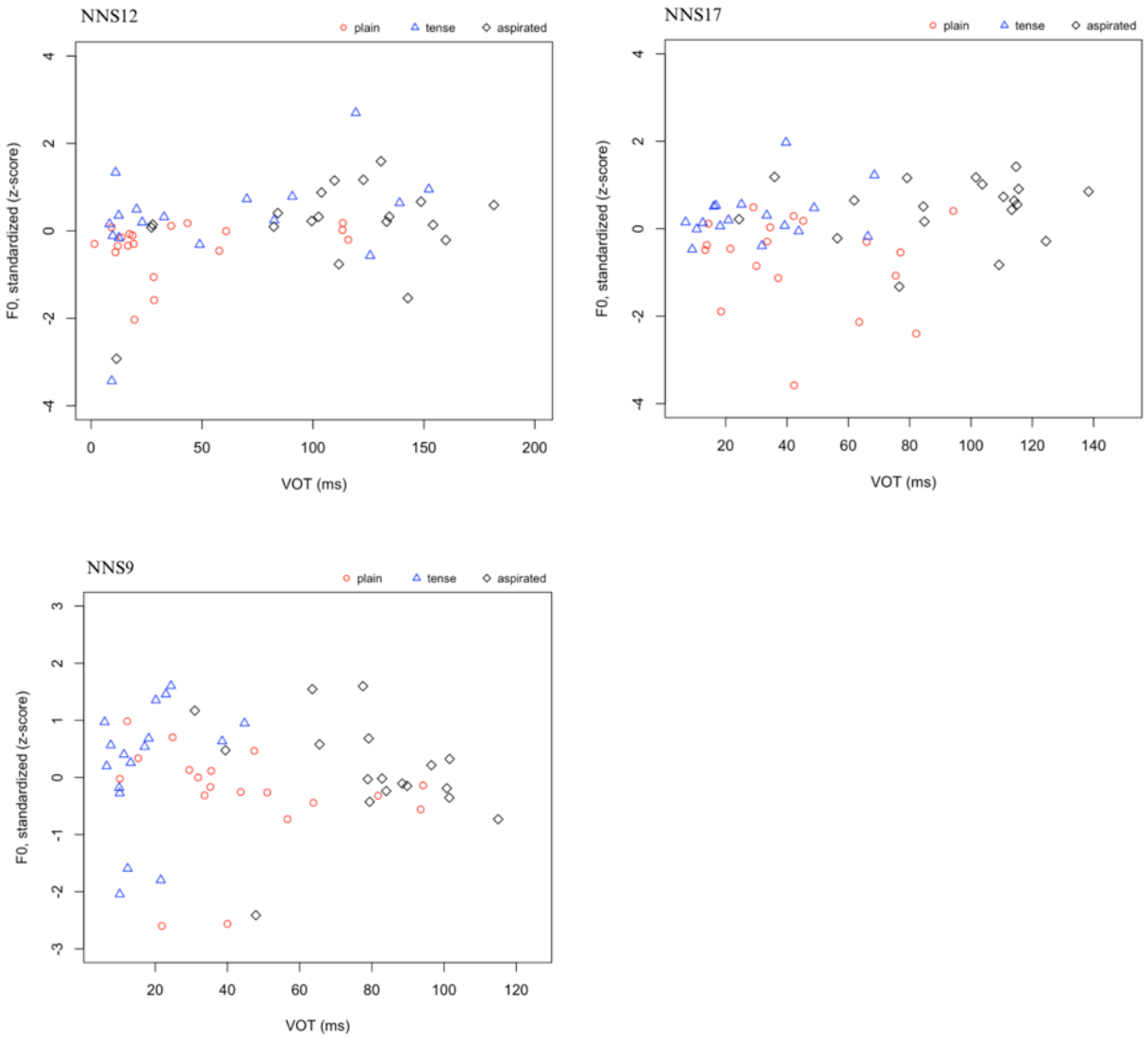


Table 11: LDA Confusion Matrix of NNS12

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	16	3	4
Plain	3	15	7
Tense	0	1	6

Table 12: LDA Confusion Matrix of NNS17

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	15	1	1
Plain	2	11	3
Tense	2	7	13

Table 13: LDA Confusion Matrix of NNS9

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	15	3	0
Plain	2	11	3
Tense	1	5	14

The phonetic spacings in Group 5 (Figure 9) indicate that these speakers (NNS12, NNS17, and NNS9) employ a single two-way distinction using VOT as the acoustic cue to differentiate aspirated plosives from plain and tense plosives. Aspirated plosives exhibit long VOT values, while plain and tense plosives overlap with shorter and mid-range VOT values. The predictions outlined in the confusion matrices for NNS12 (Table 11), NNS17 (Table) 12, and NNS9 (Table 13) highlight this, showing considerable overlap between tense and plain plosives. Additionally, there is no apparent use of F0 as an acoustic cue to differentiate between plain, tense, or aspirated plosives.

Figure 10: Representative scatter plot of productions of Group 6 speakers (NNS3 and NNS13)

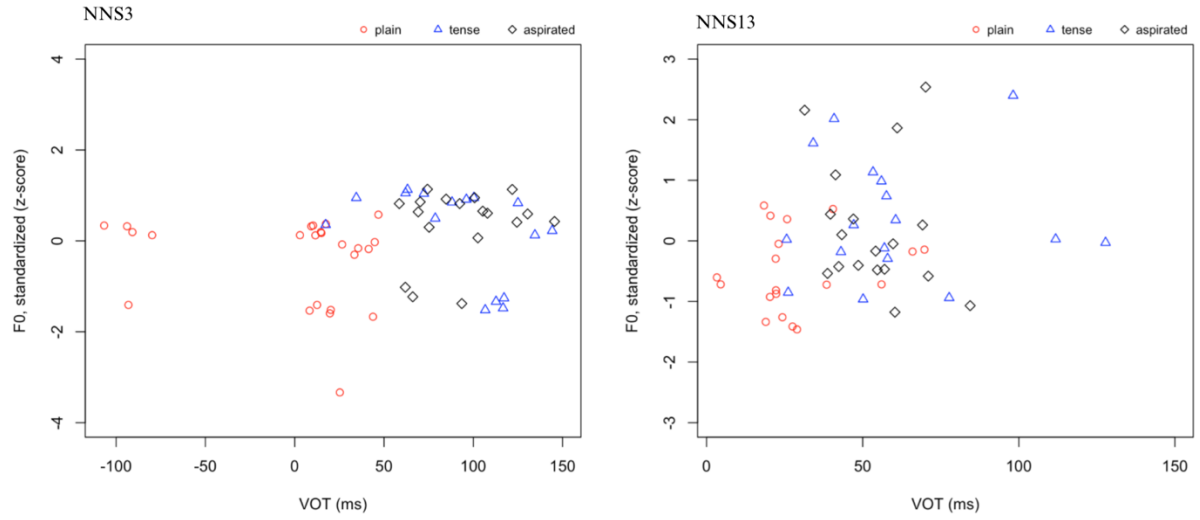


Table 14: LDA Confusion Matrix of NNS3

LDA Prediction	*Aspirated*	*Plain*	*Tense*

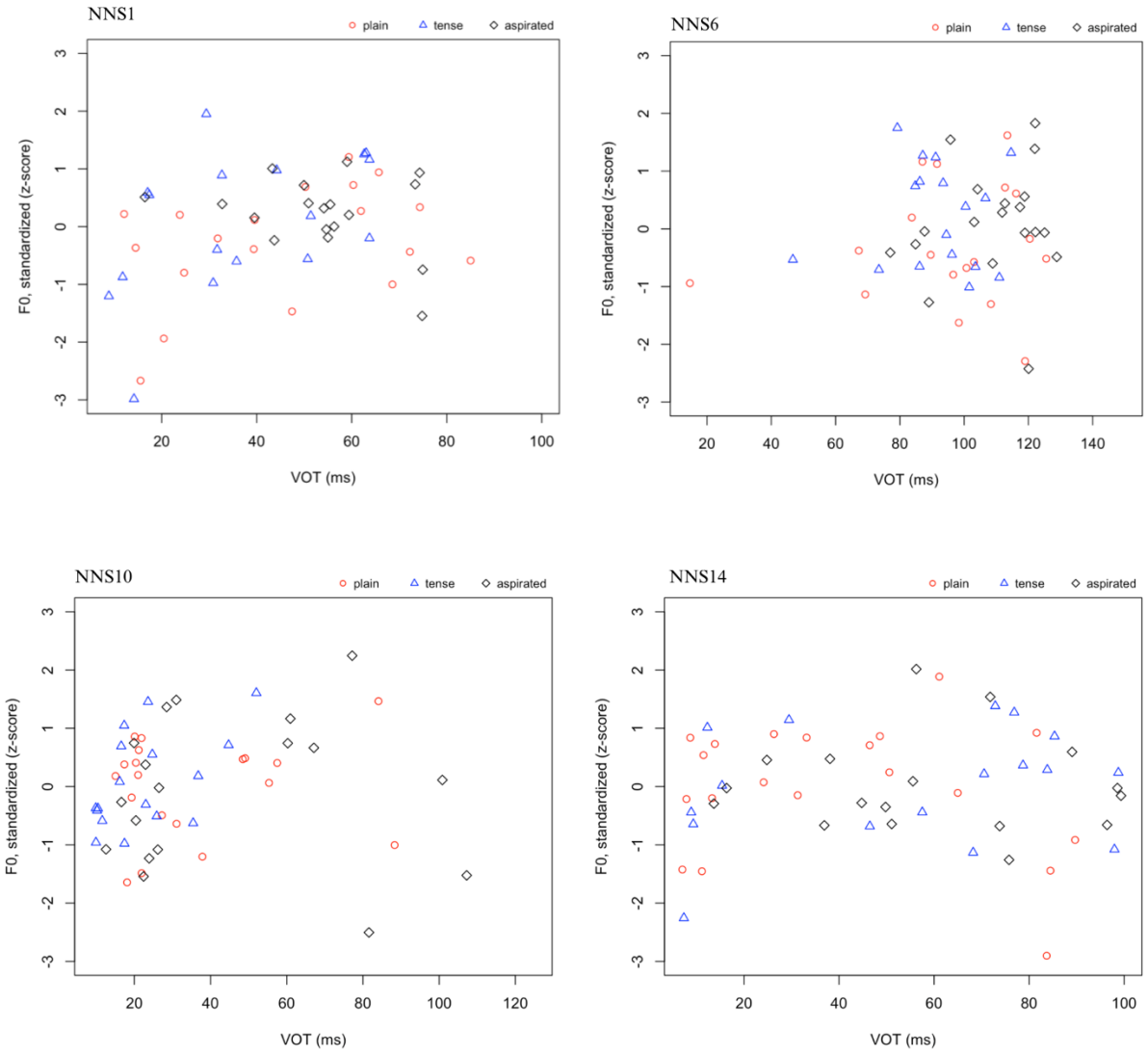
Aspirated	15	1	11
Plain	2	23	1
Tense	1	0	4

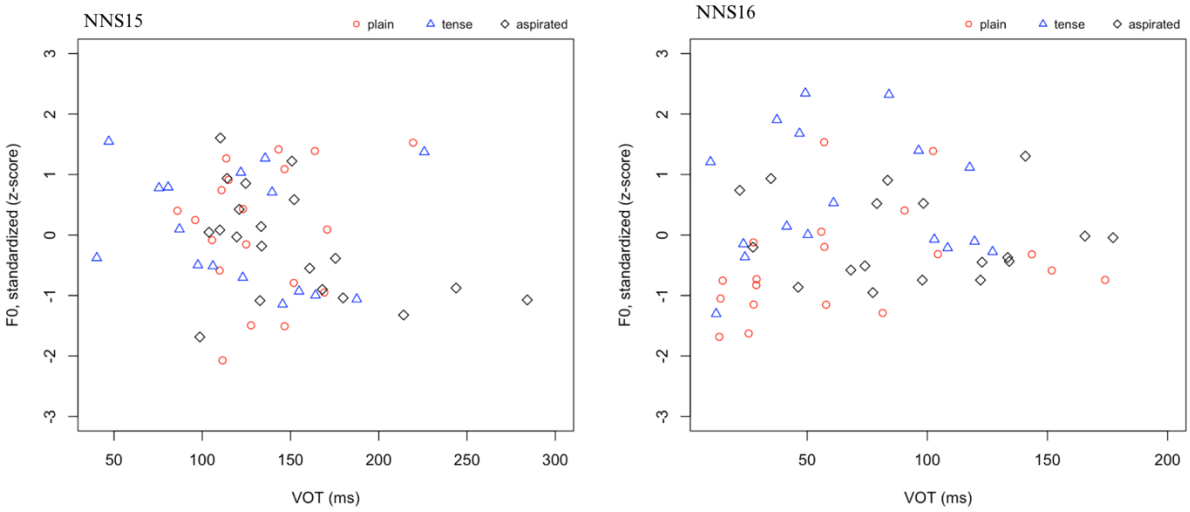
Table 15: LDA Confusion Matrix of NNS13

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	12	3	6
Plain	2	15	3
Tense	4	1	8

Group 6 comprised two speakers, NNS3 and NNS13 (Figure 10). Both speakers demonstrated a reversal in the expected VOT values for plain and tense plosives. Plain plosives were produced with shorter VOT values, while tense plosives exhibited mid-range to long VOT values, overlapping with those of aspirated plosives. The LDA predictions for both speakers (Table 14; NNS3 and Table 15; NNS13) reflect this overlap, as tense plosives were incorrectly predicted as aspirated plosives with higher frequency. Consistent with Group 5, no evident use of F0 as a distinguishing acoustic cue was observed. Additionally, a noteworthy aspect of speaker NNS3's phonetic spacing was observed: this speaker was the only one to produce any of the three laryngeal types with negative VOT values, indicating that voicing occurred prior to the release of the plosive.

Figure 11: Representative scatter plot of productions of Group 7 speakers (first row: NNS1, NNS6; second row: NNS10, NNS14; third row: NNS15, NNS16; fourth row; NNS18)





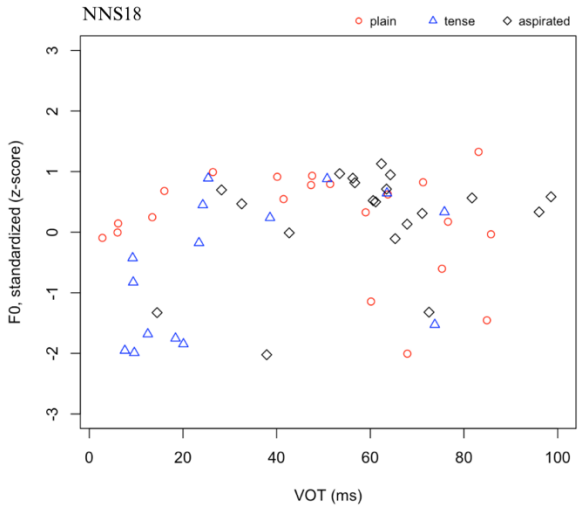


Table 16: LDA Confusion Matrix of NNS1

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	11	8	5
Plain	3	5	4
Tense	4	6	8

Table 17: LDA Confusion Matrix of NNS6

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	13	5	4
Plain	4	8	6
Tense	2	5	7

Table 18: LDA Confusion Matrix of NNS10

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	7	7	1
Plain	4	4	5
Tense	7	8	11

Table 19: LDA Confusion Matrix of NNS14

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	10	7	6
Plain	9	10	3
Tense	1	2	7

Table 20: LDA Confusion Matrix of NNS15

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	8	6	9
Plain	9	14	8
Tense	0	0	0

Table 21: LDA Confusion Matrix of NNS16

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	8	5	4
Plain	6	12	4
Tense	4	2	9

Table 22: LDA Confusion Matrix of NNS18

LDA Prediction	*Aspirated*	*Plain*	*Tense*
Aspirated	12	9	3
Plain	6	8	4
Tense	2	4	8

Group 7 (see Figure 11) is composed of the remaining seven speakers (NNS1, NNS6, NNS10, NNS14, NNS15, NNS16, and NNS18). The phonetic spacings within this group indicate

that these speakers do not maintain distinct categories for plain, tense, and aspirated plosives in terms of VOT and/or F0. Instead, the production of these sounds occurs over the same broad phonetic space, without clear differentiation between the categories. These overlaps are also evident in the variations in prediction accuracy within the confusion matrices for these speakers (Tables 16, 17, 18, 19, 20, 21, and 22).

As previously mentioned, both native Korean speakers and non-native Korean speakers were asked to complete background language questionnaires. The qualitative results from the non-native speakers' responses will be examined in the following section.

4.2 Qualitative Results

The non-native Korean participants in this study completed a language background questionnaire that inquired about the duration of their Korean study, their participation in formal or classroom-based Korean lessons (including the duration of such lessons if applicable), and their weekly Korean usage in hours. Given the open-ended nature of these questions, responses varied significantly in type and length. To ensure clarity and conciseness, the responses were categorized into four main groups representing the primary distinctions within the collected data. The summarized results of the participants' duration of Korean learning, duration of classroom-based Korean learning, and weekly Korean usage are presented below in Table 5, Table 6, and Table 7, respectively.

Table 24: Summary of duration of Korean learning by non-native speakers

Duration of Korean learning (years)	*Number of Participants*
0.5 – 1	2
3 – 5	9
6 – 9	3
10+	4

Table 25: Summary of duration of Korean learning in a classroom setting by non-native speakers

Duration of Korean learning in a classroom setting (years)	*Number of Participants*
0	3
1– 2	7
3– 4	8
5+	1

Table 26: Summary of weekly usage of Korean by non-native speakers

Weekly Korean usage (hours)	*Number of Participants*
0	3
1– 2	9
4– 9	4
10+	2

Following the earlier grouping schema of non-native speakers based on their use of VOT and/or F0 as acoustic cues, a summarized chart of the three aforementioned aspects (duration of

Korean learning, duration of classroom-based Korean learning, and weekly Korean usage) in addition to Korean speaking proficiency sorted by grouping is presented below in Table 8.

Table 27: Summary of duration of Korean learning, duration of Korean learning in a classroom setting, weekly usage of Korean and speaking proficiency, sorted by participant grouping (Groups 1-7)

	Duration of Korean learning (years)	*Duration of Korean learning in a classroom setting (years)*	*Weekly Korean usage (Hours)*	*Speaking Proficiency*
Group 1	10+	1-2	6-8	Intermediate
Group 2	3 - 5	3-4	0	Intermediate
Group 2	6 - 9	1-2	1-2	Intermediate
Group 3	10+	1-2	4-9	Beginner
Group 4	3 - 5	n/a	1-2	Beginner
Group 4	3 - 5	3-4	4-9	Intermediate
Group 5	0.5 - 1	1-2	0	Beginner
Group 5	6 - 9	3-4	1-2	Advanced
Group 5	10+	3-4	1-2	Intermediate
Group 6	3 - 5	1-2	1-2	Intermediate
Group 7	0.5 -1	n/a	1-2	Beginner
Group 7	3 - 5	1-2	0	Intermediate
Group 7	3 - 5	3-4	1-2	Beginner
Group 7	3 - 5	3-4	1-2	Intermediate
Group 7	3 - 5	3-4	10+	Intermediate
Group 7	3 - 5	3-4	10+	Intermediate
Group 7	6 - 9	1-2	4-9	Advanced
Group 7	10+	5+	1-2	Advanced

5 Discussion

To summarize the results of this study, there is considerable variation in non-native speakers' production of the Korean three-way laryngeal distinction. This finding aligns with previous research documenting similar variability (Chang, 2009; Lee, 2022; Seo et al., 2020; Ryu, 2017). Regarding research question 1, which examines whether non-native speakers utilize F0 and/or VOT to distinguish the three-way laryngeal contrast in Korean, the results indicate that more than half of the participants (n = 10) utilize at least one acoustic cue to differentiate between the laryngeal categories, with VOT being the most commonly used cue. This reliance on VOT is consistent with the participants' L1 phonetic systems, which often use VOT as a primary cue for plosive distinctions. Notably, only one participant (NNS11) demonstrated a native-like representation of the Korean three-way distinction in their production, effectively utilizing both F0 and VOT to differentiate between the plain, tense, and aspirated plosives (Figure 5). In contrast, the majority of participants (n = 17) either relied on an L1-like two-way distinction (Figures 6, 7, 8, 9, 10) or failed to make any clear distinction between the Korean three-way contrast (Figure 11). This suggests that while some non-native speakers can approximate native-like distinctions, many struggle to move beyond the phonetic categories of their L1.

These findings both support and challenge theoretical frameworks like SLM-r and PAM. Participants appear to produce L2 laryngeal contrasts through a lens similar to their L1 phonemic categories, which is consistent with these frameworks. However, a significant portion of the participants (n = 8) made no discernible distinctions between the laryngeal contrasts. This large proportion making no distinctions, even within an L1-like two-way framework, presents seemingly counterevidence against theories positing that L2 phoneme acquisition is strongly influenced by L1 phonetic categories.

Regarding research question 2, which explores whether increased L2 proficiency affects the use of F0 and VOT in the production of the three-way laryngeal contrast in Korean, the analysis of proficiency ratings, as self-reported by the participants, revealed no clear correlation between proficiency levels and the ability to produce the Korean three-way distinction. As shown in Table 7, there is a mix of beginner, intermediate, and advanced speakers within each grouping of phonetic distinctions. This suggests that proficiency in the results of this current study does not have a discernable influence correlating with the production of the three-way laryngeal contrast. This finding contrasts with Ryu (2017), who found a positive correlation between increased proficiency and better phonemic distinction. The lack of correlation in this study might be attributed to the subjective nature of self-reported proficiency or other factors not accounted for in the current analysis, which will be further discussed in section 5.1.

The variation observed in this study—both in the production of the Korean three-way laryngeal distinction and in the proficiency of speakers who do or do not differentiate between the Korean plosives—raises questions about the underlying reasons for such variation among speakers. One significant factor to consider is the method by which these sounds are introduced to learners. The complexity of the Korean three-way laryngeal distinction presents a challenge for Korean language instructors and learners alike, as effectively teaching this nuanced phonetic feature requires clear and precise methods. This prompts the question: how do the majority of L2 Korean learners first encounter and learn these sounds?

Many beginner-level Korean learning materials utilize the romanization of Hangul. This approach is understandable, given that introducing learners to an unfamiliar orthography can be daunting. Romanization simplifies the learning process by representing Korean sounds with familiar English letters, thereby reducing the intimidation associated with new and complex

aspects of L2 learning. However, the use of romanization can also present challenges in acquiring L2 phonetic categories such as the three-way laryngeal distinction. For instance, if learners are taught that Korean plain plosives correspond to /b/, /d/, and /g/ in English, and aspirated plosives correspond to /p/, /t/, and /k/, the tense plosives may become problematic. Romanization often categorizes tense plosives with double letters, such as /pp/, /tt/, or /kk/. This approach raises a question: what does /pp/ signify to a learner? The lack of direct phonetic correlation between these Romanized forms and their actual Korean counterparts may hinder learners' understanding and accurate production of tense plosives. These double letter romanization representations are often accompanied by explanations of how to pronounce tense or aspirated sounds, which can be as confusing, if not more so, for learners. For example, the popular free Korean learning website HowToStudyKorean.com describes the pronunciation of tense plosives as sounding similar to plain plosives but with "more force at the beginning of the pronunciation" (HowToStudyKorean, 2024). Similarly, on the language learning app Duolingo, aspirated plosives are described as being "pronounced with a strong puff of air," while tense consonants are characterized as being "pronounced with greater force" (Duolingo, 2024). These descriptions, while aiming to guide learners, are often vague and may not provide sufficient clarity on how to produce the sounds accurately. The notion of producing a sound with "greater force" is particularly problematic because it lacks specificity. Learners may find it difficult to translate this abstract concept into practical articulation techniques. Furthermore, these resources do not address the role of fundamental frequency, or pitch, in the production of these phonemes. Although it might be inferred that a sound produced with "greater force" could also involve a higher pitch, such assumptions are not explicitly stated. This omission is significant because, as shown in Figure 1, native speakers produce tense plosives across a wide range of F0 values,

while aspirated plosives are typically associated with a higher F0. By neglecting to mention F0 as an acoustic cue, resources like Duolingo and HowToStudyKorean.com may inadvertently hinder learners' ability to fully grasp and produce the three-way laryngeal distinction in Korean plosives. As a result, learners may continue to default to a two-way distinction, similar to the results observed in this study.

It is important to recognize that my intention is not to be critical of these valuable and accessible educational resources. These tools are designed to assist learners in navigating the complexities of a new language, and the challenge of conveying such nuanced phonetic distinctions is substantial. The intricate nature of the Korean three-way laryngeal distinction poses a considerable difficulty not only for learners but also for educators who strive to present these concepts in an accessible and comprehensible manner. The current explanations provided by these resources, while helpful, may not fully capture the subtlety required to grasp and produce these sounds accurately. This underscores the need for enhanced instructional strategies and more detailed resources to support learners in mastering the nuances of Korean phonetic categories.

Another significant factor that may contribute to the variation observed among speakers in the results is the amount and quality of input that learners receive in the target L2. In many language learning contexts, students are exposed to the target language for only a few hours each week, which may limit their opportunities to engage with the language comprehensively. This issue is particularly evident in environments where the target language is not widely spoken, leading to infrequent or minimal real-world interaction with native speakers.

According to the SLM-r, learners can potentially re-categorize their initial understanding of L2 sounds into categories that align more closely with those of the L2, as opposed to their L1,

through increased exposure and practice (Flege & Bohn, 2021). This model posits that, over time, with sufficient and repeated input, learners can adjust their perceptual and production systems to accommodate the nuances of the L2 phonetic system (Flege & Bohn, 2021). This principle suggests that learners who receive extensive and high-quality input are more likely to develop a more refined and native-like understanding of L2 phonetic categories (Flege & Bohn, 2021). However, the practical application of this principle is often complex. When the L2 is spoken infrequently in the learner's community or when language instruction is limited to periodic classes, learners may struggle to achieve the level of exposure necessary for significant phonetic reorganization. The challenge of acquiring a new phonetic system becomes even more pronounced when learners do not have consistent opportunities to hear and practice the language in authentic contexts.

Interestingly, the data from NNS11, the participant who demonstrated the most native-like production of the three-way laryngeal distinction, provides some insight into this issue. NNS11 reported having spent an extended period living in Korea, which likely afforded them substantial daily exposure to the language. This immersive experience would have provided ample opportunities to interact with native speakers and hear the L2 in a variety of contexts, potentially facilitating a more accurate reorganization of their phonetic categories. The observation that NNS11 achieved a high degree of native-like distinction supports Flege and Bohn's (2021) assertion that increased L2 input can significantly impact the perception and production of L2 phonemes. Although this conclusion is based on data from a single participant, it highlights the importance of extensive and meaningful exposure in achieving native-like phonetic accuracy. This finding suggests that future research could benefit from exploring the relationship between L2 input and phonetic development more systematically, as will be

discussed in Section 5.2. Such research could provide further insights into how varying levels of input influence the acquisition of complex phonetic distinctions and inform strategies for more effective language instruction.

5.1 Limitations of the Current Study

Unfortunately, this study encountered several limitations that affected its overall effectiveness, and the quality of the data obtained.

The most significant limitation of this study was the relatively small number of non-native participants who were recruited. Due to time constraints and logistical challenges, the final participant pool was considerably smaller than initially anticipated. This limited sample size constrained the range and depth of the data collected, potentially impacting the strength and generalizability of the study's findings. A larger participant pool would have provided a more extensive dataset, facilitating a more comprehensive analysis of the research questions and potentially leading to more reliable and broadly applicable conclusions.

Another significant limitation was the inability to use specialized recording equipment, such as soundproof booths and high-quality microphones, for capturing participants' speech. Instead, recordings were made using smartphones or laptops. While these devices offered convenience and allowed participants to record themselves in accessible locations, they introduced several issues that compromised audio quality. Despite efforts to mitigate these issues by processing the audio files with Audacity sound editing software to reduce background and environmental noise, some recordings remained problematic due to residual noise overlapping with the speech signal. This residual noise further complicated the analysis process.

Additionally, the self-recording method employed in the study introduced several challenges. This approach led to issues with the recorded utterances, such as incorrect word usage, unintelligible pronunciation, and excessively rapid speech. These problems were common in many of the audio recordings and significantly increased the difficulty of accurately analyzing the data. As a result, some data points had to be excluded from analysis. The variability in recording quality due to these issues compromised the overall precision and reliability of the study's findings.

Furthermore, the limitation of not being able to conduct in-person participation also affected the language background questionnaire used in the study. Although the questionnaire was designed to be user-friendly and accessible, its open-ended questions led to a wide range of response types. While some participants provided detailed and informative answers, others offered responses that were vague or lacked sufficient detail. This variability in responses impacted the consistency and depth of the data collected, potentially influencing the comprehensiveness of the insights gained from the questionnaire.

Overall, these limitations underscore the need for larger participant pools, improved recording conditions, and more controlled data collection methods in future research. Addressing these issues will enhance the accuracy and reliability of both auditory analysis and questionnaire responses. This topic will be examined in greater detail in the following section.

5.2 Exploring Future Research Directions

Given the limitations of this study, it is advisable to consider it as a pilot study. This perspective suggests that there are numerous directions for future research to build upon the current findings. By addressing the identified limitations and exploring new avenues, future studies have the potential to yield more robust and conclusive data, thereby advancing our understanding of the research topic.

The most evident recommendation for future research would be to encompass larger participant populations. This expansion could be approached in two significant ways. First, incorporating a broader range of speakers from diverse L2 backgrounds would be beneficial. This study specifically focused on participants from L1 backgrounds with a two-way distinction in their phonemic inventory. Replicating the study with participants from L1 backgrounds that exhibit three-way, or four-way distinctions could provide valuable insights into the support or lack thereof for theoretical frameworks such as the SLM-r and the PAM. Second, increasing the participant pool to include individuals with varying levels of proficiency would enhance the results of the research. Larger groups representing different proficiency levels would not only facilitate the collection and analysis of more comprehensive datasets but also allow for the examination of proficiency effects. This approach would enable a more detailed investigation into how varying levels of language proficiency influence the acquisition and production of L2 phonemes.

This leads to the recommendation of incorporating an additional method for assessing participants' proficiency. The current study relied solely on a self-reported proficiency scale, which, while accessible and convenient for participants, had its limitations. Since participants were rating themselves, there was no way to objectively verify the proficiency levels they

reported. Future research that includes participants with varying levels of proficiency should employ an objective proficiency test alongside a self-reported proficiency test. For instance, to assess participants' speaking proficiency more thoroughly, recorded speech samples could be evaluated by native Korean speakers or Korean language instructors, and the results could be scaled according to standardized criteria. This dual approach would not only provide a more accurate and reliable measure of participants' proficiency levels but also offer a more comprehensive understanding of their language abilities. By combining self-reported data with objective assessments, future studies can ensure a higher degree of accuracy and validity in their findings, ultimately contributing to a deeper understanding of proficiency effects in SLA processes.

Continuing with recommendations regarding the qualitative data collection, if this study were to be conducted again, a more detailed language background questionnaire should be implemented. Additionally, administering the questionnaire verbally by the primary researcher could enhance data quality. These recommendations stem from the need for more analyzable qualitative data from participants. Utilizing a more in-depth questionnaire could potentially lead to richer qualitative data. For instance, in this study, participants were asked to report their weekly Korean usage. A more detailed version of this question could ask participants to break down their weekly reading, writing, speaking, and listening usage separately. This granular approach would provide a clearer picture of the participants' language exposure and use, contributing to a more nuanced analysis.

Moreover, as discussed earlier, a potential factor behind the variation in production among speakers is the way in which they learn Korean. Therefore, incorporating questions that ask participants to detail how they first encountered these novel phonemes could provide valuable

insights. For example, rather than simply asking if they took Korean lessons, the questionnaire could ask them to elaborate on their methods of learning, such as through self-study, formal education, or immersion.

Administering the questionnaire in person (or via Zoom) rather than being completed by the participants on their own would allow the primary researcher the ability to ask follow-up questions in real-time. This approach would enable the researcher to probe deeper into participants' responses, ensuring a more thorough understanding of their language learning experiences. For example, if a participant mentions that they learned Korean through self-study, the researcher could ask them to elaborate on the specific methods and resources they used. This real-time interaction could lead to richer and more detailed responses, enhancing the overall quality of the qualitative data collected.

In addition to administering the questionnaire in person, conducting a post-test interview or questionnaire with participants could reveal valuable insights. This follow-up interaction would allow researchers to gather detailed information about participants' perceptions and production of the Korean three-way distinction. During the post-test interview, participants could be asked to explain how they perceive and produce the distinction between the Korean plosives. Specifically, they could be questioned on whether they consciously use certain phonetic cues such as voice onset time or fundamental frequency to make these distinctions. Understanding whether their production is motivated by conscious effort or unconscious processes could shed light on the underlying mechanisms of L2 phoneme acquisition. Moreover, directly asking participants if they consciously produce the three-way distinction in specific ways may uncover additional phonetic cues they might be using. While VOT and F0 are commonly studied, participants might rely on other cues, such as closure duration or other acoustic features. By exploring these

aspects, researchers can gain a more comprehensive understanding of the strategies L2 learners employ in distinguishing Korean plosives.

6 Conclusion

This study investigated the production of the Korean three-way laryngeal plosive distinction by non-native speakers, with a focus on the use of voice onset time and fundamental frequency as acoustic cues and examined the effects of language proficiency. The findings revealed substantial variation in how non-native speakers produce the Korean three-way distinction. Consistent with previous research (Chang, 2009; Lee, 2022; Seo et al., 2020; Ryu, 2017), most participants either relied on an L1-like two-way distinction or failed to make clear distinctions between the Korean plosives. Notably, only one participant demonstrated a native-like use of both F0 and VOT to achieve the three-way contrast, underscoring the challenge of acquiring accurate L2 phonemic distinctions, especially when the L1 phonetic system lacks direct parallels. In contrast to Ryu (2017), which found a positive relationship between proficiency and phonemic accuracy, this study did not observe a significant impact of proficiency levels on production.

Future research should address the limitations of this study by incorporating a larger and more diverse participant pool, including objective proficiency measures alongside self-reports, and employing detailed language background questionnaires. Expanding the participant base to include various L1 backgrounds and proficiency levels will enhance our understanding of L2 phoneme acquisition. Additionally, exploring different learning contexts and methods could provide valuable insights into effective instructional strategies and refine theoretical frameworks such as SLM-r and PAM.

In conclusion, this study has offered insights into the production of the Korean three-way laryngeal distinction by non-native speakers. While the findings contribute to our understanding of this topic, they also highlight the need for further research to address existing limitations and explore new avenues to improve L2 learning outcomes. Building on these findings and addressing identified gaps has the potential to advance knowledge in second language phoneme acquisition and contribute to more effective language learning strategies.

Appendix A: Word List

Word	Meaning	Word Type	Phrase Structure
간	Liver (n.)	TOPIK Intermediate	간이에요
겅	_____	Pseudoword	겅이에요
곰	Bear (n.)	TOPIK Intermediate	곰이에요
군	Military (n.)	TOPIK Intermediate	군이에요
근	about (det.)	TOPIK Intermediate	근이에요
금	Gold (n.)	TOPIK Intermediate	금이에요
깜	_____	Pseudoword	깜이에요
껌	Gum (n.)	TOPIK Elementary	껌이에요
꽁	_____	Pseudoword	꽁이에요
꿈	Dream (n.)	TOPIK Elementary	꿈이에요
끈	String (n.)	TOPIK Intermediate	끈이에요
끔	_____	Pseudoword	끔이에요
달	Moon (n.)	TOPIK Elementary	달이에요
당	Political Party (n.)	TOPIK Intermediate	당이에요
돈	Money (n.)	TOPIK Elementary	돈이에요
돔	Dome (n.)	TOPIK Intermediate	돔이에요
둘	Two (n.)	TOPIK Elementary	둘이에요
둠	_____	Pseudoword	둠이에요
딸	Daughter (n.)	TOPIK Elementary	딸이에요

땅	Ground/Land (n.)	TOPIK Elementary	땅이에요
떡	Rice cake (n.)	TOPIK Elementary	떡이에요
뜸	_____	Pseudoword	뜸이에요
뜬	Floating (v.)	TOPIK Intermediate	뜬이에요
만	_____	Filler	만이에요
맘	_____	Filler	맘이에요
먼	_____	Filler	먼이에요
믄	_____	Filler	믄이에요
민	_____	Filler	민이에요
반	Half (n.)	TOPIK Elementary	반이에요
발	Foot (n.)	TOPIK Elementary	발이에요
번	_____	Pseudoword	번이에요
분	Person (n.)	TOPIK Elementary	분이에요
븐	_____	Pseudoword	븐이에요
빈	_____	Pseudoword	빈이에요
빨	Red (adj.)	TOPIK Elementary	빨이에요
빰	_____	Pseudoword	빰이에요
빵	Bread (n.)	TOPIK Elementary	빵이에요
뿔	Horn (n.)	TOPIK Elementary	뿔이에요
쁑	_____	Pseudoword	쁑이에요
삥	Round (adv.)	TOPIK Intermediate	삥이에요

산	_____	Filler	산이에요
삼	_____	Filler	삼이에요
선	_____	Filler	선이에요
손	_____	Filler	손이에요
술	_____	Filler	술이에요
슨	_____	Filler	슨이에요
잔	_____	Filler	잔이에요
잠	_____	Filler	잠이에요
준	_____	Filler	준이에요
진	_____	Filler	진이에요
집	_____	Filler	집이에요
칸	Box, partition (n.)	TOPIK Intermediate	칸이에요
캄	_____	Pseudoword	캄이에요
캉	_____	Pseudoword	캉이에요
컨	Room (n.)	TOPIK Elementary	컨이에요
콩	Bean (n.)	TOPIK Elementary	콩이에요
큰	Big (adj.)	TOPIK Elementary	큰이에요
탕	Thick soup (n.)	TOPIK Intermediate	탕이에요
텀	_____	Pseudoword	텀이에요
톤	Tone (n.)	TOPIK Elementary	톤이에요
톰	_____	Pseudoword	돔이에요

튼	Strong, Healthy	TOPIK Elementary	튼이에요
틈	Crack/Gap (n.)	TOPIK Intermediate	틈이에요
팀	Team (n.)	TOPIK Elementary	팀이에요
팍	Out loud (adv.)	TOPIK Intermediate	팍이에요
판	Plate (n.)	TOPIK Elementary	판이에요
팔	Arm (n.)	TOPIK Elementary	팔이에요
팡	_____	Pseudoword	팡이에요
픈	_____	Pseudoword	픈이에요
핀	Pin (n.)	TOPIK Elementary	핀이에요

Appendix B: Questionnaire Questions

Native Speaker Questionnaire

1. What is your age?
2. What is your place of birth?
3. What are the first languages of your parents?
4. What is the language mainly used at your home?
5. Do you speak any other languages other than Korean and English?
    a. If yes, please list the languages(s) and indicate your proficiency level in each (beginner, intermediate, advanced, near native)

Non-Native Speaker Questionnaire

1. What is your age?
2. What is your native/first language?
3. How long have you been learning Korean?
4. Please rate your proficiency level in Korean:

    Reading [Beginner, Intermediate, Advanced, Near-Native]

    Speaking [Beginner, Intermediate, Advanced, Near-Native]

    Listening [Beginner, Intermediate, Advanced, Near-Native]

    Writing [Beginner, Intermediate, Advanced, Near-Native]
5. What motivated or led you to learn Korean?
6. Have you taken formal Korean lessons?
    a. If yes, how long have you received Korean instruction?

7. Please briefly describe your method of learning Korean. (e.g. self-taught, teacher-led group language courses, private tutoring, language apps, etc.)
8. Approximately, how many hours a week do you use Korean?
9. Do you speak any languages other than your native language and Korean?
    a. If yes, please list the language(s) and indicate your proficiency level in each (beginner, intermediate, advanced, near-native).

References

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words*, *167*(224), 233-277.

Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (version 6.4) [Computer program]. Retrieved from http://www.praat.org/

Chang, CB. "The implementation of laryngeal contrast in Korean as a second language." *UC Berkeley PhonLab Annual Report* 5.5 (2009).

Chang, S.-E., & Mandock, K. (2019). A Phonetic Study of Korean Heritage Learners' Production of Korean Word-initial Stops. *Heritage Language Journal*, *16*(3), 273–295. https://doi.org/10.46538/hlj.16.3.1

Cho, T., Jun, S. A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. Journal of phonetics, 30(2), 193-228

Choi, J., Kim, S., & Cho, T. (2020). An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account. *PLoS ONE*, *15*(10 October). https://doi.org/10.1371/journal.pone.0240682

Duolingo (2024). *Duolingo* (Version 7.33.0) [Mobile software app]. https://www.duolingo.com

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.

Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). *Second language speech learning: Theoretical and empirical progress*, *10*(9781108886901.002).

Goto, H. (1971). Auditory perception by normal Japanese of the sounds L and R. *Neuropsychologia, 9*, 317-323.

Kang, K.-H., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *The Journal of the Acoustical Society of America*, *119*(3), 1672–1683. https://doi.org/10.1121/1.2166607

Kang, Y., & Han, S. (2013). Tonogenesis in early Contemporary Seoul Korean: A longitudinal case study. *Lingua*, *134*, 62–74. https://doi.org/10.1016/j.lingua.2013.06.002

Kang, Y., & Nagy, N. (2016). VOT merger in Heritage Korean in Toronto. *Language Variation and Change*, *28*(2), 249–272. https://doi.org/10.1017/S095439451600003X

Kim, M.-R., & Duanmu, S. (2004). “Tense” and “lax” stops in Korean. *Journal of East Asian Languages*, *13*(1), 59–104.

Lee, H., & Jongman, A. (2012). Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association*, *42*(2), 145–169. https://doi.org/10.2307/26351858

Lee, B., Yamaguchi, N., & Fougeron, C. (2022, September). Why is Korean lenis stop difficult to perceive for L2 Korean learners?. In *INTERSPEECH* (pp. 1861-1865).

Lisker, L., & Abramson, A. S. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *WORD*, *20*(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

Lotto, A. J., Sato, M., & Diehl, R. L. (2004). Mapping the task for the second language learner: The case of Japanese acquisition of/r/and/l. *From sound to sense*, *50*(2004), C381-C386.

Lusin, N., Peterson, T., Sulewski, C., & Zafer, R. (2023). *Enrollments in languages other than English in US institutions of higher education, fall 2021*. Modern Language Association of America.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*(5), 331-340.

Narayan, C. R. (2023). Speaking rate, oro-laryngeal timing, and place of articulation effects on burst amplitude: Evidence From English and Tamil. *Language and Speech*, *66*(4), 851-869.

Ryu, N. Y. (2017). Perception of Korean contrasts by Mandarin learners: The role of L2 proficiency. Toronto Working Papers in Linguistics, 38.

Seo, Y., Dmitrieva, O., & Cuza, A. (2022). Crosslinguistic Influence in the Discrimination of Korean Stop Contrast by Heritage Speakers and Second Language Learners. *Languages*, *7*(1). https://doi.org/10.3390/languages7010006

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, *23*(2), 287-308.

Tranter, N. (2012). The languages of Japan and Korea. Routledge.

Wang, H. S., & Pyun, D. O. (2020). Hallyu and Korean language learning: Gender and ethnicity factors. *The Korean Language in America*, *24*(2), 30-59.