
THE JOURNAL OF THE UNDERGRADUATE LINGUISTICS ASSOCIATION OF BRITAIN

ISSN 2754-0820

# ACQUISITION OF PLOSIVE PERCEPTION IN KOREAN L2 LEARNERS

Emma Laird

University of Glasgow
emmalaird123@yahoo.co.uk

**Keywords:** L2 acquisition; Korean; English; perception; plosives; stop consonants

**JoULAB URL:** https://www.ulab.org.uk/journal/volumes/1/issues/2/articles/3

**Citation:** Laird, E. (2022). Acquisition of Plosive Perception in Korean L2 Learners. *Journal of The Undergraduate Linguistics Association of Britain*, 1(2), pp. 66-80.

Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

# Acquisition of Plosive Perception in Korean L2 Learners

**Emma Laird**
*University of Glasgow*

**Abstract.** Korean as an L2 has increased in popularity over the last decade, and with it has opened up a great interest into studying the linguistic nature of Korean L2 acquisition. The goal of this research was to explore the ability of Korean L2 learners to acquire the ability to perceive between Korean stop consonants, in both word-initial and intervocalic positions. Best's Perceptual Assimilation Model and Flege's Speech Learning Model are utilised to provide the theoretical framework for discussing naïve listener perception of non-native sound and L2 learner perception of L2 sounds respectively. While English utilises the voiced vs. voiceless contrast, Korean has a three-way distinction between stops. An AX discrimination task was performed amongst 24 participants making up three groups — naïve listeners of Korean, English L1-Korean L2 learners, and Korean L1 speakers. It was hypothesised that the accuracy scores of participants in correctly discriminating between plosives would be lowest in naïve listeners, and highest in native speakers, and that participants would have higher accuracy ratings for the intervocalic tokens as opposed to the word-initial tokens. Surprisingly, the results showed that English L1-Korean L2 learners had the lowest accuracy ratings out of all three groups, as opposed to the prediction that naïve listeners would have the lowest accuracy ratings due to their lack of interaction with Korean phonology. However, there were a number of methodological issues (discussed in the latter Sections of this paper) that may explain the discrepancies between the predicted results and the results attained.

**Plain English Abstract.** This research looked at how well people can hear the difference between Korean plosive consonants in word-initial and intervocalic positions. There are different theories in place for how people learn to hear the difference between sounds in foreign languages, and two of these theories (Best's Perceptual Assimilation Model and Flege's Speech Learning Model) were used to guide the predictions for this research. English primarily uses voicing to show the difference between stop consonants, but it is more complex in Korean, which raises the question: how does a native English speaker learn how to hear the difference between three very similar consonants in Korean, when they are only hearing the difference between two in English? Twenty-four participants took part in an AX discrimination task across three groups: native English speakers who had never heard Korean before (naïve listeners); native English speakers who were learning Korean; and native Korean speakers. It was predicted that the naïve listeners would find hearing the difference between these Korean consonants the most difficult, the Korean learners would perform slightly better in the task, and the native speakers would have the greatest accuracy in the task. It was also predicted that everyone would perform better when the consonants appeared intervocalically, as opposed to word-initially. Surprisingly, the Korean learners had the least accuracy out of all three groups. However, there were some issues with the way the experiment was designed and set up that may explain why the results were different from the predictions.

**Keywords: L2 acquisition; Korean; English; perception; plosives; stop consonants**

## 1 Introduction

Various perceptual models have been created to account for non-native and L2 perception capabilities. Best's Perceptual Assimilation Model (Best, 1995; Best & Tyler, 2007) (hereafter PAM) and Flege's Speech Learning Model (Flege, 1995) (hereafter SLM) are utilised in this research to provide a theoretical framework on which to frame the predictions and discussions throughout this research.

Korean is particularly of interest when approaching L2 perception due to the 3-way laryngeal distinction present in their plosives and affricates. While English utilises (primarily) the voiced vs. voiceless distinction in stop perception, Korean word-initial stops are all voiceless. Due to this, other acoustic cues are required in order to segregate the three plosive types — lax, tense, and aspirate. Outwith the laryngeal configuration utilised to describe the phonation types of plosives utilised by Kang


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

and Lee (2002), VOT and the F0 of the vowel following the plosive are the main discriminatory factors utilised in perception.

While previous research has investigated the perception of Korean plosives, the majority of academic focus has been on the word-initial plosives in Korean (see Kang & Lee, 2002; Holliday, 2018), despite there being distinct phonetic characteristics amongst plosives depending on their position within a word. This research will compare the ability of participants to accurately discriminate between plosives in both word-initial and intervocalic settings.

# 2 Phonology of Korean Stop Consonants

There has been significant research into the three-way laryngeal distinction, from both articulatory and acoustic perspectives (Hardcastle, 1973; Cho et al., 2002). The table below provides a comprehensive overview of the terms utilised to describe the three-way laryngeal distinction, as well as the transcriptions utilised; the terms captured below are found in Kim (1997), Silva (1991), Cho et al. (2002), Kim and Lotto (2002), Kim and Duanmu (2004), and Shin et al. (2013). The *hangul* (Korean alphabet) examples and romanisations are presented in order of articulation; that is, bilabial, followed by alveolar and then velar.

**Table 1:** *Korean consonant labelling terms with examples and transcriptions*

| **Terminology** | **Variants of Terminology** | **Examples (*Hangul*)** | **Romanisation (Revised Romanisation)** | **IPA (word-initial)** |
|---|---|---|---|---|
| Lax | Lenis, Plain, Unaspirated, Slightly Aspirated, Occlusive | ㅂ, ㄷ, ㄱ | b, d, g | [p] [t] [k] |
| Tense | Fortis, Hard, Glottalized, Reinforced | ㅃ, ㄸ, ㄲ | bb, dd, gg | [p̬] [t̬] [k̬]<br>[p’] [t’] [k’]<br>[p*] [t*] [k*]<br>[P] [T] [K] |
| Aspirate | Voiceless, Aspirated, Heavily Aspirated | ㅍ, ㅌ, ㅋ | p, t, k | [$p^{h]}$] [$t^{h}$] [$k^{h}$] |

In word-initial position, Korean has no voiced plosives. The lax, tense, and aspirate plosives are all voiceless, and thus voicing is not an acoustic cue that learners can utilise to categorise their perception of these plosives. Instead, listeners primarily utilise VOT and F0 as the acoustic cues that indicate which plosive is being perceived. Originally, there existed both a three-way VOT distinction *and* a three-way F0 distinction of the laryngeal contrasts in standard South Korean plosives. However, a tonogenesis-like change is emerging with a shift to a two-way VOT distinction; in younger Seoul-dialect speakers (particularly female speakers), the VOT values between lax and aspirate plosives are levelling out, while the F0 distinction between plosive types is becoming even more defined (Kang, 2014). The newly-observed two-way VOT distinction is now emerging due to the overlap in VOT values for lax and aspirated plosives; this was observed in the speech data utilised in this study, with word-initial lax and aspirate plosives only having a difference of a few milliseconds in VOT, while F0 values maintained an audible distinction with a difference of over 100Hz (see Table 2 in Section 5.4 for these figures).

This was also further established by Lee and Jongman (2018), who have observed younger Seoul speakers favouring F0 as an acoustic cue, while both older and young South Kyungsang speakers still utilise VOT and F0 as perceptual cues (although the young South Kyungsang speakers did place greater emphasis on F0 as a cue than their older counterparts). This is not to say that VOT and F0 are the only acoustic cues utilised to aid perception of word-initial Korean plosives, however; phonation type (Shin et al., 2013), H1–H2 of the following vowel (Holliday & Kong, 2011), aerodynamic mechanisms and supralaryngeal phonation between Seoul and Cheju speakers (Cho et al., 2002), and phonemic vowel length contrast in the case of the Chonnam dialect (Choi, 2002) have also been observed as acoustic cues utilised for plosive perception.

Comparative to the word-initial plosives, there exists a greater number of acoustic cues, and variation amongst these cues, that can aid perception of the intervocalic plosives. While English learners of Korean typically cannot rely on the voiced/voiceless distinction for word-initial plosives that exist in their L1, they can utilise this distinction for the intervocalic stops. As well as this, Table 2 in Section 5.4 provides the mean VOT and F0 values for the word-initial and intervocalic tokens recorded for this research — one can observe the relatively similar VOT values for word-initial lax and aspirate consonants, in contrast to the very distinct F0, hold period and VOT values for the intervocalic plosives. This is further evidence of the tonogenesis-like sound change happening in Seoul Korean, where VOT values for lax and aspirate word-initial plosives are levelling, and F0 is emerging as a distinct perceptual factor (Kang, 2014).

Word-final plosives were not included in this study, as a plosive in such a position is reduced to its unreleased form, signified with the diacritic [p̚]. This also occurs with other types of consonants — for example, the sibilant /s/ is reduced to [t̚], i.e., to the same place of articulation.

# 3 L2 Acquisition

L2 perception models attempt to delineate the perception, acquisition, and organisation of non-native or L2 phonology in relation to the speaker's L1, particularly in relation to the comparability of sounds between the two languages (Best & Tyler, 2007).

Best's Perceptual Assimilation Model (Best, 1995) theorises that naïve speakers categorise non-native sounds in relation to their similarity (or lack of) to phones in their native language. Study of non-native perception has established that naïve listeners have difficulty in both categorising and discriminating phonemes in non-native languages, particularly when the contrasts to be made for discrimination purposes do not exist in their L1. This is relational to the naïve listener's own L1, with non-native stimuli less similar to the listener's L1 phonology being easier to discriminate or categorise as it does not overlap pre-existing phonological categories the naïve listener has from their L1. Applying this to the following perception experiment, it is expected that discrimination ability is poor in naïve participants as they have not developed the perceptual ability to discriminate the Korean plosives due to their perceived similarity to previously established L1 (in this case, English) categories.

Flege's (1995) Speech Learning Model (SLM) postulates that L2 learners can establish new phonetic categories for L2 sounds if the said sound differs phonetically from the closest L1 sound, and if this sound is discernible as an L2 learner. However, as Korean plosives are relatively similar to English plosives (as opposed to a language that utilises, for example, clicks) it can be assumed based on SLM that new phonetic categories would not be created. However, SLM also says that L2 learners have the ability to create long-term memory categories concerning the identification of 'language-specific aspects' (Flege, 1995). If this is taken as true, L2 Korean learners would be able to create specific phonetic categories to accommodate the perception of varying VOT and/or F0 values of Korean plosives.


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

# 4 Research Questions and Hypotheses

The following research questions and hypotheses were collated based on the research outlined in previous Sections.

> **$RQ_1$** – How does perceptual ability vary between naïve listeners of Korean, English L1-Korean L2 learners and Korean L1 speakers?
> **$H_1$** – Naïve listeners will show the greatest difficulty in distinguishing between word-initial plosives in natural (non-manipulated) tokens. Learners may have difficulty but not to the extent of naïve listeners.
> **$H_2$** – Naïve listeners will show the greatest inability in distinguishing between intervocalic plosives in natural (non-manipulated) tokens. Learners may have difficulty but not to the extent of naïve listeners. However, these results will not be as extreme as those for the word-initial discrimination task.
>
> **$RQ_2$** – What acoustic cues do learners favour to aid perception, focussing on VOT and F0 in this research?
> **$H_3$** – Naïve, learners, and native listeners will have difficulty in distinguishing between word-initial plosives when F0 has been manipulated. F0 manipulation will have the greatest effect on correct perception (as opposed to VOT).
> **$H_4$** – Naïve, learners, and native listeners may have difficulty in distinguishing between manipulated intervocalic tokens (such as between intervocalic lax and tense with VOT manipulation), but not to the extent of the word-initial manipulated discrimination tasks.

# 5 Methodology

## 5.1 Participants

All subjects were recruited through snowball sampling and online calls for participants shared by email and on various Korean Culture/Language online groups based at Scottish universities.

The speaker who volunteered to record their speech for the experimental stimuli was a 21-year-old South Korean female. She had spent over half her life in the Gangnam area of Seoul, the capital city of South Korea, and has had significant exposure and use of the standard Seoul Korean dialect.

Twenty-four participants took part in the perception experiment: seven native Korean speakers (3M/4F, mean age 29.1, S.D. 5.3), nine English L1-Korean L2 learners (0M/9F, mean age 22, S.D. 3.5) and eight English L1 with limited-to-no exposure to the Korean language (4M/4F, mean age 23.6, S.D. 2.8). All participants were over 18 years old to comply with ethical guidance, and no participants reported having hearing difficulties.

## 5.2 Materials

Materials consisted of a world list comprised of 18 CVC and VCV words (3 places of articulation × 3 laryngeal contrasts × 2 word positions) as shown in Appendices One and Two. The CVC words were chosen as they all share the same VC context /an/ and are all words that exist in Korean. The VCV words /a_a/, where _ is replaced by the chosen plosive, were chosen as this context produced the greatest number of real Korean words. A small number of VCV words were pseudowords, but this was

unavoidable. This allowed for consistency between the CVC and VCV words as they share the same vowel context.

## 5.3 Recording

Recordings were made in a sound-proof booth using a Beyerdynamic Opus 55.18 MK II neck-worn microphone connected to an ART USB Mix Three Channel pre-amplifier. Recordings were made stereophonically directly into Praat 6.1.09 (Boersma & Weenik, 2020) and were recording at a 44.1kHz sampling rate. The recordings were then converted to monoaural in Praat. Word list recordings were then cut into individual tokens, present 108 tokens. Naturally-produced tokens were favoured over utilising artificially generated tokens for this experiment, as Thomas (2002) observed that participants perform better in speech perception experiments when natural tokens are utilised.

The participant was given the opportunity to read the word list provided beforehand to acquaint herself with the pseudowords (see Appendices One and Two). She was then instructed to read the word list a total of 6 times — three recordings prompted to be natural speech, and three prompted to have the speaker level their tone. While the tone-levelling attempt was made in order to observe whether the participant could produce tokens with F0 naturalised between the varying laryngeal contrasts, it was unsuccessful, and these tokens were later excluded from the experiment data.

## 5.4 Analysis of Recordings

VOT was measured for both word-initial and intervocalic tokens. The hold period between the end of the first vowel and the burst of the plosive was measured; this was only necessary for the intervocalic tokens. F0 was measured for all tokens and was measured at the first glottal pulse as seen on the spectrogram in Praat. The mean measurements for VOT, F0 and, where applicable, hold period length are shown below:

**Table 2:** *Mean acoustic measurements*

| Consonant Position | Consonant Type | Hold Period (ms) | VOT (ms) | F0 (Hz) at onset of following vowel |
|---|---|---|---|---|
| **Word-initial** | Lax | - | 103 | 161 |
| | Tense | - | 15 | 251 |
| | Aspirate | - | 99 | 266 |
| **Intervocalic** | Lax | 93 | 17 | 207 |
| | Tense | 271 | 15 | 230 |
| | Aspirate | 233 | 58 | 245 |

## 5.5 Manipulation of Recordings

Manipulated tokens were created to allow for the use of a token that had one property (either VOT or F0) consistent with another type of token, e.g., a lax base token with the VOT of a tense token. This would allow for the observation of the manipulation of particular acoustic cues on perceptual ability in


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

comparison to perceptual ability of natural tokens. Tokens to be manipulated were chosen based on how close they lay to the average VOT and F0 figures for tokens collected for this experiment.

VOT was manipulated utilising the 'cross-splicing' method; that is, copying the VOT of Token A, and deleting the original VOT of Token B, and replacing the original VOT of Token B with the new VOT of Token A. When cutting out the VOT, the start and end of each selected section was moved to the nearest zero crossing.

F0 was manipulated following the directions in Will Styler's Praat manual (Styler, 2020). Where Token A was shorter in duration than Token B, the duration difference between the two files was calculated and added into the Token A as silence, so the pitch tiers would match the onset of voicing. Once the manipulation was complete, this silence was removed, and the token reverted to its original duration. Where Token A to be manipulated was longer than Token B, Token B had the difference in duration calculated and added in as silence, as above. For the intervocalic tokens, where factors such as the hold period had to be taken into account, the difference between the hold period *plus* the VOT was calculated between Token A and B. In whatever token this was shortest, the difference was then added in as silence, as above, and the same manipulation technique was followed as in the Praat manual mentioned at the beginning of this paragraph (Styler, 2020). However, as in the VCV context the vowel preceding the consonant varies in length dependant on the consonant used, the manipulated pitch tiers were manually adjusted to fit the length of the vowel where required — the frequency of each pitch pulse to be manipulated was locked on its vertical axis, so the frequency would not be altered, but the pulse could be moved horizontally to be aligned within the vowel duration.

## 5.6 Experiment Procedure

Participants were provided with an information sheet detailing the experiment procedure and their data rights. They were then given a consent from and questionnaires (see Appendices Three–Five). All participants were asked to fill out a demographic and linguistic questionnaire before the experiment began — demographic information was included in Section 5.1 (see Appendices Three–Five for questionnaires).

Korean L1 participants were asked about the locations where they had previously grown up and lived in South Korea, in order to allow for any discussion of perceptual variation that may have arisen with participants being exposed to non-standard varieties of Korean. Previous acoustic studies (Choi, 2002; Holliday & Kong, 2011) have shown that cues utilised in discriminating Korean plosives are subject to dialectal variation. However, the majority of participants came from the region in and around Seoul, where the standard form of Korean is spoken. As well as this, as the standard Seoul dialect is utilised in most forms of standard media, participants will have had significant exposure to perception of this dialect.

The English L1-Korean L2 learners had a number of questions to answer pertaining to their individual experiences of learning Korean (see Appendix Four). A simplified form of the Interagency Language Roundtable scale was utilised (excluding the '0' score, as it was assumed that learners would at least be above the 'no proficiency' rating), and the participants rated themselves from 1–5, 1 being 'beginner' and 5 being 'fluent' (Interagency Language Roundtable, n.d.). No participant rated themselves above '3'. Four participants rated their proficiency as '1', three participants are '2' and two participants as '3'.

The Korean L2 learners' learning duration varied from five months to four years, however one participant did not provide a numerical answer on their questionnaire (they answered that they studied 'a little bit, on and off'). Learning modes included being self-taught, formal university tuition, and language immersion in South Korea, with a number of participants selecting a combination of learning

modes — for example, one participant chose a combination of self-taught, formal (in-person) language classes, and language immersion.

The experiment was run on a MacBook Air running OpenSesame ver. 3.2.8 (Mathôt et al., 2011). The experiment consisted of a six-part AX discrimination task, with an additional practice section at the start of the experiment to allow participants to acclimatise themselves to both the controls and format of the experiment. The keys used to indicate responses from participants (the 'a' key and the 'l' key) were identified with red and green stickers respectively, to allow participants to have a visual cue for their choice. In the event of a participant being colour-blind, the red sticker had a cross to indicate the participant perceived the tokens as 'different' and the green sticker had a 'tick' to indicate their perception of the tokens as the same. The experiment was divided into six sections as follows:

**Table 3:** *Experiment sections*

| Section Number | Token Type |
|---|---|
| 1 | Word-initial (CVC) Natural Tokens |
| 2 | Word-initial VOT Manipulated Tokens |
| 3 | Word-initial F0 Manipulated Tokens |
| 4 | Intervocalic (VCV) Natural Tokens |
| 5 | Intervocalic VOT Manipulated Tokens |
| 6 | Intervocalic F0 Manipulated Tokens |

Participants were instructed to listen to each pair of words through a Sennheiser GSP 302 noise-isolating headset and indicate whether the pair of tokens they heard sounded the same or different by pressing one of two stickered keys mentioned above. Each pair of tokens was played once, resulting in 126 responses. Within each section, all pairs of tokens were randomised. The OpenSesame program recorded participants' response, reaction time, number of correct answers, overall accuracy score, as well as the details for each response such as whether the tokens for that particular response were natural or manipulated etc.

## 5.7 Statistical Methodology

The following statistical analyses were run in R Ver 1.2.5033. (R Core Team, 2017). Due to the unbalanced nature of the subject groups in this study, Levene's Test was performed on the accuracy ratings extracted from each participant's dataset (Field et al., 2012). All data visualisation was created using ggplot2 (Wickham, 2012).

# 6 Results

As previously mentioned, it was predicted that naïve listeners would have the greatest difficulty with the discrimination task due to the postulates set out in the perception models in Section 3. The table below contains a collation of the average results for each group of participants, organised by token type and word position of the plosive. The accuracy rating (%) is the number of correct responses divided by the total response count, then multiplied by 100 to achieve a percentage score.


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

**Table 4:** *Accuracy ratings (%) from AX discrimination task*

| | Whole Experiment Accuracy (%) | | | Natural Tokens Only Accuracy (%) | | | VOT Manipulated Only Accuracy (%) | | | F0 Manipulated Only Accuracy (%) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | All | CVC | VCV | All | CVC | VCV | All | CVC | VCV | All | CVC | VCV |
| **Native** | 77.0 | 74.1 | 79.8 | 93.9 | 90.4 | 96.3 | 63.9 | 63.3 | 64.3 | 63.1 | 60.3 | 65.9 |
| **L2 Learners** | 55.5 | 47.4 | 63.5 | 65.2 | 57.6 | 72.8 | 39.5 | 22.8 | 56.2 | 60.2 | 63.6 | 56.8 |
| **Naïve Listeners** | 67.7 | 65.5 | 69.8 | 78.9 | 75.9 | 81.9 | 52.1 | 43.1 | 61.1 | 66.0 | 71.5 | 60.4 |

Looking at the results for the whole experiment accuracy for the participants, a surprising result emerges. The Korean L1 participants, as expected, have the highest accuracy ratings (averaging 77% over the entire experiment), but naïve listeners average at 67.7% while Korean L2 learners average at the lower score of 55.5%. As predicted, participants across all groups achieved higher accuracy ratings when discriminating intervocalic tokens as opposed to word-initial plosives. Please note that the y-axis in Figure 1 does not start at 0.

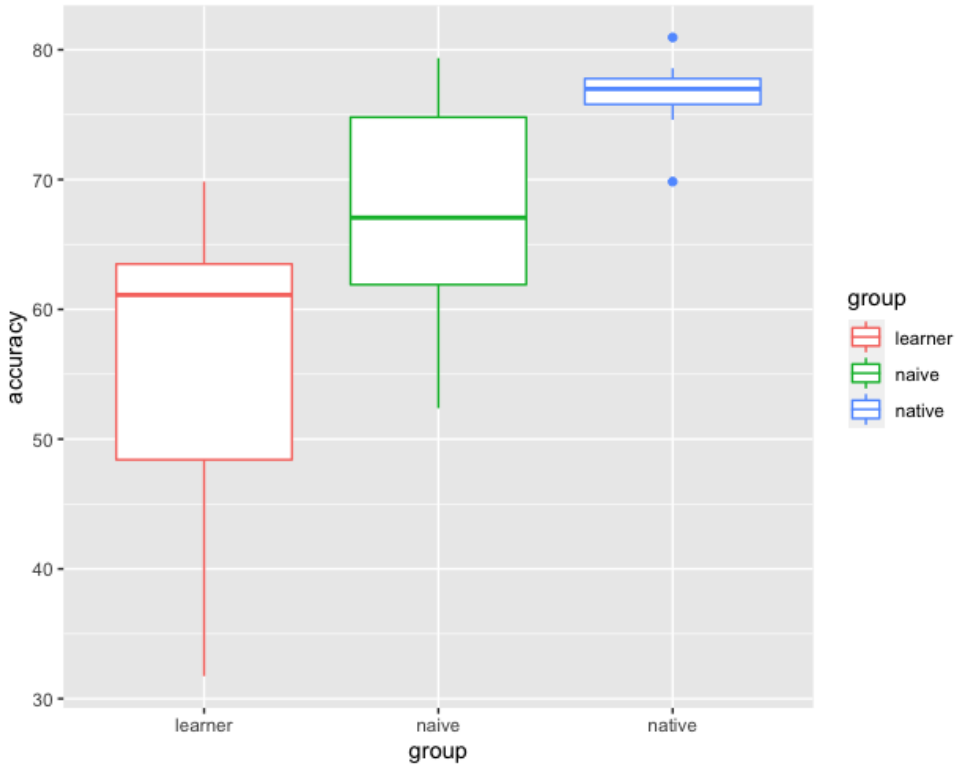


**Figure 1:** *Boxplot of accuracy ratings (%).*

As the participant groups had unbalanced numbers, Levene's Test was performed on the accuracy ratings extracted from each participant's dataset. Assuming a 95% confidence interval, regarding the accuracy percentages in the discrimination task, the variances were similar, $F(2, 21) = 3.47$, $p = 0.1689$. As the homogeneity of variance was met, a one-way independent ANOVA was then ran and it was

discovered that there was a significant effect of participant type on accuracy scores, $F(2, 21) = 3.47$, $p < 0.01$, $p = 0.00147$.

# 7 Discussion

## 7.1 Evaluation of Main Predictions

This study set out to observe the perceptual ability of English L1-Korean L2 learners in comparison to naïve and native participants when discriminating Korean word-initial and intervocalic plosives. By comparing these learners to native speakers and naïve listeners, it was hoped that the progression from naïve to learner would highlight the acquisition of this perceptual ability, with the native group acting as a control group to observe normal levels of perceptual ability in Korean. As well as this, the acoustic factors of VOT and F0 were to be manipulated to observe their effect on perceptual accuracy. While it was expected that Korean L2 learners would perform better than naïve listeners, this was not the case — Korean L2 learners had overall lower accuracy ratings than all other groups.

The earlier discussion of L2 perception models guided the predictions that naïve listeners would face the greatest difficulty in discrimination between tokens, and that L2 learners would perform better. As seen by the results in Section 6, this was not the case: naïve listeners had a higher average accuracy score than L2 learners, as well as less variation with their results. L2 learners overall had the lowest average accuracy scores, as well as having significant variation with their accuracy results, as seen in Figure 1.

Flege's (1995) Speech Learning Model (SLM) guided the predictions that L2 learners would have developed the categorisation of 'language-specific aspects' to utilise when discriminating acoustic cues such as VOT and F0. The accuracy ratings however showed that this was not the case; L2 learners had the weakest perceptual ability out all of participants. Flege discussed that when the age of the learner is great, i.e., the learner is older (particularly in this study as all learners were adults), the greater difficulty the learner has in creating phonetic categories. It may be that, with the participants in this research, their categories are still in development, particularly as some participants have only been studying for a very short period.

Due to the unexpected but nevertheless interesting results, it can be stated that $H_1$ and $H_2$ were not supported, as the results did not follow the accuracy 'slope' predicted with naïve listeners at the bottom and native listeners at the top. This study has shown that adult Korean L2 learners have the greatest difficulty in discriminating between both word-initial and intervocalic plosives.

$H_3$ stated that manipulation of F0 would have a greater adverse effect on perceptual ability when discriminating word-initial plosives than manipulation of VOT. Korean L1 participants performed relatively similarly between the two manipulation categories (with only a 3% difference in accuracy ratings, seen in Table 4 in Section 6.2). Korean L2 learners and naïve listeners both performed better when F0 was manipulated than when VOT was manipulated, most likely as VOT is utilised more often in English perception, as opposed to utilising the F0 of the proceeding vowel as is used in Korean.

Similarly, $H_4$ predicted that manipulation of F0 would have a greater adverse effect on perceptual ability when discriminating between tokens in an intervocalic context. All participants performed better in the intervocalic context than in the word-initial context, and again performed better with F0 manipulated tokens than when faced with VOT manipulated tokens. Overall, all participants performed better in the intervocalic section of the experiment than in the word-initial section, most likely due to the extra acoustic features present in an intervocalic context that can influence perception. That is, the vowel length preceding the plosive, and the hold period between the initial vowel and the plosive are both perceptual cues that do not exist in the word-initial context. Thus, $H_4$ was partially supported, as


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

participants performed better when discriminating between intervocalic plosives as opposed to word-initial plosives, but they did not follow the pattern predicted that they would have lower accuracy scores when discriminating between F0 manipulated tokens as opposed to VOT manipulated tokens.

## 7.2 Methodological Issues

A number of methodological limitations may have affected these results.

While the original plan for this research set out to have three equal groups of 10 participants, this was not reflected in the participant recruitment numbers. Due to time constraints and difficulty in recruiting participants, these numbers were not reached. To compensate for this Levene's Test was utilised during the statistical analyses as it provides the ability to analyse groups of varying sample size.

A number of the naïve listener participants were recruited from a mass email sent to English Language and Linguistics students at the University of Glasgow. As some of these students come from a linguistic or specifically phonetic background, they may have been influenced and biased by the more finely detailed perception required from their studies and thus may have been attuned to noticing discrepancies between tokens. As well as this, a few naïve listeners were reported to having listened to/watch Korean pop culture, and this may have provided them with just enough experience of Korean to be able to perceive the discrimination analysed in this study.

The English L1-Korean L2 learners came from a variety of language learning backgrounds, inviting a lot of variation into their participant group. The methods and duration of learning varied a lot between participants, and this lack of consistency may explain the wide variety of accuracy results observed in Table 4 in Section 6.2.

While VOT and F0 are referred to as the primary acoustic cues utilised for Korean plosive perception, they are by no means the sole cues utilised. Other cues, such as H1-H2 of the following vowel (Holliday & Kong, 2011), vowel length, hold period preceding the burst in intervocalic tokens, voicing presence in intervocalic plosives etc., may also have influenced perception and the discrimination of plosives in this study.

# 8 Conclusion

The collation and discussion of results has shown through this particular piece some surprising yet nevertheless interesting observations of perception capabilities between varying participant groups. While the research carried out in Sections 2 and 3 guided the predictions that naïve listeners would have the lowest accuracy when undertaking the discrimination task, this was not the case. Native listeners performed best in terms of accuracy of perception, which is to be expected, but Korean L2 learners had the lowest accuracy ratings out of all participants with both word-initial and intervocalic stimuli. While the predicted hypotheses were not supported by these results, the results open the doors to further examination of adult L2 perception — if these results are indicative of a large issue in adult L2 acquisition, this is most certainly an obstacle for L2 learners that could be further explored. By placing the L2 learners' results in comparison to both naïve and native listeners, it highlights an intriguing pattern in perceptual ability development.

# 9 References

Best, C. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues* (pp. 171–204). New York: Timonium.

Best, C. & Tyler, M. D. (2007). Nonnative and second language speech perception. In O-S Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning* (pp. 13–34). Amsterdam: John Benjamins Publishing.

Boersma, P., & Weenik, D. (2020). Praat: doing phonetics by computer [Computer Program]. Version 6.1.09. Available at: <http://www.praat.org/>.

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

Choi, H. (2002). Acoustic Cues for the Korean Stop Contrast – Dialectal Variation. *ZAS Papers in Linguistics*, 28, 1–12.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London: SAGE Publications Ltd.

Flege, J. E. (1995). Second Language Speech Learning: Theory, Findings and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues* (pp. 233–277). Baltimore: York Press.

Hardcastle, W. J. (1973). Some Observations on the *tense-lax* distinction in initial stops in Korean. *Journal of Phonetics*, 1(3), 263–272.

Holliday, J. (2018). The Perception and Production of Word-Initial Korean Stops by Native Speakers of Japanese. *Language and Speech*, 62(3), 494–508.

Holliday, J., & Kong, E. J. (2011). Dialectal Variation in the Acoustic Correlates of Korean Stops. *International Congress of Phonetic Science*, 17, 878–881.

Interagency Language Roundtable (n.d.). Language Skill Description. Available at: <https://govtilr.org/Skills/ILRscale2.htm>.

Kang, S., & Lee, S. (2002). Acoustic Properties of Word-Initial Korean Stops in Speech Perception. *Korean Association of Language Sciences: Journal of Language Science*, 9(2), 1–20.

Kang, Y. (2014). Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 76–90.

Kim, C. W. (1997). The structure of phonological units in Han'gul. *The Korean alphabet*, 145–160.

Kim, M., & Duanmu, S. (2004). "Tense" and "Lax" stops in Korean. *Journal of East Asian Linguistics*, 13(1), 59–104.

Kim, M. C., & Lotto, A. J. (2002). An Investigation of Acoustic Characteristics of Korean Stop Consonants Produced by Non-Heritage Learners. *The Korean Language in America*, 7, 177–187.

Lee, H., & Jongman, A. (2018). Effects of Sound Change on the Weighting of Acoustic Cues to the Three-Way Laryngeal Stop Contrast in Korean: Diachronic and Dialect Comparisons. *Language and Speech*, 62(3), 509–530.

Mathôt, S., Schrejj, D., & Theeuwes, J. (2011). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behaviour Research Methods*, 44(2), 314–324.

R Core Team. (2020). R: A language and environment for statistical computing. Available at: <https://www.R-project.org/>.

Shin, J., Kiaer, J., & Cha, J. (2013). *The Sounds of Korean*. Cambridge: Cambridge University Press.

Silva, D. J. (1991). A phonetically based analysis of [voice] and [fortis] in Korean. *Japanese/Korean Linguistics*, 2, 164–174.


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

Styler, W. (2020). Using Praat for Linguistic Research. Available at: <http://wstyler.ucsd.edu/praat/UsingPraatforLinguisticResearchLatest.pdf>.

Thomas, E. R. (2002). Sociophonetic approaches of speech perception experiments. *American Speech*, 77(2), 115–147.

Wickham, H. (2012). ggplot2: Elegant Graphics for Data Analysis. Available at: <https://ggplot2.tidyverse.org>.

# 10 Appendices

## 10.1 Appendix One: Word List with Transcriptions and Translations

Word-initial words adapted from Shin et al. (2013). Words with no translation are nonsense words.

**Table 5:** *Word list for recording session*

| **Word (*hangul*)** | **Transcription** | **Romanisation (RR)** | **Translation(s)** |
|---|---|---|---|
| 반 | [pan] | ban | Half, class |
| 빤 | [p*an] | bban | To suck/wash - adnominal |
| 판 | [p$^{h}$an] | pan | Board |
| 단 | [tan] | dan | Just, but, sweet, column, gear |
| 딴 | [t*an] | ddan | To pick - adnominal |
| 탄 | [t$^{h}$an] | tan | To ride - adnominal |
| 간 | [kan] | gan | Liver |
| 깐 | [k*an] | ggan | Estimation |
| 칸 | [k$^{h}$an] | kan | Box, blank |
| | | | |
| 아바 | [aba] | aba | - |
| 아빠 | [ap*a] | abba | Dad, daddy (inf.) |
| 아파 | [ap$^{h}$a] | apa | It hurts! (from아프다 – to be hurt, sick, in pain) |
| 아다 | [ada] | ada | I know (inf., from알다) |
| 아따 | [at*a] | adda | Well, Ey, Oh! (exclamation) |
| 아타 | [at$^{h}$a] | ata | - |
| 아가 | [aga] | aga | Baby |
| 아까 | [ak*a] | agga | Earlier |
| 아카 | [ak$^{h}$a] | aka | - |

## 10.2 Appendix Two: Word List as Presented to Participant

(1) 반

(2) 딴

(3) 아카

(4) 판

(5) 아따

(6) 아파

(7) 단

(8) 아가

(9) 탄

(10) 아다

(11) 간

(12) 아빠

(13) 깐

(14) 아바

(15) 칸

(16) 빤

(17) 아까

(18) 아타

## 10.3 Appendix Three: Participant Questionnaire — Korean Native Speakers

Age:

Gender: Male / Female / Other / Prefer not to say

Where exactly in Korea did you live? How long did you live there?
*For example:*
*Namyangju-si, Gyeonggi (from birth until age 17)*
*Gwanak-gu, Seoul (from age 17 until age 21)*

Do you speak any other language(s)?

Do you have any hearing difficulties?


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022

Thank you for your time and support!

## 10.4 Appendix Four: Participant Questionnaire — English L1-Korean L2 Learners

Age:

Gender: Male / Female / Other / Prefer not to say

Where are you from (region of country)?

How long have you been learning Korean?

How do you/did you learn Korean? (Select the answer closest to your own experience):

1. Self-taught
2. 1-to-1 tuition
3. Online classes
4. Formal language classes (such as at school/university)
5. Immersion
6. Other (please write below)

How would you rate your Korean proficiency on a scale of 1-5, 1 being an absolute beginner and 5 being fluent? (Circle your answer)

1 2 3 4 5

Do you have any hearing difficulties?

## 10.5 Appendix Five: Participant Questionnaire — English L1 Korean Naïve Participants

Age:

Gender: Male / Female / Other / Prefer not to say

Where are you from (region of country)?
*e.g., Clydebank, Glasgow*

Is English your first language?

Do you speak any other language(s)?

Do you watch/listen to/read any Korean pop culture/literature?
e.g., TV Shows, music etc.

Do you have any hearing difficulties?

# About the Author

Emma Laird graduated from the University of Glasgow in 2020 with an MA in English Language and Linguistics, going on to complete an MSc in Speech, Language and Sociolinguistics in September 2021 at the University of Glasgow. Her research interests include Scottish sociophonetics (her MSc dissertation discussed rhotic variation in the Central Belt), Korean phonology, accent bias, and neurolinguistics.

# Acknowledgements

Ethical permission was granted by the University of Glasgow's College of Arts Research Ethics Committee on the 29th of November 2019. Many thanks to every participant who took the time to take part in this piece of research, and to Dr Rachel Smith for her irreplaceable guidance and enthusiasm. Thank you to the University of Edinburgh for allowing me to utilise their perception labs while testing participants based at their campus.


Received: 17/07/2020
Revisions: 30/01/2021
Accepted: 09/02/2022