The Effects of Pitch on Second Language Learners’ Categorical Perception of Korean Alveolar Lax and Tense Stops*

Ho Eun Park†
University of Wisconsin-Milwaukee

Because English speakers use voice onset time (VOT) as their main cue to discriminate L1-English stops, it would be interesting to see if L1-English learners of L2-Korean use fundamental frequency (f0) effectively to distinguish Korean contrastive stops, for which f0 is known to be important. Thirteen English learners of Korean participated in an AX discrimination task. Results showed that f0 significantly affected the learners’ perception of Korean alveolar lax-tense stops, demonstrating that English learners were sensitive to f0 in L2 despite its absence as a main cue in L1.

Keywords: Korean stop contrast, Korean alveolar lax-tense stops, pitch, f0, VOT, L1-English learners of L2-Korean, AX discrimination task, pitch effects

1 Introduction

In Korean, there are three laryngeal distinctions—lax, tense, and aspirated—in stops. In order to distinguish these Korean stops, pitch, or fundamental frequency (f0), has become known as a pivotal feature in modern Korean (Kong 2012; Silva 2006) as the voice onset times (VOTs) of each Korean stop are overlapping. Although it is controversial whether pitch is a distinctive feature, Korean has been known for its consistent pitch differences related to those initial stop consonants. Tense and aspirated stops tend to have higher pitch than lax stops (Ahn & Iverson 2004). Thus, the pitch effects on categorical perception of differentiation between Korean stops should be carefully investigated. Furthermore, since the distinction between lax, tense, and

* I would like to express my deepest appreciation to professor Pycha in linguistics department at University of Wisconsin-Milwaukee for providing me lots of comments and advice to complete this study. Also, I thank my colleague, Mary Clinkenbeard, who proofread my study and gave a lot of comments, and Dr. Dmitrieva at Purdue University who offered me useful articles.
† Address correspondence to hoeun@uwm.edu

Proceedings of the 9th Purdue Linguistics Association Symposium (2014)
aspirated stops does not exist in English, the contrast seems to be difficult to either pronounce or perceive for L1-English learners of L2-Korean (hereafter, KSL English learners). In order to resolve these difficulties and to learn this Korean stop contrast, learning to use pitch differences effectively is essential. That is, KSL English learners need to be able to use important features of Korean stops, in particular, pitch differences, effectively, to learn Korean even if their L1-main cue seems to be VOT (Lisker & Abramson 1964).

The purpose of this research is to investigate how novice learners of Korean are sensitive to pitch differences in discriminating L2-Korean alveolar lax stops from alveolar tense stops.

This paper’s main research question is as follows:

How does categorical perception change depending on different pitch variations for KSL English learners? That is, how do KSL English learners approach Korean lax-tense stops with pitch variation?

Different pitch variations of Korean lax and tense stops will affect categorical perception of ‘ㄷ[t]’ and ‘ㄸ[t’]’ for KSL English learners. Specifically, if the pitch level of tense ‘ㄸ[t’]’ is lower than its natural pitch level, it will be more difficult to distinguish ‘ㄸ[t’]’ from ‘ㄷ[t]’ compared to natural pitch variation because it is known that tense stops have a higher pitch than lax stops. That is, KSL English learners will need to be able to use f0 information and be sensitive to pitch difference to distinguish Korean contrastive stops, even if their L1-main cue is VOT.

2 Background

There exists a myriad of research on the categorical perception of phonemes (Liberman et al. 1957; Lisker & Abramson 1970). By conducting a phoneme identification and a discrimination test with /ba/, /da/, and /ga/, Liberman et al. (1957) revealed that speakers were more likely to discriminate sounds in different phonemic categories rather than within categories. Additionally, Lisker and Abramson (1970) showed that there are different aspects of categorical perception across languages depending on different VOT values. They also found that Korean stops are differentiated in production by VOT differences (Lisker & Abramson 1964).

1 Native speakers of English who learn Korean as a second language
2 We chose lax and tense stops among the three contrasts in Korean stops because the interval of pitch seems to be smaller than that of lax and aspirated stops (Ahn & Iverson, 2004). If there are significant pitch effects between lax and tense contrast, then we could expect more significant pitch effects between lax and aspirated stops. Similarly, we chose alveolar stops ‘ㄷ[t]’ and ‘ㄸ[t’]’ as stimuli, because their pitch gap seems to be the smallest among other lax-tense contrasts by three native Korean speakers, one male and two females, who are originally from Seoul and use standard Korean. That is, to emphasize my point here, a more controversial pitch relationship was selected for English speakers. If my hypothesis is proven to be correct, there should be even more significant pitch effects between other lax-tense stop relations. Also, in English, tense stops (e.g. ‘ㄸ[t’]’) do not exist, so the stops are considered to be more difficult than lax and aspirate stops. These are the main reasons why a tense stop is selected to be compared with a lax stop.
3 In this study, pitch refers to the onset fundamental frequency (f0).
However, currently there are many scholars who insist that contemporary standard Korean is developing pitch (f0) distinctions as its VOT distinction overlaps in stop consonants (Kong 2012; Kim 2000; Kim & Duanmu 2004; Silva 2006). Kong (2012) showed that Korean listeners use both VOT and f0 values while distinguishing Korean lax, tense and aspirated stop sounds, although VOT values overlap between lax and aspirated stops.

Kim (2000) strongly insisted that f0 is becoming a more important cue in modern Korean by investigating the phonetic and phonological aspects of the relationship between segment types and tone in Korean (two dialects of Seoul and Jeonnam) and English. In her research, she discovered that for the perception of phonation contrasts, Korean listeners use the tonal information of Korean words, and then she even argued that those findings are interpreted as evidence that Korean is undergoing tonogenesis. Moreover, Kim (2012) showed the merging values of VOTs between the three-laryngeal stops in Korean from 1960’s to present. The VOTs are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Tense</th>
<th>Lax</th>
<th>Aspirated</th>
<th>Asp-lax Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960’s–1970’s</td>
<td>Mean</td>
<td>11ms</td>
<td>32ms</td>
<td>104ms</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0–52ms</td>
<td>15–100ms</td>
<td>30–210ms</td>
</tr>
<tr>
<td>1990’s–2002’s</td>
<td>Mean</td>
<td>14ms</td>
<td>49ms</td>
<td>91ms</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>9–50ms</td>
<td>15–89ms</td>
<td>75–121ms</td>
</tr>
<tr>
<td>2004’s–present</td>
<td>Mean</td>
<td>15ms</td>
<td>63ms</td>
<td>77ms</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>2–26ms</td>
<td>17–171ms</td>
<td>22–196ms</td>
</tr>
</tbody>
</table>

Table 1: Mean VOTs (ms) and ranges from 1960’s to present (Kim 2012).

As we can see from Table 1, the VOT overlap between lax and aspirated stops seems to be clear after 2004, whereas the VOT of tense stop remains similar. Because of the VOT overlap between lax and aspirated stops, Korean speakers might use other cues such as f0 to distinguish these stops.

Ahn and Iverson (2004) note that pitch is not a contrastive property (distinctive feature) in Korean phonation. However, there is consistently higher pitch in vowels following initial tense and aspirated consonants as compared to lax or sonorant consonants, suggesting that there is a constant pitch difference in distinguishing the stops.

Based on the findings described above, this present study will examine the effects of pitch on discriminating Korean stops in the context of L2-Korean learning, namely, how English learners of L2-Korean acquire L2-Korean using different cues such as VOT and f0.

3 Methods and Procedures

An AX discrimination task was designed to see whether there would be mean score differences between four different pitch conditions of the target sounds, as well as between two different word conditions, CV non-word and CVC minimal pairs.
3.1 Participants

Thirteen native English speakers, six of which were male and seven of which were female, were recruited from an elementary Korean course—Korean 101—at the University of Wisconsin-Milwaukee (UWM). They were between 18 and 23 years old, and this course was their first experience in learning Korean. Also, this experiment was done when they had been learning Korean for about three to four months. The reason why novice learners were chosen was to see whether the learners picked up f0 cues in the first stage of learning. The students who participated in the experiment were given extra credit.

3.2 Stimuli

There was one continuum which contains four different pitch relationships as stimuli in this experiment. The base relationship of stimuli—a pair of Korean lax and tense stops—was spoken by a Korean female speaker. A pair of alveolar lax and tense stops, ‘ㄸ[ti]’ and ‘ㄷ[ti]’, were selected as the base relation of stimuli, because the pitch gap between these stops was at least 20Hz and was considered to be the smallest compared to other Korean lax and tense stops produced by three native speakers of Korean including the researcher. Their sounds were used only to compare the intervals of pitch between lax and tense stops using Praat.

After comparing the intervals and selecting the target stimuli, each stimulus of a Korean female was audio-recorded in a sound booth of a phonetics lab at UWM. The stimuli were controlled to have the highest amplitude setting in Praat.

A pair of ‘따[t’a]’ and ‘다[t’a]’ was selected as a CV non-word relation, while a pair of ‘달[tal]’ which means ‘moon’ and ‘딸[t’al]’ which means ‘daughter’ in Korean was chosen as an exemplar of a CVC minimal pair in Korean. Both CV non-word and CVC word conditions were used in order to see whether there are coda effects or lexical effects in the listeners’ perception.

They were manipulated by five different pitch conditions in Praat. The first two points of pitch variation were considered to be manipulated, and the values of VOT were kept intact.

As is already explained, we wanted to select the most controversial relationship to see whether there are pitch effects in categorical perception of Korean stops for KSL English learners.

All three native Korean speakers, one male and two females, use standard Korean. Their length of residence in the U.S. varies from 2 to 6 years. It did not seem meaningful to obtain the average pitch across speakers because each individual had at least 100 Hz difference in average pitch level. Furthermore, the average pitch level was variable depending on each speaker’s condition or surroundings. Therefore, in the present study, their overall pitch averages for lax and tense stops were not used, but the pitch intervals between the stops were compared.

Teasing apart the coda and lexical effects was out of the scope of this study. The participants could use both coda and lexical information at the same time.

Among the five pitch conditions, one condition was a foil to prevent subjects’ bias in the discrimination task.

The pitch manipulation function of Praat was used.

In this study, only the first two points of the pitch variation were manipulated by Praat. This is because only the first two points of pitch seem to be related to the pitch levels of lax and tense stops. The other points of pitch showed the similar shape and level of pitch variations as a pair of stimuli sharing the same surroundings except the initial sound of stop consonants, so it was not necessary to manipulate them.

The VOT of each stimulus was not manipulated to be the same. The detailed values of VOT are as follows:
Figure 1: natural condition: a pair of natural lax-tense stops

Figure 2: raised high condition: the tense stop in a pair is increased by 20Hz

Figure 3: lowered high condition: the tense stop in a pair is reduced by 20Hz

Figure 4: neutralized condition: a pair of neutralized lax-tense stops

<table>
<thead>
<tr>
<th>Stimuli: a continuum</th>
<th>VOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV non-word /tə/-/t’a/</td>
<td>/tə/: 53ms /t’a/: 13ms</td>
</tr>
<tr>
<td>CVC word /tal/-/t’al/</td>
<td>/tal/: 36ms /t’al/: 11ms</td>
</tr>
</tbody>
</table>
The first condition, “natural condition”, was included as a control group in order to obtain baseline data for the other conditions. This condition included a pair of the lax and tense stops which were natural and non-manipulated. The second condition was a “raised high condition” in which a natural, non-manipulated lax stop was compared to a tense stop increased by 20Hz. The third condition was a “lowered high condition” in which a natural lax stop was compared to a tense stop which was reduced by 20Hz. Those two conditions were included in order to see whether there are either increased or reduced pitch effects compared to the natural condition. The fourth condition was a “neutralized condition” in which a pair of lax and tense stops were neutralized at 200Hz. The fifth condition was a “same stimuli condition” of ‘다[ta]’/ ‘다[ta]’, and ‘달[tal]’/‘달[tal]’. This condition functioned as a foil of the experiment to prevent certain bias in an AX discrimination test, and thus, this condition was not analyzed.

The purpose of the AX task was to reveal the effects of pitch on the distinction between lax and tense stops ‘ㄷ[t]’ and ‘ㄸ[t’]’ by comparing different pitch relationships in a continuum compared to the natural baseline. Figures 1 to 4 show the four pitch relationships of a continuum where the lax stop is on the left side and the tense stop is on the right side in each pair.

3.3 Procedure

All thirteen participants participated in the AX discrimination task as a group in a quiet classroom. The AX task script by Praat was conducted by a researcher using a laptop, and the laptop was connected to a sound amplifier so that all participants could listen to the sounds and answer according to what they heard. Each participant was given a piece of paper where she or he could choose “Same” or “Different”. They were told to check “Same” if two stimuli were the same, or check “Different” if two stimuli were different sounds. Specifically, they were told to focus on the initial stop sounds of the pairs which could be the same or different. The five conditions of each pair of ‘다[ta]’/‘따[t’]’ and ‘달[tal]’/‘딸[t’al]’ were controlled to be repeated three times so that thirty questions were randomly assigned by Praat. The overall methods and procedures are summarized in Table 2.

---

11 This condition is called raised high condition based on the assumption that a Korean tense stop already has a higher pitch level than that of a lax stop. Only the tense stop is manipulated in each pair. The same explanation can be applied to lowered high condition. Also, 20Hz difference was chosen to be manipulated both in the raised high and lowered high conditions, as the pitch gap between alveolar lax and tense stops was considered to be at least 20Hz while comparing the intervals of stops.

12 The values of f0 are not shown exactly in these figures (Error range of ±5Hz). The f0 of tense stops is manipulated by increasing or reducing it by 20Hz in Praat.
Table 2: overall methods and procedures

4 Results

Only data of ten out of the thirteen participants, five male and five female, were analyzed. The responses of three participants were discarded because they skipped one or two questions. In total, 240 questions were analyzed (four pitch conditions*two lexical types*ten listeners*three repetitions). One point was given for each correct answer, while zero points were given for each wrong answer. For the analysis of the responses, a two-way repeated analysis of variance (ANOVA) was conducted by SPSS in order to determine whether there were significant effects of the four different pitch conditions (e.g., natural condition, raised high condition, lowered high condition, and neutralized condition) as well as two lexical types (e.g., non-word and word) or two different syllable structures of CV and CVC, respectively (four pitch conditions*two lexical or syllable types), and also to see if there was an interaction between pitch conditions and lexical or syllable types.

Recall that the fifth condition, the foil of the experiment, was not analyzed.
The two-way repeated ANOVA revealed a significant main effect of pitch differences (F(2.208, 19.872) = 9.115, p=0.001, two-tailed). Namely, the mean score differed significantly depending on four pitch conditions. On the other hand, there were no significant effects of lexical or syllable type (F (1, 9) = 0.184, p=0.678, two-tailed). Nor was there significant interaction between pitch conditions and lexical types (F (2.304, 20.733) =1.181, p=0.332, two-tailed). Therefore, only one factor, pitch differences, meaningfully affected the outcome variable, the mean scores of the task. Also, the effects of pitch were the same regardless of lexical or syllable type. In other words, there were no significant lexical or coda effects, in this case. The results indicated that pitch differences affected categorical perception of Korean lax-tense contrast for KSL English learners.

![Figure 5: mean score differences depending on four pitch conditions](image)

Figure 5: mean score differences depending on four pitch conditions

As can be seen in figure 5, participants obtained the highest mean score in the natural condition in both CV non-word and CVC word types. They also obtained the same highest score as the natural condition at raised high condition in CVC word type, while having the second highest mean score at this condition in CV non-word type, followed by the lowered high condition for both lexical types. Neutralized condition showed the lowest score in the task in both lexical types. This condition appeared to be the most confusing and difficult pitch condition for KSL English learners to identify. Similar results were shown both for CV non-word and CVC
word types. Figure 6 provides the ten KSL English learners’ overall percentage of correct responses in the AX discrimination task.

![AX discrimination task](image)

**Figure 6: percentage (%) of correct responses**

A Tukey HSD post hoc test was conducted to see specifically which pitch condition significantly affected other pitch conditions. It was revealed that the mean score of the lowered high condition was significantly lower than the natural condition, which was a baseline (p=0.048). This means that when the pitch level of a tense stop is lowered by 20Hz, KSL English learners had great difficulty to discriminate Korean lax from tense, perceiving tense stops as lax stops, and vice versa. From this result, we can see that KSL English learners are sensitive to pitch differences and use pitch differences to distinguish the Korean lax-tense contrast.

Moreover, the mean score of the neutralized condition was significantly lower than the raised high condition (p=0.002) and the natural condition which was a control condition (p=0.003). The learners seemed to have the most difficulty in this condition. In this case, we could clearly recognize that KSL English learners were not sensitive to their L1-main cue, VOT, because KSL English learners were not able to distinguish Korean stops using VOT, even if the values of VOT were kept intact in the condition. They focused more on the pitch difference, although they were expected to use VOT as their main cue to distinguish Korean stops. It is even more noticeable, considering that they are only novice learners of Korean. Therefore, KSL English learners seemed to use f0, effectively, while they did not seem to use their L1-main cue VOT effectively.
Conclusion

In this study, we found that L1-English speakers of L2-Korean used f0 in discriminating the Korean contrastive stops, indicating that pitch mainly affected categorical perception of the Korean stops for KSL English learners. When the natural pitch of the tense stop was reduced, KSL English learners could not distinguish Korean tense from lax stops, showing that KSL English learners of beginning level were sensitive to f0 and could use f0 effectively to perceive and distinguish Korean stop ‘ㄱ[t]’ from ‘ㄲ[t]’. When the pitch of both stops was neutralized, KSL English learners could not distinguish tense from lax stops. From this, we can see that they did not use their L1-main cue VOT effectively, considering that VOT was in the natural condition. Thus, it seems that KSL English learners are more sensitive to f0 than VOT in discriminating Korean stops. These results imply that using f0 effectively might be necessary in order to learn Korean stop contrast for English speakers.

There could be several reasons to explain the results. One of the reasons could be L1-English f0 cue effects. As f0 is considered to be a secondary cue in English (Llanos et al. 2013) and English speakers may already know how to use f0 effectively, the learners of Korean could be sensitive to f0 as well. Llanos et al. (2013) said that when assimilated to the same native category according to a main cue, learners may give a greater weight to secondary cues in order to discriminate between non-native categories. In our case, the learners could use their secondary cue, f0, more effectively because Korean lax and tense stops, which have positive VOT values,
could be mapped onto the English voiced and voiceless stops in terms of VOT. This possibility should be looked at thoroughly by investigating whether, or to what extent, Korean alveolar stop contrasts could be perceptually mapped onto English alveolar voiced or voiceless stops.

Moreover, KSL English learners’ effective use of f0 could be due to L2-Korean pitch effects because f0 may be becoming a more important cue in the Korean stop contrast as there is the VOT overlap between Korean lax and aspirated stops (Kim 2012). Because a tense stop can still be distinguished by VOT, the learners need to know how to use VOT, their main L1 cue, effectively to differentiate Korean lax stop from tense stop. However, in the present study, KSL English learners did not use VOT even when VOTs in the continuum were kept intact; rather they used f0. The results clearly showed that the learners were more sensitive to f0 than VOT to distinguish L2-lax stop from tense stop. In order to find out the exact reason of this phenomenon, we should do a more in depth analysis of whether the VOTs of Korean lax and aspirated stops are truly overlapping, and how tense stop can be differentiated from lax and aspirated stops phonetically and phonologically. The results of the present study should be compared to those of Korean speakers and English speakers who have not been exposed to Korean in the future. In this case, we can determine whether English speakers are affected by L2 Korean pitch effects or if the speakers can naturally use f0 as well when distinguishing Korean stops.

In the current study, the VOTs were not manipulated in the pitch conditions in the continuum, so it was difficult to determine the relationship between f0 and VOT in Korean stops. In order to investigate the sole effect of pitch as well as the interaction between the cues, both f0 and VOT should be manipulated properly in the continuum: there should be no VOT differences in the continuum, for an instance. Or, we can compare participants’ performance of two pairs of stimuli in which the VOT difference for one pair is smaller and the VOT difference for the other pair is larger. From this, we will be able to see more thoroughly how KSL English learners use the cues in discriminating Korean contrastive stops.

There were no significant lexical or coda effects in this study. However, there seemed to be certain lexical significance in the neutralized condition as CVC-word relation tended to have a higher mean score compared to the CV-non word relation in this condition. That is, KSL English learners might more easily discriminate Korean lax-tense stops in the CVC-word with neutralized condition in which there was no pitch variation between lax and tense stops. It seemed that the learners use lexical or coda information in this condition to distinguish the Korean stops. We should look at these lexical or coda effects in future studies.
References


