

Acoustic observation for English speakers' perception of a three-way laryngeal contrast of Korean stops

Nahyun Kwon

University of Queensland

n.kwon@uq.edu.au

While the two-way voicing contrast of English stops can be distinguished by VOT alone, the three-way laryngeal contrast of Korean stops requires additional acoustic parameter, f_0 , together with VOT for its realization (Chang, 2010; M. Kim, 2004). The distinct acoustic characteristics of the Korean and English stops may create difficulties in English speakers' discrimination of the non-native Korean contrasts. To confirm this hypothesis, the current study examines English speakers' discrimination of a three-way laryngeal distinction of Korean stops /p t k/ in the word-initial position of disyllabic minimal pairs. The result supports the hypothetical link between acoustic patterns and perceptual discrimination to a large extent by displaying a relatively low correct discrimination level on the lenis-fortis contrast. This leads to a conclusion that f_0 is as important as VOT for non-native listeners to fully perceive the three-way contrast of Korean stops.

Keywords: VOT, f_0 , Korean stops, L2 speech perception, acoustic phonetics

1. Introduction

Best, McRoberts, and Sithole (1988)'s Perceptual Assimilation Model (PAM) explains various levels of naïve listeners' discrimination of L2 contrasts in relation

Selected Papers from the 44th Conference of the Australian Linguistic Society, 2013

edited by Lauren Gawne and Jill Vaughan

to their perceptual similarities with L1 categories. Under this model, phones that fall into two different categories of L1 (“two categories”) are predicted to be the easiest to discriminate whereas phones that are assimilated to a single L1 category (“single category”) are predicted to be the hardest. Furthermore, phones that are assimilated to a single category yet unequally (“category goodness difference”) and that are not assimilated to any of L1 category (“non-assimilable”) are claimed to lie on the intermediate level of difficulty. On this basis, Schmidt (2007)’s study, where English speakers closely perceived word-initial Korean lenis /p t k/ and aspirated stops /p^h t^h k^h/ as homorganic English voiceless stops /p t k/, and fortis stops /p’ t’ k’/ as voiced ones /b d g/, suggests that English listeners would discriminate lenis vs. fortis and aspirated vs. fortis Korean stops better than they would the lenis vs. aspirated contrast. However, the present study, which examined English speakers’ discrimination of the corresponding Korean stops, showed a rather contradicting result to the PAM expectation because the lowest discrimination was found in the lenis-fortis contrast. To account for this rather surprising result, I adopt a different approach from the PAM model in this paper: I measure acoustic-phonetic differences between Korean and English stops, and examine their roles as perceptual cues. Let us first examine the previously reported acoustic characteristics of Korean and English stops.

The phonetic characteristics of a three-way distinction of Korean stops are cross-linguistically “atypical” (M-R Kim, Beddor, & Horrocks, 2002). They do not show a “default” contrast in voicing and aspiration category (Kong, Beckman, & Edwards, 2011). Instead, they exhibit contrast among different laryngeal activities¹. The three-way laryngeal contrast, namely fortis, lenis, and aspirated, in Korean stops appears at three different places of articulation: bilabial, alveolar and velar (see Table 1).

¹ The three-way contrast becomes neutralized in syllable final position. Also, although they do not show a voicing contrast, lenis stops become voiced when they appear in inter-vocalic position.

| Place category | Laryngeal feature | | |
|----------------|-------------------|-------|-------------------|
| | Fortis | Lenis | Aspirated |
| Bilabial | /pʔ/ | /p/ | /p ^h / |
| Alveolar | /tʔ/ | /t/ | /t ^h / |
| Velar | /kʔ/ | /k/ | /k ^h / |

Table 1. *Three-way contrastive stops in Korean.*

To distinguish this typologically unusual distinction of Korean stops, many previous studies introduced multiple acoustic parameters, such as VOT (Cho, Jun & Ladefoged, 2002; Hardcastle, 1973; Kagaya, 1974; C-W Kim, 1965; M-R Kim et al., 2002; Lisker & Abramson, 1964), f_0 (Cho et al., 2002; Dart, 1987; Hardcastle, 1973; Kagaya, 1974; M-R Kim et al., 2002; M Kim, 2004), H1-H2 differences at vowel onset (Cho et al., 2002; Han, 1998; M-R Kim et al., 2002), and stop closure duration (Cho et al., 2002; M-R C Kim, 1994). Among which, VOT and f_0 together have been widely accepted as major acoustic cues for a successful distinction (Chang, 2010; M Kim, 2004; Lee & Iverson, 2012): Previously, VOT has been claimed to be longest in aspirated followed by lenis and fortis with a possible overlap between lenis and fortis (i.e., aspirated > lenis = fortis). Thus, it has been considered that VOT provides salient information for the fortis vs. aspirated and lenis vs. aspirated, but not necessarily for the lenis vs. fortis stops. f_0 of the following vowel onset, on the other hand, has been claimed to be highest in aspirated followed by fortis and lenis with a possible overlap between aspirated and fortis (i.e., aspirated = fortis > lenis). Thus, it has been considered to best distinguish lenis vs. fortis contrast where VOT overlaps.

Now, let us turn our attention to English stops. Unlike Korean stops, they only have a two-way phonological voicing contrast, namely voiced and voiceless. Therefore, it has been claimed that VOT alone can be a sufficient cue to differentiate the contrast (Francis & Nusbaum, 2002; M-R Kim et al., 2002). At this point, we can clearly see that there are apparent differences in the acoustic properties of Korean and English stops. And, they could create difficulties in English speakers' discrimination of the non-native Korean contrasts. That is, English listeners may not be used to attending to the additional primary acoustic cue of Korean stops, f_0 . Thus, they may rely on VOT only, which is also distinctive in English, when making their perceptual decisions across the Korean

stimuli. This may result in a relatively poor discrimination of the lenis vs. fortis contrast, where f_0 plays a salient role, compared to the other two contrasts. Francis and Nusbaum (2002), in particular, support this hypothesis by showing that naïve English listeners mostly attended to VOT in their perception of the Korean stops. However, in the sense that their study consisted of only artificial stimuli, it poses a methodological weakness that their results might not have captured subjects' natural responses. Furthermore, it is not also assured that their hypothesis will be effective in words of different syllable length because they only used monosyllabic stimuli. Thus, the present study conducts a perceptual experiment with existing disyllabic minimal pairs and examine whether naïve English listeners display the same tendency towards the acoustic cues when discriminating the three-way Korean contrastive stops /p t k/ in the word-initial position. The following sections outline the description of the experiment design and analysis of its result with further discussion.

2. Method

2.1 Subjects

English-speaking students, who are over 18, enrolled in an introductory linguistics course at the University of Queensland served as participants. The number of participants was 122 and they were given course credit for participation.

2.2 Materials

Stimuli consisted of two repetitions of nine existing Korean disyllabic minimal pairs which contrast laryngeal features in word initial stops (2 repetitions x 3 laryngeal contrastive pairs x 3 places of articulation), yielding a total of 18 stimulus items (see Table 2). Tokens at each place of articulation had different vowel contexts (i.e., /i/ in bilabial, /ə/ in alveolar and /ɛ/ in velar) because the availability of items that share the identical vowel context across place of articulation was restricted in Korean. In terms of syllable structure, all of the items were in CVCV sequences except for the alveolar stops. Inclusion of coda consonant in the first syllable of the alveolar context was unavoidable due to the

limited availability of existing Korean words. Despite such inconsistent phonetic contexts between conditions, stimuli contained real words because they were expected to play a better role in drawing subjects' natural responses than artificial words.

| Place category | Laryngeal feature | | |
|----------------|-----------------------|-------------------------------------|------------------------------------|
| | Lenis vs. Fortis | Fortis vs. Aspirated | Lenis vs. Aspirated |
| Bilabial | /pi.ta/ vs /p'i.ta/ | /p'i.ta/ vs /p ^h i.ta/ | /pi.ta/ vs /p ^h i.ta/ |
| Alveolar | /təl.ta/ vs /t'əl.ta/ | /t'əl.ta/ vs /t ^h əl.ta/ | /təl.ta/ vs /t ^h əl.ta/ |
| Velar | /kɛ.ta/ vs /k'ɛ.ta/ | /k'ɛ.ta/ vs /k ^h ɛ.ta/ | /kɛ.ta/ vs /k ^h ɛ.ta/ |

Table 2. Korean stimuli for a perception test.

The minimal pairs, which show the different laryngeal contrast at different places of articulation, were automatically shuffled in a random order.

2.3 Recording

Each minimal pair was read two times in the same order by the author, a 30 year old female native speaker of Korean language who had lived in Seoul for 22 years. She read the stimuli, written in Korean Hangeul orthography, with a simple falling pitch accent. The recording was digitally made in a quiet room with a Shure SM10A microphone connected to a Marantz PMD660 portable recorder at a sampling rate of 48 kHz.

2.4 Procedure

The experiment was conducted on the university online survey system, blackboard. A group of English speakers were asked to listen to the hyperlinked audio files for the two repetitions of each minimal pair and answer yes-no questions for its discrimination without time limit. The instruction, which was given to the participants, is as follows:

In this experiment, you will be asked to click sound files which are hyperlinked to each yes-no question and tick on:

'Yes' if you think they are the same words with the same sounds or

‘No’ if you think they are different words with different sounds

Each pair will be repeated twice at a 3-second interval in the sound file. Make sure to tick on ‘Yes’ or ‘No’ before you move on to the next question. Do not replay it as I am interested in your initial impression of foreign sounds.

2.5 Acoustic measurement

VOT and f_0 in all stimuli including the repeated items were measured with free acoustic analysis software, Praat. First, VOT refers to the time between the start of the release burst and the onset of voicing (see Figure 1). Thus, it was identified as the time, in milliseconds, from the beginning of the release burst to the first point of the regular glottal periods and a clear voicing bar in the spectrogram (Broersma, 2010; Chang, 2010; Francis & Nusbaum, 2002; M-R Kim et al., 2002; M Kim, 2004).

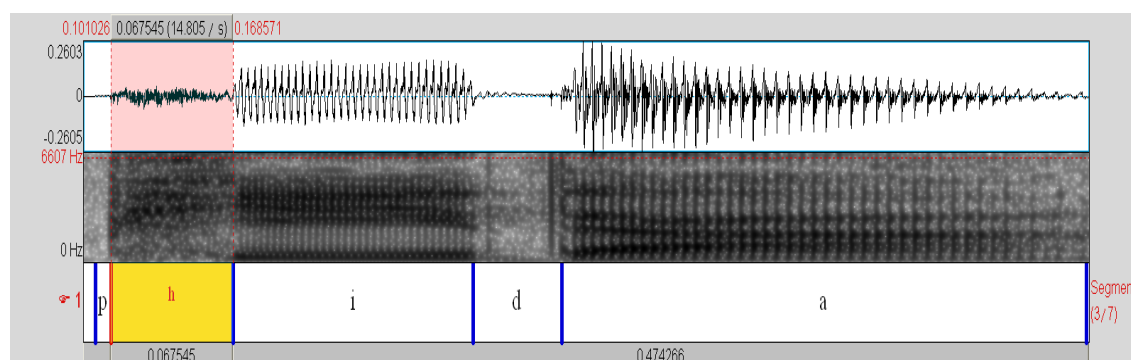


Figure 1. VOT value of Korean bilabial aspirated stop in /p^hi.ta/².

Second, fundamental frequency of the following vowel, f_0 , was obtained by measuring the average duration of the first three glottal pulses in the vowel and converting to a frequency value (Chang, 2010). And, it was ensured that the initial periods of all stimuli were not shorter or longer than 30% of the following period (see Figure 2).

² Voiceless lenis stop /t/ became voiced /d/ in inter-vocalic position by a Korean phonological rule known as “Intervocalic Lenis Stop Voicing” (Cho, Jun, & Ladefoged, 2002).

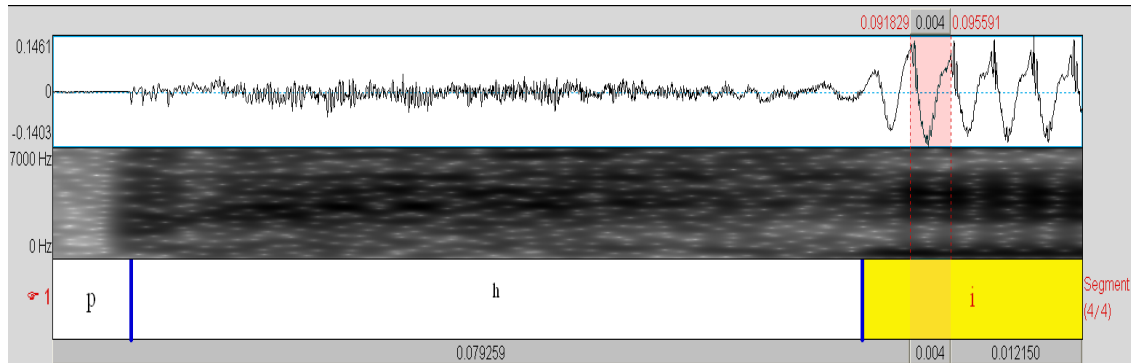


Figure 2. f_0 value of Korean bilabial aspirated stop in /p^hi.ta/.

3. Result

The result of acoustic measurement showed that the values of the two primary acoustic cues for the recorded Korean stimuli were in general consistent with the previous findings (Cho et al., 2002; M-R Kim et al., 2002; M Kim, 2004)³.

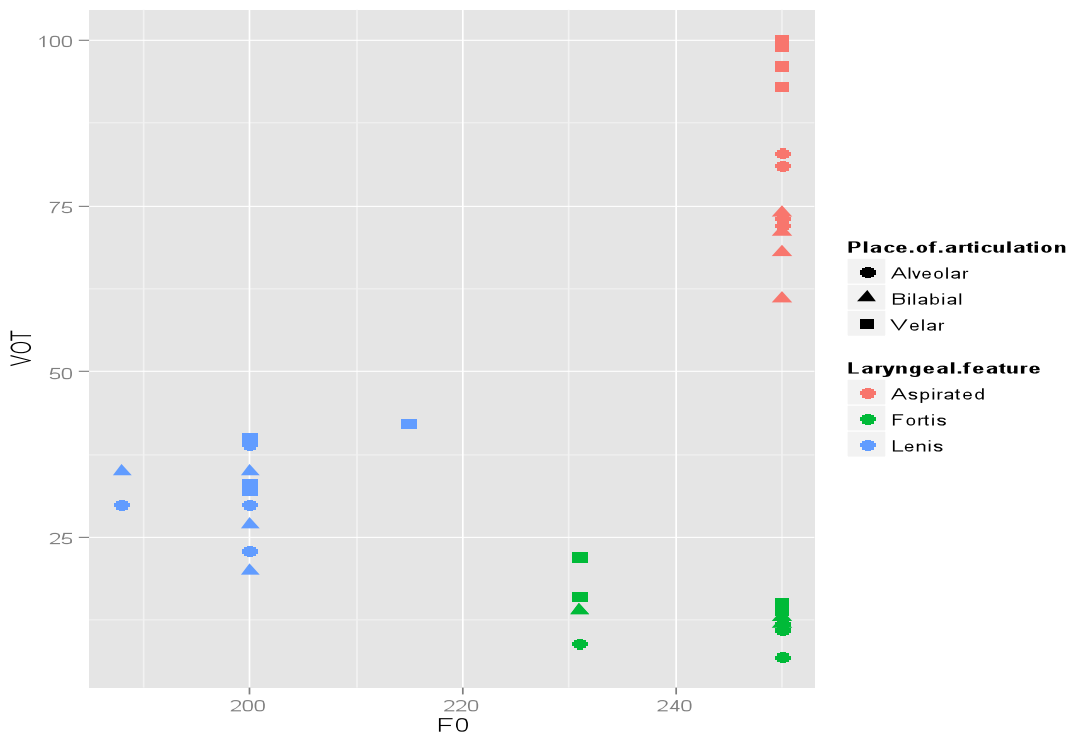


Figure 3. VOT(ms) and f_0 (Hz) values of Korean stimuli.

³ Note that the production results are preliminary in the sense that the number of speaker and repetition per token was limited, and any potential microprosodic effects such as intrinsic vowel pitch were not taken into account.

Consider Figure 3 above. Fortis - lenis stops and aspirated - fortis stops formed two separate clusters based on their VOT and f0 values, respectively. This indicates that VOT overlaps between fortis and lenis, and f0 overlaps between aspirated and fortis as expected. The detailed reports are found in the following subsections.

3.1 Voice Onset Time (VOT)

Similar to the previous findings, VOT was longest in aspirated stops followed by lenis and fortis at all three places of articulation with one overlapping point between bilabial lenis and velar fortis (see Table 3 and Figure 4 below). The mean VOTs for aspirated, lenis and fortis category were 80.9 (sd=13.2), 32.1 (sd=6.6) and 13.2 (sd=3.7), respectively. And, there were significant difference at least for one pair of stop categories according to one-way ANOVA, $F(2,15)=106.05$, $p < 0.001$.

| | Lenis | Fortis | Fortis | Aspirated | Lenis | Aspirated |
|----------|----------------|------------------|-----------------|------------------|------------------|------------------|
| Bilabial | 31 (27- 35) | 14 (14-14) | 12.5 (12-13) | 69.5 (68 -71) | 27.5 (20 -35) | 67.5 (61 -74) |
| Alveolar | 31 (23 -39) | 9.5 (7 -12) | 10 (9 -11) | 82 (81 -83) | 30 (30 -30) | 72.5 (72 -73) |
| Velar | 36 (32 -40) | 14.5 (14 -15) | 19 (16 -22) | 98 (96-100) | 37.5 (33 -42) | 96 (93 -99) |

Note: VOT ranges for the two tokens in each minimal pair are given in parentheses.

Table 3. Mean VOT values in milliseconds.

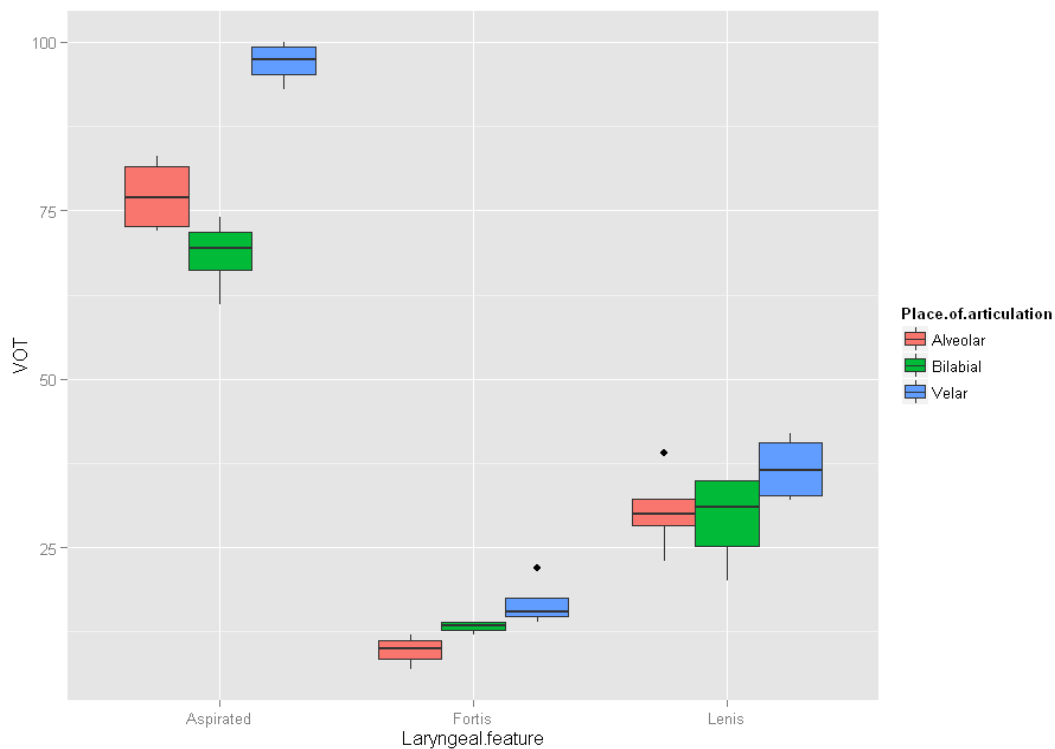


Figure 4. VOT ranges for different laryngeal categories.

To figure out where the most significant VOT difference is laid in the possible combinations, I further conducted Tukey's multi pair-wise comparisons. As a result, the VOT differences for the fortis-aspirated contrast and lenis-aspirated contrast were more significant than the difference for the lenis-fortis contrast (i.e., the confidence interval for lenis-fortis is closest to the dashed zero line as shown in Figure 5). It helps the hypothesis that English listeners are more likely to show lesser sensitivity to the lenis – fortis contrast, where VOT difference is least significant, than they do to the other two contrasts. This prediction that English subjects will find it hard to clearly differentiate the lenis and fortis contrast is further supported by Ladefoged & Johnstone (2001) that their mean VOTs range (9.5 ms – 37.5 ms) belongs to one English category, voiced stops (10 ms – 50 ms).

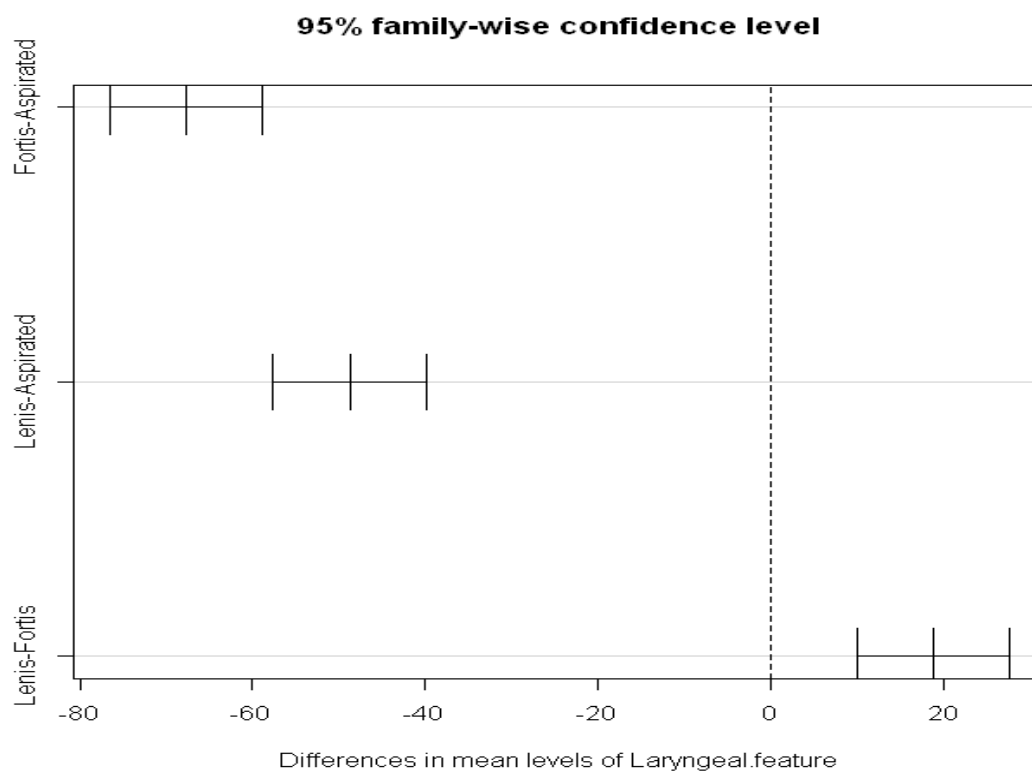


Figure 5. Tukey's multi pair-wise comparisons for mean VOTs.

3.2 Fundamental frequency of the following vowel (f_0)

Again, similar to the previous literature, f_0 of the following vowel was highest in aspirated stops followed by fortis and lenis with some overlaps between aspirated and fortis (see Table 4 and Figure 6)⁴.

In detail, the averaged mean f_0 s for aspirated, fortis and lenis category were 250.0 Hz (sd=0), 242.0 Hz (sd=9.7) and 199.2 Hz (sd=6.7), respectively. The mean of each laryngeal category above was calculated by averaging the means of the corresponding laryngeal category in different places of articulation which appear in different pairs (e.g., Mean f_0 for lenis category was calculated by averaging means of bilabial, alveolar, and velar lenis stops which appear in both lenis-fortis pair and lenis-aspirated pair). Like mean VOT values above, they were also

⁴ A potential high intrinsic voice pitch for /i/ does not seem to influence the results as shown in Figure 6 where bilabial stimuli do not have a particularly high f_0 range.

significantly different at least for one pair of stop categories, $F(2, 15) = 119.58$, $p < .001$.

| | Lenis | Fortis | Fortis | Aspirated | Lenis | Aspirated |
|----------|-------|--------|--------------------|-----------|--------------------|-----------|
| Bilabial | 200 | 231 | 250 | 250 | 194 (188-200) | 250 |
| Alveolar | 200 | 250 | 240.5 (231-250) | 250 | 194 (188-200) | 250 |
| Velar | 200 | 250 | 231 | 250 | 207.5 (200-215) | 250 |

Note: f_0 ranges for the two tokens in each item are given in parentheses. No parentheses indicate that the f_0 values for the two tokens are identical.

Table 4. Mean f_0 values of the following vowels in Hz.

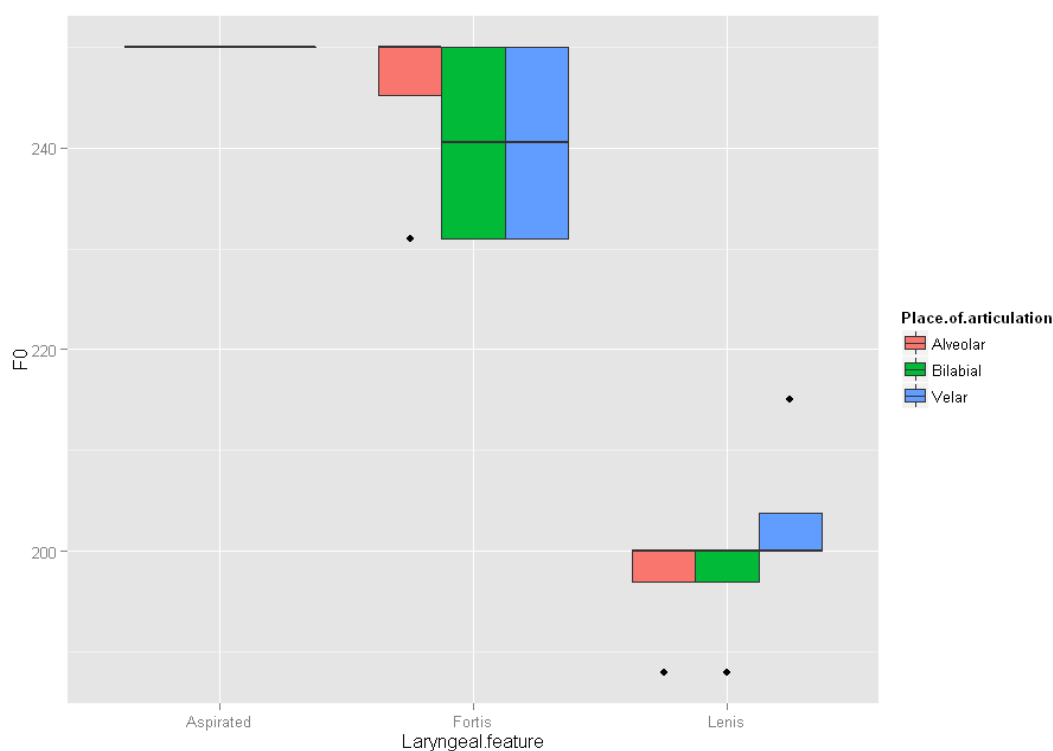


Figure 6. f_0 ranges for different laryngeal categories.

Tukey's test (see Figure 7) further revealed that the significant f_0 differences reside in the lenis-aspirated and lenis-fortis contrast but not in the fortis-aspirated contrast ($p = 0.09$). The significant f_0 difference for the lenis-fortis contrast confirms our assertion that f_0 is an essential acoustic cue together with VOT to fully distinguish the three-way Korean contrast: f_0 differentiates the lenis-fortis contrast where VOT overlaps.

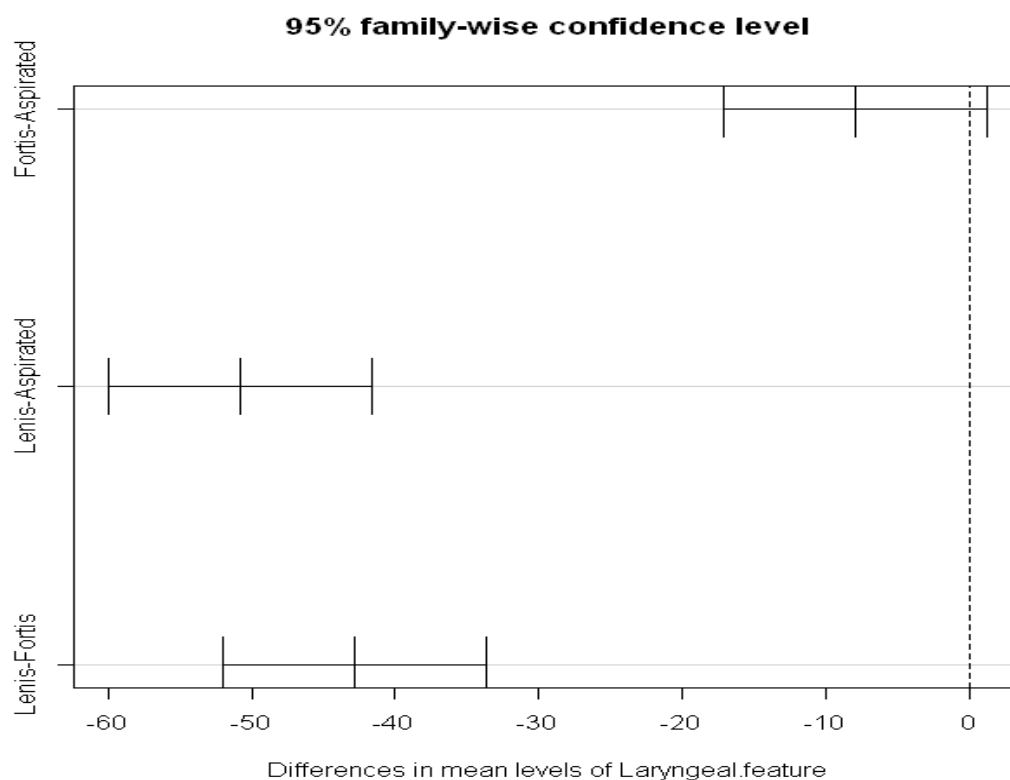


Figure 7. Tukey's multi pair-wise comparisons for mean f0s.

However, under the assumption that English listeners may not be used to attending to the secondary acoustic cue, f_0 , for their stop distinction⁵, it is less clear that there would be any strong advantages of the significant f_0 differences for the lenis-fortis contrast in English speakers' discrimination. Likewise, the non-significant f_0 difference for the fortis-aspirated would not negatively affect English listeners' discrimination of the corresponding contrast. In sum, English speakers would not show poor discrimination in the fortis-aspirated contrast (i.e., significant VOT difference, non-significant f_0 difference) in particular but would do so for the lenis-fortis contrast (i.e., less significant VOT difference, significant f_0 difference).

⁵ f_0 can be a perceptual cue to stop voicing contrast in English but its function is secondary because it is easily affected by VOT, not vice versa (M-R Kim et al., 2002).

3.3 English listeners' discrimination of the Korean stop contrasts

Each subject's same-different judgements of the stimuli were converted to binary numbers (i.e., 0 = wrong guess, 1 = correct guess). Then, a binomial test was applied to examine whether the probability of subjects' correct guess is statistically different from chance (H_0 = true probability of success is equal to 0.5) (Dalgaard, 2008). Table 5 shows the results that confirm our working hypothesis to a great extent: English speakers performed the best discrimination of the fortis-aspirated contrast followed by the lenis-aspirated and lenis-fortis contrast.

| Place category | Manner category | Laryngeal feature | | |
|----------------|-----------------|-------------------|----------------------|---------------------|
| | | Lenis vs. Fortis | Fortis vs. Aspirated | Lenis vs. Aspirated |
| Bilabial | Stop | 46% | 96%*** | 83%*** |
| Alveolar | Stop | 75%*** | 83%*** | 18%*** |
| Velar | Stop | 50% | 98%*** | 69%** |

p < 0.05 '**'

p < 0.01 '***'

p < 0.001 '****'

Table 5. English listeners' correct guess level of the Korean stop contrasts.

In detail, the above chance level of correct discrimination was found in all places of articulation for the fortis-aspirated contrast as expected by its significant VOT differences. The remaining two pairs of the laryngeal features of stops also follow our prediction in general except at the alveolar region. At alveolar, they exhibited discrimination patterns opposite to those predicted: the correct discrimination level was higher in the lenis-fortis than lenis-aspirated contrast. Furthermore, the wrong discrimination of the lenis-aspirated contrast did reach significance. This clearly indicates that the magnitude of the differences in VOT here is inversely related to English speakers' discrimination level. Thus, before one radically concludes that English speakers only rely on VOT, the primary acoustic cue of English, when making their perceptual decisions across the three-way contrast of Korean stops, it is important to account for this rather unusual behaviour of the alveolar stops at the present discrimination task.

The unusual behaviour of the alveolar stops may be explicable in terms of acoustic-perceptual salience of alveolar versus labial and velar release bursts.

According to Stevens, Keyser, and Kawasaki (1986: 431-441), alveolar stops have greater burst energy associated with a relatively faster tongue release because they have a smaller contact between tongue and palatal region than bilabial (no linguopalatal contact) and velar stops (larger linguopalatal contact). Thus, it is possible that the salient alveolar burst energy interferes with listeners' access to VOT for the alveolar lenis-aspirated contrast resulting in a significantly low level of correct discrimination. Conversely, the burst energy, as additional auditory cue, helps English listeners to differentiate the alveolar lenis-fortis contrast where only f_0 plays a major role for its distinction. The alveolar fortis-aspirated contrast, on the other hand, remains untouched by the strong burst energy because the size of their VOT difference is extremely large. To determine the validity of the energy burst effect, I conducted additional acoustic measurement for relative burst energy of the stimuli at three different places of articulation. Following Cho et al. (2002), 10ms at the burst and 20 ms at the middle of the vowel were sampled for each stimulus item. Then, the percentage of the average energy at the sampled burst relative to the average energy at the sampled vowel midpoint was employed. As a result, contrary to our expectation, burst energy was not particularly stronger for alveolar stops and in fact it was stronger for bilabial stops in general although the differences were non-significant (see Table 6 and Figure 8). This indicates that the stop burst measurements do not assist in explaining the perceptual data.

| | Lenis | Fortis | Fortis | Aspirated | Lenis | Aspirated |
|----------|------------|-----------|-----------|------------|-----------|------------|
| Bilabial | 81-74.9 | 89.5 -91 | 99 -91.8 | 87.2 -85.7 | 86.4-79.1 | 84.9 -80.8 |
| Alveolar | 81.2- 82.8 | 89.2 - 90 | 94.6-90.4 | 79.5 -80.1 | 82.2-84.6 | 82.9 -76.8 |
| Velar | 84.7-79.1 | 85.6-91.9 | 82 - 91.9 | 84.4 - 84 | 82.6-73.9 | 79 - 80.1 |

Table 6. Relative burst energy (%)

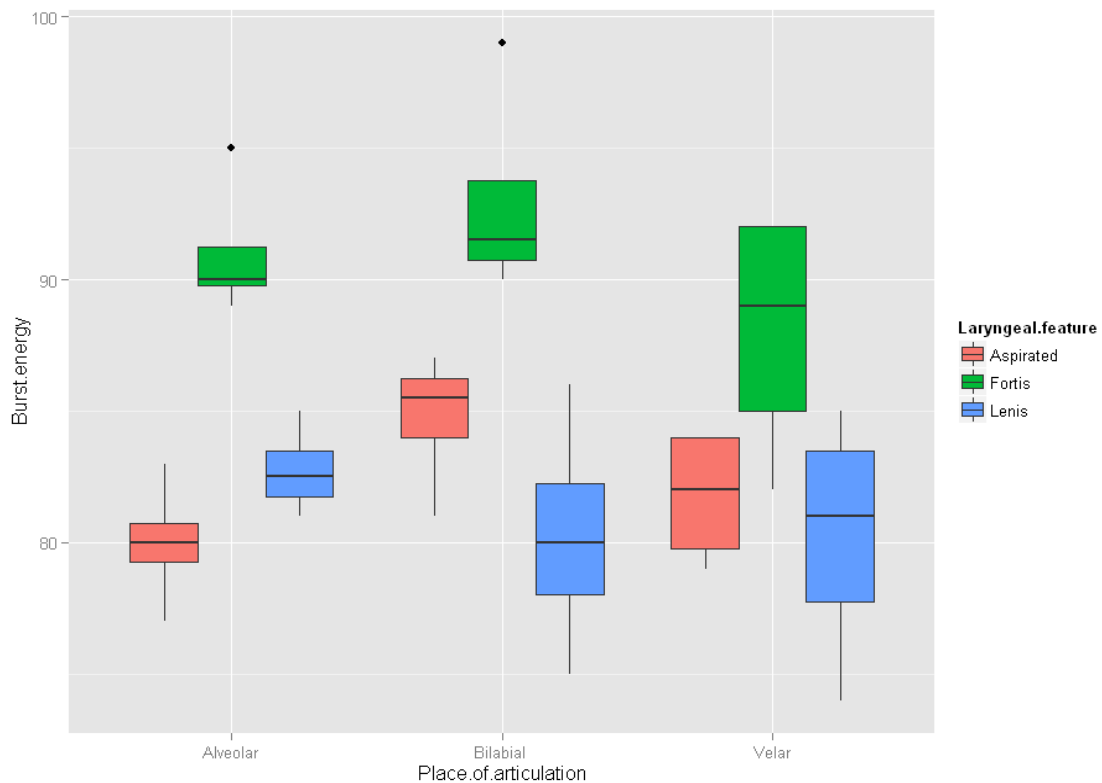


Figure 8. Burst energy range at different places of articulation.

4. Discussion

Setting aside this unanswered conflicting behaviour of the alveolar stop, the current study makes two general acoustic observations regarding the three-way laryngeal contrast of Korean stops. First, VOT in combination with f_0 can successfully distinguish the three stop categories. As shown in Figure 1, f_0 separates a cluster of the lenis and fortis stops which shares similar VOT values, and VOT separates the f_0 cluster of the fortis and aspirated stops. Second, VOT is the primary cue that English speakers attend to discriminate the Korean stops as shown in our discrimination task where the lenis-fortis contrast was discriminated worse than the other two contrasts. The two points above support our prediction that English speakers rely on the acoustic cue that is strongly distinctive in English only when making perceptual decisions across the three Korean stop categories.

To strengthen the hypothesis, let us now re-address all of the present counterexamples that are found at the alveolar region. As noted in section 3.3, English listeners surprisingly showed above chance level of correct discrimination on the alveolar lenis–fortis stops, which do not have a significant VOT difference, $t(1) = 2.68$, $p = 0.22$. The unusual judgements to the alveolar stops is also found on the lenis-aspirated contrast, $t(1) = 85$, $p < 0.01$, in a reversed manner. The discrimination levels of the alveolar lenis-fortis and lenis-aspirated stops clearly contradict our hypothesis and suggest that VOT difference and English listeners' correct discrimination are inversely related. Why only alveolar shows this conflicting behaviour? It could be that the immediate phonetic environment of the present stimuli may be interfering English listeners' perceptual judgements⁶. As is apparent in the stimulus set, the alveolar stops have additional syllable-final /l/ due to the limited availability of existing Korean words. In practice, when /l/ occurs syllable finally in English, the back of the tongue is raised toward the soft palate and the second formant (F_2) is decreased (Johnson, 2011; Ladefoged & Johnstone, 2011). However, Korean syllable final /l/ does not undergo velarisation, and in fact retains F_2 frequencies even higher than those for syllable-initial English /l/, as a comparison of the F_2 range of the Korean syllable-final /l/ in the stimuli (1739 Hz - 2087 Hz) with mean F_2 of English syllable-initial /l/ in earlier works (Nolan, 1983, 1350 Hz; Lehiste, 1962, 980 Hz; Ladefoged & Johnstone, 2011, 1100Hz) clearly demonstrates. On this basis, the high tongue dorsum position required by clear /l/ in Korean may contribute the high dorsum position of lenis /t/ and this may strengthen the release characteristics of lenis category. In this high dorsum context, English listeners may hear the VOT of the lenis category more clearly and it possibly results in rather unusual perceptual judgments at the alveolar region. Alternatively, it seems also possible that English listeners considered Korean syllable-final /l/ as an additional foreign sound that they need to focus on and it disturbed their judgement by directing their attention away from the syllable-initial stop contrasts. In other words, non-native listeners may not make most effective use of VOT for the stop distinction when the

⁶ The different vowel context does not seem to be a factor because Korean /ə/ was not significantly different from the other two vowels, /i/ and /ε/, in terms of vowel duration.

contrasts in question are in foreign phonetic contexts. For example, for the alveolar lenis-aspirated contrast, where VOT can be the primary cue, the discrimination level may be significantly decreased due to a masking effect caused by more than one phonetic cue accessible: a conflict between VOT of the syllable-initial stops and F_2 of the syllable-final /l/. On the other hand, the above chance level of correct discrimination on the alveolar lenis-fortis contrast, where VOT is not a salient acoustic cue, may be increased because the additional phonetic cue in the immediate environment (i.e., F_2 of the syllable-final /l/) does not create a conflict with VOT. In terms of the fortis-aspirated contrast, English listeners' discrimination was not affected by the presence of the foreign syllable-final /l/ unlike in the other two alveolar contrasts. This may be possible because the alveolar fortis-aspirated stops have the greatest magnitude of VOT differences and thus, cannot be disturbed. Accordingly, it seems plausible that the unusual discrimination level of the alveolar stops may relate to the influence of the syllable-final /l/. However, its direct evidence is lacking because the current experiment was limited to eliciting same-different responses only. Clearly, more work is needed to explore this interesting possibility more thoroughly. For example, in the future study, one can ask English listeners to write the closest English alphabetic symbol for each consonantal segment in the alveolar stimuli and examine their labelling responses. By doing so, the influence of /l/ can be clearly separated out. Without this direct measurement from orthographic labelling task, it will not be possible to verify the influence of word-final /l/ on English listeners' perceptual decision.

5. Conclusion

As verified in the discriminant function analysis in section 3, VOT in combination with f_0 onset can discriminate perfectly between the three Korean stop categories. On this basis, it was hypothesised that English listeners may not be sensitised to the f_0 cue. This successfully explains the current perceptual results that contradict the expectation of the Perceptual Assimilation Model (PAM), where the lowest discrimination was predicted to be found in the lenis-aspirated contrast based on Schmidt (2007)'s study. But, the present hypothesis still does not fully account for why English listeners did better than chance on the alveolar fortis-lenis contrast.

To search for the reasonable explanations, the relative intensity of the burst at release was measured but it did not help to explain this unpredicted perceptual pattern. Another possible reason was sought under the assumption that non-native listeners do not make most effective use of the cue available (i.e., VOT) when exposed to foreign sound contrasts. Their discriminations may be disturbed by the additional foreign sounds in context. However, it is not yet known on which segment they made their perceptual decisions. To ascertain this syllable-final /l/ effect, one needs to collect identification responses from subjects as a follow-up study.

In sum, despite the limitation of this study, which dealt with perceptions only based on single-token contrasts, it agrees with the results of Francis & Nusbaum (2002)'s study where English listeners' discrimination of the corresponding Korean stops in artificial monosyllabic words was examined. Also, it generally answers the question of whether English listeners attend to VOT only when discriminating the three-way contrast of Korean stops. Bear in mind that there were exceptions in the alveolar stimuli; It can be concluded that English listeners are not sensitized to the f₀ cue of the Korean stops from the fact that English listeners showed the highest level of correct discrimination on the aspirated-fortis contrast and the opposite to the fortis-lenis contrast.

Acknowledgement

I am deeply indebted to Dr. John Ingram for his helpful comments and advice on earlier versions of this work. I also thank to two anonymous reviewers for their comments and suggestions in the preparation of this article. The responsibility for any remaining errors is entirely mine.

References

- Best CT, GW McRoberts & N Sithole 1988 'The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants' *Journal of Experimental Psychology: Human Perception and Performance* 14(3): 345-360.
- Broersma M 2010 'Korean lenis, fortis, and aspirated stops: Effect of place of articulation on acoustic realization' Paper presented at the Proceedings of the 11th Annual Conference

- of the International Speech Communication Association (Interspeech 2010). Makuhari, Japan.
- Chang CB 2010 ‘The implementation of laryngeal contrast in Korean as a second language’ *Harvard studies in Korean linguistics* 13: 91-104.
- Cho T, SA Jun & P Ladefoged 2002 ‘Acoustic and aerodynamic correlates of Korean stops and fricatives’ *Journal of Phonetics* 30(2): 193-228.
- Dalgaard P 2008 *Introductory statistics with R* New York: Springer.
- Dart SN 1987 ‘An aerodynamic study of Korean stop consonants: Measurements and modeling’ *The Journal of the Acoustical Society of America* 81(1):138-147.
- Francis AL & HC Nusbaum 2002 ‘Selective attention and the acquisition of new phonetic categories’ *Journal of Experimental Psychology: Human Perception and Performance* 28(2):349-366.
- Han N 1998 *A comparative acoustic study of Korean by native Korean children and Korean-American children* Unpublished PhD thesis, University of California, Los Angeles.
- Hardcastle WJ 1973 ‘Some observations on the tense-lax distinction in initial stops in Korean’ *Journal of Phonetics* 1(3): 263-272.
- Johnson K 2011 *Acoustic and auditory phonetics* Cambridge, Mass: Blackwell Publishers:
- Kagaya R 1974 ‘A fiberoptic and acoustic study of the Korean stops, affricates and fricatives’ *Journal of Phonetics* 2(2): 161-180.
- Kim CW 1965 ‘On the autonomy of the tensity feature in stop classification (with special reference to Korean stops)’ *Word* 21(3): 339-359.
- Kim MR, PS Beddor & J Horrocks 2002 ‘The contribution of consonantal and vocalic information to the perception of Korean initial stops’ *Journal of Phonetics* 30(1): 77-100.
- Kim MRC 1994 *Acoustic characteristics of Korean stops and perception of English stop consonants* Unpublished PhD thesis, University of Wisconsin, Madison.
- Kim M 2004 ‘Correlation between VOT and F0 in the perception of Korean stops and affricates’ Paper presented at the INTERSPEECH 2004. Jeju Island, 4-8 October.
- Kong EJ, ME Beckman & J Edwards 2011 ‘Why are Korean tense stops acquired so early?: The role of acoustic properties’ *Journal of Phonetics* 39(2): 196-211.
- Ladefoged P & K Johnstone 2011 *A course in phonetics* Boston, MA: Wadsworth/Cengage Learning.
- Lee SAS & GK Iverson 2012 ‘Stop consonant productions of Korean–English bilingual children’ *Bilingualism: Language and cognition* 15(02): 275-287.
- Lehiste I 1962 *Acoustical characteristics of selected English consonants* Bloomington: Indiana University.
- Lisker L & AS Abramson 1964 ‘A cross-language study of voicing in initial stops: Acoustical measurements’ *Word* 20(3): 384-422.
- Nolan F 1983 *The phonetic bases of speaker recognition* Cambridge; New York: Cambridge University Press.
- Schmidt AM 2007 ‘Cross-language consonant identification: English and Korean’ in OS Bohn & MJ Munro (eds) *Language experience in second language speech learning: In honor of James Emil Flege* Amsterdam: John Benjamins Publishing. pp. 185-200.
- Stevens KN, SJ Keyser & H Kawasaki 1986 ‘Toward a phonetic and phonological theory of redundant features’ in J Perkell & DH Klatt (eds) *Invariance and variability in speech processes* Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 426-449.



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Kwon, Nahyun

Title:

Acoustic observation for English speakers perception of a three-way laryngeal contrast of Korean stops

Date:

2014

Publication Status:

Published

Persistent Link:

<http://hdl.handle.net/11343/40957>

File Description:

Acoustic observation for English speakers' perception of a three-way laryngeal contrast of Korean stops