말소리와 음성과학 제2권 제3호 (2010)
pp. 3~10

Generational Differences in the Perception of Korean Stops

Kang, Kyoung-Ho[1)]

ABSTRACT

The proposal that a sound change is occurring in Korean stops was evidenced in this study through identification experiments on Korean stops. Perceptual weight of acoustic correlates to Korean stop manner contrast [VOT (Voice Onset Time), H1-H2 (amplitude difference between the first and second harmonics), and F0 (Fundamental frequency)] was examined with re-synthesized /t^{h}a/, /ta/, and /t*a/ syllables for younger and older Seoul speakers of Korean. For the identification of the aspirated and lenis stops, F0 cue weight relative to VOT was greater for the younger listeners than the older listeners. For H1-H2 cue weight, the two listener groups were more or less the same. These findings were parallel to the production differences found in the earlier work of the author. Combined with production differences, these perception differences between younger and older generations of Seoul speakers suggested that there are generational differences in the phonetic targets of Korean aspirated and lenis stops and such differences are realized in the perception of the stops.

Keywords: generational difference, Korean stops, perceptual cue weight

1. Introduction

As a follow-up study of Kang & Guion (2008), the current study aims to investigate whether speech production relates to speech perception with respect to the sound change in the Korean stop system. As reported in Silva et al. (2004), Silva (2006), Wright (2007), and Kang & Guion (2008), it has been proposed that for young Seoul speakers, VOT for aspirated stops has reduced over the past two generations(also, VOT for lenis increased) and the distinction between aspirated and lenis stops, which has traditionally been rendered, at least partly, by VOT, has come to be coded primarily by a F0 difference on the vowels following the stops.

In the literature on Korean stops, the insufficiency of VOT alone to the distinction of the three-way contrast of Korean has been noted in the earlier studies and in turn, the voice characteristics of the following vowels, such as F0, intensity (indexed by H1-H2) have been suggested as supplementary cues to the stop contrast (Lisker & Abramson, 1964; C.-W. Kim, 1965; Han & Weitzman, 1970). More recently, M.-R. Kim et al. (2002) investigated the importance of consonantal and vocalic information in the perception of Korean initial stops. The finding was that neither consonant information alone nor vocalic information alone provided sufficient cues for the three-way distinctions, but low F0 of the vowel following lenis stop provided a dominant cue over H1-H2 or VOT for the two-way distinction of lenis fromfortis or aspirated stops. Also, in a study investigating attention weight of VOT and F0 for the perception of /t^{h}a/ and /ta/ syllables by Seoul speakers in the early 20s of age, Y.-H. Kim (2007) reported that the attention weight was greater for F0 than VOT.

As perception studies on Korean stops have shown until recently, attention to the vocalic information (that is, F0 and/or H1-H2) for the Korean stop manner distinction is not new. However, to my best knowledge, attention to the diachronic change in VOT values is only recent, and a few investigations of this matter are available to date (Silva et al., 2004; Silva, 2006; Wright, 2007). The primary goal of the current study is to further

1) Kyungsung University kkang245@gmail.com, corresponding author
Portions of this paper are based on the author's doctoral dissertation (Kang, 2009)

Received: August 1, 2010
Revision: September 1, 2010
Accepted: September 2, 2010

examine the findings from the production experiments in Kang & Guion (2008) as they relate to speech perception. The results of thetwo production experiments suggested that F0 is a stronger correlate than VOT for the aspirated-lenis stop distinction for younger Seoul speakers, whereas F0 is a secondary correlate, with VOT being primary, for older Seoul speakers. In hyperarticulated, clear speech production of aspirated and lenis stops, the F0 difference was more enhanced than the VOT difference for younger speakers, whereas only VOT was enhanced for older speakers. These production differences were interpreted to mean that the two groups had different VOT and F0 phonetic targets for the production of aspirated and lenis stops and that the different phonetic targets induced different clear speech enhancement patterns.

Considering the previously reported production differences related to VOT and F0 between the younger and older Korean speakers (Kang & Guion, 2008), the focus of the current study was directed to the aspirated and lenis stop contrast. Given that F0 was a stronger correlate than VOT for younger speakers whereas VOT was stronger for older speakers in the production data, this study examined whether the production differences between the two speaker groups lead to perceptual differences in a parallel manner. Two logical predictions were consideredfor the examination of potentially differing perceptual weight of VOT and F0. One was that F0 is a more heavily weighted perceptual correlate than VOT for younger listeners for the aspirated-lenis stop distinction, whereas VOTis more heavily weighted than F0 for older listeners. The other was that F0 is relatively more weighted for younger listeners than older listeners for the stop distinction without a weighting reversal between F0 and VOT as stated in the first prediction. In order to respond to these questions, the speech tokens of the stop productions were manipulated for the acoustic correlates of Korean stop manner contrasts (VOT, H1-H2, and F0). The three acoustic correlates were included for the analyses of the current study as a set, although VOT and F0 is more closely relevant for the distinction as long as the aspirated and lenis contrast is concerned (see Cho et al, 2002; Kang & Guion, 2006, 2008 for the similar H1-H2 values of Korean aspirated and lenis stops). Also, as was the case for the production experiments, listeners were separated into two age groups: younger and older.

2. Perception experiment with re-synthesized stimuli

2.1 Participants

A total of 20 native speakers of Korean participated in a forced-choice identification task. 10 listeners, 20 to 29 years old were assigned to "younger group" (mean = 20.2 years), and another 10 listeners over 40 years old were assigned to "older group" (mean = 48.5 years). All of the listeners were speakers of the Seoul or Kyung-gi province dialect. The younger listeners were students at Hanyang University (HU) in Korea, and the older listeners were staff members at HU or their family members. At the time of the task, none of the listeners reported any history of language or hearing disorders.

2.2 Perception stimuli

A 36 year-old male speaker of the Seoul dialect produced the source recordings for the re-synthesized stimuli. 11 tokens of 다 /ta/, 타 /t^{h}a/, and 따 /t*a/ syllables were produced. The source speech signals were recorded using a high-quality headset microphone and a Marantz digital recorder (PMD 670) at the sampling rate of 22,500 Hz. The recordings were made in a sound-attenuated booth at the University of Oregon, U.S.A. Adobe Audition 2.0 was used for the editing work.

The alveolar stops plus the low vowel /a/ syllables were submitted to the acoustic measurement of VOT, H1-H2, and F0. Three tokens of /ta/, /t^{h}a/, and /t*a/ were chosen as bases for re-synthesis based on the typicality of H1-H2 values, which are a primary acoustic correlate of voice quality. The selected /ta/, /t^{h}a/, and /t*a/ tokens had the values of 1.5, 6.5, and -3.6 dB at the vowel onset, respectively. These values demonstrate distinctive voice qualities of modal or breathy voicing for lenis /ta/ and aspirated /t^{h}a/ types (1.5 and 6.5 dB), and tense or creaky voicing for fortis /t*a/ type (-3.6 dB). So, in terms of an H1-H2 scale, there were three levels.

To equate the amplitude difference among the bases, the peak intensity of vowel portion of three base tokens were normalized to 50% using Audacity (ver.1.2.6). The entire stretch between stop release burst and the beginning of the first periodic cycle of the vowel was removed from the three base tokens. Then, the vowel portions were equated to 260 ms for the three base tokens by truncating the end of the vowel durations. 260 ms was chosen because the shortest duration of clear formant structure among the three tokens was 260 ms (for the /t*a/ base). The original vowel durations were 303 ms, 286 ms, and 267 ms for the /ta/, /t^{h}a/, and /t*a/ bases, respectively. The vowel for the /t*a/ base showed

a good amplitude envelope in that it tapered off at the end of the vowel and sounded natural. The other two base tokens were manipulated so that they had quite similar amplitude envelopes and tapered off at the end of the vowel. With the amplitude envelope function of Audacity (ver.1.2.6), the last 60 ms of the vowel was ramped down from 100% to 50 % (/t^ha/ base) or 40% (/ta/ base) of the intensity of the original stimuli.

Following this, consonantal portions (release burst and aspiration to the beginning of the first periodic cycle) were conjoined to the three vowel bases prepared. The base for the consonantal portion was taken from another /t^ha/ token, which had an original VOT of 91 ms. Six different consonant portions were prepared by shortening the 91 ms VOT from the end (i.e. right-hand side) with a 15 ms interval. Thus, the stimuli had six different VOT levels: 10, 25, 40, 55, 70, and 85ms.

As the last step, the vowel portion was manipulated for F0 at four levels: 110, 120, 130, and 140 Hz (using Praat, ver. 4.6). The F0 levels were held constant throughout the entire duration of the vowel. The original F0 was 115, 144, and 128 Hz at the vowel onset for the base /ta/, /t^ha/, and /t*a/ tokens, respectively. Through the entire re-synthesis process, 72 stimuli types (three H1-H2 levels × six VOT levels × four F0 levels) were created in total.

2.3 Procedure

Each of the 72 stimulus type was presented 10 times, 1 time each in 10 randomized blocks, eliciting 720 responses from each of the 20 listeners. The three response categories were provided in Korean orthography, such that 다 /ta/, 타 /t^ha/, and 따 /t*a/ were displayed on a computer screen for each of the 720 stimuli. Listeners were instructed to choose one of these three categories by clicking a mouseafter listening to the stimuli delivered to both ears through a headset. Because most of the test stimuli were phonetically ambiguous, the listeners were told that the stimuli may not sound like any of the three response choices and they use their instant impressions to make a decision. The Multiple Forced Choice (MFC) function of Praat for speech perception experiment was used to administer the perception task. The task progressed at the listeners' own pace and the participants were informedthat there was no time limit to complete the task. Listeners were offered short breaks between the blocks. 30 stimuli were presented during a practice session. On average each listener used about 35 to 40 minutes to complete the task.

2.4 Statistical design

In order to assess potential differences between younger and older listeners in the perceptual weight of VOT, H1-H2, and F0 for the distinction of aspirated and lenis stops, two sets of statistical analyses were performed. First, a multivariate analysis of variance (MANOVA) was performed to examine the overall group (younger and older) differences for the distinction of the aspirated and lenis stops. The number of /ta/ response and the number of /t^ha/ response for each of the 720 stimuli for each listener were entered as dependent variables. The number of fortis /t*a/ response was not used as a dependent variable. The interest of analysis was in the /t^ha/ - /ta/ contrast. Moreover, the addition of /t*a/ response makes the sum of 10 for the entire 10 times-repetition over the three response categories for each stimulus, and this made the statistics senseless. The three acoustic dimensions of the test stimuli, VOT, H1-H2, and F0, in addition to group, were entered as independent variables. As a result, a four-way multivariate analysis with two dependent variables was performed.

After investigating overall potential group differences in the perceptual weight of the three acoustic properties, logistic regression analyses were performed to further examine group differences (Morrison, 2007). Three binary logistic regression analyses were performed for each younger and older listener group for the three response categories in a pairwise manner. The purpose of the analyses was to determinethe relative perceptual weights of VOT, H1-H2,and F0 in the distinction of the aspirated and lenis contrast of Korean stops. The number of /t^ha/ or /ta/ response for the 720 stimuli was entered as a dependent variable for each listener. That is, the binary values, /ta/ or /t^ha/ response (/ta/ coded as 0 and /t^ha/ coded as 1) were entered as an outcome variable in the regression analysis. The predictor variables were VOT, H1-H2, and F0 (the six levels of VOT were encoded in the sequence of 10, 25, 40, 55, 70, and 85 ms; the three levels of H1-H2 were encoded in the sequence of 6.5, 1.5, and -3.6 dB; the four levels of F0 were encoded in the sequence of 110, 120, 130, and 140 Hz).

3. Results

3.1 Results of MANOVA

A MANOVA with factors, VOT (10, 25, 40, 55, 70, 85ms), H1-H2 (-3.6, 1.5, 6.5 dB), F0 (110, 120, 130, 140 Hz), and group (younger, older) returned significant main effects for all of the four independent variables (VOT [$F(10, 2590) = 128.391$, $p <$

0.001]; H1-H2 [F(4, 2590) = 513.847, p < 0.001]; F0 [F(6, 2590) = 315.172, p < 0.001], and group [F(2, 1295) = 28.393, p < 0.001]). In addition, three-way interactions of group × VOT × F0 [F(30, 2590) = 1.485, p = 0.044] and group × H1-H2 × VOT [F(20, 2590) = 1.826, p = 0.014] were found, as well as significant interactions for all of the two-way combinations of the four factors. The four-way interaction of the four factors was not significant (p = 0.986). The three-way interactions indicate that the younger and older groups differed in the manner in which the level of VOT affected response choice depending on the level of F0. Likewise, the two groups differed in the manner in which the level of VOT affected response choice depending on the level of H1-H2.

Since the focus of the analysis was on the investigation of potential group difference in VOT and F0 weightfor the distinction of the aspirated and lenis stops with respect to the sound change proposal, the group × VOT × F0 interaction was only further examined in this study. The content of the significant group × VOT × F0 interaction was plotted in the following four figures.The four figures were plotted individually according to the four F0 levels. Each figure shows the number of /t^{h}a/ and /ta/ responses varying as function of VOT levels for both younger and older groups. Thus, the four figures, as a set, show an interaction between VOT and F0 for the choice of /t^{h}a/ and /ta/ response with the two speaker groups compared with each other. Specifically, Figure 1 shows the mean number of /t^{h}a/ and /ta/ response for each of the 10 younger and older listeners for the stimuli with 110 Hz varying in VOT values. Figures 2, 3, and 4 represent the mean number of responses in the same manner for the stimuli with 120, 130, 140 Hz, respectively.

Let us begin with Figure 1, which represents responses for stimuli with 110 Hz. With this low F0, the number of /ta/ response was greater than /t^{h}a/ response at all of the six VOT levels for both younger and older listeners.For stimuli with long VOTs, 55, 70, 85ms, /ta/ response decreased, whereas /t^{h}a/ response increased. However, /ta/ response was still greater than /t^{h}a/ response for the stimuli with long VOTs. These results suggest that the effect of the low F0 held throughout the entire VOT range.

In addition to the role of F0 distinguishing aspirated and lenis stops, Figure 1 also shows a group difference in the degree of F0 effect on the distinction of the two stop types. Younger listeners (solid line with filled circles) had greater /ta/ response than older listeners (dotted line with empty circles) for all VOT levels. This suggests that younger listeners were more affected by the low F0 than older listeners and identified more tokens as lenis stops overall at this low F0.

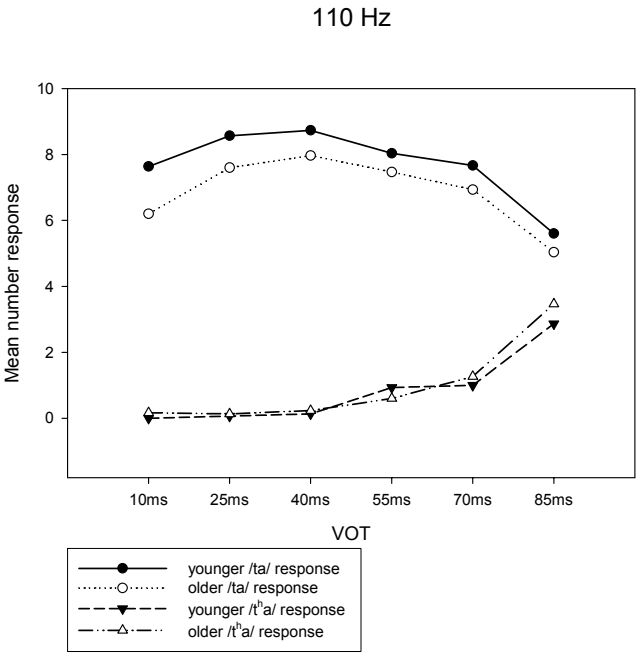

Figure 1. Mean number of response, collapsed across H1-H2, to /ta/ and /t^{h}a/ stimuli with 110 Hz F0 for younger (n=10) and older (n=10) listeners.

For the response to stimuli with 120 Hz (Figure 2), the number of /ta/ response was lesser and the number of /t^{h}a/ response was greater overall, compared with the responses to stimuli with 110 Hz. The responses were more strongly affected by VOT than was the case with the 110 Hz stimuli. In addition, the two listener groups differed in the extent to which they responded to the increased F0. For the younger listeners, as VOT increased, /t^{h}a/ response increased more rapidly and at the same time, /ta/ response decreased more rapidly than was the case for older listeners. This suggests that the younger listeners were more greatly affected by the F0 increase. Also, the crossover points for /t^{h}a/ vs. /ta/ response were different between the two groups. The crossover point for younger listeners was at a lower VOT value (about 55 ms) than for older listeners (about 80 ms). These results indicate that for stimuli with shorter VOTs, the younger listeners gave more /t^{h}a/ responses. In other words, older listeners needed stimuli with longer VOTs to give /t^{h}a/ response.

For the stimuli with 130 Hz (Figure 3), the patterns seen for stimuli with 120 Hz were largely repeated. Overall, with this increased F0, more /t^{h}a/ responsesand less /ta/ responses were given by both listener groups compared with the stimuli with the lower F0 values. Also, the increase of /t^{h}a/ response and the decrease of the /ta/ response over the VOT range were more rapid. As for the

group difference, the crossover point was earlier for the younger listeners than for the older listeners, indicating that the younger listeners were more greatly affected by the increased F0. For both listener groups, the crossover points were made at lower VOT values compared with the responses to stimuli of 120 Hz. However, as was the case of stimuli with 120 Hz, the crossover point for younger listeners was at a lower VOT value (about 30 ms) than older listeners (50 ms).

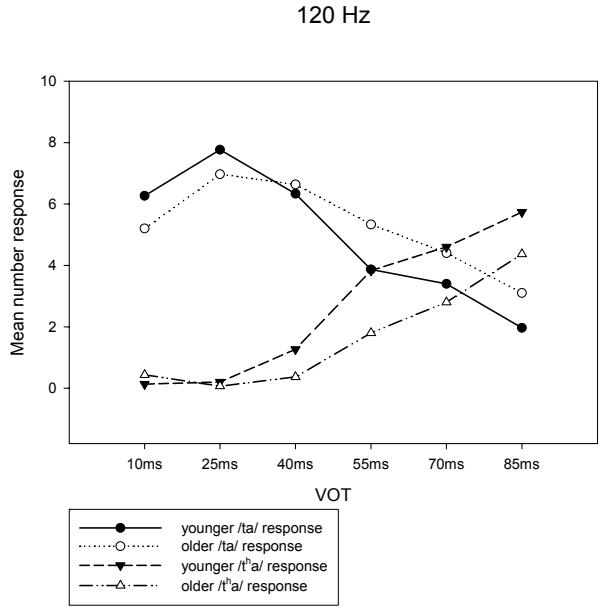

Figure 2. Mean number of response, collapsed across H1-H2, to /ta/ and /t^{h}a/ stimuli with 120 Hz F0 for younger (n=10) and older (n=10) listeners.

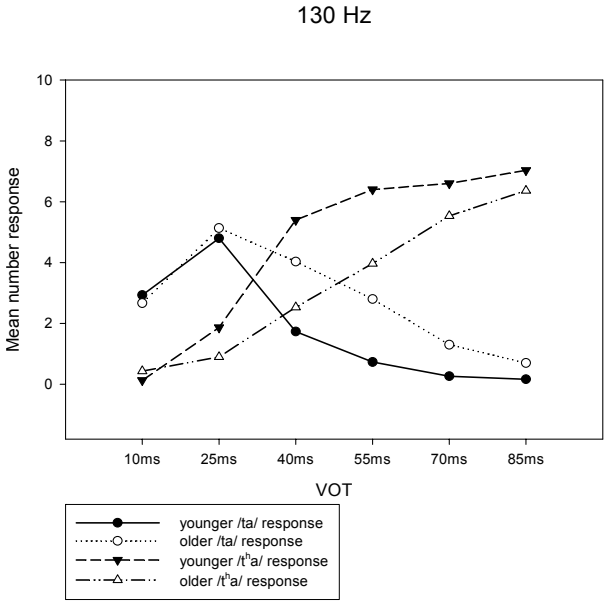

Figure 3. Mean number of response, collapsed across H1-H2, to /ta/ and /t^{h}a/ stimuli with 130 Hz F0 for younger (n=10) and older (n=10) listeners.

The response to stimuli with 140 Hz (Figure 4) was more or less the inverse of the response to stimuli with 110 Hz. The number of /t^{h}a/ response was greater than /ta/ response for nearly the entire VOT range, especially for younger listeners. This result replicates the F0 effect found with stimuli with 110 Hz in an inverse way. With stimuli of 110 Hz, /ta/ response was greater than /t^{h}a/ response at all VOT levels. In addition, just as the other three Figures did, Figure 4 also shows a group difference. Note that, for both younger and older listeners, for the stimuli with 120 and 130 Hz, the 25 ms was a point where the greatest /ta/ response was induced and thus, /ta/ responses were distinctively greater than /t^{h}a/ responses at 25 ms (See Figures 2 and 3). However, for stimuli with 140 Hz, even at 25 ms, /t^{h}a/ responses were greater than /ta/ responses for younger listeners. In addition, the crossover point at a lower VOT level compared with older listeners was repeated with the stimuli with 140 Hz. These results suggest that younger listeners gave more /t^{h}a/ responses to the stimuli with short VOTs than older listeners for the stimuli of 140 Hz.

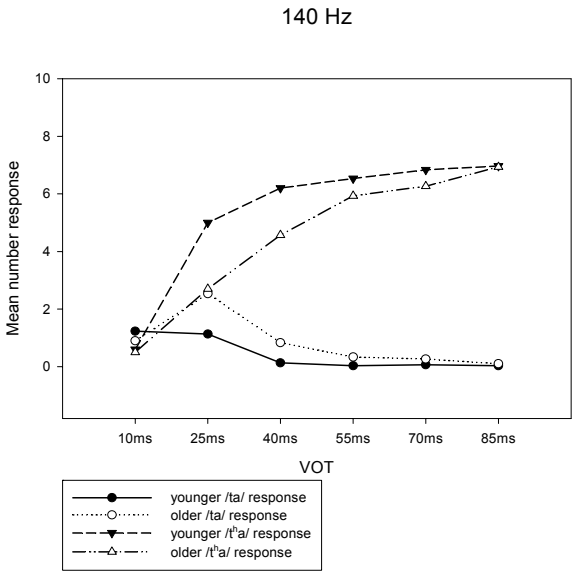

Figure 4. Mean number of response, collapsed across H1-H2, to /ta/ and /t^{h}a/ stimuli with 140 Hz F0 for younger (n=10) and older (n=10) listeners.

Examining all of the four figures together, one notes that, as F0 increased, the younger listeners gave more /t^{h}a/ responses than older listeners for the stimuli with short VOTs. As can be seen from Figures 2, 3, and 4, as F0 increased the crossover points (where more /t^{h}a/ responses began) shifted left, to the lower VOTs, for both groups. This indicates that, as the F0 of stimuli increased, the listeners gave more and more /t^{h}a/ responses for the stimuli with shorter VOTs. However, at the same time, the

crossover points for younger listeners were consistently to the left of those for older listeners, that is, at the lower VOT levels. These results suggest that younger listeners were more greatly affected by the F0 modulation than older listeners for the distinction of Korean aspirated and lenis stops.

3.2 Results of logistic regression analyses

The results from the multivariate analysis indicated that the younger and older speakers were affected by VOT and F0 modulation to a different degree with respect to the distinction of the aspirated and lenis stops. Younger speakers were more affected by F0 modulation than the older speakers for the /t^{h}a/ and /ta/ distinction. Logistic regression analyses exhibit this group difference in more detail by showing the relative contribution of the three acoustic correlates.

Let us begin with younger speakers. A test of the full model with all of the three predictors (VOT, H1-H2, and F0) against a constant-only model was significant, χ^2 (3) = 3870, $p < .05$. This indicates that the three predictors reliably distinguished /t^{h}a/ and /ta/ responses with a good model fit, Nagelkerke R^2 = .71. With all of the three predictor variables, correct classification rates were 88% for /ta/ response and 85% for /t^{h}a/ response. The overall correct classification rate was 86%.

For older speakers, as was the case of younger speakers, a test of the full model with all three predictors against a constant-only model was significant, χ^2(3) = 2677, $p < .05$. This indicates that the three predictors reliably distinguished /t^{h}a/ and /ta/ responses with a good model fit, Nagelkerke R^2 = .59. With all of the three predictor variables, correct classification rates were 87% for /ta/ response and 75% for /t^{h}a/ response, with an overall correct classification rate of 82%.

Table 1 shows the relative contribution of the individual predictors to the model. For both younger and older listeners, VOT and F0 were significant predictors of the outcome. H1-H2 was only significant for younger listeners. The interesting difference was that the odds ratio for F0 was 9.6 for younger listeners and was 4.8 for older listeners. This result indicates that for the younger listeners, the likelihood of /t^{h}a/ response was 9.6 times greater after one-unit increase inF0 (10 Hz in the current stimuli) compared with before the change. By the same token, it means that the likelihood of /t^{h}a/ response for the older listeners was 4.8 times greater after one-unit increase in F0. Therefore, for a given increase in F0, the likelihood of /t^{h}a/ response by younger listeners was by far greater than the likelihood of /t^{h}a/ response by older listeners.

The results also showed that for both groups, the likelihood of /t^{h}a/ response after one-unit increase in VOT increased. The odds ratio was 3.2 for younger listeners and 2.5 for older listeners, indicating that the likelihood of /t^{h}a/ response was 3.2 times greater for younger listeners and 2.5 times greater for older listeners after one-unit VOT increase. These odds ratios for VOT are not as different from each other as the odds ratios for F0. This indicates that compared with the effect of F0 modulation, younger and older listeners were not much different from each other for the effect of VOT modulation on the perception of the /t^{h}a/ and /ta/ types, even though both groups were affected by the VOT modulation.

In addition, older listeners were not affected by the H1-H2 modulation for the distinction of the /t^{h}a/ and /ta/ categories, whereas younger listeners were affected. Table 1 (the Wald statistic column) shows that H1-H2 was not a significant contributor to the distinction of the /t^{h}a/ and /ta/ categories for older listeners. However, the odds ratio for H1-H2 was around 1 for both listener groups, which indicates that the likelihood of choosing one response category over the other almost stays the same. In addition, the difference was minimal: 0.7 for younger and 0.9 for older listeners. These results indicate that the younger and older listeners were not much different from eachother for the likelihood of /t^{h}a/ response after one-unit H1-H2 decrease in the stimuliThis result coincides with the production data reported in Kang & Guion (2008), which showed no difference between younger and older groups for the H1-H2 enhancement patterns in hyperarticulated, clear speech production of aspirated and lenis stops.. Thus, the combined results from VOT, H1-H2, and F0 indicate that the most prominent difference between younger and older listeners resides on the contribution power of F0 to the distinction of /ta/ and /t^{h}a/ types.

Table 1. Contribution of three predictors to the perception of /ta/ and /t^{h}a/

Group	Predictor variables	*B* (*SE*)	Wald (*df* = 1)	Odds Ratio (*Exp*(*B*))
Younger	VOT	1.17 (.04)	834.3*	3.2
	H1-H2	-.32 (.06)	26.9*	0.7
	F0	2.26 (.07)	1207.1*	9.6
	Constant	-9.25 (.29)	1001.5*	0.000
Older	VOT	0.91 (.03)	753.8*	2.5
	H1-H2	0.09 (.06)	1.8	0.9
	F0	1.57 (.05)	1040.2*	4.8
	Constant	-7.73 (.25)	972.7*	0.000

(*B*, the logistic coefficient represents the log of the odds of an event occurrence for one-unit change in the predictor variable. Wald statistic tells whether the predictor in question make a significant contribution to the outcome. The odds ratio (*Exp* (*B*)) represents the relative strength of the variables for the prediction of the outcome. An odds ratio of greater than 1 indicates that the odds of event occurrence increases when one-unit change is made in a predictor variable. An odds ratio of less than 1 indicates that the odds of event occurrence decreases when one-unit change is made in a predictor variable. Asterisks indicate significant effects, $p < .05$.)

3.3 Discussion

The current studyinvestigated whether the three acoustic correlates of manner contrast in Korean stops (VOT, H1-H2, and F0) were perceptually differently weighted between younger and older listeners of the Seoul dialect, focusing on the aspirated and lenis stop contrast. The major finding is that F0 was relatively more heavily weighted for younger listeners than older listeners. The MANOVA analyses (see Figures 2, 3, and 4) showed that the VOT threshold for /t^{h}a/ response was lower for younger listeners than older listeners consistently at the F0 levels, indicating a greater F0 effect on the /t^{h}a/ - /ta/ identification for younger listeners. The logistic regression analyses returned similar results. For the distinction of /t^{h}a/ - /ta/ contrast, the contribution of F0 was greater for younger listeners than older listeners. The findings also included that the contribution of H1-H2 was more or less the same to the perceptual distinction of the /t^{h}a/ - /ta/ contrast.

The findings reported in this study can be considered in light of the production differences between the younger and olderspeakers. In Kang & Guion (2008), based on the sound change proposed to be occurring among younger generation of the Seoul dialect of Korean (Silva, 2006; Wright, 2007), different clear speech enhancement patterns were predicted for the production of Korean aspirated and lenis stops: younger speakers werepredicted to enhance F0 difference in clear speech, whereas older speakers were predicted to enhance VOT difference. The results of the production experiments revealed that younger speakers primarily enhanced F0 difference between the two stops with a marginal VOT enhancement, whereas older speakers enhanced the VOT difference with no enhancement in F0. The findings from the current study suggest that these production differences between the two groups are related to different perceptual weights of F0 between the two listener groups.The younger speakers primarily enhanced F0 difference of aspirated and lenis stops in clear speech production, and their perception of the aspirated and lenis stops were more greatly affected by the F0 modulation compared with the older listeners.The older speakers, who showed the enhancement of VOT difference in clear speech production, were far less affected by the F0 modulation than younger listeners for the perception of Korean aspirated and lenis stops.

Another point that meritsfurther discussion is the relative F0 weight between younger and older groups in the distinction of lenis /ta/ from aspirated /t^{h}a/ contrast. Both MANOVA and logistic regression analyses suggested that younger listeners had the greater dependency on F0 relative to older listeners for the distinction of the stop contrast. Let us consider these results in relation to the production results reported in Kang & Guion (2008). The younger listeners who, in clear speech production, enhanced F0 difference with a marginal VOT enhancement had a greater contribution of F0 compared with VOT in the distinction of the /ta/ - /t^{h}a/ contrast (see Table 1 for odds ratio of F0 and VOT). These results indicate a matched relationship between production and perception. However, the older listeners who only enhanced VOT difference also had a greater contribution of F0 over VOT to the /ta/ and /t^{h}a/ distinction (odds ratios were 2.5 for VOT and 4.8 for F0 in Table 1), even thoughthe contribution difference between F0 and VOT was not as large as the case of the younger speakers (odds ratios were 3.2 for VOT and 9.6 for F0. See Table 1 for odds ratio of F0 and VOT for both listener groups). These findings suggest that F0 is more heavily weighted than VOT for both younger and older speakers for the distinction of /ta/ and /t^{h}a/ contrast. For younger speakers, the weight of F0 might be heavy enough to induce the enhancement of this feature in clear speech production. For older listeners, the F0 weight might not be prominently heavier over VOT weight. Both younger and older listeners are exposed to the use of VOT and F0 by other speakers, young and old, all the time. They both might use F0 for the stop identification. However, F0 weightmight be prominently heavy only for younger listeners in comparison to VOT weight.

4. Summary and conclusion

The current study showed that there are generational differences in the perception of Korean stops. For the perceptual distinction of Korean aspirated and lenis stop contrast, younger Seoul speakers had relatively greater F0 cue weight than older

Seoul speakers, and this finding accords with the second prediction stated in Introduction. The results of the productionexperiments in Kang & Guion (2008) and perception experiments in the current study are taken together, it can be suggested that younger and older Korean speakers of the Seoul dialecthave different phonetic targets for the production of Korean aspirated and lenis stops, such that the VOT distinction between the two stop types has diminished and F0 difference between the two stop types arose more salient for younger speakers and accordingly, for the identification of aspirated and lenis stops, younger speakers are more greatly resort to the F0 contrast of the two stop types compared with older speakers.

The current study also leaves limits, especially for the experiment. First, the participants for the production data and the participants for the current perception data are not identical, although their age and dialectal orientation were carefully controlled and matched between the two studies. Second, the perception data of the current study was obtained only from the alveolar stops for the comprehensive three places of articulation of Korean stops, whereas the production data were extracted from all over the three place contrasts. Future studies that remedy these issues would warrant more solid evidence for the generation differences in the production and perception of Korean stops.

References

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

Han, M. S. & Weitzman, R. S. (1970). "Acoustic features of Korean /P, T, K/ /p, t, k/ and /p^h, t^h, k^h/", *Phonetica,* Vol. 22, pp. 112-128.

Kang, K.-H. (2009). "Phonological systems in bilinguals: Age of learning effects on the stop systems of Korean-English bilinguals", Unpublished Doctoral dissertation. University of Oregon.

Kang, K.-H. & Guion, S. G. (2006). "Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies", *Journal of the Acoustical Society of America,* Vol. 119, pp. 1673-1683.

Kang, K.-H. & Guion, S. G. (2008). "Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies", *Journal of the Acoustical Society of America,* Vol. 124, pp. 3909-3917.

Kim, C.-W. (1965). "On the autonomy of the tensity feature in stop classification (with special reference to Korean stops)", *Word,* Vol. 21, pp. 339-359.

Kim, M.-R. Cho. (1994). "Acoustic characteristics of Korean stops and perception of English stop consonants", Doctoral dissertation, University of Wisconsin at Madison.

Kim, M.-R., Beddor, P. S. & Horrocks, J. (2002). "The contribution of consonantal and vocalic information to the perception of Korean initial stops", *Journal of Phonetics,* Vol. 30, pp. 77-100.

Kim, Y.-H. (2007). "The effect of attention on discrimination between lenis /t/ and aspirated /t^h/ of Korean alveolar stops as shown by the Korean and the Japanese". Doctoral dissertation, Seoul National University.
(김윤현 (2007). "선택적 주의가한국어 치조 연음과 격음의변별에 미치는 영향: 한국어모국어 화자 대 일본어모국어 화자", 서울대학교 박사학위논문).

Lisker, L. & Abramson, A. S. (1964). "A cross-language study of voicing in initial stops: acoustical measurement," *Word,* Vol. 20, pp. 384-422.

Morrison, G. S. (2007). “Logistic regression modeling for first-and second-language perception data”, in P. Prieto, J. Mascaro & M.-J. Sole (eds.), *Amsterdam Studies in the Theory and History of Linguistic Science. Series IV, Current Issues in Linguistic Theory,* Vol. 282, pp. 219-236.

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

Silva, D. J., Choi, Y.-J. & J. E. Kim. (2004). "Diachronic shift in VOT values for Korean stop consonants", in S. Kuno, I.-H. Lee, J. Whitman, J. Maling, Y.-S. Kang & Y.-J. Kim (eds.), *Harvard Studies in Korean Linguistics X.S*, pp. 173-185. Cambridge, MA: Department of Linguistics, Harvard University.

Wright, J. D. (2007). "Laryngeal contrast in Seoul Korean", Doctoral dissertation, University of Pennsylvania.

• **Kang, Kyoung-Ho,** Corresponding author
Department of English Literature and Language
Kyungsung University
Email: kkang245@gmail.com
Interests: phonetics, first/second language acquisition