말소리와 음성과학 제4권 제1호 (2012)
pp. 13~20

Perception of Korean stops with a three-way laryngeal contrast

Kong, Eun Jong[1)]

ABSTRACT

A lax stop in Korean, one of the three laryngeal contrastive stops, has undergone a sound change in terms of its acoustic properties. Prior production studies described this recent lax stop as being differentiated from tense and aspirated stops primarily by fundamental frequencies (f0). And, the acoustic property of voice onset time (VOT) further separates tense stops from lax and aspirated stops. The current research explores how these two major acoustic parameters of f0 and VOT cue the three stop categories in Korean adult listeners' perception. Thirty-one native speakers of Korean participated in two experimental tasks: categorization judgment and within-category goodness ratings. Two sets of audio stimuli were prepared by synthesizing English and Korean male speakers' CV productions. The findings showed that while f0 cues listeners to lax stops as production patterns would predict, VOT were closely related to listeners' categorization and goodness ratings of lax stops. This suggests that accurate characterizations of the recent lax stop category need to be based on Korean speakers' perceptual behavior as well as production patterns.

Keywords: laryngeal contrast, categorization, category goodness, Korean stops, speech perception

1. Introduction

Korean stops with a three-way laryngeal contrast have received an amount of attention from linguists due to their typologically rare phonological status and a recent sound change in acoustic properties of the three categories occurring in Seoul dialect (Kim, 1965; Han & Weitzman, 1970; Silva, 2006, Wright, 2007; Kang & Guion, 2008). Korean stops are contrastive among tense (or fortis), lax (or lenis) and aspirated stops. In a seminal acoustic approach to stop voicing contrast in 11 languages by Lisker & Abramson (1964), lax (or lenis) stops in Korean were described as a mildly aspirated category having a voice onset time (VOT) intermediate between a short lag range (for tense stops) and a long lag range (for aspirated stops). While many languages with either/both a voicing or/and an aspiration contrast could differentiate the categories based on VOT ranges (e.g., a short vs. long lag VOT for an aspiration contrast as in Cantonese, and a lead vs. short lag VOT for a voicing contrast as in French), there were a few languages such as Hindi where VOT alone failed to clearly separate the categories. Korean was one of the languages.

Follow-up acoustic studies over the past couple of decades have revealed that Korean speakers' use of fundamental frequency (f0) became primary in differentiating lax stops from tense and aspirated stops in their productions (e.g., Kim & Duanmu, 2004; Silva, 2006). Specifically, the sound change described that VOT values of lax stops have increased to overlap with aspirated stops at a long lag VOT range, and a lower f0 property of lax stops differentiates themselves from aspirated stops. In productions, tense stops are associated with a relatively higher f0 values and a shorter VOT. This important role of f0 in Korean stops can be considered distinguished from what it does for languages such as English, where VOT alone sufficiently separates voiced and voiceless stop categories. Word-initially, English voiceless stops are phonetically aspirated and are acoustically differentiated from voiced stops by having long lag VOT values: Voiced stops are associated with short lag VOT or lead VOT values (e.g., Jacewicz, Fox and Lyle, 2009). Studies treated f0 in English as a secondary or redundant

1) Korea Aerospace University, ekong@kau.ac.kr

접수일자: 2012년 2월 16일
수정일자: 2012년 3월 10일
게재결정: 2012년 3월 13일

dimension that distinguishes voiced stops (with lower f0 values) from voiceless stops (with higher f0 values), as the first formant trajectory are known to cue the voicing contrast in English stops besides VOT (e.g., Stevens & Klatt, 1974; Summerfield & Haggard, 1977; Whalen, Abramson, Lisker & Mody, 1993).

With respect to the two major acoustic parameters (i.e., f0 and VOT), research has focused on how they pattern in Korean speakers' productions to differentiate the three laryngeal categories. Topics include how a segmental property of f0 associated with stop consonants interacts supra-segmental features in Korean (e.g., Jun, 1998), how older/younger generations of Korean manipulate acoustic parameters as an enhancement feature in their speech (Kang & Guion, 2009), in what order Korean children master the three stop categories (Kong, Beckman & Edwards, 2011), and how the sound change affects loanword adaptation patterns between younger/older generations (Kang, Han, Kochetov & Kong, 2012).

Compared to a number of studies dealing with productions, only a few studies exist which investigated how Korean listeners utilize the two major acoustic cues in judging tense, lax and aspirated stops in perception (e.g., Kim, Beddor & Horrocks, 2002; Kim, 2004). Although the pattern can be derived in part by findings from prior production studies, it is not always accurate to assume that listeners' perceptual tendency directly mirrors production patterns. For example, Kim (2004) showed that f0 did not affect Korean listener's perception of tense stops, while the category is known to have relatively higher f0 values resulting in H(igh) tone at accentual phrase initial position. The study also found a trading relationship (i.e., a negative correlation) between VOT and f0 in a perception of lax stops vs. aspirated stops. Similarly, Kong et al., (2011) examined relative cue weights among VOT, f0 and H1-H2 (breathiness measure) that affected Korean adult listeners in judging production accuracies of stop laryngeal categories of children's productions. Findings included that adult listeners relied on a lower f0 in judging children' tokens as lax stops from aspirated stops, as predicted by an acoustic characteristics of lax stops in productions. Interestingly, the listeners also put a similar amount of weight on VOT in perceiving children's productions as lax stops from aspirated stops. This was unexpected given that productions of lax stops have a wide overlap with aspirated stops at VOT dimension. They pointed out that this complex combination of VOT and f0 in perceptual differentiation between lax stops and aspirated stops might explain children's later mastery of lax stop category. As shown in the above perception studies, the mismatch between production and perception makes it necessary to have independent evidence of how Korean listeners utilize the two major acoustic cues, VOT and f0, in perception.

The current study aimed to meet this need by exploring two specific questions. First, we are curious to know how each of acoustic dimensions of VOT and f0 is partitioned in Korean adult listeners' perceptual space to correspond to tense, lax and aspirated stop categories. Unlike previous perception studies (e.g., Kim et al., 2002; Kim, 2004), we pursue this question with acoustically controlled sets of stimuli. Second, we aimed to examine how each acoustic property affects the listeners in judging a within-category goodness of the stops.

To explore the research questions, we conducted experiments testing Korean adult listeners' categorization and their goodness rating of stimuli synthesized to vary VOT and f0. We expected that listeners' categorization tendency would reveal how the three stop categories correspond the acoustic dimensions of VOT and f0. And, the listeners' judgment of within-category goodness ratings would help to define the prototypes of the categories in terms of the multiple acoustic dimensions, as Kuhl & Iverson (1995) among others employed the goodness rating method to characterize within-category properties. We included a sets of auditory stimuli manipulated from English (listeners' non-native language) in addition to a set of Korean source stimuli (listeners' native language) to ensure that listeners' exposures to foreign languages (English) did not alter their perceptual judgment of category goodness with respect to utilizations of the two acoustic cues.

2. Method

2.1 Procedure

2.1.1 materials

Audio stimuli were synthesized using natural productions of male speakers, who are a single native speaker of Korean and a single native speaker of English. An English male speaker in his twenties made multiple productions of 'tat' /t^{h}at/ and 'dat' /dat/ and a Korean male speaker in his thirties produced words beginning with /t'at/ '땃', /tat/ '닷' and /t^{h}at/ '탓'. Among those repetitions of words, we selected a pair/triple of tokens in each language that carry the least different F1 and F2 onset at the vocalic onset. Vowel durations were adjusted to be identical between the selected tokens by shrinking or stretching their original duration to an average duration of two vowels in

English and three vowels in Korean, respectively. This process of duration adjustment did not result in noteworthy quality changes in the tokens impressionistically. On the basis of naturalness of normalized vowels, we decided to use vowels in English 'dat' /dat/ and Korean /tat/ '닷' productions as a vocalic source. The coda consonants were removed.

Of these selected vocalic sources in English and Korean, f0 values at the vocalic onset were manipulated to vary in five linear steps from 98 Hz to 130 Hz (i.e., 98 Hz, 106 Hz, 114 Hz, 122 Hz and 130 Hz). These specific values of F0 variations were adapted from Whalen et al. (1993). The onset f0 values gradually increased/decreased to 114Hz at 20ms into the vowel and were maintained as 114Hz for the rest of the source duration. VOT values were varied in six steps in English stimuli (i.e., 6 ms, 13ms, 19ms, 28ms, 40ms, 59ms) and seven steps in Korean stimuli (i.e., 9ms, 11ms, 17ms, 25ms, 38ms, 59ms, 85ms). The two end values of VOT were determined based on the original VOT of a pair/triple of natural productions, and the four intermediate values were chosen by dividing the VOT interval (i.e., longest VOT – shortest VOT) into 4 equal steps in a log scale. These intermediate steps differed between English and Korean because of the VOT differences in the original tokens, which the stimuli were made based on.

2.1.2 Participants and tasks

The experiment was conducted in Madison, Wisconsin, USA by recruiting 31 Korean listeners who were in their twenties and thirties. Since different degrees of participants' English exposures might confound the results, we divided the listeners into two groups according to a length of staying in English speaking countries. Seventeen of them had lived in English speaking countries less than a year (low English exposure group: Most of them had less than three months of stay in English speaking countries), while 14 of them have lived in US for more than 5 consecutive years (high English exposure group).

There were two tasks for the listeners. For one task, they were instructed to choose a single consonant out of the three Korean consonant categories (/t/, /t'/ and /t^h/ presented as "ㄷ", "ㄸ" and "ㅌ" on a computer screen). Immediately after subjects had made a categorical decision, the same stimulus was played again asking them to judge a goodness of the category (e.g., asking "how good 'ㄷ' was the token" given the consonant category 'ㄷ' chosen.) Responses were collected using an arrow displayed on the computer monitor. The left end of arrow was labeled as 'very good', and the right end of arrow was labeled as 'very bad'. The subjects judged goodness of the selected category by clicking any location at the arrow in a continuous manner (see Massaro & Cohen, 1983; Munson, Edwards, Schellinger, Beckman & Meyer, 2010 for more details on the visual analogue scaling method.)

A block of stimuli whose vocalic source was Korean words was given before a block of stimuli that came from English words. A block of Korean stimuli consisted of 105 trials and a English stimulus block of 90 trials which were three repetitions of 35 (Korean: 7 VOT steps * 5 f0 steps) and 30 (English: 6 VOT steps * 5 f0 steps) unique stimuli. The total duration of the session was less than 20 minutes. The low English exposure group completed both blocks of Korean source stimuli and English source stimuli. However, the high English exposure group did the block of Korean source stimuli only - which was not deliberately but unfortunately done so. At the time of the participant's visit, they were given another type of perception experiment on vowels. The order of the experiments was counterbalanced so that a half of participants did the consonant perception tasks before the other vowel task. The participants received monetary compensation.

2.1.3 Statistical analysis

The listeners' categorical responses were analyzed using a multinomial logistic regression model, using *nnet* package Venables & Ripley, 2002) in R (R Development Core Team, 2011). In this regression model, a dependent variable was three discrete categories chosen by the listeners and independent variables were VOT and f0 of the stimuli, treated as continuous variables, and interactions between the two acoustic parameters (VOT and f0) and the two stimulus sources (Korean source vs. English source). This multinomial logistic regression model predicted the odds of being categorized as each of the three categories as a function of VOT and f0.

A mixed-effects regression model was made to examine a relationship between goodness ratings of the selected categories and each acoustic parameter (*lme4* package in R). The click locations at x-axis of goodness judgment were transformed into logit values to deal with different end points of the best and the worst examples from individual listeners. The transformed values were entered as a dependent variable of the model. Fixed effects of the model were acoustic parameters of VOT and f0 as a continuous variable and the source languages and also

included an interaction between the source language and each acoustic parameter. A random effect was individual listeners and an intercept and a slope of each acoustic parameter were allowed to vary at the listener level. Six separate regression models were made for the three response categories in each acoustic parameter (3 response categories * 2 acoustic parameters) to examine whether VOT and f0 can explain the goodness ratings of the chosen category. Since the current version of *lmer* command in *lme4* package does not provide p-values for t-values of fixed effect variable, we will rely on the estimated t-values of the model and separate likelihood-ratio tests to determine statistical significances of each fixed effect variable (see Bayaan, 2008). A likelihood-ratio test compares a regression model with a target fixed effect and a model without it in terms of model fits. The fixed effect variable is considered to be significant if a complex model (the former one with a target variable) significantly improves a fit based on this test.

3. Results

3.1 Categorical judgment

The categorical responses were summed across the listeners in each stimulus condition of f0 and VOT. <Figure 1> shows the response categories in a matrix of f0 and VOT, in which the labels represent the type of consonants that was agreed by more than 21 trials out of 51 trials by low-English exposure group (18 trials out of total 42 trials by high-English exposure group) – 21 and 18 were the minimum counts that returned the binomial probability of less than .05 when the probability of each event was .33 (i.e., 1/3). As shown in top panels of <Figure 1>, the acoustic space of f0 and VOT was carved in a way that stimuli of lower f0 were categorized as a lax stop /t/, while those of higher f0 as either a tense stop /t'/ or an aspirated stop /t^h/. When f0 was higher (> 114 Hz), listeners heard stimuli of shorter VOT as a tense stop /t'/ and those of longer VOT as an aspirated stop /t^h/. While the overall patterns of carving the acoustic spaces with /t/, /t'/ and /t^h/ were similar between Korean stimuli and English stimuli, the exact boundaries of f0 and VOT slightly differed between the two sets. Compared to the responses to Korean stimuli (the top left panels of <Figure 1>), listeners were conservative to categorize the English stimuli as a lax type /t/. Only a small set of English source stimuli were chosen as a lax stop /t/ whose f0 and VOT values were lower than those of Korean source stimuli.

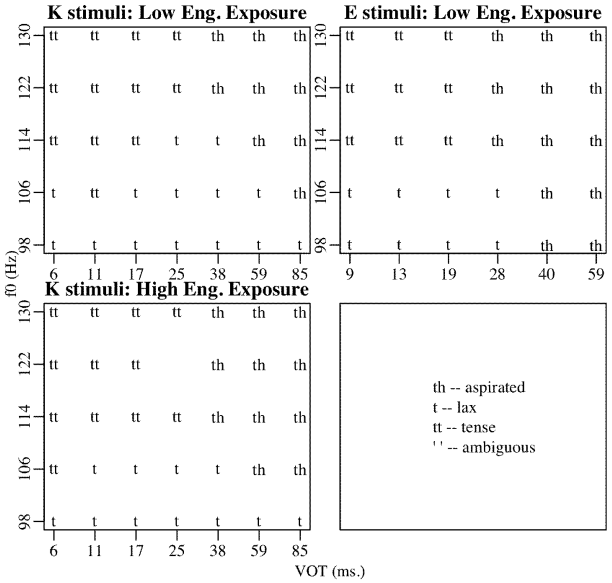

Figure 1. Distributions of response categories in a two-dimensional acoustic dimension. Labels ("tt" for /t'/, "t" for /t/ and "th" for /th/) indicate the categories agreed by 21 trials out of total 51 trials from low English exposure group and 18 trials out of total 42 trials from high English exposure group.

The multinomial regression models were made to explain the variations of categorical responses (a dependent variable) as a function of VOT or f0. Six different regression models were constructed according to the experimental conditions. For the listener group of low English exposure, four different regression models were made for the two stimuli languages and the two acoustic parameters (f0 and VOT). For the listener group of high English exposure, two different regression models were constructed for the two acoustic parameters of VOT and f0. <Figure 2> illustrates a probability density of three categorical responses as a function of VOT (left column) and f0 (right column). The predicted values by the six multinomial regression models are overlaid as lines on top of density plots. <Table 1> summarizes probabilities of the three response categories estimated by the regression models. As in the top left panel of <Figure 2>, probabilities of being categorized as tense and aspirated types changed dramatically as VOT values increased. As VOT values increased, the probability of being categorized as /t'/ decreased, but the probability of /t^h/ categorization increased. By contrast, a probability of being categorized as /t/ did not change much as VOT values change in both sets of stimuli. It is also noteworthy that the probability of /t/ responses as a function of VOT exhibited great differences between

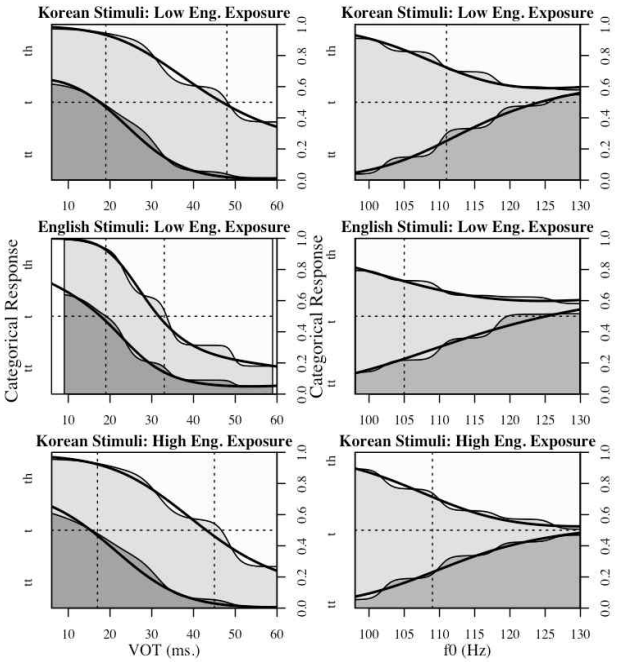

Figure 2. Probability density plots of response categories /t'/, /t/ and /t^h/ separated by listener groups (low vs. high English exposure listeners), acoustic parameters (VOT and f0).

Korean stimuli and English stimuli (see the top vs. middle panels in the left column of <Figure 2>): The probability of /t/ categorization was generally lower in English stimuli than in Korean stimuli given the VOT values. This difference is graphically shown as a narrower range of /t/ in the mid panel than that in the top panel in the left column of <Figure 2>. The horizontal line was drawn at the probability of .5, and two vertical lines were drawn to indicate the exact VOT values at which either tense or aspirated type exceeded the probability of .5. Compared to Korean source stimuli at the top panel, a criterion VOT value above which the stimuli were categorized as aspirated stops was shorter, indicating that listeners categorized English source stimuli as lax stops only when their VOT values were sufficiently short.

In a f0 dimension, probabilities of being categorized as a tense stop /t'/ and a aspirated stop /t^h/ increased as f0 values increased. By contrast, a probability of being categorized as a lax stop /t/ decreased, as f0 increased. These general tendencies were almost similar between English source stimuli and Korean source stimuli (compare the top and the mid panels of the right column in <Figure 2>). When the vertical lines indicating a .5 probability of lax stop responses were drawn, the criterion f0 value for English source stimuli (i.e., slightly above 110Hz) was lower than the f0 value for Korean source stimuli (i.e., about 105Hz). Consistent with a tendency in a VOT dimension, a conservative judgment criterion of f0 was observed in categorizing the sound as lax stops for the set of English source stimuli.

A listener group difference (i.e., low Eng. exposure vs. high Eng. Exposure) was not noticeable in the category identification task (compare the top and bottom panels of <Figure 2>.)

Table 1. Probability of categorical responses at VOT values estimated by multinomial logistic regression models (Low English exposure group): Values in parenthesis indicate High English exposure group.

	Korean Stimuli			English Stimuli		
VOT	t'	t	t^h	t'	t	t^h
6	66.1(65.3)	32.5(31.7)	1.3(3.1)	73.6	26.2	0.2
9	62.3(61.3)	35.7(34.9)	2.0(3.8)	67.6	31.9	0.5
11	59.4(58.2)	38(37.3)	2.6(4.5)	63.5	35.6	0.9
13	56.2(54.6)	40.4(40.0)	3.4(5.3)	59.2	39.2	1.6
17	48.9(46.6)	45.4(45.7)	5.7(7.7)	50.2	45.1	4.7
19	44.8(42.2)	47.9(48.6)	7.2(9.2)	45.5	47	7.5
25	32.0(28.6)	54.3(55.6)	13.7(15.8)	30.9	46.1	23.0
28	25.7(22.3)	56.3(57.5)	18(20.3)	24.1	41.8	34.1
38	9.5(7.5)	55 (53.0)	35.5(39.4)	9.4	23.3	67.3
40	7.5(5.8)	53.5(50.5)	39(43.7)	7.9	20.6	71.5
59	0.4(0.4)	35.9(24.5)	63.6(75.1)	5.1	13.2	81.7
85	0 (0.1)	27.9(16.4)	72.1(83.5)	50.1	34.3	15.6

Table 2. Probability of categorical responses at f0 values estimated by multinomial logistic regression models (Low English exposure group). Values in parenthesis indicate High English exposure group.

	Korean Stimuli			English Stimuli		
f0	tt	t	th	tt	t	th
98	4.7(7.3)	88.3(82.3)	7.0(10.3)	13.4	67.8	18.8
106	14.7(17.9)	67.0(59.1)	18.3(23.0)	24.2	47.0	28.8
114	31.8(31.7)	35.2(31.0)	33.0(37.3)	36.2	27.1	36.7
122	47.0(42.3)	12.6(12.2)	40.4(45.5)	46.5	13.4	40.1
130	56.2(48.4)	3.6(4.1)	40.1(47.5)	54.2	6.0	39.8

3.2 Within-category goodness rating

Using a mixed effects regression, logit-transformed values of goodness ratings were regressed as a function of VOT and f0 in each selected consonant category (/t'/, /t/ and /t^h/). The listener groups of low vs. high English exposure were collapsed due to a null main effect of group distinction. This created six linear regression models that considered stimulus source as a factor interacting with VOT and f0, respectively in each model.

<Figure 3> plots the predicted goodness ratings of the

regression models. Overall, it was observed that category goodness of English-source stimuli was rated lower than that of Korean-source stimuli for the three phonetic categories in each of acoustic parameters. This tendency is indicated by locations of filled circles (representing Korean stimuli) almost always higher than crosses in the panels (representing English stimuli). According to likelihood tests, the effects of stimulus source language were significant at a significance level of $p < 0.05$ in all regression models except the f0 model for /t'/ (see the top right panel of <Figure 3>): Chi.Sq = 0.503, p = 0.48 in f0 model. Completely overlapped plotting characters in the f0 model for /t'/ (top right panel) suggest that goodness ratings for /t'/ category were little affected by the source languages. The lack of source language effect in f0 model for /t'/ might support that Korean listeners are not sensitive to f0 in perceiving tense category as found in Kim (2004).

For English source stimuli, listeners' goodness ratings were significantly correlated with both VOT and f0: coefficients of slopes were estimated to have absolute t-values greater than 2 in mixed-effects models. Each panel in <Figure 3> displays either negative or positive direction of linear slopes with filled circle plotting characters. As indicated by a negative slope of VOT (top left panel) and a positive slope of f0 (top right panel), listeners tended to judge non-native source tokens of lower VOT and higher f0 values as better examples of a /t'/ category. Similarly, they judged tokens of higher VOT and higher f0 values as better examples of a /t^h/ category, illustrated by a positive direction of linear slopes in the two bottom panels. Finally, for a /t/ category, listeners considered tokens of non-native source with lower f0 values as better examples of a lax stop. What requires a careful interpretation was an association of lower VOT with a better /t/ category, indicated by a negative linear slope in the mid left panel. One way of understanding this negative correlation between VOT and /t/ category is that longer VOT values made lax stops sound less good because /t/ competed with /t^h/ category. This pattern seems relevant to a trade-off between VOT and f0 in lax vs. aspiration distinction observed in Kim (2004).

Linear slope patterns of goodness judgments for Korean stimuli were not greatly different from those for English stimuli. Slope directions were identical between the two source conditions except that in VOT model for /t/, a slope was no different from a flat line, indicating a change of VOT did not explain goodness ratings of /t/ category: slope coefficient= -0.0024, t= -0.076 (see the mid left panel of <Figure 3>).

Given the slope patterns in the VOT models (left panels of <Figure 3>), it was English stimuli (non-native source stimuli) that showed a stronger relationship between within-category goodness judgments and acoustic values of VOT: Overall, English source stimuli had steeper slopes than Korean ones. Those steeper slopes indicate that Korean listeners were more sensitive to the changes of VOT values when listening to non-native source tokens. According to likelihood-ratio tests, the effects of interaction between source language and VOT were significant: Slope coefficients for English stimuli were statistically different from those for Korean stimuli. In contrast, none of the f0 models for /t'/, /t/ and /t^h/ had significant effect of source language interaction: slope coefficients between English stimuli and Korean stimuli were not statistically different.

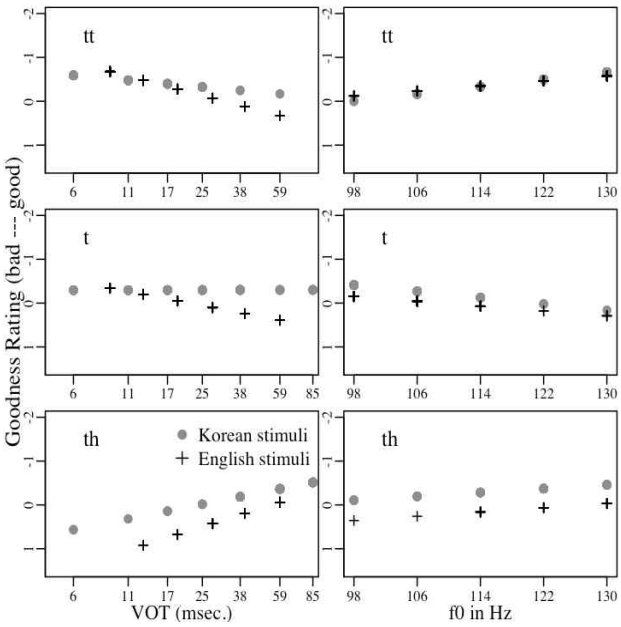

Figure 3. Model predicted values of goodness ratings as to English and Korean source stimuli. Panels were separated by the three response categories (rows) and acoustic parameters (columns). 'tt', 't' and 'th' represent /t', t, /t^h/, respectively.

4. Discussion & Conclusion

We explored Korean adult listeners' perception of acoustically varying stimuli in order to examine how the acoustic properties of VOT and f0 cue the listeners to categorize the Korean laryngeal stops. Perceptually defined acoustic characteristics of the three stop categories resembled what earlier production studies have described in terms of VOT and f0. The listeners

utilized VOT in categorizing stimulus tokens in a way that tokens with shorter VOT values were perceived as tense stops. Also, longer VOT values cued the listeners to perceive them as either lax or aspirated stops. Lower f0 values tended to be categorized as lax stops whereas higher f0 values as aspirated stops.

Consistent with findings from prior studies (e.g., Kim, 2004; Kong et al., 2011), perception of lax stops required not only the f0 values are lower but also that VOT are intermediate. When a lax category competed with a tense category at a shorter VOT range, relatively longer VOT tokens within the shorter range could be heard as lax stops with non-low f0 values. Likewise, when a lax category competed with an aspirated category at a relatively longer VOT range, non-low f0 values tended to be heard as lax stops if their VOT values were shorter. This perceptual evidence of trading relationship describes a lax stop as a category of complex acoustic characteristics: A VOT property is not ignorable in perception of lax category even though f0 primarily characterizes it. Perhaps it might serve as supporting evidence of this complex acoustic nature of lax stop category that the lax stop response in the categorization task was influenced by stimulus source languages more than the two other categories.

Response patterns of within-category goodness rating further helped to describe the prototype of each category in terms of VOT and f0. As predicted by the categorization pattern, good tense stops were composed of a shorter VOT and a higher f0, and good aspirated stops were composed of a longer VOT and a higher f0. The patterns did not differ between stimulus source languages. Unlike tense and aspirated stops, good lax stops were correlated with shorter VOT values only when the stimulus source was English. This direction of correlation was compatible to categorization responses, supporting that a VOT property is still an important cue to a perception of lax stops. As for the lack of correlation between lax stops and VOT for Korean stimuli, it might be ascribed to the fact that perceptually good lax stops have somewhat intermediate VOT values, bounded not only by tense stops at a shorter VOT range but also by aspirated stops at a longer VOT range (refer to the triangle-shape distribution of the lax stop /t/ categorization in <Figure 1>. Goodness ratings of this category may not be captured effectively by a simple linear regression – more complex form of models such as quadratic or cubic terms are needed to accommodate the patterns.

The study showed that the stimulus source type affected defining the prototypes of stop laryngeal categories. For one, the native stimuli were rated as better examples of the category than the non-native stimuli -- the exception was observed in tense stops. The source type did not affect the goodness ratings of tense stops, although listeners judged higher f0 as good tense stops. As Kim (2004) suggested, f0 appears to be only weakly related to tense stop perception. In addition, listeners' sensitivities to VOT were greater in English source stimulus, although the directions of how VOT or f0 were correlated with each stop category did not differ by source stimulus types. The same amount of changes in VOT influenced Korean listeners' category goodness judgment more greatly for non-native source stimuli. These tendencies are congruent with what research on prototypes (or perceptual magnetic effects, in a broader sense) would predict. Listeners were more likely to rely on lower level of sound representations (i.e., acoustic cues), when processing non-native stimuli, which are distant from prototypes of speech categories.

Finally, the current study did not find significant effect of listeners' English exposure in perception patterns. We acknowledge that given that the experiments were conducted in English speaking environment, it might not be an ideal condition to observe the targeted effect. Also, the current study could not afford to administer the block of English-source stimuli to the listeners of the high English exposure group. It may be possible that the effect of listeners' English exposure exists only for those who have a high degree of English experience. A follow-up study is needed to test whether there exist effects of speakers' English exposure and whether the degree of English exposure matters in revealing the effect. These questions will carry an important implication on second language acquisition and bilingualism research in discussing interactive influences of learners' native/non-native languages on each other.

To conclude, the study explored Korean listeners' use of VOT and f0 cues in speech perception by examining how sounds are judged as stop laryngeal categories and as better or worse examples of the selected category. Findings suggest that in large part, perceptual characteristics of the laryngeal categories mirrored the patterns observed in prior production studies. However, listeners' responses of categorizations and goodness ratings further characterized that VOT still cues lax stops, even though lower f0 values already clearly separate it from tense and aspirated stops on the basis of production patterns. Findings of the study emphasize a need of thorough

research on speech perception independently of production studies in order to understand the acoustic characteristics of Korean stops with a laryngeal contrast.

References

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*, New York: Cambridge University Press.

Han, M. & Weitzman, R. (1970). Acoustic features of Korean /p', t', k'/, /p, t, k/ and /p^h, t^h, k^h/, *Phonetica*, Vol. 22, 112-128.

Jun, S.A. (1998). The accentual phrase in the Korean prosodic hierarchy, *Phonology*, Vol. 15, 189-226.

Jacewicz, E., Fox, R.A., & Lyle, S. (2009). Variation in stop consonant voicing in two regional varieties of American English, *Journal of the International Phonetic Association,* Vol. 39, 313-334.

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, 3909-3917.

Kang, Y., Han, S., Kochetov, A., & Kong, E. (2012). Cross-language perception and loanword adaptation under sound change in progress, A paper presented at Linguistic Society of America, Portland, Oregon, USA.

Kong, E., Beckman, M. E., & Edwards, J. (2011). The role of acoustic properties in predicting the mastery patterns of Korean laryngeal contrast, *Journal of Phonetics*, Vol. 39, 196-211.

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

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

Kim, M., & Duanmu, S. (2004). 'Tense' and 'lax' stops in Korean, *Journal of East Asian Linguistics*, Vol. 13, 59-104.

Kim, M. (2004). Correlation between VOT and F0 in the perception of Korean stops and affricates, *Proceedings of The 8th International Conference on Spoken Language Processing*, 49-52.

Kuhl, P.K. & Iverson, P. (1995). Linguistic experience and the 'perceptual magnet effect', *Speech Perception and Linguistic Experience*, Winifred Strange (ed.), Baltimore: York Press.

Lisker, L. & Abramson, A. (1964). A cross-language study of voicing in initial stops: acoustical measurements, *Words*, Vol. 20, 384-442.

Massaro, D. W. and Cohen, M. M. (1983). Integration of visual and auditory information in speech perception, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 9, 753-771.

Munson, B., Edwards, J., Schellinger, S., Beckman, M.E., & Meyer, M. (2010). Deconstructing phonetic transcription, *Clin. Linguist. Phon*., Vol. 24, 245-260.

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

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

Stevens, K.N. & Klatt, D.H. (1974). Role of formant transitions in the voiced/voiceless distinction for stops, *Journal of the Acoustical Society of America*, Vol. 55, 653-659.

Summerfield, Q. & Haggard, M. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants, *Journal of the Acoustical Society of America*, Vol. 62, 453-448.

Venables, W.N. & Ripley, B.D. (2002). *Modern Applied Statistics with S*. Fourth Edition. New York: Springer.

Wright, J.D. (2007). *Laryngeal contrast in Seoul Korean*. Ph.D. Dissertation, University of Pennsylvania, Philadelphia, PA.

Whalen, D.H., Abramson, A.S., Lisker, L. & Mody, M. (1993). F0 gives voicing information even with unambiguous voice onset times, *Journal of the Acoustical Society of America*, Vol. 93, 2152-2159.

• **Kong, Eun Jong**
Korea Aerospace University
100 Hanggongdae-gil, Hwajeon-dong, Deokyang-gu,
Goyang-City, Gyeonggi-do, 412-791, Korea
Tel: 1-82-2-300-0338 Fax: 1-82-2-300-0493
E-mail: ekong@kau.ac.kr
Area of interest: Acoustic phonetics, Language acquisition, Speech perception