

Individual variability in processing English source input into Korean stops

Eun Jong Kong*
(Korea Aerospace University)

Soyoung Kang**
(Carleton University)

Kong, Eun Jong and Kang, Soyoung. 2017. Individual variability in processing English source input into Korean stops. *Studies in Phonetics, Phonology and Morphology* 23.2. 201-219. The current study investigates whether and how Korean listeners are influenced by English source input in identifying three laryngeal categories of Korean stops (/tʰ/, /t/, and /tʰ/). We relate identification patterns of Korean and English source input to individuals' utilizations of two acoustic cues (VOT and *f0*) under the English source condition. Korean adult listeners (N=53) participated in three-way forced-choice identification tasks where the auditory stimuli of VOT×*f0* combinations synthesized from Korean and English speakers' productions were presented in two separate sessions. Both group-averaged and individual-level analyses showed varying degrees of source effects on listeners' identifications of the Korean stops, revealing that the response category distributions are meaningfully related to individual listeners' modifications of cue-weighting strategies confronted by the English-source signals. Our results suggest that non-native source input modifies individual strategies in dealing with sub-phonemic acoustic cues, ultimately yielding different category identification patterns from native input identifications. (Korea Aerospace University and Carleton University)

Keywords: Korean stop perception, individual differences, multiple acoustic cues, VOT, *f0*, non-native input, stop laryngeal contrast

1. Introduction

The quality of the speech signal is one of the fundamental factors in processing speech sounds. Unnaturalness in sound input has been known to hinder sound identification. Clearer categorization patterns were observed in the identification task when the stimulus sets were manipulated to be natural by having full spectrum of acoustic parameters in steps over a number of time windows (e.g., Schouten and Van

* First author

** Corresponding author

Hessen 1992, Van Hessen and Schouten 1999, Schouten et al. 2003, ter Schure et al. 2011). Similarly, noise in speech perception challenges listeners to focus on relevant linguistic cues. In external environmental noise condition, listeners' robust linguistic representations determined successful perceptual message transmission (e.g., Assmann and Summerfield 2004, Shi 2009). Internal noise in the signals such as cognitive loads or information degradations interfered with listeners' perceptual reliance on primary acoustic cue (Gordon et al. 1993, Francis and Nusbaum 2002). Also, child speech samples are not fully accommodated by a definite set of native sound inventories in transcriptions, yielding continuous accuracy ratings by adult transcribers due to children's less adult-like and often covertly contrastive acoustic information present in their speech productions (Schellinger et al. 2017).

The current study explored one such situation that the non-canonical quality of the speech input affects the identification of the segmental categories, namely the effect of non-native input source on the identification of native category. Also given that non-native source input played as a noise in sentential level processing (as foreign- or regional-accented speech slowed processing time with low comprehensibility [e.g., Munro and Derwing 1995, Floccia et al. 2009]), we questioned whether and how listeners' perceptual strategies of utilizing multiple acoustic cues would be modified to deal with non-native input source conditions. If a confrontation of processing non-native speech source renders listeners to reorder relative informativity of multiple acoustic cues - as their attention to primary cues is disturbed and other redundant cues are enhanced for compensation (Mattys et al. 2005), these processing modifications would return not just quantitatively degraded goodness ratings but qualitatively discrepant category identification patterns compared to response patterns when the input source was not compromised.

Locating this question on more particular linguistic contexts, we investigated how Korean adult listeners would be affected by input source conditions (Korean vs. English source signals) in identifying Korean stop categories with a three-way laryngeal contrast. The three types of Korean laryngeal stops, namely tense, lax and aspirated ones are known to be differentiated based on VOT, f_0 and other acoustic cues (Han and Weitzman 1970, Cho et al. 2002, Silva 2006, Kong et al. 2011, Kang 2014). Prior perception studies showed that, in Seoul Korean, the tense type with short-lag VOTs was distinguished from the lax and aspirated types with long-lag VOTs, and the lax type with lower f_0 range was differentiated from the tense and aspirated types with relatively higher f_0 values (e.g., Kim et al. 2002, Kong et al.

2011, Lee et al. 2013, among others). If English input source were to affect Korean listeners in weighing the multiple acoustic cues, Korean listeners would identify the sounds of Korean source and those of English source differently despite the same acoustic values of the input. We are interested in comparing the response categories in the two source conditions and in relating those categorical decisions to Korean listeners' modified perceptual utilizations of the acoustic cues due to English source input. Broadly, potential contributions of the study can be made to better understanding how English stops are adapted into the Korean categories since Korean listeners' patterns of using acoustic cues have been pointed out as one of important factors to predict the adaptation outputs, but scarcely experimented (e.g., Kang 2003, Oh 2009).

The current research has taken individual differences into a methodological consideration in estimating the effect of input source on utilizing multiple acoustic cues in the identification of the Korean stops for reasons. Besides general benefits of looking into individual variations in cognitive processing of a language (Hagaard et al. 1970, Hazan and Rosen 1991, Idemaru et al. 2012, Kong and Edwards 2016, Kapnoula et al. 2017), there are language specific reasons to consider individual differences in processing multiple acoustic cues. With the three Korean stops undergoing a sound change of sub-phonemic acoustic specifications (Silva 2006, Kang 2014, among others), listeners are likely to be exposed to both conservative and progressive phonetic forms of the stops with the latter emphasizing the increased role of *f0* (e.g., Kong et al. 2011, Schertz et al. 2016). Unlike productions, individual listeners would have to have perceptual flexibility of cue-weightings even in processing native source inputs let alone non-native ones in order to successfully communicate with both conservative and progressive speakers (Beddor 2009, 2015, Stevens and Harrington 2014). Given the linguistic environment, close examinations of individual variations in stop laryngeal perception would better reflect how Korean listeners are influenced by the input source conditions in processing the multiple acoustic cues.

Given the background, specific goals of the current study are to explore the effect of English source inputs on Korean listeners' identification of the Korean laryngeal stops and to examine the relationship between identification patterns and individual processing strategies of multiple acoustic cues confronted by a foreign input source.

2. Methods (perception experiment)

2.1 Stimuli

An English set of auditory stimuli were prepared by manipulating acoustic values of VOT and *f0* over the natural productions of /da/ and /ta/ by a English-speaking male speaker (born and raised in Wisconsin, USA). With one production of /da/ as a base, six-step VOTs were created by adding release portion of a /ta/ production: 9ms [original /d/ VOT], 13ms, 19ms, 28ms, 40ms, and 59ms [original /t/ VOT] – intermediate VOT values between the two original end-VOTs were defined by dividing the VOT differences in log-steps considering non-linear nature of psycho-acoustics. In the stimuli of each VOT step, a five-different sustained *f0*s were overlaid (98Hz, 106Hz, 114Hz, 122Hz, 130Hz) on the vowel. This synthesizing procedure was done by using pitch tier manipulation function in *Praat* (Boersma and Weenink 2013), yielding 30 different syllables (stop-/a/). The same set of stimuli based on English source was used in previous experiments published in Kong et al. (2011) and Kong and Edwards (2016). The Korean set of stimuli was prepared by the same procedure to the English one except that the base syllable was obtained from a Korean-speaking male speaker. He produced three Korean syllables (‘따’ /t’a/, ‘타’ /ta/, and ‘타’ /t^ha/), among which the syllable of lax-vowel served as a base to manipulate the acoustic dimensions.

2.2 Participants, task and procedures

53 Seoul Korean-speaking young adults (aged from 20 to 27) participated in a three-alternative forced-choice (3AFC) task with monetary compensation. None of them reported speech, hearing or language problems.

Participants listened to audio stimuli through a headphone and clicked on a consonant category that they heard among the three given choices of Korean laryngeal stops (‘ㄷ’, ‘ㄸ’, ‘ㅌ’) shown in the monitor. The sessions were blocked by the stimulus source languages, with the Korean set of stimulus given before the English one. The experiment was programmed in E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) to present the stimuli in a randomized order and to record the response automatically.

TOEIC scores were surveyed to be used as a measure of English proficiency (score

range: 300~965; mean score: 759.1). On average, it took participants approximately 20 minutes to complete the sessions – The participants also completed another mode of speech tasks such as English- and Korean-word reading tasks when they visited for the experiment.

2.3 Analysis

With the responses from both stimulus sessions (K-source and E-source), three mixed-effects logistic regression models were constructed where the binary response categories are considered: (1) tense vs. aspirated, (2) tense vs. lax, and (3) lax vs. aspirated responses (implemented in R using *lme4* package [Bates et al. 2015]). Each model had the two acoustic dimensions (VOT and f_0 , continuous variables) as fixed effect variables so as to estimate group-typical relative perceptual weights of VOT and f_0 in explaining the likelihood of choosing one category over the other stop category. The stimulus source languages (*K-/E-SourceCondition*) and its interaction with each of VOT and f_0 dimensions were another fixed effects considered in the model to assess the group-averaged effect of source language difference in processing the stimuli.

The model had a random intercept and random slopes of VOT and f_0 varying both at the *listener* level and the *listener-by-source* language level. Given the random effect specifications of the regression model, individual listeners' coefficients of each source-language condition were calculated by adding the by-subject random effect coefficients and the subject-by-source language condition (K-source or E-source) to the fixed effect coefficients of VOT and f_0 (in standardized unit). This numeric index would represent individual listeners' sensitivity to each cue when they process the Korean- or English-source stimuli to identify the Korean stop categories. In addition, the effects of English source on individual listeners' L1 processing was calculated by subtracting the coefficients of the K-source condition from those of the E-source condition obtained from the listener-by-source language random effect coefficients.

3. Results

3.1 Response category distribution

The distributions of the three Korean stop response categories were calculated by

stimulus source language condition (Figure 1-left). The effect of stimulus-source was observed despite the same acoustics of VOT and f_0 , as the distributions of the response categories differed between the two conditions of the source language (K-source, and E-source). While the identification rates of the lax stop were similar between the two source language conditions, the identification rates for the aspirated type increased and those for the tense type decreased in the E-source condition. These patterns based on group-aggregated counts were confirmed when individuals' identification rates were examined. Fig.1-right plots the individual rates of aspirated and tense types in the English stimulus source condition against those in the Korean stimulus condition. The data-points of the individuals' aspirated response counts were mostly distributed above the diagonal line ($y=x$) suggesting more response counts in the E-source condition, while those of the tense response counts were mostly located below the diagonal line indicating fewer response counts of the tense type in the E-source condition.

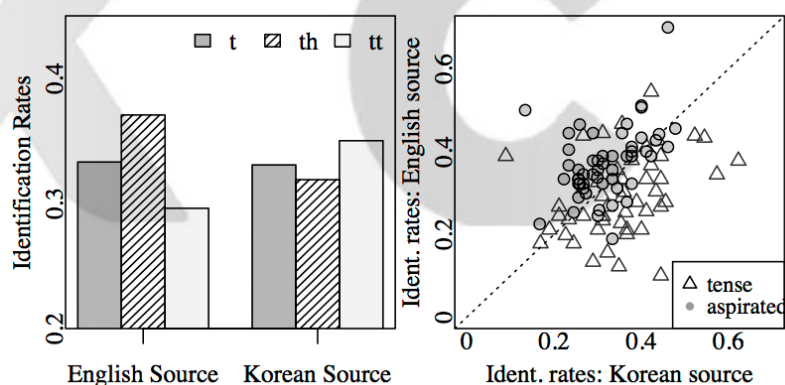


Figure 1. [Left] Identification rates of tense (/t/ 'tt'), lax (/t/ 't') and aspirated (/t^h/, 't^h') stops in each stimulus source language condition (N=53). [Right] Individual listeners' identification rates of the tense and aspirated stops in the English source condition against those in the Korean source condition.

Individual changes of the response counts were not consistently accounted for by their English proficiency measured by TOEIC scores. Correlation tests revealed that the decreased tense counts and increased aspirated stop counts were not correlated with their English test scores (English proficiency with tense count changes: $r=0.15$,

$t[51]=1.10$, $p=0.27$; with asp. count changes: $r=0.17$, $t[51]=1.27$, $p=0.208$), whereas the count changes of the lax type were correlated with their English proficiency in a negative direction (English proficiency with lax count changes: $r=-0.29$, $t[51]=-2.1$, $p<0.05$), suggesting that the more counts of the lax types in the E-source condition were associated with listeners with lower English proficiency.

3.2 Effects of the English source: Group characteristics

The mixed-effects logistic regression models were built with the Korean-source stimuli condition as a reference category (Table 1) and, based on the coefficients, inverse-logit curves were made for each model of contrast pairs as presented in Figure 2.

Table 1. Parameter estimations of the mixed effects regression models: Tense-asp., tense-lax, and lax-asp pairs. The K-source stimuli condition was set as a reference level. Bold indicates $p<0.05$, and italic indicates $p<0.1$.

	Tense-Asp.		Tense-Lax		Lax-Asp.	
	Estimate	S.E	Estimate	S.E	Estimate	S.E
Intercept	-1.324	0.297	0.136	0.201	-1.465	0.232
VOT	6.332	0.42	1.81	0.15	4.12	0.299
f_0	0.188	0.164	-2.916	0.177	3.206	0.195
E-SourceCond.	0.739	0.32	0.334	0.233	<i>0.466</i>	<i>0.244</i>
VOT×E-SourceCond.	-0.391	0.413	-0.61	0.164	0.402	0.33
f_0 ×E-SourceCond.	0.012	0.175	0.121	0.164	<i>-0.35</i>	<i>0.213</i>

When the stimuli were made of the native source (solid lines), the regression curves of VOT dramatically increased across the three models as VOT increased, indicating that the changes of VOT values meaningfully influenced listeners in selecting one type of the stops over the other in each model of stop contrasts. The model summaries in Table 1 show that the VOT coefficients of the three models were statistically significant ($p<0.001$). When the VOT cue interacted with the E-source condition, the tense-lax model yielded a statistically significant interaction term between VOT and stimulus source condition ($\beta=-0.61$, $S.E=0.16$, $p<0.001$), indicating that the listeners' perceptual dependency on VOT, on average, was weakened in the E-source condition. The VOT slope in Fig.2-mid corresponding to the tense-lax model was shallower in the E-source condition than in the native source

condition. The two other models did not reveal statistically significant interaction terms between VOT and source conditions. Interacting with E-source condition, the reliance on VOT either increased (lax-asp. model: $\beta=0.40$) or decreased (tense-asp. model: $\beta=-0.39$).

The model coefficient summary showed that f_0 meaningfully distinguished the lax from the others while it was not a significant acoustic cue in the tense-aspirated model ($\beta=0.19$, S.E.=0.16, $p=0.25$) as in Table 1. The regression curves of f_0 in the K-source condition dramatically fell in perceiving the lax from the tense stops (Figure 2-middle) or abruptly rose in perceiving the lax from the aspirated stops (Figure 2-right). By contrast, the curve was relatively flat in the tense-aspirated model as the f_0 values increased. The role of f_0 in the tense-lax model was even greater than that of VOT according to the model-estimated beta coefficients ($\beta_{f_0}=-2.9$, $\beta_{VOT}=1.8$). The interaction terms of f_0 with the E-source condition were significant only in the lax-aspirated model, although marginal ($\beta=-0.35$, S.E.=0.21, $p=0.09$). The negative coefficient of the interaction term indicates that the role of f_0 was weakened in the E-source condition compared to the role in the K-source condition.

Overall, the group-averaged estimations of the acoustic variables in the K-source condition were consistent with previous findings of the perceptual cue weighting between VOT and f_0 in perceiving the Korean stops in that VOT was a sole acoustic cue for the tense vs. aspirated contrast, and also was a primary cue in differentiating the lax from the aspirated stops (Kong et al. 2011). Reflecting the enhanced role of f_0 in characterizing the lax from the other types (Silva 2006, Kang 2014, among others), f_0 played a secondary role for the lax vs. aspirated contrast and a primary role for the lax vs. tense contrast in the K-source condition. When the listeners processed the E-source stimuli, the group-averaged interaction effects of the stimulus source condition on utilizations of VOT and f_0 were not consistently significant across the three regression models of stop contrast pairs. Interestingly, it was consistent that the E-source condition weakened perceptual reliance on acoustic cues, when the stimulus source condition affected the use of acoustic cue with a statistical significance in the regression model.

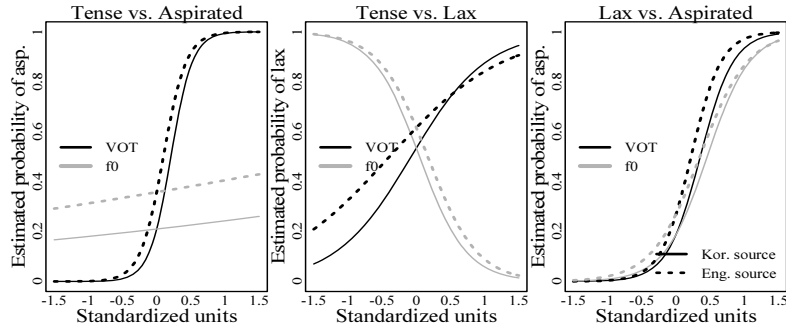


Figure 2. Inverse-logit curves based on the three mixed effects regression models: Tense-aspirated (left), tense-lax (middle) and lax-aspirated (right). The likelihood of the aspirated over the tense (left), the lax over the tense (middle) and the lax over the aspirated stops (right) is plotted as a function of VOT (black) and f_0 (grey) changes. Two stimulus source conditions were indicated by the line type (K-source: solid, E-source: dashed).

3.3 Effects of the English source: Individual characteristics

In order to further examine the effects of stimulus source condition on the use of the acoustic cue at the individual level, we conducted partial correlation tests between individual coefficient differences in each acoustic parameter ($\text{coefficients}_{[\text{E-source}]} - \text{coefficients}_{[\text{K-source}]}$) and individual coefficients of VOT and f_0 in the K-source condition. Tests were implemented in R using the package *ppcor* (Kim 2015). Individual English proficiency was considered as a control variable since their proficiency was correlated with the coefficient increases or decreases of VOT and f_0 between the source conditions in the models: Correlations with proficiency: VOT coef. difference_[tense-aspirated]: $r=0.23$, $p=0.097$; f_0 coef. diff._[tense-lax]: $r=0.33$, $p=0.015$; VOT coef. diff._[lax-aspirated]: $r=0.23$, $p=0.094$, f_0 coef. diff._[lax-aspirated]: $r=-0.25$, $p=0.064$.

Across the three contrast pairs of the stops, the variations of individual VOT coefficient differences between the source conditions were significantly correlated with the individual VOT coefficients of the native-source stimuli condition (Table 2 for the test summary). The top two rows of Figure 3 visualize the distributions of VOT coefficients. As in scatter-plots in the first row, the negative signs of the relationship suggest that listeners who depended on VOT more in the native source condition tended to rely on the cue less in the non-native source condition. This pattern is congruent with the findings from group-averaged model estimation presented in Table 1.

The second row panels in Figure 3 showed that individual VOT coefficients in the K-source condition also were correlated with their $f0$ coefficient differences between the stimulus source conditions for the tense-aspirated and lax-aspirated models. The correlation coefficients had positive signs exhibiting an association of greater individual VOT coefficients of the K-source condition with greater $f0$ coefficient differences between the source conditions (or greater increase of $f0$ coefficient in the E-source condition than in the K-source condition). This indicates that listeners who were more sensitive to VOT responding to the K-source stimuli tended to depend on $f0$ more in the E-source stimuli condition than in the native condition.

Similar to individual patterns of the VOT use in the E-source condition, individual $f0$ coefficients in the K-stimulus source condition were negatively correlated with the $f0$ coefficient differences between the source conditions (bottom two rows in Figure 3). The patterns were observed in the tense-lax, and lax-aspirated pairs but not in the tense-aspirated stop pair. The negative relationship shown in the second-to-bottom panels of Figure 3 indicates that individual listeners who relied on $f0$ more than others in the native source stimulus condition tended to depend on the $f0$ cue less in the E-source condition, and conversely listeners dependent on $f0$ less than others in the native source stimulus condition tended to depend on the $f0$ cue more in the E-source condition. Bottom panels showed that individual $f0$ coefficients of the K-source in the lax-aspirated pair were correlated with the individual VOT coefficient differences between the source conditions, showing that listeners with greater reliance on $f0$ in the K-source condition tended to depend on VOT more in the E-source condition than in the K-source condition.

Table 2. Summary of the partial correlation tests between individual VOT or $f0$ coefficients in the K-source condition and coefficient differences between the source conditions (E-source minus K-source).

		Tense-Asp.		Tense-Lax		Lax-Asp.	
VOT (K)	VOT Diff. (E-K)	-0.493	***	-0.421	***	-0.51	***
	$f0$ Diff. (E-K)	0.375	**	0.088		0.417	***
$f0$ (K)	VOT Diff. (E-K)	0.187		-0.11		0.385	***
	$f0$ Diff. (E-K)	-0.212		-0.396	***	-0.451	***

*** $p < 0.0001$, ** $p < 0.005$, * $p < 0.05$, . $p < 0.1$

To summarize, individual-level analysis confirmed the group-averaged interaction effects of the source conditions in that the E-source condition affected individual

listeners in utilizing the acoustic parameters, and further revealed that the magnitudes of the stimulus source effects across listeners were related to their reliance on VOT and f_0 in the native source condition. Across the contrast pairs, listeners who depended on one acoustic cue more than others in the K-source condition were negatively affected by the non-typicality of the stimulus source (i.e., E-source condition) in utilizing the cue, and they employed the other available cue more actively in the E-source condition than in the K-source condition, presumably to redistribute the attention.

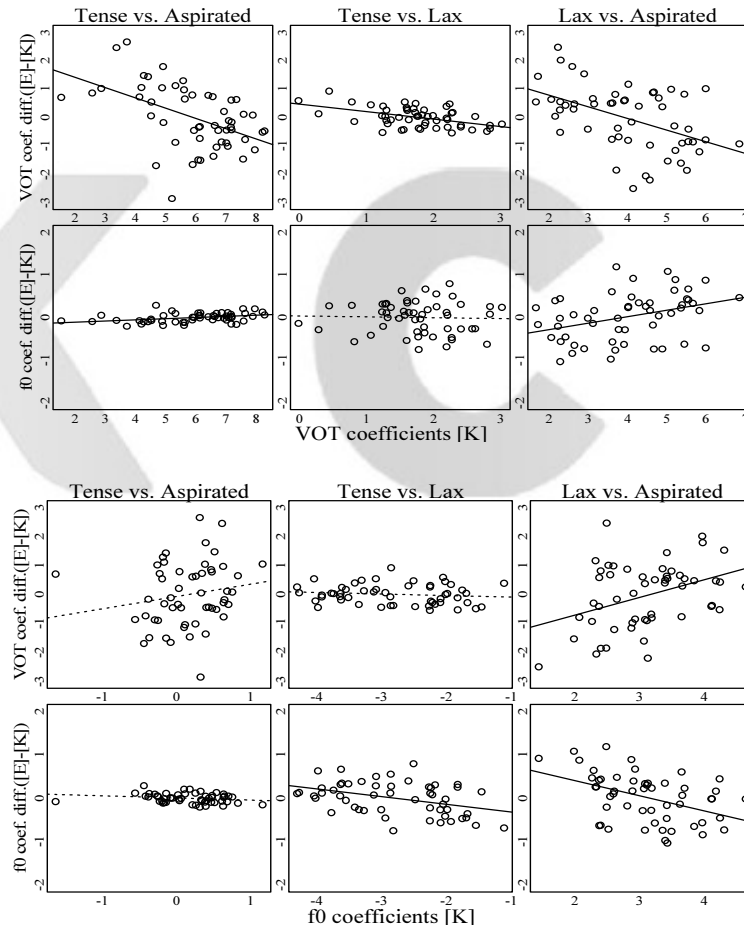


Figure 3. Scatter-plots of VOT and f_0 coefficient differences between the source types (y-axis) as a function of individual coefficients of VOT (top two panels) and f_0 (bottom two panels) in the K-source condition (x-axis).

3.4 Relationship between stimulus source effects and response categories

We explored whether the individual differences in utilizing VOT or f_0 between the stimulus source conditions can account for the response count differences between the conditions. Partial correlation tests were performed between individual coefficient differences of each acoustic cue between the conditions ($\text{coefficients}_{[\text{E-source}]} - \text{coefficients}_{[\text{K-source}]}$) and count differences of tense, lax and aspirated stops between the conditions ($\text{counts}_{[\text{E-source}]} - \text{counts}_{[\text{K-source}]}$). The tests considered individuals' English proficiency (TOEIC scores), and individual coefficients of acoustic variables (VOT and f_0) in the K-source condition as control variables.

Table 3. Partial correlation coefficients between category count differences (E-source minus K-source) and coefficient differences (E-source minus K-source).

		Tense-Asp.		Tense-Lax		Lax-Asp.
Tense Count Diff. (E-K)	VOT Diff. (E-K)	0.454	***	0.53	***	0.054
	f_0 Diff. (E-K)	-0.638	***	-0.354	*	0.104
Lax Count Diff. (E-K)	VOT Diff. (E-K)	-0.162		-0.397	**	-0.164
	f_0 Diff. (E-K)	0.168		0.123		0.266
Asp. Count Diff. (E-K)	VOT Diff. (E-K)	-0.407	**	-0.18		0.11
	f_0 Diff. (E-K)	0.64	***	0.292	*	-0.419

*** $p < 0.0001$, ** $p < 0.005$, * $p < 0.05$, . $p < 0.1$

As presented in Table 3, correlation test outputs showed that the count changes of each stop category between the stimulus source conditions were closely related to how greatly individual listeners were affected by the E-source condition in utilizing VOT and f_0 cues for the stop laryngeal perception. The correlation coefficients between tense count differences and acoustic coefficient differences were significant in the two relevant models (i.e., tense-asp., and tense-lax pairs) positively for the VOT cue, and negatively for the f_0 cue. That is, the decreased response counts of the tense type in the E-source condition were associated with increased use of f_0 and reduced use of VOT in the E-source condition. It is noted that the individuals' sensitivity differences to the f_0 cue between the source conditions were meaningfully correlated with their categorical decisions of tense responses, even though the group-averaged f_0 coefficient was not significant in the tense-asp model (see Table 1).

Similarly, the count differences of the aspirated type between the source conditions

were significantly correlated with individual coefficient differences of VOT and f_0 in the tense-aspirated model. The increased counts of the aspirated type were associated with reduced reliance on VOT and increased use of f_0 in the model. In the lax-aspirated model, the count differences of the aspirated type response were negatively related to the f_0 coefficient differences but not to those of VOT between the source conditions, although both variables were significant fixed effects according to the group estimation (Table 1).

The count differences of the lax type response between the source conditions, which were minimal among the three types, were negatively correlated with the VOT coefficient differences in the tense-lax model. The lax responses increased in the E-source condition as individual perceptual dependency on VOT decreased in perceiving the lax over the tense stops. In the lax-aspirated model, increased response counts of the lax category were associated with greater f_0 coefficients in the E-source condition than in the K-source condition, although the level of significance turned out to be marginal in the correlation test.

Overall, the findings of the correlation tests support that the listeners' use of acoustic cues affected by the non-canonical stimulus source condition was closely related to the count differences of the category types between the stimulus conditions. This set of individual-level analysis was useful in supplementing the group-averaged effects of the stimulus source conditions on the acoustic cue utilizations by showing that the individual magnitudes of perceptual sensitivities to acoustic variables in the E-source condition explained the response distributions of the stop categorizations in the perception of the E-source stimuli.

4. Discussion and Conclusion

The current study explored whether and how the stimulus alternations between Korean-source and English-source input conditions would affect Korean listeners' categorization patterns of the native stops, by examining listeners' utilizations of the two acoustic cues, VOT and f_0 . Despite the numerically identical conditions of the two target acoustic variables, the Korean listeners patterned differently between the two source language conditions in making categorical decisions of the stimuli into the native stops, yielding fewer responses of the tense stop and more responses of the aspirated stop in the English source stimuli condition compared to the native source condition. The study further found that listeners' differential uses of the acoustic cues

were responsible for these response distributions between the two stimulus conditions. Both in the group- and individual-level analyses, the perceptual reliance on the cues was mitigated in the English source condition and the response category distributions in the English source condition were correlated with individuals' alleviated or enhanced reliance on the acoustic cues. The findings suggest that non-canonical source information influenced listeners in processing speech input into the native categories, preventing them from utilizing the usual effective acoustic cues to the maximum and/or redistributing the attention to other available cues.

In exploring the effect of stimuli source in the categorization process, individual-level analyses could capture the gradient nature of the source effect on individuals' perceptual reliance on the acoustic cues and on their categorical judgment of the stimuli sets. Consistent with group-averaged estimations of the source language effects on the use of VOT and *f*0 in the category decisions, correlation relationship among individual coefficients confirmed the tendency that non-canonical stimulus source condition hindered the utilization of the acoustic cues.

Looking closely into individual variations, we further found that the magnitudes of the source effect on the acoustic cue utilization systematically varied across individuals in a way that listeners with greater reliance on one cue (either VOT or *f*0) in the Korean stimulus source condition not only depended less on the same cue in the English stimulus source condition but also relied on the other available cue more in the English source condition than in the Korean source condition. (Conversely, listeners who relied on one cue less than other listeners did in the Korean stimulus source condition attended to the same cue more in the English stimulus source condition.) Individual analyses were effective in showing the internal structures of the two integrated cues, which the group-level analyses of each fixed variable did not assess. Supporting the group-averaged patterns, individual tendencies showed that the non-canonical information in the speech signal leads to perceptual modifications of the acoustic cue uses in identifying the Korean stops.

With respect to modified uses of the acoustic cues in the English source condition, we reasoned that the modifications were made less likely to activate the English (the source language) perception module. Against the perception module of the English language that would encourage more emphasis on VOT than *f*0 in differentiating the voiceless-voiced stops, our results exhibited the *f*0 cue enhanced in the English source condition (It is noted that the *f*0 use was enhanced but did not outweigh the VOT use in the English source condition). Instead of language-specific accounts,

more general mechanism of attention modulation seems to play a role in explaining these modified uses of the acoustic cues in categorizing the non-canonical source stimuli into native stops (Gordon et al. 1993, Mattys et al. 2005, Francis et al. 2008). That is, general attentional mechanism works in a way that, in a non-ideal listening condition, listeners' reliance on usual primary cues that operate with high degrees of attention is weakened because attentional resources are limited. The role of secondary cues gets strengthened and less affected in this condition of limited attentional resources. In the tense-aspirated model where VOT was a primary cue (i.e., VOT but not *f*₀ was important acoustic information), listeners who depended on VOT more than other listeners in the native source condition tended to utilize the *f*₀ cue more in the English source condition while their reliance on the VOT cue decreased. This direction of cue modifications is opposite to the source language (English) characteristics in realizing the laryngeal contrast of the stops. Our interpretations of the current findings based on the general attention modulation mechanism also appear to work for the other two models of lax-tense and lax-aspirated contrasts when we regard VOT as a primary cue in the Korean stop perception (see conflict findings of cue-weightings in Kong et al. 2011, Lee et al. 2012, and Schurtz et al. 2015). In the current task of native stop processing, the general mechanism of attentional modulation under adverse listening condition (i.e., non-native source input) can provide a sensible account of the acoustic cue modifications in the current results.

While the current study provided experimental evidence that English source input affected listeners in the speech perception in a categorical (e.g., tense or lax or aspirated stops) manner, more studies need to follow in order to generalize the nature of perceptual modifications of sub-phonemic cues in processing non-native input. Would Japanese or French source input influence Korean listeners in processing the acoustic cues in the same way to English source input, as they are non-native (language internal noise) anyways? Then, to what degrees would language-specific or language-general processing mechanisms be involved in modifying the processing strategies and making category identifications? We leave these questions for future research.

To conclude, the present study showed the effects of English language source in Korean listeners' category identifications of the stops. Individual-level analysis was useful in confirming that Korean listeners' modified perceptual cue-weighting strategies were responsible for categorical decisions of the English source input into

Korean stops which are different from the ones of native source input.

REFERENCES

- ASSMANN, PETER and QUENTIN SUMMERFIELD. 2004. The perception of speech under adverse conditions. *Speech Processing in the Auditory System*, 231-308. New York: Springer.
- BATES, DOUGLAS, MARTIN MAECHLER, BEN BOLKER and STEVE WALKER. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67.1, 1-48.
- BEDDOR, PATRICE. 2009. A coarticulatory path to sound change. *Language* 85.4, 785-821.
- _____. 2015. The relation between language users' perception and production repertoires. *Proceedings of the 18th ICPHS*, 171-204.
- BOERSMA, PAUL and DAVID WEENINK. 2013. *Praat*: Doing phonetics by computer (Version 5.3.39) [Computer program].
- CHO, TAEHONG, SUN-AH JUN and PETER LADEFOGED. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of phonetics* 30.2, 193-228.
- FLOCCIA, CAROLINE, JESEPH BUTLER, JEREMY GOSLIN and LUCY ELLIS. 2009. Regional and foreign accent processing in English: Can listeners adapt?. *Journal of Psycholinguistic Research* 38.4, 379-412.
- FRANCIS, ALEXANDER and HOWARD NUSBAUM. 2002. Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance* 28.2, 349.
- FRANCIS, ALEXANDER, NATALYA KAGANOVICH and COURTNEY DRISCOLL-HUBER. 2008. Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *The Journal of the Acoustical Society of America* 124.2, 1234-1251.
- GORDON, PETER, JENNIFER EBERHARDT and JAY RUECKL. 1993. Attentional modulation of the phonetic significance of acoustic cues. *Cognitive Psychology* 25.1, 1-42.
- HAGGARD, MARK, STEPHEN AMBLER and MO CALLOW. 1970. Pitch as a voicing cue. *The Journal of the Acoustical Society of America* 47.2B, 613-617.

- HAN, MEIKO and REYMOND WEITZMAN. 1970. Acoustic features of Korean/P, T, K/, /p, t, k/ and /p^h, t^h, k^h. *Phonetica* 22.2, 112-128.
- HAZAN, VALARIE and STUART ROSEN. 1991. Individual variability in the perception of cues to place contrasts in initial stops. *Attention, Perception, & Psychophysics* 49.2, 187-200.
- IDEMARU, KORI, LORI HOLT and HOWARD SELTMAN. 2012. Individual differences in cue weights are stable across time: The case of Japanese stop lengths a. *The Journal of the Acoustical Society of America* 132.6, 3950-3964.
- KANG, YOONJUNG. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20.2, 219-273.
- _____. 2014. Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics* 45, 76-90.
- KAPNOULA, EFTHYMIA, MATTHEW WINN, EUN JONG KONG, JAN EDWARDS and BOB MCMURRAY. 2017. Evaluating the sources and functions of gradiency in phoneme categorization: an individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. <http://dx.doi.org/10.1037/xhp0000410>
- KIM, MI-RYOUNG, PATRICE BEDDOR and JULIE HORROCKS. 2002. The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30.1, 77-100.
- KIM, SEONGHO. 2015. ppcor: an R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods* 22.6, 665. <https://CRAN.R-project.org/package=ppcor>
- KONG, EUN JONG, MARY BECKMAN and JAN EDWARDS. 2011. Why are Korean tense stops acquired so early?: The role of acoustic properties. *Journal of Phonetics* 39.2, 196-211.
- KONG, EUN JONG and JAN EDWARDS. 2016. Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics* 59, 40-57.
- LEE, HYUNJUNG, STEPHAN POLITZER-AHLES and ALLARD JONGMAN. 2012. Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *Journal of Phonetics* 41.2, 117-132.
- MATTYS, SVEN, LAURENCE WHITE and JAMES MELHORN. 2005. Integration of

- multiple speech segmentation cues: a hierarchical framework. *Journal of Experimental Psychology: General* 134.4, 477.
- MUNRO, MURRAY and TRACEY DERWING. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45.1, 73-97.
- OH, MIRA. 2009. The phonetic approach to English stop adaptation in Korean and Thai. *Studies in Phonetics, Phonology and Morphology* 15.2, 283-299.
- PSYCHOLOGY SOFTWARE TOOLS, INC. [E-Prime 2.0]. 2012. [Computer program]. <http://www.pstnet.com>.
- SHELLINGER, SARAH, BENJAMIN MUNSON and JAN EDWARDS. 2017. Gradient perception of children's productions of /s/ and /θ/: A comparative study of rating methods. *Clinical Linguistics & Phonetics* 31.1, 80-103.
- SCHERTZ, JESSAMYN, TAEHONG CHO, ANDREW LOTTO and NATASHA WARNER. 2016. Individual differences in perceptual adaptability of foreign sound categories. *Attention, Perception, & Psychophysics* 78.1, 355-367.
- SCHOUTEN, BERT and ARJAN VAN HESSEN. 1992. Modeling phoneme perception. I: Categorical perception. *The Journal of the Acoustical Society of America* 92.4, 1841-1855.
- SCHOUTEN, BERT, ELLEN GERRITS and ARJAN VAN HESSEN. 2003. The end of categorical perception as we know it. *Speech Communication* 41.1, 71-80.
- SHI, LU FENG. 2009. Normal-hearing English-as-a-second-language listeners' recognition of English words in competing signals. *International Journal of Audiology* 48.5, 260-270.
- SILVA, DAVID. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23.2, 287-308.
- STEVENS, MARY and JONATHAN HARRINGTON. 2014. The individual and the actuation of sound change. *Loquens* 1.1, 003.
- TEAM, R CORE. 2016. *A Language and Environment for Statistical Computing* (Version 3.2.4) [Computer program]. <http://www.R-project.org>
- TER SCHURE, SOPHIE, KATEŘINA CHLÁDKOVÁ and JAN-WILLEM VAN LEUSSEN. 2011. Comparing identification of artificial and natural vowels. *Proceedings of the 17th International Congress of Phonetic Sciences*, 1770-1773.
- VAN HESSEN, ARJAN VAN and BERT SCHOUTEN. 1999. Categorical perception as a function of stimulus quality. *Phonetica* 56.1-2, 56-72.

Eun Jong Kong
Department of English
Korea Aerospace University
76 Hanggongdaehangro Deogyang-gu Goyang-si Gyeonggi-do
Korea 412-791
e-mail: ekong@kau.ac.kr

Soyoung Kang
School of Linguistics and Language Studies
Carleton University
1125 Colonel By Dr. Ottawa, ON K1S 5B6
Canada
e-mail: kang139@gmail.com

received: July 10, 2017
revised: August 6, 2017
accepted: August 15, 2017

K C I