Effects of dialectal differences in the use of native-language acoustic cues on the production and perception of second language stops*

Sooa Lim· Jeong-Im Han**
(Konkuk University)

Lim, Sooa and Jeong-Im Han. **2014.Effects of dialectal differences in the use of native-language acoustic cues on the production and perception of second language stops**. *Studies in Phonetics, Phonology and Morphology.* 20.3. 403-426. The present study investigated whether the relative importance of specific acoustic properties in native language dialects affects the extent to which acoustic properties are used in the production and perception of a second language. The case at hand centered on acquisition of English voiced and voiceless stops by two groups of Korean dialect speakers, Kyungsang and Seoul Korean. Both dialect speakers use the same acoustic cues for the Korean stops, but their relative importance is different: Specifically Seoul speakers employ VOT and F0, and Kyungsang speakers primarily use VOT due to their use of F0 for lexical tone contrast. Given such dialectal differences, this study explored whether Kyungsang and Seoul Korean speakers made distinct use of these two cues to signal the word-initial stops in English where VOT and, less importantly, onset F0 are additionally used for stop contrast. The results showed that in production, both dialect speakers showed a parallel pattern for VOT and F0; however, in perception, Kyungsang listeners had greater reliance on VOT but less on F0 compared with Seoul listeners. These results partly support the feature hypothesis (McAllister et al., 2002) and provide new insight into the relationship between speech production and perception (**Konkuk University**).

Keywords: L1 transfer, VOT, F0, English stops, Korean dialects, speech production and perception

1. Introduction

Adults who learn a second language (L2) have difficulties in producing and perceiving L2 phonetic segments despite exposure to the target language for a substantial amount of time. This challenge primarily exists because after acquisition of the phonology of one language (a native language, L1), adult learners tend to interpret sounds heard in L2 in terms of L1 phonological units (See Strange, 1995). Based on the assumption that the learner's native language plays a role in acquisition, most previous studies of L2 speech acquisition of phonetic contrasts have focused on the similarities and differences between segments of L1 and L2 (Best 1995, Best et al. 2001,Flege 1987, 1995,Flege et al., 1995, among others). While these studies have contributed significantly to the understanding of the importance of the

---

*This work was supported by Konkuk University in 2014. We are grateful to three anonymous reviewers for their valuable comments and constructive suggestions.
** Corresponding author

native language experience, they primarily addressed cross-linguistic relations at the segmental level.

A line of research has tried to define the role of the learner's L1 in the acquisition of L2 phonemes beyond a segmental level, showing that L2 perceptual cue weighting strategies are dependent on the learners' L1 (Escudero 2001,Escudero et al. 2009,Flege and Hillenbrand 1986, Gottfried and Suiter 1997, Holt and Lotto 2006, Iverson et al. 2003 among others). These studies argue that adult learners are likely to rely on their L1 experience when they learn how to integrate the information across acoustic cues in L2. As an example, take the classic case of English /l/ and /r/ perception by Japanese listeners (Iverson et al. 2003). For native English listeners, the third formant values (F3) carry most of the perceptual weight, whereas Japanese listeners appear to rely on the second formant (F2) values when they categorize the English /l/ and /r/. Due to this non-optimal cue weighting, Japanese listeners have great difficulty learning these categories in English. The influence of the mismatch between weighting functions appropriate for L1 and L2 was also shown in the acquisition of vowels. Escuderoet al. (2009) examined how Spanish and German learners of Dutch use the two acoustic cues, vowel duration and spectrum, for the categorization of Dutch /a:/-/ɑ/ contrast. It was found that both learners of Dutch were sensitive to vowel duration and spectrum in the vowel categorization, but German learners weight spectral information more heavily than duration, whereas Spanish learners prefer duration. The authors argued that the preference for one acoustic cue over the other between duration and spectrum in their L1 vowel perception leads to the differences in the perceptual cue weightings.

Lee et al. (2006) also noted phonetic features signaling vowel reduction in English, including comparatively lower F0, shorter duration, weaker intensity, and midrange formant frequencies, compared with the corresponding full vowels tested in the Korean and Japanese L2 learners to identify any difficulties in producing the English reduced vowels. Korean and Japanese differ in the prosodic system to realize the four phonetic features used to signal an English unstressed vowel: Korean speakers use the features of F0 but not the other features, whereas Japanese speakers use the features of F0 and duration but not the other features. Given the cross-linguistic difference in prosodic system, phonetic features used to signal a phonological difference in the native language were predicted to be produced accurately, and most predictions were upheld: Except for the intensity feature for Japanese speakers, both Korean and Japanese learners showed low production accuracy for the features which were not relevant to the phonology of their native language.

Similarly, McAllister et al. (2002) proposed a specific hypothesis that "the relative importance of a feature in the L1 will determine the extent to which the feature is successfully used in producing and perceiving phonological contrasts in the L2" (McAllister et al. 2002:254) (called "feature hypothesis"). To evaluate this hypothesis, McAllister et al. (2002) examined the acquisition of Swedish vowel length by Spanish, English, and Estonian L2 learners who differed in degrees of overall prominence of the quantity feature. Namely, the duration distinctions are salient in Estonian but not in Spanish. English is regarded as a language that does not exploit duration in its phonology, but native speakers of English are known to be sensitive to the duration feature (Whalen 1989). Thus, duration appeared to be most prominent in Estonian followed by English, and it was least prominent in Spanish. The results of an identification experiment revealed that the identification scores were obtained in the order of Estonian, English, and Spanish learners. A subsequent production experiment measured the size of differences in vowel-to-consonant ratios when L2 learners produced the words with long or short vowels. Only the Spanish learners were found to produce significantly smaller ratio differences in quantity distinctions compared with the native Swedish speakers. These results suggest that the Estonian learners who had vowel duration as a salient phonological feature were successful in facilitating their learning of Swedish quantity contrasts. On the contrary, the Spanish learners who had no phonological feature of duration had difficulty learning the Swedish quantity contrasts. Notably, the English learners showed intermediate level of performance attributed to their sensitivity to the duration distinctions, even though duration was not phonological.

However, some previous studies showed that adult L2 learners can incorporate cues to distinguish categories in their target language which are not in their L1, through exposure to the target language or universal saliency of specific acoustic cues (Bohn 1995,Escudero and Boersma 2004,Kondaurova and Francis 2008, Shea and Curtin 2010 among others). For instance, in Escudero and Boersma (2004), Spanish learners of English learned to use the duration cue in their perception of the English /i/ -/ɪ/ contrast, although their L2 relies primarily on spectral information to distinguish among vowels. Spanish learners were shown to notice that English vowels /i/ and /ɪ/ are differentiated by the duration cue and with increased language experience they learned to distinguish these two vowels in a way specific to their target language. Bohn (1995) explained similar findings with a universal saliency of duration. He showed that Mandarin Chinese as well as Spanish learners of English differentiated the English /i/ - /ɪ/ contrast by relying heavily on duration differences, even though neither language learners have had experience with them. To explain this language-independent perception, Bohn proposed a “desensitization hypothesis” that whenever spectral differences are insufficient to differentiate vowel contrasts because L1 experience do not sensitize listeners to these spectral differences, duration differences are used to differentiate the L2 vowel contrast. As native speakers of Spanish and Mandarin Chinese have only one vowel in the high-front area of the vowel space, they might be desensitized to spectral differences. This causes L2 listeners to differentiate the L2 vowels on the basis of duration differences, because duration cues are easy to access in vowel perception. Thus these studies suggest that adult L2 learners can use acoustic cues which are not in their L1.

In the present study, we further investigate the role of phonetic features in the native language on the production and/or perception of a second language. We specifically investigate whether the relative importance of specific acoustic properties in L1 dialects will affect the extent to which those acoustic properties are used in the production and perception of L2 contrast. The case at hand centers on acquisition of English voiced and voiceless stops in word-initial position as in *pie* versus *buy*, by two dialect groups of Korean speakers, Kyungsang and Seoul Korean. The Kyungsang and Seoul dialect speakers use the same acoustic cues to signal the Korean stops, namely VOT and F0, but their relative importance differs between these two dialects (Kenstowicz and Park 2006, Lee and Jongman 2012, Lee et al. 2013). Given dialectal differences, this study will explore whether Kyungsang and Seoul Korean speakers make distinct use of these two cues to signal the word-initial stops in English.

1.1Korean and English stops

Korean is widely known to have a three-way laryngeal contrast among voiceless stops such as lenis (e.g., *tal* 'moon), fortis (e.g., *t'al*'daughter'), and aspirated (e.g., t^h*al* 'mask') stops. Voice Onset Time (VOT), fundamental frequency of the onset of the following vowel (onset F0) have been identified as key acoustic correlates for such a 3-way contrast (Cho et al. 2002). However, the phonetic manifestation of the Korean stop contrast has been varying in recent years between these two acoustic cues. In the mid-1900s, VOT was identified as a key acoustic correlate for each stop category: fortis stops manifested with short VOTs, aspirated stops manifested with long VOTs, and lenis stops manifested with intermediate VOTs (Han and Weitzman, 1970; Kim, 1965). While older speakers still maintain clear VOT distinctions between lenis and aspirated stops, younger speakers tend to neutralize those differences (Silva 2006). This change in VOT patterns indicates that younger speakers likely use differences in the onset F0 as the primary acoustic cue to differentiate the Korean stops; the mean F0 of the vowel onset following a lenis stop is significantly lower than after an aspirated or fortis stop.

However, these descriptions generally fit standard Korean, namely, the Seoul dialect. The Kyungsang dialect of Korean, which is spoken in the south eastern part of the Korean peninsula, shows somewhat different pattern on the use of these two acoustic cues, VOT and onset F0 because this dialect preserves the pitch contrast. In this dialect, the pitch is primarily used to cue lexical tone contrast, High and Low, and consequently the use of pitch for the laryngeal contrast may be more constrained than in non-tonal dialects (Francis et al. 2006). Consequently, Kyungsangdialact speakers primarily use VOT, whereas Seoul dialect speakers use both VOT and onset F0 to distinguish between the three stops. In fact, Kenstowicz and Park (2006), Lee and Jongman (2012), and Leeet al. (2013) provided experimental results supporting this hypothesis. For example, Lee and Jongman (2012) examined the 3-way distinction among stops in Kyungsang and Seoul Korean, and found that ranges of F0 were well separated across the three stops for Seoul, but overlapped for Kyungsang, whereas VOT clearly distinguished them for Kyungsang, but overlapped between lenis and aspirated stops for Seoul. In a subsequent study by Lee et al. (2013), the perception of the 3-way distinction among Korean stops in these two dialects was explored. The VOT and F0 cues were shown to trade off each other for the perception of the three stops, but Kyungsang and Seoul Korean listeners were different in the use of these two cues: Seoul listeners used F0 cues primarily, whereas the role of this cue was more limited for Kyungsang than Seoul listeners. Summarizing these studies, different tonal systems between Kyungsang and Seoul Korean dialect speakers led to different cue weightings for the 3-way stop distinction in Korean.

In comparison to Korean, English stop consonants show a two-way contrast such as voiced (e.g., *buy*) versus voiceless (e.g., *pie*) stops. Previous research showed that VOT is typically considered the most salient cue among various acoustic cues (Abramson and Lisker 1985, Gordon et al. 1993,Lisker 1978, Whalen et al. 1993; Raphael 2005). In word-initial position the voiceless stops are produced with large amounts of aspiration after the release of the closure, while voiced stops show no aspiration or only a small amount of aspiration. Thus phonologically voiceless versus voiced stops in English are phonetically realized as aspirated (with long-lag VOT) versus unaspirated (with short-lag VOT) stops. Unlike VOT, onset F0 has been shown to be a less employed, secondary, redundant cue (Abramson and Lisker 1985, Gordon et al. 1993,Benki 2003, Kingston 1986, Whalen et al. 1993). The onset F0 is correlated with the phonological feature of voicing in initial stops such that the frequency following voiceless stops is higher than after voiced stops (Haggard et al. 1970,Ohde 1984, Whalen 1989). In non-tonal languages such as English, the influence of consonants on F0 is less constrained than in tonal languages because the pitch contour is not used for tonal judgments; however, individual differences have been shown in the use of this cue, which is not used by some speakers to cue the perception of voiced and voiceless stops in English (Haggard et al. 1970, 1981). Under conditions of high cognitive load, listeners are less likely to rely on VOT as a cue to voicing, and the relative weight to onset F0 correspondingly increases; however, onset F0 under ideal listening conditions is more likely to be ignored if VOT is unambiguous (Gordon et al. 1993). Thus English speakers appear to use onset F0 perturbations primarily to enhance a voicing contrast as a secondary cue.

1.2 The present study

The present study further examines the role of the L1 transfer. In addition to the influences of the native language segments (consonants and vowels), prosody, and distinctive features to the L2 contrasts as shown in previous studies, we examine whether relative importance of specific acoustic properties in the L1 dialect will affect the extent to which those acoustic properties are used in the production and perception of an L2 contrast. Given that VOT is preferred over onset F0 between these two cues to the English stop contrast in word-initial position, two dialect groups of Korean learners (Kyungsang versus Seoul) might use VOT more heavily than onset F0 for English stop contrast. However, it is not known whether Kyungsang dialect speakers make lesser use of the onset F0 than Seoul dialect speakers signaling the English stops in word-initial position. This question is based on the fact that F0 is not a reliable cue to distinguish the three-way Korean stop contrast for Kyungsang dialect speakers due to their use of F0 as a lexical tone contrast. We can hypothesize that the native language dialect pattern might be transferred to the production and perception of the L2 stop contrast. According to the hypothesis, the Kyungsang speakers might produce smaller frequency differences between voiced and voiceless stops than Seoul speakers; and the Kyungsang listeners might make less use of the frequency differences to perceive the English voicing contrast than Seoul listeners. To test the hypothesis, the present study reports the results of production of the English voiced and voiceless stops Kyungsang and Seoul Korean speakers, and their perception of the English stops.

2. Method

2.1Production

a.Participants:Two groups of Korean native speakers were recruited from Seoul and Kyungsang dialectal groups. Each group had six male speakers.[1]The age in the Seoul dialectal group ranged from 22 to 27 years old (mean = 24). The age in the Kyungsang dialect group ranged from 22 to 26 years old (mean = 24). All speakers in either dialect group had lived and had been educated in the target dialect area, and 58% of the Seoul dialect speakers' parents spoke the Seoul dialect and 75% of the Kyungsang dialect speakers' parents spoke the Kyungsang dialect. All speakers in each dialect group began to study English after the critical period and lacked experience

---

[1]The experiment was conducted in Seoul, the area for the standard Korean, when the Kyungsang dialect speakers moved to this city to enter the college. Previous sociolinguistic studies showed that men are known to be less influenced by the social stigma directed against the nonstandard forms, whereas women are likely to respond to overt prestige associated with them (Labov 1991). Thus only male speakers were tested in this study to manifest the dialectal effects on L2 acquisition more clearly.

living and studying in English speaking countries for more than 6 months. To make the English proficiency comparable between these two groups of speakers, each participant took the paper-based TOEFL practice test (PBT) (listening section). The mean scores of the listening test were 186 (sd = 9.1) out of 226 for the Seoul speaker group and 185 (sd = 17.6) out of 226 for the Kyungsang speaker group. Their scores were not significantly different [t(10)=.103, *p*>.05]. None of the speakers in either dialect group reported any speech or hearing disorders, and all speakers were paidfor their participation. The detailed biographical and languagebackground information of participants was given in Appendix A.

b. Stimuli:The stimuli consisted of a set of 50 English words (20 target words and 30 fillers). The target words were all monosyllabic real words with phonological voiced or voiceless initial stops as in Table 1. The voiceless stops occurring in word-initial position were pronounced as corresponding aspirated stops phonetically; and the corresponding voiced stops were realized as unaspirated. All words had CV(C) structure with various vowels such as /i/, /ɪ/, /æ/, /ʌ/, and /aɪ/ to control for possible effects of following vowels. The target stops were either bilabial or alveolar. Only real words were used to avoid the difficulty of eliciting nonsense syllables, even though the lexical frequency of some words was rather low. The filler words included various consonants except the bilabial and alveolar stops, and vowels in word-initial position and all other conditions were the same as the target words. Each word was included in a frame sentence 'The word is ____'.

Table 1. Target words

	bilabial		alveolar	
	/p/>[p^h]	/b/>[p]	/t/>[t^h]	/d/>[t]
[i]	peak	beak	team	deem
[ɪ]	pit	bit	tip	dip
[æ]	pat	bat	tab	dab
[ʌ]	pun	bun	tub	dub
[aɪ]	pie	buy	tie	die

c. Procedure: Eachparticipant was recorded in a sound-proof booth using a Tascam HD-P2 solid-state recorder and a Shure KSM 44 microphone. During the recording, individual PowerPoint files with sentences was shown to the participants through a window in the soundproof booth. Each sentence was displayed at a regular rate, and the subject was instructed to read each sentence shown. Each participant read the randomized sentences twice. The time to complete the recording was approximately 12 minutes. The recorded material was sampled at 22,050 Hz with 16-bit quantization. The VOT and F0 values of the target stops were measured using Praat. The VOT was measured from the consonant release to the onset of the second formant frequency of the following vowel in the waveform. The F0 was measured at the onset of voicing (beginning of the second formant frequency) and subsequently at 10, 20, 30, 40, 50, 60, 70, and 80 ms into the vowel. Because the inherent duration of each vowel in the target tokens might be different, and the most prominent F0 changes related to the preceding consonant were expected to be found during the first 100 ms after voicing onset (Hombert, 1978), only the first 80 ms after voicing onset were examined for F0 changes related to the stop consonants. The F0 values were extracted automatically at the start of a discernible second formant for each token, using Praat. The exact position for the onset of F2 was determined manually by researchers.

2.2 Perception

a. Participants:Two groups of Korean native speakers were recruited from Seoul and Kyungsang dialectal groups. Each group had 15 male speakers. The perception task was administered after the production task. The twelve speakers who participated in the production experiment additionally participated in the perception experiment.There was no significant difference in the results between the participants who had taken part in the production task (n=6) and those who had not (n=9) [$F(1,13)=1.10$, $p=.313$ for Seoul listeners; $F(1,13)=.05$, $p=.827$ for Kyungsang listeners]. The age in the Seoul dialectal group ranged from 20 to 27 years old (mean = 24). The age in the Kyungsang dialect group ranged from 19 to 27 years old (mean = 23). All speakers in each dialect group had lived and had been educated in the target dialect area, and their parents spoke the same target dialect at the rate of 60% for the Seoul dialect speakers, and 86% for the Kyungsang dialect speakers. All speakers in each dialect group began to study English after the critical period and lacked any experience living and studying in English speaking countries for more than 6 months. To make the English proficiency comparable between these two groups of speakers, each participant took the paper-based TOEFL practice test (PBT) (listening section). The mean scores of the listening test were 176 (sd = 17) out of 226 for the Seoul speaker group and 174 (sd = 16.7) out of 226 for Kyungsang speaker group. Their scores were not significantly different [$t(28)=.328$, $p>.05$].No speakers in either dialect group reported any speech or hearing disorders, and all speakers were paid for their participation. The detailed biographical and language background information of participants was given in Appendix B.

b. Stimuli:All stimuli were derived from a single CV syllable created by digitally splicing naturally produced, but resynthesized, burst and aspiration noise onto a synthetic vowel. To select the natural token for synthesis, 60 each of 'pea' and 'bee' examples were produced by a male native speaker of American English (Columbus, Ohio). He recorded the two examples in a frame sentence 'I say _____', in a sound-proof booth using Tascam cassette recorder (HD-P2) and Shure KSM 44 microphone. The mean VOT value for the /p/ tokens was 62.5 ms (40-82), and the onset F0 was 143.2Hz (126.3-

159.8), whereas the mean VOT for the /b/ tokens was 5.1 ms (4-6), and the onset F0 was 120.4Hz (109.6-130.7). Among the many 'pea' and 'bee' examples recorded, a voiceless token (/pi/) was chosen for the base rather than a voiced token (/bi/) because it is easier to reduce the aspiration noise already present than to add in aspiration noise that is not there. The aspiration noise and F0 values from this token were subsequently manipulated to give rise to approximately equal numbers of 'pea' and 'bee' percepts. Resynthesis was conducted with the "pitch synchronous overlap and add" (PSOLA) algorithm (Moulines and Charpentier 1990) as implemented in Praat. To make the VOT-related cues of the base word ambiguous, VOT was compressed in eight steps (60 ms, 50 ms, 40 ms, 30 ms, 20 ms, 10 ms, 5 ms, 0 ms). Sixty-four different syllables were generated from eight base tokens with the specific VOT values by fully crossing eight levels of onset F0 (110 Hz, 120 Hz, 130 Hz, 140 Hz, 150 Hz, 160 Hz, 170 Hz, and 180 Hz). The minimum and maximum F0 values were based on the range of F0 of production of /pi/ and /bi/ by the above male speaker. Each token began at specific F0 onset values, fell or rose over the course of the first 80 ms following the vowel onset, and ended at a point corresponding to 119 Hz (the mean F0 value of the original token at this point). All other properties remained the same across all tokens. In an informal testing, the resulting words were expected to be heard approximately equally as /p/ and /b/.

c. Procedure: The stimuli were presented binaurally in randomized order via a Sennheiser HD-590 headphone. Stimulus presentation and response collection was controlled by SuperLab Pro (Cedrus). Each listener was asked to hear a single syllable and identify it as either 'pea' or 'bee' by clicking on one of two buttons designated with the appropriate pictures (SnodgrassandVanderwart1980). Participants completed a total of four repetitions of 256 trials (8 levels of VOT x 8 levels of F0 perturbations x 4 repetitions) with a randomized stimulus order. Before testing, the example tokens with the maximum/minimum values of VOT and onset F0 were presented to the listeners to familiarize themselves (60 ms of VOT with 170 Hz F0 or 180 Hz F0 for 'pea' token; 0 ms of VOT with 110 Hz F0 or 120 Hz F0 for 'bee' token). The ISI was 1 second, and the maximum time for response was 3 seconds. The perception experiment took approximately 10 minutes. Following the experiment, each listener took the listening test of sample TOEFL (PBT) to evaluate their English proficiency, and the test lasted 29 minutes as explained as above.

3. Results

3.1Production

The first analysis was conducted on VOT. The results for the production of VOT were submitted to a mixed ANOVA with dialect (Kyungsang vs. Seoul) as between-subject variables and place (labial vs. alveolar) and voicing (voiceless vs. voiced) as within-subject variables. There were significant effects of voicing [F(1,10) = 201.2, *p*<.001], place [F(1,10) = 103.8, *p*<.001], and group [F(1,10) = 6.6, *p*<.05]. The interaction of voicing and group was significant [[F(1,10) = 17.4, *p*<.05]. However, there were no other significant interaction effects ([F(1,10) = .3, *p*>.05] for place*group, [F(1,10) = 1.5, *p*>.05] for voicing*place, [F(1,10) = .3, *p*>.05] for voicing*place*group).

Figure 1 shows the mean VOT values of the English bilabial and alveolar stops in two laryngeal contrasts, specifically word-initial voiced and word-initial voiceless produced by Kyungsang and Seoul dialect speakers.

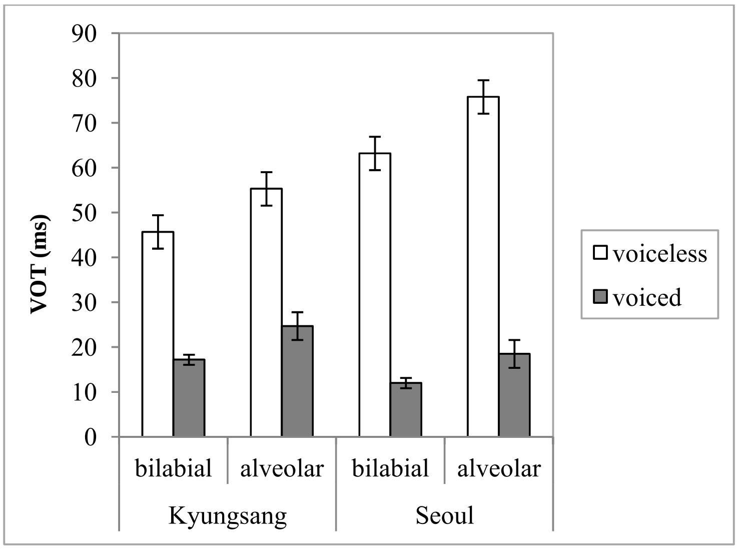

Figure 1. Mean VOT for English voiceless and voiced stops produced by Kyungsang and Seoul dialect speakers

Notably, the VOT values produced by Korean learners differentiated the voiced from voiceless stops in English, regardless of the dialect. In both groups of speakers, the VOT of the voiceless stop was significantly longer than the voiced stop. Both groups of speakers produced alveolar stops with longer VOT values than bilabial stops, which is common in other languages. To consider the interaction of voicing and group, a simple effect analysis was conducted and showed that the VOT values of voiceless stops were significantly different between Kyungsang and Seoul speakers [F(1,10) = 15, *p*<.05], but no significant difference was observed for voiced stops [F(1,10) = 5, *p*>.05]. These results indicate that Seoul speakers produced English voiceless stops with larger amount of aspiration than Kyungsang speakers, whereas both groups of speakers produced English voiced stops with similar VOT values.

To consider onset F0 values, Figure 2 and Figure 3 summarize the mean F0 values of voiced and voiceless stops in English from the onset to the first 80 ms point of the following vowel for Kyungsang and Seoul dialect speakers respectively.

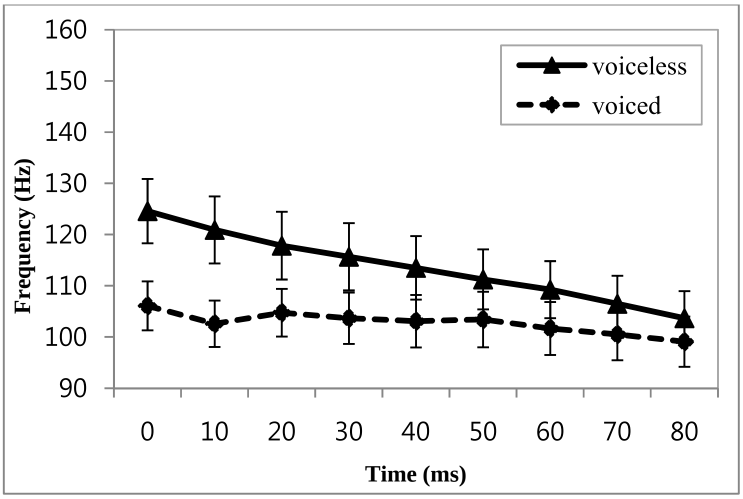

Figure 2. Mean F0 for English voiceless and voiced stops produced by Kyungsang dialect speakers

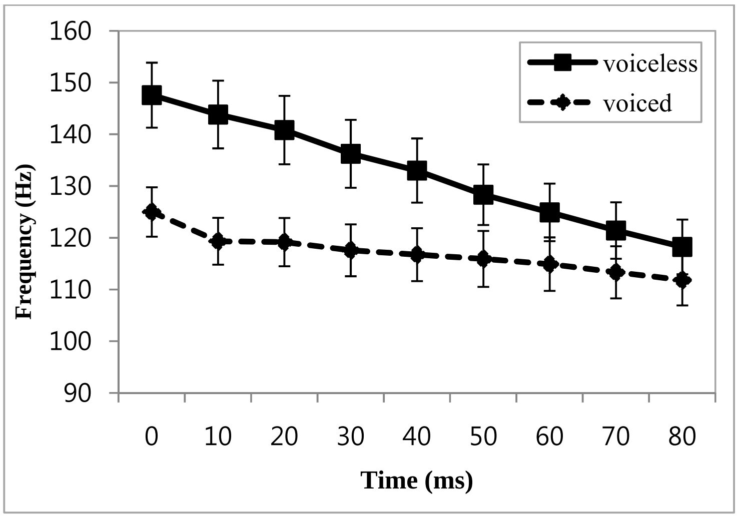

Figure 3. Mean F0 for English voiceless and voiced stops produced by Seoul dialect speakers

In both groups of speakers, the onset F0 values of the word-initial voiced stops were lower than the voiceless stops, and the frequency of the voiceless stops appeared to be rapidly falling over the 80 ms of the vowel, whereas the voiced stops appeared to be slowly falling or remain relatively steady in the midpoint of the vowel. The speakers' frequency values were submitted to a mixed ANOVA in which dialect (Kyungsang vs. Seoul) was a between-subject variable, and voicing (voiceless vs. voiced), place (labial vs. alveolar) and time (0 ms to 80 ms) were within-subject variables. The main effects of voicing [F(1,10) = 72, *p*<.001], time [F(8,80) = 37.8, *p*<.001] were significant, as was interaction [F(8,80) = 19.5, *p*<.001]. However, the main effects or interaction effects of all other factors were not significant ([F(1,10) = 5.1, *p* = .05 for dialect, [F(1,10) = 1.4, *p*>.05] for place, [F(1,10) = 2.3, *p*>.05] for voicing*dialect; [F(1,10) = 0, *p*>.05] for place*dialect; [F(8,80) = 2.1, *p*>.05] for time*dialect; [F(1,10) = 1.4, *p*>.05] for voicing*place; [F(1,10) = 2.1, *p*>.05] for voicing*place*dialect; [F(8,80) = .8, *p*>.05] for voicing*time*dialect; [F(8,80) = .7, *p*>.05] for place*time; [F(8,80) = 1.2, *p*>.05] for place*time*dialect; [F(8,80) = 1.7, *p*>.05] for voicing*place*time; [F(8,80) = .7, *p*>.05] for voicing*place*time*dialect). These results indicate that both dialect groups of speakers similarly used F0 to categorize the voiced and voiceless stops in English.

The frequency contour was shown to be extended to the steady state of the vowel, but the onset F0 more directly reflected the difference between voiced and voiceless stops. Thus, the difference of the onset F0 values (for 0 ms point) between voiced and voiceless stops was further compared between two groups of dialect speakers. The results showed that this comparison was not significantly different [F(1,10) = .96, *p*>.05], contrary to the prediction that Kyungsang dialect speakers may have less reliance on the frequency values to categorize voiced and voiceless stops in English than Seoul speakers.

3.2 Perception

Listener identification scores were submitted to a mixed ANOVA in which dialect (Kyungsang vs. Seoul) was a between-subject variable and VOT and F0 were within-subject variables. There were significant effects of VOT [F(4, 95) = 611, *p*<.001], F0 [F(4, 102) = 53, *p*<.001], and their interaction with the dialect group ([F(4, 95) = 5, *p*<.05] for VOT*dialect; [F(4, 102) = 3, *p*<.05] for F0*dialect), but the effect of dialect was not significant [F(1,27) = .0, *p*>.05]. Figures 4 and 5 present the identification patterns as a function of VOT and F0, respectively, for Kyungsang and Seoul dialect speakers. First, Figure 4 presents the function of VOT and responses were averaged across the 9 different F0 values.

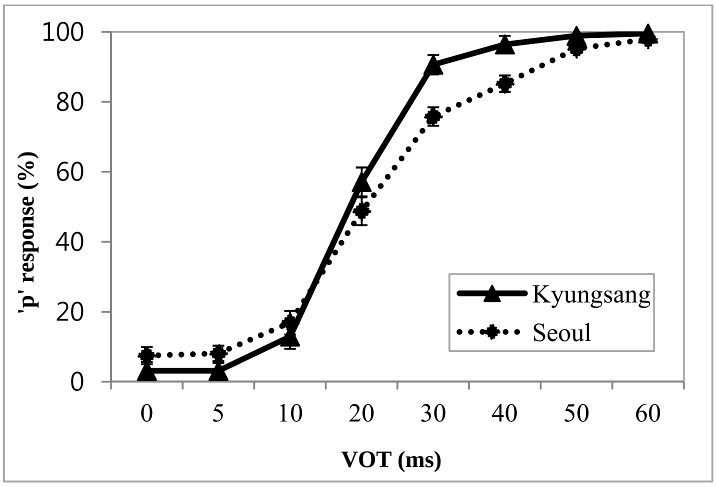

Figure 4. Proportion of ‘p’ identification responses as a function of VOT for Kyungsang and Seoul dialect listeners pooled across F0 transition

In Figure 4, an extremely short VOT (e.g., 5 or 10 ms) provided either dialect listener with a salient perceptual cue for the voiced stop, whereas a rather long VOT (e.g., 50 or 60 ms) was identified as the voiceless stop by both groups of listeners. More importantly, more than 90% of ‘p’ responses were obtained from Kyungsang listeners in the VOT range from 20 to 30 ms, whereas the same VOT range elicited fewer ‘p’ responses from Seoul listeners. The resulting curve representing the perceptual responses from Kyungsang listeners looked similar to the canonical type of categorical perception along the continuum, and Seoul listeners showed a relatively gradual increase in ‘p’ responses with a VOT increase. To break the interaction effect between the dialect and VOT, simple effect analysis was conducted and showed that two groups of listeners showed a significant difference in responses when the VOT values were 30 ms [$F(1,28) = 14$, $p<.05$] and 40 ms [$F(1,28) = 14$, $p<.05$]. This result demonstrates that the VOT values were more effective for the distinction between the English voiced and voiceless stops in Kyungsang listeners than in Seoul listeners.

However, the F0 did not seem to provide a strong perceptual cue for the perception of English voiced and voiceless stops.Figure 5 represents the function of F0, averaging the response across VOT values.

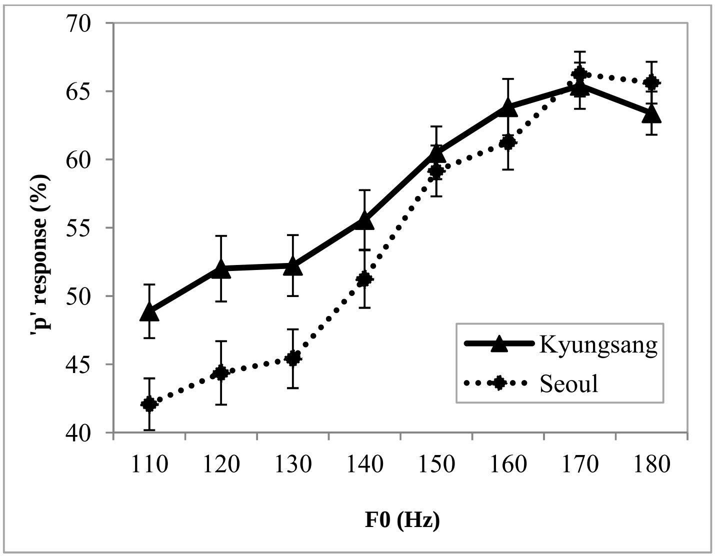

Figure 5. Proportion of 'p' identification responses as a function of F0 for Kyungsang and Seoul dialect listeners pooled across VOT transition

Figure 5 showed that both dialect groups of listeners elicited slightly increased 'p' responses as F0 increased. Even in F0 values, the two dialect group of listeners appeared to use F0 in a different way as suggested in the interaction effect of dialect and F0. In low F0 ranges, the proportions of 'pea' responses were approximately 8% greater for Kyungsang listeners than Seoul listeners. In the highest F0 value, Seoul listeners perceived 66% of the stimuli as a word of 'pea', whereas the proportion was only up to 63% for Kyungsang listeners. Simple effects showed that two groups of listeners showed a significant difference in responses when the F0 values were 110 Hz [$F(1,28) = 7$, $p<.05$], 120 Hz [$F(1,28) = 6$, $p<.05$], 130 Hz [$F(1,28) = 6$, $p<.05$], and 180 Hz [$F(1,28) = 6$, $p<.05$]. Thus Seoul listeners showed a relatively sharper perceptual curve than did Kyungsang listeners.This result indicates that the same F0 range was more effective for Seoul listeners than Kyungsang listeners in the categorization of English voiced and voiceless stops.

The overall identification pattern appeared to be similar between Kyungsang and Seoul dialect listeners: both groups of listeners were likely to use VOT as a primary cue, and F0 as a secondary cue to distinguish word-initial voiced from voiceless stops in English. However, a closer inspection revealed a difference between two dialect listeners.

4. Discussion

The present study examined the production and perception of the English stops in word-initial position by two dialect groups of Korean speakers (Kyungsang versus Seoul). Kyungsang speakers use VOT as a primary cue and onset F0 as a secondary cue to distinguish Korean stops, while Seoul speakers rely more on onset F0 than VOT (Kenstowicz and Park 2006, Lee and Jongman 2012, Lee et al. 2013). This study tested whether relative importance of specific acoustic properties between two dialect groups of speakers was reflected in the extent to which those acoustic properties were used in the production and perception of an L2 contrast. As mentioned in the Introduction, VOT is a primary cue and F0 is a secondary, redundant cue to distinguish English voiceless and voiced stops. We predicted that both groups of Korean learners of English might use VOT more heavily than onset F0 for English stop contrast; however, Kyungsang dialect speakers might make less use of the onset F0 than Seoul dialect speakers due to their use of F0 for a lexical tone in their native language dialect.

In production, it was observed that the VOT variation showed a parallel pattern between the two dialect groups of speakers: the VOT of the voiceless stop was significantly longer than the voiced stop, even though the VOT difference between voiced and voiceless stops was greater in Seoul than Kyungsang speakers, which is a somewhat unexpected result. With regard to the F0 values, the two dialect groups of speakers similarly used the frequency change to distinguish the voiced and voiceless stop contrast. The onset F0 was lower for the voiced stop than the voiceless stop, and the F0 contour in the following 80 ms of vowel intervals consequently fell rapidly in the case of voiceless stops, whereas the frequency fell rather slowly or even remained steady in the case of voiced stops. The results of the production experiment indicate that both dialect speakers similarly used VOT and F0 to categorize English voiced and voiceless stops.

Kyungsang and Seoul dialect listeners exhibited differences in perceptions of VOT and frequency changes. Both dialect listeners used VOT as a primary cue and F0 as a redundant cue to distinguish the English voiced from voiceless stops. However, Kyungsang listeners showed a sharper curve in the responses as a function of VOT than Seoul listeners, suggesting that Kyungsang listeners had more reliance on the VOT values to categorize the stop contrast in English than Seoul listeners. In the responses as a function of F0, Seoul listeners showed a relatively sharper curve than Kyungsang listeners, although both listeners showed only a little change in the responses as the frequency increased. The results of the perception experiment indicate that both dialect groups of listeners appeared to primarily depend on VOT to discriminate the voiceless and voiced stops in English, and the role of F0 might be redundant. More importantly, two dialect listeners' use of two acoustic cues for the English stop voicing contrast was influenced by their own phonetic constraints during performance of perception. Thus, these two listeners were not listening exactly like English listeners.

Combining the results of production and perception experiments indicates that the relative importance of the acoustic properties, VOT and F0, between two dialect groups of speakers was not greatly reflected in the extent to which those acoustic properties were used in the production and perception of the English stop contrast. This result might be related to the fact that VOT played a stronger role than F0 to distinguish English voiced and voiceless stops. All participants in the present study had learned English for more than 10 years at the time of experiment, even though participants had no experience living in English-speaking countries, and they appeared to acquire use of acoustic cues to the laryngeal contrast in English. With enough experience, the VOT was already a robust and significant cue to the voicing contrast sufficient to distinguish English voiced and voiceless stops. The F0 was simultaneously only used as a secondary cue to enhance the voicing contrast. Furthermore, the onset F0 differences in English stops were shown to be relatively subtle in comparison to the F0 differences between Korean stops. At vowel onset, the frequency differences were 15-20 Hz for English stops and 50-60 Hz for Korean stops (Chang 2010, 2012). Additionally, the frequency differences in Korean were shown to be further extended to the steady state of the vowel, indicating that the frequency differences in the steady state of the vowel between Korean lenis and aspirated stops were larger than the F0 differences between English stops (Oh 2011). These facts lead us to conclude that Korean learners did not have to use a secondary, redundant cue such as (onset) F0 for economic reason. To obtain a clearer understanding of the acquisition of VOT and F0 values, it is necessary to test the Korean learners at much earlier stage of learning.

Notably, two dialect groups of Korean speakers with sufficient L2 experience showed a slightly different perceptual pattern, where their own use of two acoustic cues in their native language dialect appeared to be reflected in the identification of English voiced and voiceless stops. When this finding is considered in terms of the feature hypothesis (McAllister et al. 2002), the relative importance of the acoustic cues in the native language dialect directly determined the extent to which those cues were successfully used in perceiving phonological contrast in the second language. Considering that speakers of the two Korean dialects showed very similar English proficiency as shown in the test scores, it seems plausible that the participants were transferring their native experience with two acoustic cues, VOT and F0, to the acquisition of the English voicing contrast. The Kyungsang dialect speakers appeared to have difficulties in manipulating frequency as a voicing cue because they use F0 to cue lexical tone contrast in their native language. For this reason, these speakers tended to have greater reliance on VOT.

It must be noted that the production and perception experiments exhibited non-parallel results; differences between Kyungsang and Seoul dialect speakers in terms of their use of the two acoustic cues to the English voiced and voiceless stops were observed in perception, but not in production. The present results are contrary to the recent findings that show positive and even strong correlation between production and perception in the second language acquisition (Borden et al.1983,Flege 1993,Bradlow et al. 1997,Perkell et al. 2004 among others). In Bradlow et al. (1997), Japanese learners of English participated in a period of identification training of English /r/ and /l/ and showed a significant improvement in the production of these two segments. Without any production training, the knowledge gained from the perceptual training was shown to be transferred to the production domain.

Shultz (2011) and Shultz et al. (2012) attempted to explain this discrepancy with distinct functional goals in speech production and speech perception. Shultz (2011) and Shultz et al. (2012) investigated individual variation in English native speakers' relative weighting of two acoustic cues to the English stop voicing contrast, VOT and onset F0 in production and perception tasks, in order to explore the connections between speech production and perception. The results of the production task showed a significant negative correlation between VOT and onset F0, whereas the results of the perception task showed a nearly significant tendency toward a positive correlation between these two cues. The production results suggest that there is a trade-off in the use of these two cues: all of the participants were primary users of VOT, but some placed greater emphasis on VOT and de-emphasized onset F0. This result is consistent with the hypothesis by Repp (1982), who stated that the VOT and onset F0 exist in a trading relationship, and is in agreement with Kingston et al. (2008), who argued that speakers deliberately manipulate acoustic cues that give rise to similar consequences to integrate and enhance the perceptual difference for a certain contrast. On the contrary, the perception results suggest that participants show high or low weightings for both acoustic cues and indicate that only some people can accurately perceive the target phonetic cues.

Shultz (2011) and Shultz et al. (2012) argued that the non-parallel results of production and perception may reflect different goals in production and perception. In production, each person may show different weightings of acoustic cues and thus there can be multiple ways to accurately produce his/her message to other people, whereas in perception, people should be able to notice multiple cues to accurately perceive the speech sounds produced by people who use different combinations of acoustic cues. Consequently, speakers tend to show a trade-off in the use of acoustic cues to preserve the amount of effort needed to convey a message, while listeners try to use as many acoustic cues as possible to understand the speaker's message. Similar interpretation can be applied to the results of the present study. In production, two groups of Korean learners of English employed the VOT primarily to distinguish the voiced from voiceless stops in a similar fashion because this cue seemed to be sufficient to distinguish these two stops. However, in perception, both Kyungsang and Seoul Korean listeners tried to use any cue available to distinguish voiced from voiceless stops, regardless of VOT or F0. In this process, participant experience of native language dialect directly influenced the perception of the stop categories.

Overall, the present findings might not provide conclusive evidence to support the feature hypothesis, but the present results deepen our understanding of the L1 transfer and address the question of what aspects of the native language are actually transferred. In addition to the transfer at the segmental and prosodic level, the results show that relative importance of specific acoustic cues in the L1 dialect could affect the extent to which those acoustic properties are used in the perception (and production) of an L2 contrast. Based on the present findings, a complete model of cross-linguistic phonetic influence must account for cross-language developments that occur at a sub-segmental level.

Appendix A. Characteristics of the speakers in the production experiment

	age	gender	AOL	listening test scores /226	dialect
1	22	m	13	190	Seoul
2	25	m	14	186	Seoul
3	23	m	8	186	Seoul
4	24	m	9	173	Seoul
5	27	m	12	180	Seoul
6	23	m	10	200	Seoul
7	23	m	10	167	Kyungsang
8	22	m	14	173	Kyungsang
9	23	m	9	193	Kyungsang
10	24	m	9	183	Kyungsang
11	26	m	14	178	Kyungsang
12	25	m	14	216	Kyungsang

AOL=age of learning onset (years)

Appendix B. Characteristics of the listeners in the perception experiment

	age	gender	AOL	listening test scores /226	dialect
1	22	m	13	190	Seoul
2	25	m	14	186	Seoul
3	23	m	8	186	Seoul
4	24	m	9	173	Seoul
5	27	m	12	180	Seoul

6	23	m	10	200	Seoul
7	24	m	8	167	Seoul
8	25	m	13	160	Seoul
9	20	m	10	173	Seoul
10	23	m	11	167	Seoul
11	23	m	10	160	Seoul
12	20	m	9	160	Seoul
13	26	m	11	216	Seoul
14	24	m	10	160	Seoul
15	27	m	10	160	Seoul
16	23	m	10	167	Kyungsang
17	22	m	14	173	Kyungsang
18	23	m	9	193	Kyungsang
19	24	m	9	183	Kyungsang
20	26	m	14	178	Kyungsang
21	25	m	14	216	Kyungsang
22	27	m	14	176	Kyungsang
23	23	m	12	167	Kyungsang
24	25	m	13	158	Kyungsang
25	23	m	13	167	Kyungsang
26	19	m	8	157	Kyungsang
27	20	m	11	157	Kyungsang
28	24	m	10	183	Kyungsang
29	20	m	9	180	Kyungsang
30	25	m	10	153	Kyungsang

AOL=age of learning onset (years)

REFERENCES

ABRAMSON,ARTHURS. and LEIGHLISKER. 1985. Relative power of cues: F0 shift versus voice timing. In V. Fromkin(ed.).*Phonetic Linguistics*, 25-33. New York: Academic.

BENKI, JOSE R.2003. Place of articulation and first formant transition type both affect perception of voicing in English. *Journal of Phonetics* 29, 1-22.

BEST, CATHERINE T.1995.A direct realist view of cross-languagespeech perception. InW. Strange (ed.).*Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 171-204. Baltimore: York Press.

BEST, CATHERINE T., GERALD W. MCROBERTS and ELIZABETH GOODELL. 2001. Discrimination ofnonnative consonant contrasts varying in perceptualassimilation to the listener's native phonological system.*Journal of the Acoustical Society of America* 109, 775-794.

BORDEN, GLORIA, ADELE GERBER and GARY MILSARK. 1983. Production and perception of the /r/-/l/ contrast in Korean adults learning English. *Language Learning* 33(4), 499-526.

BORN, O.-S. 1995. Cross language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 279-304. Baltimore: York Press.

BRADLOW, ANN R., DAVID B. PISONI, REIKO A. YAMADAand YOH'ICHITOHKURA.1997. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production.*Journal of the Acoustical Society of America* 101, 2299-2310.

CHANG, CHARLES B.2010.The implementation of laryngeal contrast in Korean as a second language.*Harvard Studies in KoreanLinguistics* 13, 91-104.

.2012. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal ofPhonetics* 40, 249-268.

CHO, TAEHONG, SUN-AHJUN and PETER LADEFOGED.2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30, 198-228.

ESCUDERO, PAOLA.2001. The role of the input in the development of L1 and l2 sound contrasts: Language-specific cue weighting for vowels. In A. H.-J. Do, O. Dominguez and A. Johansen (eds.). *Proceedings of the 25th annual Boston University conference on language development*, 250-261. Somerville, MA: Cascadilla Press.

ESCUDERO, PAOLA andPAULBOERSMA. 2004. Bridging the gap between L2 speech perception research and phonological theory. Studies in Second Language Acquisition 26, 551-585.

ESCUDERO, PAOLA,TITIABENDERS andSILVIALIPSKI. 2009. Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners.*Journal of Phonetics* 37, 452-466.

FLEGE, JAMES EMIL.1987. The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15, 47-65.

. 1993. Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America* 93.3, 1589-1608.

. 1995. Second language speech learning: Theory, findings and problems. InW. Strange (ed.).*Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 233-273. Timonium, MD: York Press.

FLEGE, JAMES EMILand JAMESHILLENBRAND. 1986. Differential use of temporal cues to the /s/-/z/contrast by native and non-native speakers of English. *The Journal of the Acoustical Society of America* 79.2, 508-517.

FLEGE, JAMES EMIL, MURRAY J. MUNRO and LAN RA MACKAY.1995.

Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*97, 1-26.

FRANCIS,ALEXANDER L., NATALYAKAGANOVIC, and COURTNEY DRISCOLL-HUBER, J. 2008.Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English.*Journal of the Acoustical Society of America* 124.2, 1234-1251.

FRANCIS, ALEXANDER L., VIRGINIA VALTER, KA MANWONG and JESSKA LAM CHAN. 2006. Is fundamental frequency a cue to aspiration in initial stops? *Journal of the Acoustical Society of America* 120.5, 2884-2895.

GORDON, PETER C., JENNIFER L.EBERHARDT and JAY G. RUECKL.1993. Attentional modulation of the phonetic significance of acoustic cues. *Cognitive Psychology* 25, 1-42.

GOTTFRIED, TERRY L. and TRACI L. SUITER.1997. Effect of linguistic experience on the identification of Mandarin tones. *Journal of Phonetics* 25, 207-231.

HAGGARD,MARK, STEPHEN AMBLER andMO CALLOW.1970. Pitch as voicing cue. *Journal of the Acoustical Society of America* 47, 613-617.

HAGGARD, MARK P., QUENTIN SUMMERFIELD and MARTIN ROBERTS. 1981. Psychoacoustical and cultural determinants of phoneme boundaries: Evidence from trading F0 cues in the voiced-voiceless distinction. *Journal of Phonetics* 9, 49-62.

HAN, MIEKO S. and RAYMOND S. WEITZMAN.1970. Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica*22, 112-128.

HOLT, LORI and ANDREW J.LOTTO. 2006. Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*119, 3059-3071.

HOMBERT, JEAN-MARIE. 1978. Consonant types, vowel quality and tone. In V. Fromkin (ed.).*Tone: A Linguistic Survey*, 77-111. New York: Academic.

IVERSON, PAUL, PATRICIA K. KUHL, REIKO AKAHANE-YAMADA, EUGEN DIESCH, YOH'ICH TOHKURA, ANDREAS KETTERMANN and CLAUDIA SIEBERT. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87.1, B47-B57.

KENSTOWICZ, MICHAEL and CHIYOUN PARK.2006. Laryngeal features and tone in Kyungsang Korean: A phonetic study. *Studies in Phonetics, Phonology and Morphology* 12.2, 247-264.

KIM, CHIN-WU.1965. On the autonomy of the tensity feature in stop classification. *Word* 21, 339-359.

KINGSTON, JOHN.1986. Are F0 differences after stops deliberate or accidental? *Journal of the Acoustical Society of America* 79, S27 (abstract).

KINGSTON, JOHN, RANDY L. DIEHL, CECILIA J.KIRKand WENDY A.CASTLEMAN,2008. On the internal perceptual structure of distinctive features: The [voice] contrast. *Journal of Phonetics* 36.3, 28-54.

KONDAUROVA, MARIA V. and ALEXANDER L. FRANCIS. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America* 124.6, 3959-3971.

LEE, BORIM, SUSAN G. GUION and TETSUOHARADA. 2006. Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition* 28.3, 487-513.

LEE, HYUNJUNG and ALLARDJONGMAN. 2012. Effects of tone on the three-way laryngeal distinction in Korean: An acoustic and aerodynamic comparison of the Seoul and South Kyungsang dialects. *Journal of the International Phonetic Association* 42.2, 145-169.

LEE, HYUNJUNG, STEPHEN POLITZER-AHLESand ALLARD JONGMAN. 2013. 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, 117-132.

LISKER, LEIGH.1978. In qualified defense of VOT. *Language and Speech* 21, 375-383.

MCALLISTER, ROBERT, JAMES E. FLEGEand THORSTEN PISKE. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30, 229-258.

MOULINES, ERIC and FRANCIS CHARPENTIER. 1990.Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones.*Speech Communication*9, 453-467.

OH, EUNJIN.2011. Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics* 39.1, 59-67.

OHDE, RALPH N.1984.Fundamental frequency as an acoustic correlate of stop consonant voicing.*Journal of the Acoustical Society of America* 75,224-230.

PERKELL, JOSEPH S.,FRANK H. GUENTHER,HARLAN LANE, MELANIE L. MATTHIES, ELLEN STOCKMANN, MARK TIEDE andMAJIDZANDIPOUR2004. The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *Journal of the Acoustical Society of America* 116, 2338-2344.

RAPHAEL, LAWRENCE J.2005.Acoustic cues to the perception of segmental phonemes.InD. B.Pisoni and R. E.Remez (eds.).*The Handbook of Speech Perception*, 182-206. Malden:Blackwell.

REPP, BRUNO H.1982.Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological*

Bulletin 92, 81-110.

SILVA, DAVID J.2006.Acoustic evidence for the emergence of tonal contrast in contemporary Korean.*Phonology* 23, 287-308.

SNODGRASS, JOAN G. and MARY VANDERWART.1980. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory* 6.2, 174-215.

SHEA,CRISTINA andSUZANNECURTIN. 2010. Discovering the relationship between context and allophones in a second language: Evidence for distribution-based learning. *Studies in Second Language Acquisition*, 32.4,581-606.

SHULTZ, AMANDA A. 2011.*Individual Differences in Cue Weighting of Stop Consonant Voicing in Perception and Production*. MA Thesis. Purdue University.

SHULTZ, AMANDA A., ALEXANDER L. FRANCISand FERNANDO LLANOS. 2012. Differential cue weighting in perception and production of consonant voicing. *Journal of the Acoustical Society of America* 132.2, EL95-101.

STRANGE, WINIFRED.1995.*Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York Press.

WEINBERGER, STEVEN H.1990.Minimal segments in second language phonology.InJ.Leather and A.James (eds.).*New Sounds 90, Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-Language Speech*, 72-97. Amsterdam: University of Amsterdam.

WHALEN, DOUG H. 1989. Vowel and consonant judgments are not independent when cued by the same information. *Perception and Psychophysics* 46, 284-292.

WHALEN, DOUG H.,ARTHUR S. ABRAMSON, LEIGH LISKERandMARIAMODY. 1993. F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America 93*, 2152-2159.

Sooa Lim
Department of English
Neungdong-ro 120, Gwangjin-gu
Seoul, 143-701, Korea
Email: lssua@hanmail.net

Jeong-Im Han
Department of English
Neungdong-ro 120, Gwangjin-gu
Seoul, 143-701, Korea
Email: jhan@konkuk.ac.kr

received: October 22, 2014
revised: November 28, 2014
accepted: December 8, 2014