말소리와 음성과학 제3권 제2호 (2011)
pp. 53~61

The interlanguage Speech Intelligibility Benefit for Korean Learners of English: Production of English Front Vowels

Jeong-Im Han[1] · Tae-Hwan Choi[2] · Injae Lim[3] · Joo-Kyeong Lee[4]

ABSTRACT

The present work is a follow-up study to that of Han, Choi, Lim and Lee (2011), where an asymmetry in the source segments eliciting the interlanguage speech intelligibility benefit (ISIB) was found such that the vowels which did not match any vowel of the Korean language were likely to elicit more ISIB than matched vowels. In order to identify the source of the stronger ISIB in non-matched vowels, acoustic analyses of the stimuli were performed. Two pairs of English front vowels [i] vs. [ɪ], and [ɛ] vs. [æ] were recorded by English native talkers and two groups of Korean learners according to their English proficiency, and then their vowel duration and the frequencies of the first two formants (F1, F2) were measured. The results demonstrated that the non-matched vowels such as [ɪ], and [æ] produced by Korean talkers seemed to show more deviated acoustic characteristics from those of the natives, with longer duration and with closer formant values to the matched vowels, [i] and [ɛ], than those of the English natives. Combining the results of acoustic measurements in the present study and those of word identification in Han et al. (2011), we suggest that relatively better performance in word identification by Korean talkers/listeners than the native English talkers/listeners is associated with the shared interlanguage of Korean talkers and listeners.

Key words: the interlanguage speech intelligibility benefit (ISIB), Korean front vowels

1. Introduction

It is very important to see how mutual intelligibility is maintained in the communication among non-native speakers of English as well as that between native speakers and non-native speakers, given that non-native speakers of English now outnumber native speakers (Crystal, 2003). Investigations concerning the intelligibility of native and non-native speech for native and non-native listeners have shown that a variety of factors play a role such as rate of speech, signal-to-noise ratio, speaking style, language background, proficiency of speakers and listeners, word frequency, neighborhood density, and availability of contexts (Derwing & Munro, 2001; van Wijngaarden, Steeneken & Houtgast, 2002; Bradlow & Bent, 2002; Bradlow & Pisoni, 1999; Imai, Walley & Flege, 2005; Mayo, Florentine & Buus, 1997).

Among them the language background of talkers and listeners has been shown to be crucial in determining the intelligibility of speech (Munro & Derwing, 2005; Bent & Bradlow, 2003; van Wijngaarden, 2001; Hayes-Harb, Smith, Bent & Bradlow, 2008 among others). In general, it has been assumed that native listeners find native speech more intelligible than non-native speech, but more interestingly, non-native talkers' speech to non-native listeners who share the L1 is sometimes as intelligible as or even more intelligible than native talkers' speech, because they share the same interlanguage (similar L2 phonological representations). Bent and Bradlow (2003) called this the 'interlanguage speech intelligibility benefit' (subsequently referred to as the ISIB), and Hayes-Harb et al. (2008) further put forth more refined definitions such as ISIB-L (listeners) vs. ISIB-T

1) Konkuk University, jhan@konkuk.ac.kr
2) Konkuk University, bak2san@hanmail.net
3) Konkuk University, injaelim@konkuk.ac.kr
4) University of Seoul, jookyeong@uos.ac.kr
This work was supported by the National Research Foundation of Korean Grant funded by the Korean Government (NRF-2009-32A-A00137). We would like to thank three anonymous reviewers for valuable comments on earlier version of the paper.

(talkers). The ISIB-T compares the intelligibility of native vs. non-native talkers for non-native listeners, and ISIB-L compares the intelligibility of non-native talkers for native vs. non-native listeners as in Figure 1.

'ISIB-T'

native talker　　　　non-native talker

↘　　↙

non-native listener

'ISIB-L'

non-native talker

↙　　↘

non-native listener　　native listener

Figure 1. ISIB-T (upper) and ISIB-L (lower)

The ISIB-T is concerned with cases in which speech by non-native talkers is more intelligible to non-native listeners than speech by native talkers, while the ISIB-L refers to cases where non-native speech is more intelligible to non-native listeners than it is to native listeners.

It is noteworthy that there have been two approaches to the definition of the ISIB. Bent and Bradlow (2003) originally defined the ISIB as the case non-native listeners found non-native speech at least as intelligible as native speech, whereas Stibbard and Lee (2006) and Hayes-Harb et al. (2008) questioned this definition, arguing that the word 'benefit' should be used only for the case where performance by non-native listeners exceeds that by native listeners. In this paper, we will adopt both, namely, the definition applied literally (Stibbard & Lee, 2006; Hayes-Harb et al., 2008) and that applied less strictly (Bent & Bradlow, 2003), considering implications from both definitions.

In Han et al. (2011), we evaluated the ISIB-T and ISIB-L with Korean learners of English in an EFL context, and more specifically examined their identification of English front vowels, focusing on the effects of learners' proficiency and task types on the ISIB. It was shown that Korean listeners found the Korean talkers' speech at least as intelligible as the native speech, and the Korean talkers' speech was equally intelligible to Korean and native listeners (the 'ISIB-T' and the 'ISIB-L' in a less strict sense). But there was no indication that the performance by Korean talkers/listeners surpassed that by native talkers/listeners (the 'ISIB-T' and the 'ISIB-L' in a literal sense). When the proficiency of talkers/listeners and the availability of contexts were manipulated, these two factors were not shown to be crucial to either ISIB-T or ISIB-L.

Another important question raised in this study was if there is any difference in the speech intelligibility between the L1-matched vowel and the non-matched, new vowel in L2. Among English vowels, front vowels such as [i], [ɪ], [ɛ] and [æ], have been reported to give much difficulty to the Korean learners of English (Ingram & Park, 1997; Flege, Bohn & Jang, 1997 among others). These are all front vowels, [i] and [ɪ] as high vowels, [ɛ] as a mid, and [æ] as a low vowel. The two high vowels are distinguished from themselves by tenseness, [i] as a tense vowel and [ɪ] as lax, and [ɛ] and [æ] are both lax vowels. These vowels are hard to process because the two high front vowels, [i] and [ɪ], are matched to the Korean vowel [i], and the mid and low front vowels, [ɛ] and [æ], are matched to the Korean [ɛ], showing a typical 1-to-2 mapping between L1 and L2 (Ingram & Park, 1997; Flege et al., 1997). Originally the Korean language had [ɛ] and [æ] as separate phonemes, but they have been merged and represented as [ɛ] (Hong, 1991). Based on the findings obtained from earlier acoustic measurements (F1 and F2) by Yang (1996), the Korean [i] is reported to be closer to the English [i] than [ɪ] and that the Korean [ɛ] is closer to the English [ɛ] than [æ] (Flege et al., 1997). In addition to the lower number of phonemes, the distinction between phonemically long and short vowels ([nun] 'eye' vs. [nu:n] 'snow') is not maintained in Korean except in some Southern dialects (Magen & Blumstein, 1993; Kim & Han, 1998). This might lead the Korean learners not to be attuned to the spectral and temporal differences associated with the tense-lax distinction to distinguish [i] and [ɪ].[5)]

All these facts led us to predict that Korean learners might have much difficulty in processing the two pairs of front vowels in English. As expected, our previous results of word identification demonstrated an asymmetry between the L1 matched vowels and non-matched vowels in the intelligibility, and the ISIB was observed more frequently in the L1 non-matched vowels than in the matched vowels. As for the case where the performance of Korean learners surpassed that of the natives, only the ISIB-L for HP talkers was observed in the English [i] vowel, while the ISIB-L for LP talkers and the ISIB-T for both HP and

5) As a reviewer pointed out, it is not necessarily the case that the L2 contrasts which are not in the native language of listeners entail a great deal of difficulty in perception, because certain contrasts are salient enough to be perceived by L2 listeners.

LP listeners was found in the case of [ɪ]. Even in the case of the ISIB in the less strict sense, only [ɪ] showed the ISIB-L for HP talkers. Similar to the results for high front vowels, more ISIB was observed for [æ] than [ɛ]. Simply the vowel [ɛ] did not show the ISIB in the literal sense, and only the ISIB-T for LP listeners was observed in the less strictly applied case. On the other hand, the low vowel [æ] showed a strong case of ISIB-L and a less strong case of ISIB-T for LP listeners (See Han et al. (2011) for more details).

The main goal of the present study is to identify the source of the stronger ISIB in non-matched vowels than matched vowels. To this end, acoustic analyses of the stimuli were performed. This is based on the assumption that L2 talkers' interlanguage may differ from one another (HP vs. LP learners) or from the native speech in the use of acoustic cues in the categorization of L2 sounds (Bohn, 1995; Flege et al., 1997; McAllister, Flege & Piske, 2002). Hayes-Harb et al. (2008) also examined the acoustic correlates of intelligibility for listeners and talkers from the same native language background in addition to the identification tasks, but they examined the acoustic correlates of intelligibility only with the talkers' speech. The listeners' speech as well as the talkers' speech should be included in the acoustic analysis to fully understand the source of the ISIB.

2. Method

The stimuli and participants were the same as in our previous study (Han et al., 2011).

2.1. Speech Materials

The stimuli were two sets of English minimal pairs of the CVC form, each set contrasting [i] vs. [ɪ], and [ɛ] vs. [æ] (*beat-bit*, *sheep-ship*, *peak-pick*, *feel-fill*, *read-rid* for high-vowel pairs, *bet-bat*, *pen-pan*, *lend-land*, *shell-shall*, *head-had* for non-high vowel pairs). Each token was recorded twice in a frame sentence, "I would say _____.". The same number of fillers which included different kinds of consonants and vowels from the test words was also recorded.

2.2. Participants

Forty-two participants of two different proficiency levels were selected based on the results of two diagnostic tests and a questionnaire. First, 75 voluntary Korean EFL students were recruited, all of whom were college students at the researchers' institutions with various majors. Each of the students visited the phonetics lab and took two kinds of diagnostic tests: paper-based TOEFL (PBT) practice test and Versant English Test (VET). The full-length paper-based TOEFL practice test was used to measure the subjects' proficiency of listening, vocabulary, structure and reading. In order to measure the participants' accentedness and ability in the sub-skills of speaking proficiency, which were not included in the paper-based TOEFL, the Versant English Test was conducted. VET is a computer-based oral proficiency test which measures non-native English students' pronunciation, vocabulary, sentence mastery, and fluency in approximately 15 minutes through an automated scoring system. VET has been used either in replacement of or in addition to native speakers' accentedness rating due to its convenience and cost-effectiveness. After each participant took these two diagnostic tests, he/she completed the questionnaire on his/her personal English learning history.

Based on the results of these two tests and the questionnaire, 21 high proficiency (HP) and 21 low proficiency (LP) levels of subjects were selected out of 75. Each participant's ranking was determined by the combined scores of TOEFL and VET, where those in the topmost range were assigned to the HP group and those in the bottommost range to the LP group. Anyone with an exceptionally high or low score on either of the tests was excluded as a means to maintain the groups as homogeneous as possible. Three students were excluded in the analysis, even though they were within the range of either group, because they were either found to have speaking impairments or could not complete both tests for personal reasons.

To confirm that the HP and LP subjects differed in their sub-parts of the test ratings as well as overall ratings, an independent samples T-test was conducted: it was determined that the HP and LP groups differed significantly in VET scores, PBT scores, and SumVET/PBT. The statistics showed that all these three measures were significantly different between the HP and LP groups [$t(40)=9.024$, $p=.000$ for VET; $t(40)=14.887$, $p=.000$ for PBT; $t(40)=17.138$, $p=.000$ for sum of VET/PBT]. For each proficiency group, three male and three female participants were selected randomly to form talker groups of different proficiency (HP talker and LP talker). The rest became listener groups (HP listener and LP listener). There was no statistical difference in the test scores between the talkers and the listeners for both proficiency groups [for HP, $t(19)=.84$, $p=.41$ for VET, $t(19)=-.42$, $p=.69$ for PBT, $t(19)=.175$, $p=.86$ for sum of VET/PBT; for LP, $t(19)=1.073$, $p=.30$ for VET, $t(19)=-.319$, $p=.75$ for PBT, $t(19)=.19$, $p=.85$ for sum of VET/PBT].

In addition, 21 English natives who spoke the standard

American English accent were recruited: 6 English native talkers and 15 English native listeners. The native talker group (NE talker) consisted of EFL instructors at the researchers' institutions and four of them were from the Mid-west regions of the U.S., one, from the Massachusetts, and the other, from Canada. As with the Korean EFL groups, three male and three female participants formed the talker group. The native listener group (NE listener) was recruited from the students at the University of Oregon to obtain a uniform variety of English dialects.

Table 1. Biographical and language background of participants

	NE listener	NE talker	HP listener	HP talker	LP listener	LP talker
Sex	M9, F6	M3, F3	M1, F14	M3, F3	M8, F7	M3, F3
Age	21.5 (18-39)	30.7 (18-39)	21.1 (19-24)	24.2 (20-25)	22.4 (21-27)	20.8 (19-23)
PBT	none	none	562.1 (526-593)	557.2 (546-589)	465.7 (426-506)	462.0 (420-486)
VET	none	none	136.3 (113-172)	143.7 (130-161)	93.7 (68-113)	100.3 (88-110)
LOR	N/A	N/A	2.2 month (5-13)	6.2 month (13-22)	none	none
AOA	N/A	N/A	18.6 year (5-13)	19.5 year (8-13)	N/A	N/A
AOL	N/A	N/A	10.7 year (5-13)	11.5 year (8-13)	10.7 (7-13)	10.3 (8-13)

(PBT=TOEFL paper-based test scores; VET=Versant English Test scores; LOR=length of residence; AOA=age of arrival; AOL=age of learning)

All other subject variables were as similar as possible to those of the other groups. The biographical and language backgrounds of six groups of participants are shown in Table 1.

2.3. Procedure

The recordings of six participants for each group (HP talker, HP listener, LP talker, LP listener, native talker) were randomly chosen and used for acoustic measurements. The recording was done right after the word identification task performed for our previous study (Han et al., 2001). The test scores of six listeners of both HP and LP groups were not significantly different from those of the other nine listeners out of fifteen listeners [in HP listeners, t(13)=-1.11, p=.29 for VET, t(9.46)=1.57, p=.15 for PBT; in LP listeners, t(13)=-.96, p=.36 for VET, t(13)=-.67, p=.52 for PBT]. For each word stimulus, vowel duration, and the frequencies of the first two formants (F1, F2) were estimated using waveform and spectrogram display in Praat (Boersma, 2001), based on previous findings that F1 and F2 serve as the primary determinants of vowel quality and are particularly salient among a myriad of acoustic cues signaling vowel quality (Ettlinger & Johnson, 2010).

Vowel duration was measured from the second positive peak in the periodic portion of each digitized waveform to the constriction of the post-vocalic consonant, which was signaled by a decrease in overall amplitude and waveform complexity. With the vowel duration measures, relative vowel duration (vowel duration/word duration) was calculated to prohibit the influence of the speaking speed on the duration measures. The formant values were measured at the midpoint of each vowel within its steady state. The measurements were taken by hand in the center of the bandwidth from an FFT wide-band spectrogram. However, if formant values were not clearly measurable, the LPC spectral slice display was supplemented. As there were male and female speakers, and native speakers of Korean and those of American English, direct comparison of formant values among the participants was not possible; therefore, the formant values were normalized according to Nordstrom and Lindblom's model (1975). First, a low vowel whose F1 was greater than 600 Hz was selected, which was [æ]. And then the F3 values of this vowel were measured in randomly chosen examples of both male and female speakers of Korean and English. The mean F3 of the English males was taken as the norm and was divided by the mean F3 of each of the other groups, which led to a uniform scale factor k. This scale factor was multiplied uniformly to the F1 and F2 values of the other group, producing normalized values. The total number of tokens for acoustic measurements was 600 (20 words x 5 speaker groups x 6 subjects), but several recorded tokens with lateral codas (*shell*, *shall*) or nasal codas (*land*, *lend*, *pen*, *pan*) were hard to demarcate between the vowel and the coda, which were excluded from analysis. To see the source of the ISIB explicitly, only the tokens which elicited the ISIB effects were separately examined, following Hayes-Harb et al. (2008). For each of the four vowels, tokens for ISIB-L or ISIB-T both in the literal sense and in the less strict sense were sorted and analyzed.

3. Results

Table 2 presents mean vowel/word ratio and the first two formant values (F1 and F2) for the five groups of participants, NE speakers and HP talkers and listeners, and LP talkers and listeners. To see whether there were any significant differences of the acoustic properties of the target vowels between talkers and listeners for either the HP or LP groups, an independent samples T-test was performed, and revealed no significant differences among these groups for any of the acoustic measures, except the vowel/word ratio of [i] and [ɛ] between HP talkers and listeners. For the group differences (NE, HP and LP), a one-way ANOVA was performed to see if there was a significant difference in temporal and spectral measures of the target vowels. First, there was a significant difference in the F1 values for [i] [$F(2,15)=8.178$, $p<.05$], [ɪ] [$F(2,15)=16.96$, $p<.05$], and [æ] [$F(2,15)=8.07$, $p<.05$], but not [ɛ] [$F(2,15)=.93$, $p>.05$]. A post-hoc (LSD) test showed that for all three vowels indicating significant differences, NE speakers were significantly different from the HP and LP groups, but such differences were not observed between the HP and LP groups. Second, the F2 values did not reveal any significant differences among the five groups. Finally, the ratio of vowel per word showed significant differences for the three vowels, [ɪ], [ɛ] and [æ], but not for [i] [$F(2,15)=2.704$, $p<.05$ for [ɪ]; $F(2,15)=4.05$, $p<.05$ for [ɛ]; $F(2,15)=5.96$, $p<.05$ for [æ]; $F(2,15)=3.31$, $p>.05$ for [i]]. A post-hoc (LSD) test showed that for [ɪ] and [ɛ], there were significant differences only between HP and LP learners, and the results of [æ] showed significant results between NE speakers and HP learners, and between HP and LP learners.

Next, we examined the acoustic characteristics of the vowel tokens eliciting the ISIB either in the literal sense or in the less strict sense, in order to see if there is any correlation between the intelligibility data and the acoustic characteristics of the stimuli. Specifically we are interested in the issue of whether there is any comparable asymmetry between the matched and non-matched vowels in the acoustic measures, given that non-matched vowels showed stronger ISIB than the corresponding matched vowels.

The acoustic measures eliciting the ISIB-L were first examined, following Hayes-Harb et al. (2008), based on the “ISIB tokens” (the word tokens produced by Korean talkers, and Korean listeners were better than or similar to the native listeners) and “English high-accuracy tokens” (henceforth “EH-A tokens”) (the word tokens showing the highest identification scores by NE natives). Namely, the word tokens produced by Korean talkers that resulted in the greatest discrepancies between Korean listeners’ and NE listeners’ word identification accuracy were selected. Out of 120 tokens (4 types of vowels x 5 words x 6 participants), eight items each for the ISIB tokens and EH-A tokens were included in the analysis, following Hayes-Harb et al. (2008).[6)]

The main purpose of this analysis was to see if there is any difference in the acoustic properties that the NE speakers and Korean learners may have used in identification. Table 3 shows the acoustic properties of the ISIB tokens and the EH-A tokens for the vowel tokens showing ISIB-L effects, which revealed the differences in the temporal and/or spectral cues Korean listeners and the native listeners might use in identification. In Table 3, the first three columns reveal the detailed identification results taken from our previous study only for the vowels eliciting the ISIB-L (Han et al., 2011), for cases operated both literally and less strictly. To help to understand the acoustic measures, we also presented the word identification scores of the ISIB and EH-A tokens (See "perception scores") for English native and Korean listeners.

As a statistic analysis was not feasible due to the small number of tokens, only the general pattern was examined. It can be seen that in the [i]-[ɪ] pair, the vowel [i] showed a large difference in the F2 values, with higher F2 values in the EH-A tokens (2086Hz) than the ISIB tokens (1926Hz), suggesting that Korean listeners identified the tokens in a less peripheral position (more like [ɪ]) as a high front tense vowel as compared to NE listeners. As for the [ɪ] vowel, there was a larger vowel/word

6) We selected eight tokens for each group to examine the properties of stimuli that elicited more accurate word identification by Korean learners and less accurate identification by English native listeners. However, an anonymous reviewer suggested us to conduct a mixed effects logistic regression analysis with word identification scores as a categorical dependent variable, and acoustic measures (vowel/word duration ratio, F1 and F2) as independent variables. We leave this for a future study.

duration ratio, lower F1 values, and higher F2 values in the ISIB tokens than the EH-A tokens, suggesting that Korean listeners identified the [ɪ] vowel tokens with much longer duration (tense version of [ɪ]) and produced in the region similar to [i] (less lowered) as compared to the NE listeners. On the other hand, the vowel [æ] did not show any temporal differences in the two kinds of tokens, and only F2 values showed a consistent pattern such that the ISIB tokens were shown with higher F2 values than the EH-A tokens.

Table 2. Mean acoustic measures by five different groups of speakers (NE, HP talkers/listeners, LP talkers/listeners)

Subjects	Target vowels		Vowel/ Word ratio	F1 (Hz)	F2 (Hz)
NE talkers	High	i	.30	277.0	2268.0
		ɪ	.27	442.7	1850.0
	Non-high	ɛ	.33	634.0	1746.0
		æ	.40	796.3	1581.0
HP talkers	High	i	.26	335.0	2040.3
		ɪ	.24	347.8	1957.2
	Non-high	ɛ	.30	643.7	1690.7
		æ	.31	652.3	1687.5
HP listeners	High	i	.32	336.7	2263.3
		ɪ	.26	398.2	2207.2
	Non-high	ɛ	.36	713.5	1796.5
		æ	.36	712.8	1732.2
LP talkers	High	i	.33	356.3	2122.5
		ɪ	.31	364.2	2095.7
	Non-high	ɛ	.35	572.7	1773.3
		æ	.37	593.5	1765.2
LP listeners	High	i	.31	343.5	2132.7
		ɪ	.31	341.8	2106.8
	Non-high	ɛ	.39	601.3	1760.8
		æ	.38	606.0	1783.3

Table 3. Mean acoustic measures for ISIB tokens and EH-A tokens for each vowel eliciting the ISIB-L

Target vowel	Talkers	ISIB		Token types (#)	Perception scores		Acoustic measures		
		literally→ less strictly			NE	Korean	F1 (Hz)	F2 (Hz)	V/W duration ratio
[i]	HP talker	Y		ISIB tokens (8)	5	59.1	364	1926	.20
		(LP>NE)		EH-A tokens (8)	72.4	50	324	2086	.26
[ɪ]	HP talker	N → Y		ISIB tokens (8)	15.2	56.2	330	2176	.28
		(HP=NE)		EH-A tokens (8)	95	58.1	351	1832	.26
	LP talker	Y	LP>NE	ISIB tokens (8)	23.1	62.73	348	2164	.39
		(LP=HP>NE)		EH-A tokens (8)	64.6	48.3	375	2205	.26
			HP>NE	ISIB tokens (8)	20.6	63.3	331	2180	.44
				EH-A tokens (8)	63	44.4	389	2109	.24
[æ]	HP talker	Y		ISIB tokens (8)	7.6	60.4	586	1679	.32
		(LP>NE)		EH-A tokens (8)	71.3	36.8	698	1620	.35
	LP talker	Y	LP>NE	ISIB tokens (8)	11.7	55.9	632	1750	.33
		(LP=HP>NE)		EH-A tokens (8)	55.2	44.7	609	1712	.37
			HP>NE	ISIB tokens (8)	11.7	44.1	616	1832	.39
				EH-A tokens (8)	50.7	29.3	609	1675	.34

(Y=yes, N=no; NE=native English, HP=high proficiency, LP=low proficiency; V/W=vowel/word; A<B means 'the identification accuracy scores of A are lower than those of B')

Similar to the ISIB-L, the acoustic properties of the matched vs. non-matched vowel tokens eliciting the ISIB-T between the Korean talkers and the native English talkers were compared. The same analytic procedure for the ISIB-L was not allowable for the ISIB-T in that there was only a single group of listeners. For this reason, the word tokens produced by either Korean or NE talkers and showing the highest identification accuracy were selected ("KH-A tokens": Korean high-accuracy tokens vs. "EH-A tokens": English high-accuracy tokens). Out of 360 tokens (4 types of vowels x 5 words x 6 talkers x 3 participant groups), eight items each for [i], [ɪ], [ɛ], and [æ] were selected for each of the three participant groups and analyzed.

Table 4. Mean acoustic measures for KH-A tokens and EH-A tokens for each vowel eliciting the ISIB-T

Target vowel	Listener	ISIB	Talker	Token types (#)	Perception scores	Acoustic measures		
		literally→ less strictly				F1 (Hz)	F2 (Hz)	V/W duration ratio
[ɪ]	HP listener	Y	LP	KH-A tokens (8)	92.5	365	2011	.29
		(LP=HP>NE)	HP	KH-A tokens (8)	94.9	366	1923	.23
			NE	EH-A tokens (8)	93.0	478	1891	.23
	LP listener	Y	LP	KH-A tokens (8)	90.7	366	1912	.23
		(LP=HP>NE)	HP	KH-A tokens (8)	84.9	363	1916	.24
			NE	EH-A tokens (8)	77.8	473	1886	.22
[ɛ]	LP listener	N→Y	LP	KH-A tokens (8)	58.8	563	1746	.35
		(NE=LP=HP)	HP	KH-A tokens (8)	63.7	643	1690	.27
			NE	EH-A tokens (8)	63.6	608	1759	.34
[æ]	LP listener	N→Y	LP	KH-A tokens (8)	67.0	650	1714	.31
		(NE=LP=HP)	HP	KH-A tokens (8)	75.6	652	1686	.29
			NE	EH-A tokens (8)	58.5	878	1584	.35

(Y=yes, N=no; NE=native English, HP=high proficiency, LP=low proficiency; V/W=vowel/word; A<B means 'the identification accuracy scores of A are lower than those of B')

Table 4 showed a very similar pattern to that of ISIB-T: Korean learners showed a large difference from the English natives in the non-matched vowels such as [ɪ] and [æ] rather than the matched vowel [ɛ].

As for both [ɪ] and [æ], the F1 values were very different between the Korean and English talkers. Korean learners had lower F1 values of these two vowels than the English natives, so that [ɪ] was pronounced in the region similar to [i] and the vowel [æ] was produced in the area of [ɛ]. On the other hand, the F2 values and the duration ratio seemed to be similar between these two groups of talkers.

4. Discussion

The current experiment was designed to address the question of what is the acoustic characteristics of the sounds eliciting the asymmetry in the ISIB between the L1 matched sounds and non-matched sounds.

In Han et al. (2011), there was shown to be an asymmetry in the ISIB between the L2 vowels which are matched to certain L1 vowels and non-matched L2 vowels. The vowels [ɪ] and [æ] which do not match any vowel of the Korean language were shown to elicit more ISIB than the matched vowels, [i] and [ɛ]. English has two pairs of front vowels in the high region such as [i] and [ɪ], and in the non-high region such as [ɛ] than [æ], whereas Korean has only one vowel for either region partly due to on-going sound changes: loss of contrastive vowel length and the merger of a mid-high front vowel, [ɛ] and a mid-low front vowel, [æ]. Previous acoustic measurements suggest that the Korean [i] is closer to the English [i] than [ɪ], and that the Korean [ɛ] is closer to the English [ɛ] than [æ] (Yang, 1996; Flege et al., 1997), but [ɪ] and [æ] are frequently realized as allophones of [i] and [ɛ] respectively. The results of word identification showed that the ISIB is observed more frequently in the L1 non-matched vowels than in the matched vowels.

We anticipated that these results stem from the relative similarity of the interlanguage of the Korean talkers and listeners regarding these non-matched vowels. Indeed, acoustic analyses corroborated the intelligibility test results. For example, in the ISIB-L, the Korean listeners and the native English listeners showed differences in the cues that they might have used to make their word identification decisions when listening to Korean talkers. Looking first at the high vowel pair, vowel tokens identified by [ɪ] were produced with longer duration and in the region similar to [i] (less lowered) as compared to the native English listeners. A similar spectral pattern was observed for [æ]. As for the ISIB-T, the vowel tokens of [ɪ] and [æ] produced by Korean talkers with lower F1 values than English talkers were identified as correct, suggesting that Korean talkers realized both [i] and [ɪ] as [i]-quality vowels, and they realized the English [æ] as an [ɛ]-quality vowel. Thus the tokens produced by Korean talkers showed more deviated acoustic characteristics from those of the natives, with closer formant values to the matched vowels.

Combining the results of word identification and acoustic measurements, relatively more accurate (or at least similarly accurate) performance by Korean talkers/listeners than the native English talkers/listeners can be explained with the shared interlanguage of Korean talkers and listeners. The non-matched vowels were produced with more deviated acoustic characteristics from those of the natives and Korean talkers/listeners' identification was not hindered by these deviated tokens, but the English natives had difficulties in identifying the tokens as intended by the talkers.

What is interesting is that the shared interlanguage by Korean learners related to the ISIB was more clearly shown in the spectral characteristics than the temporal ones. The non-matched vowel [ɪ] was expected to be shorter than the corresponding matched vowel [i], based on the fact that Koreans are taught in school that the high front vowels [i] and [ɪ] are distinguished by duration, not spectral quality, and dictionaries mark the distinction in the same way (Flege et al., 1997; Ingram & Park, 1987). Bohn (1995) argued that L2 listeners are less sensitive to spectral differences in regions of the vowel space that only contain one vowel in their first language, and they are likely to use the duration cue to discriminate between two L2 vowels, because durational cues are psycho-acoustically highly salient. However, the spectral differences between Korean learners and the English natives were shown to be stronger than the temporal differences. Flege et al. (1997) showed similar results to those of the present study in that Korean participants produced spectral differences in addition to length differences for making these two vowels distinguishable, contrary to expectation.

Summarizing the results, the acoustic analyses corroborated the identification results (Han et al., 2011) in that Korean talkers showed more deviated acoustic characteristics from those of the natives, and Korean listeners and English natives were different in the cues that they might have used to make their word identification decisions when listening to Korean talkers. Thus these results support the hypothesis that the ISIB mainly stems from the shared interlanguage of non-native talkers and listeners.

References

Bent, T. & Bradlow, A. R. (2003). "The interlanguage speech intelligibility benefit", *Journal of the Acoustical Society of America*, 114, 1600-1610.

Bohn, 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: Theoretical and methodological issues* (279-304). Timonium, MD: York Press.

Bradlow, A. R. & Bent, T. (2002). "The clear speech effect for non-native listeners", *Journal of the Acoustical society of America*, 112, 272-284.

Bradlow, A. R. & Pisoni, D. B. (1999). "Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item- related factors", *Journal of the Acoustical society of America*, 106, 2074-2085.

Crystal, D. (2003). *The cambridge encyclopedia of the English language*. Cambridge: Cambridge University Press.

Derwing, T. M. & Munro, M. J. (2001). "What speaking rates do non-native listeners prefer?", *Applied Linguistics*, 22, 324-337.

Ettlinger, M. & Johnson, K. (2009). "Vowel discrimination by English, French and Turkish speakers: Evidence for an exemplar-based approach to speech perception", *Phonetica*, 66(4), 222-242.

Flege, J. E., Bohn, O.-S. & Jang, S. (1997). "Effects of experience on non-native speakers' production and perception of English vowels", *Journal of Phonetics*, 25, 437-470.

Han, J.-I., Choi, T.-H., Lim, I. & Lee, J. (2011). "The effects of language background, proficiency and context on the intelligibility of foreign-accented speech", ms. Konkuk University and University of Seoul.

Hayes-Harb, R., Smith, B., Bent, T. & Bradlow, A. (2008). "The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts", *Journal of Phonetics*, 36, 664-679.

Hong, Y.-S. (1991). "A sociolinguistic study of Seoul Korean", *Korean Unification Studies*, Series 12, Seoul: Research Centre for Peace and Unification of Korea.

Imai, S., Walley, A. C. & Flege, J. E. (2005). "Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners", *Journal of the Acoustical Society of America*, 117, 896-907.

Ingram, J. C. L. & Park, S.-G. (1997). "Cross-language vowel perception and production by Japanese and Korean learners of English", *Journal of Phonetics*, 25, 343-370.

Kim, H. & Han, J.-I. (1998). "Vowel length in modern Korean: An acoustic analysis", In B.-S. Park & J.-H. Yoon (Eds.) *Proceedings of the 11th international conference on Korean linguistics*, (412-418). Seoul: Hanshin.

Magen, H. & Blumstein, S. (1993). "Effects of speech rate on the vowel length distinction in Korean", *Journal of Phonetics*, 21, 387-410.

Mayo, L. H., Florentine, M. & Buus, S. (1997). "Age of second-language acquisition and perception of speech in noise", *Journal of Speech Language Hearing Research*, 40, 686-693.

McAllister, R., Flege, J. E. & Piske, T. (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.

Munro, M. J. & Derwing, T. M. (2005). "Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech", *Language and Speech*, 38, 289-306.

Nordstrom, P. E. & B. Lindblom (1975). "A normalization procedure for vowel formant data", Paper 212 at the international congress of phonetics sciences in Leeds, August.

Stibbard, R. M. & Lee, J.-I. (2006). "Evidence against the mismatched interlanguage intelligibility benefit hypothesis", *Journal of the Acoustical Society of America*, 120, 433-442.

van Wijngaarden, S. J. (2001). "Intelligibility of native and non-native Dutch speech", *Speech Communication*, 35, 103-113.

van Wijngaarden, S. J., Steeneken, H. J. M. & Houtgast, T. (2002). "Quantifying the intelligibility of speech in noise for non-native listeners", *Journal of the Acoustical Society of America*, 111(4), 1906-1916.

Yang, B. G. (1996). "A comparative study of American English and Korean vowels produced by male and female speakers", *Journal of Phonetics*, 24, 245-261.

• **Jeong-Im Han, corresponding author**
Konkuk University
Department of English
Hwayang-dong 1, Gwangjin-gu
Seoul, 143-701, Korea
Email: jhan@konkuk.ac.kr

• **Tae-Hwan Choi**
Konkuk University
Department of English
Hwayang-dong 1, Gwangjin-gu
Seoul, 143-701, Korea
Email: bak2san@hanmail.net

• **Injae Lim**
Konkuk University
Department of English Education
Hwayang-dong 1, Gwangjin-gu
Seoul, 143-701, Korea
Email: injaelim@konkuk.ac.kr

• **Joo-Kyeong Lee**
University of Seoul
Department of English Language and Literature
Jeonnong-dong 90, Dongdaemun-gu
Seoul, 130-743, Korea
Email: jookyeong@uos.ac.kr