Acquisition of English Vowels by Korean ESL Speakers: A Phonetic Approach

Soo-Jung Kim
(Tongmyong University)

Kim, Soo-Jung. 2006. Acquisition of English Vowels by Korean ESL speakers: A Phonetic Approach. *The Linguistic Association of Korea Journal, 14*(2), 1-16. English has a phonemic distinction of tense and lax vowels that Korean does not have in its sound inventory. Discriminating tense and lax vowels in English is crucial in some cases to catch the right meaning of sentences. In this respect, this study is aimed to diagnose how accurately the tense and lax distinction is acquired by Korean ESL speakers. Specifically, the perception and production ability of Korean English learners are investigated. First, according to perception tests on three groups, high school students(early English learners) showed little difference from college students(late learners): 58 vs. 60%. The teachers group displayed a slightly higher identification rate of 71%. Second, the acoustic measurements of vowel duration and formant frequencies on the speech of Korean ESL and native English speakers suggested each group used different strategies in producing the vowels. The native speakers use both cues of vowel durations and formant frequencies. In contrast, the Korean ESL speakers showed a difference in vowel length, but no statistical difference in formant frequencies, indicating only the length feature is acquired and used in distinguishing the tense and lax vowels in English.

Key Words: English vowels, acoustics of tense/lax vowels, identification rate, vowel length, formant frequencies, Korean ESL learners

1. Introduction

A phonemic distinction of tense and lax vowels that English has (i.e. /i/ vs. /ɪ/, /u/ vs. /ʊ/) does not exist in Korean. Discriminating tense and lax vowels in English is crucial in catching the meaning of the following minimal pair sentences.

(1) a. I'm *heating* the pan.
b. I'm *hitting* the pan
(2) a. He said they *cooed.*
b. He said they *could.*

The two sentences in (1) and (2), respectively, are uttered exactly the same except for one vowel. (1a) has high front tense vowel [i] in *heating,* while (1b) has high front lax vowel [ɪ] in *hitting.* These vowels are different only in tenseness. In (2), *cooed* is pronounced with high back tense vowel [u] and *could* is with high back lax vowel [ʊ]. In many cases, the meaning of a word can be inferred through its context, but in the sentences above its meaning depends entirely on a listener's discriminability of tense and lax vowels. If a listener does not perceive each vowel correctly, he or she fails to grasp the right meaning of each sentence.

1.1. Previous research

A number of studies devoted to the perception of non-native sounds suggest that the L1 background has a strong influence in the way the sounds of the target language are perceived (Flege & Munro 1994; Ingram & Park 1997; Lado 1957). Two influential models of cross-language vowel perception (Best, 1995; Flege, 1995) are driven by considering how the foreign vowels are assimilated to native phoneme categories. According to Flege's Speech Learning Model (1987, 1995), certain non-native sounds which are phonetically close to L1 targets are more readily accommodated than others. On the other hand, Best (1995) proposed that discrimination of two non-native sounds is maximized where each is assimilated to a different native phoneme category. Where both sounds are assimilated to a single phonemic category, but where one sound constitutes a closer phonetic match to that category than the other, moderate discriminability is predicted. However, where two foreign sounds are equally good candidates for a single category, discrimination will be poor.

Standard Korean has 10 vowels, /i, e, ɛ a, o, ø, u, y, ʌ, ɨ/ without a tense-lax distinction. Kim (1972) asked Koreans with little or no knowledge of English to write the vowels and consonants in aurally presented English words using the Korean writing system. English /i/ was usually transcribed using the symbol for Korean /i/, while English /ɪ/ was identified most often with Korean /ɨ/. Similarly, based on his acoustic measurements, Yang (1996) suggested that the Korean /i/ is closer to English /i/ than /ɪ/. In a study examining the perception of English vowels by Korean adults who lived in an English-speaking country for less than a year, Tsukada et al. (2005) reported that English /i/ was always identified as Korean /i/ while English /ɪ/ was identified as Korean /ɨ/(32%) and also Korean /e/ (30%), /ɛ/(18%) and /i/(12%). Colhoun & Kim (1976) claimed that English /ɪ/ is a totally new vowel from the standpoint of Korean.

Finally, studies show that non-native speakers' accuracy in producing English vowels is related inversely to the age of the first extensive exposure to native-produced English. The experienced non-native speakers produced and perceived English vowels more accurately than did the relatively inexperienced non-native speakers (Flege, Munro & MacKay, 1995; Flege, Bohn & Jang, 1997; Tsukada, et al., 2005).

1.2. The present study

Based on the notion that early learners are believed to show more native-like performance than late learners who were first exposed to English in late adolescence or adulthood, English is officially taught from the third grade of elementary school in Korea. Early learners are expected to acquire English sounds more accurately than late learners.

This study focuses on perception and production of English tense and lax vowels by Korean ESL learners. Specifically, we examine the following questions:

1) How well can Korean ESL learners of high school and college students identify the distinction of English tense and lax vowels?

2) Did early learning help learners acquire the distinction? That is, are high school students performing better than college students?
3) What is the discriminability level for ESL teachers?
4) Finally, what are the acoustic properties of tense vs lax vowels in English produced by ESL learners?

The study is organized as follows. Section 2 reports an experiment in three groups of ESL speakers: high school, college students and English teachers. They made forced choice identifications of four English vowels /i, ɪ, u, ʊ/. Section 3 deals with a further experiment investigating acoustic properties of English tense and lax vowels produced by ESL learners. Vowel duration and formant frequencies were measured and compared.

2. Experiment One: Perception Test

2.1. Participants and Procedures

To evaluate the success of ESL speakers in perceiving these vowels, an intelligibility test was performed. Three different groups of subjects participated in the experiment: the first two groups consist of 74 college students and 37 high school students, and the third group is composed of 21 teachers/lecturers teaching English at the high school or college level. A native English speaker's production of 4 high vowels (front vowels-/i, ɪ/, back vowels-/u, ʊ/) both in isolation and within sentences was identified by the subjects. They were asked to circle the word they thought they heard and each token was heard once. The 24 tested words and sentences are demonstrated in APPENDIX 1.

2.2. Results & Discussion

The summary of mean correct identification scores, minimal and maximal identification scores out of 24 tokens and standard deviations is given in Table 1.

Table 1. Mean of correct identification scores in three listener groups

subject	Min.	Max.	mean (percentage)	s.d.
college students	6	22	14.4 (60%)	3.2
high school students	6	21	13.9 (58%)	3.1
teachers	13	21	17.0 (71%)	2.3

(min.=minimal score, max.=maximal score, s.d.=standard deviation)

Out of 24 tested words and sentences, college and high school students identified 14.4(60%) and 13.9(58%) tokens correctly, showing no significant difference in discriminating tense and lax vowels. This means that early learners who were exposed to English as early as ten or less are not outperforming late learners. The teachers group shows a slightly higher identification score, 17.0(71%). We observe a considerable variation among subjects in each group in the identifications score. The best listener had a correct identification score of 22 or 21, while the worst had a much lower score of 6 in the two student groups. 13 was the lowest score in the teachers group.

Now, let us examine whether there are any differences in the identification rate depending on vowels and their positions of utterance. Table 2 shows the correct percentage of identifications of the vowels when produced in isolation and within sentences.

Table 2. Identification rate by different groups

subject group	front vs. back	position	tense vs. lax	Identification rate (%)
college students	front	in isolation	tense	90
			lax	84
		within a sentence	tense	72
			lax	69
	back	in isolation	tense	57
			lax	55
		within a sentence	tense	47
			lax	48
high school students	front	in isolation	tense	66
			lax	58
		within a sentence	tense	66
			lax	58
	back	in isolation	tense	66
			lax	56
		within a sentence	tense	64
			lax	55
teachers	front	in isolation	tense	98
			lax	90
		within a sentence	tense	89
			lax	70
	back	in isolation	tense	81
			lax	55
		within a sentence	tense	53
			lax	49

To explore the between-vowel difference, a one-way repeated measures Analysis of Variance (ANOVA) with four vowels /i, ɪ, u, ʊ/ as a within-group factor was performed. A significant effect of vowel [$F(3, 11)=5.718$, $p=.13<.05$] was obtained, indicating that some vowels are better identified than others. A post hoc Tukey HSD test revealed that [i] was the most intelligible vowel and [ʊ] was the least well identified. No significant differences were found between [ɪ] and [u] in identification rates. The effects of speaker and front-back were significant [Speaker: $F(2,11)=6.721$, $p=.012$; Front-back: $F(1,11)=36.424$, $p=.000$]. The effect of position was not significant [$F(1,22)=3.476$, $p=.176$], indicating that whether tense and lax vowels are uttered in citation

forms or within sentences, their identification does not vary.

3. Experiment Two: Production

Acoustically, English vowels differ in vowel height and length. We would expect, therefore, that native English speakers rely on both cues when distinguishing between these vowels. Although the relationship between speech perception and production is complex and controversial, we assume that measurement of the vowels may provide some indication of which features are used in perception and production of tense and lax pairs. Specifically, we hope that acoustic analysis of their productions would show the speakers' strategy in discriminating a tense and lax distinction.

3.1. Participants and Procedures

Acoustic properties of English tense and lax vowels produced by ESL learners are examined. As a comparison group, speech of native English speakers was collected as well.

In this experiment, the acoustic properties of English vowels were measured and compared in terms of duration, and F1 and F2 frequencies. Since no difference was found in perceptibilty between college students who started learning English at their middle school years and high school students who started learning English at an earlier stage, speech data was collected only from college students. Three male and three female students participated in the experiment. As a comparison group, native English speakers' speech was collected as well. Tested words containing tense and lax vowels appear both in isolation and within sentences. A total of 288 tokens (16 words/sentences * 3 repetitions * 6 subjects) was recorded using a Sony ICD-MS525 digital recorder. Recorded data was edited using Sony Digital Voice Editor and analysed using Praat.

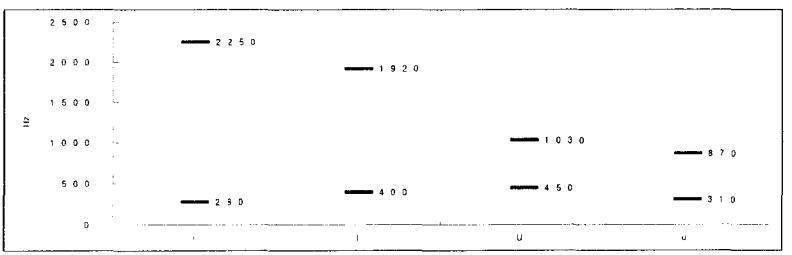
Vowel duration was measured from spectrogram and waveform displays. The frequencies of F1 and F2 were taken at the mid-point of

each vowel from the spectrogram. In the following sections, measurement results of formant frequencies and vowel duration are discussed.

3.2. Formant Frequencies

Articulatorily vowels are defined in terms of tongue height, front-back, and tenseness, and these attributes are reflected on formant structures acoustically. In the production of vowels, the filtering effect of the vocal tract produced amplitude peaks at certain frequencies by enhancing the harmonics while damping harmonics of other frequencies. These peaks in the filter function are called formants, among which the first, second and third formants are crucial in identifying vowels. In particular, the first formant (F1) is related to vowel height and the second (F2) to front/back. F1 values are increasing as the tongue moves from high to low; F2 values are decreasing as the tongue moves front to back. Figure 1 demonstrates the vowel plot for [i, ɪ, u, ʊ] in terms of formant frequencies given in Cipollone, Keiser & Vasishth (1998).

Figure 1. Formant plot for the vowels [i, ɪ, u, ʊ]



In Figure 1, we observe that the F1 values of lax vowels [ɪ] and [ʊ] are greater than their corresponding tense vowels [i] and [u], respectively. On the other hand, F2 are lower on back vowels than on front vowels.

A series of two-tailed *t*-tests were carried out on F1&F2 values of tense and lax vowels on native and ESL speakers to investigate the

spectral differences in the vowels. The results of the *t*-tests along with the means and standard deviations (in parentheses) for male speakers are summarized in Table 3.

Table 3. Mean F1 & F2 produced by male native & ESL speakers

	male	[i]	[ɪ]	t	p	[u]	[ʊ]	t	p
F1	Native	285.85 (43.87)	437.75 (36.29)	8.53	.000*	391.14 (27.08)	493.20 (38.42)	7.09	.006*
	ESL	402.19 (52.03)	389.29 (45.25)	3.34	.001	443.06 (49.99)	447.35 (54.23)	1.00	.319
F2	Native	2271.34 (71.84)	1887.40 (85.78)	14.08	.000*	913.63 (87.63)	1094.17 (98.06)	3.78	.007*
	ESL	2011.14 (240.79)	2041.12 (239.61)	1.13	.261	1101.76 (186.99)	1115.47 (190.34)	1.18	.241

* indicates the differences in frequencies were at $p<.05$.

As demonstrated in Table 2, significant differences were observed in both F1 and F2 values of tense and lax vowels for native English speakers; no statistically significant spectral differences were found for the ESL speakers. The difference in F1 between [i] and [ɪ] produced by the ESL speakers appeared to be significant, its p-value being .001. When examined closely, the F1 value of [i] is greater than that of [ɪ], while the results were opposite in the native speakers. That is, the F1 value of [ɪ] is greater than that of [i] in the native speakers. To reiterate,, F1 is related to vowel height and its values increase as the tongue from high to low. This is demonstrated by the formant plot of the vowels in Figure 1 and confirmed by the outputs of the native speakers above. Thus, the formant frequency results generated by the ESL speakers demonstrate that they did not produce the vowels accurately. In other words, they did not acquire the tense and lax distinction caused by the differences in tongue height.

Next, Table 4 shows the statical results for female speakers.

Table 4. Mean F1 & F2 produced by female native & ESL speakers

	female	[i]	[ɪ]	t	p	[u]	[ʊ]	t	p
F1	Native	317.54 (31.70)	524.83 (48.20)	18.99	.000*	364.38 (56.70)	482.697 (49.54)	9.31	.000*
	ESL	496.8 (44.4)	482.4 (40.4)	2.94	.006	515.8 (49.9)	483.9 (33.2)	5.11	.000
F2	Native	2398.84 (143.72)	2091.64 (106.03)	6.54	.000*	967.56 (79.92)	1626.85 (128.97)	21.06	.000*
	ESL	2456.6 (205.6)	2409.4 (224.5)	1.87	.070	1394.0 (188.9)	1372.2 (195.5)	1.03	.313

* indicates the differences in formant frequencies were at $p<.05$.

Again, we observe significant differences in both F1 and F2 values of tense and lax vowels for the native English speakers, but neither F1 nor F2 shows spectral differences for the ESL speakers. Just like the results of the ESL speakers in Table 3, their F1 values for and [i] vs. [ɪ], and [u] vs. [ʊ] were opposite to the native speakers'. In the ESL speakers, the F1 value of [i] is smaller than that of [ɪ], whereas in the native speakers the F1 value of [i] is greater than that of [ɪ]. Similarly, the F1 value of [u] is smaller than that of [ʊ] in the ESL speakers, while the F1 value of [u] is greater than that of [ʊ]. This indicates that the ESL speakers did not acquire the vowel height feature properly in distinguishing English tense and lax vowels.

3.3. Vowel duration

The mean vowel durations in citation forms and within sentences for native English speakers are graphically shown in Figure 1 & 2.

Figure 1. Vowel Duration in isolation-Native English speakers

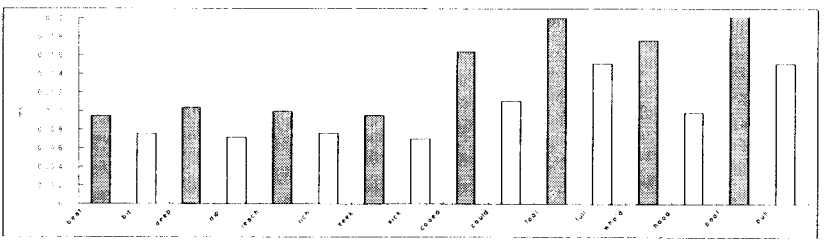
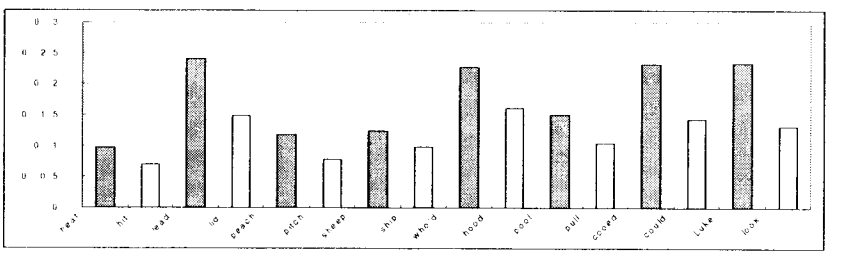


Figure 2. Vowel Duration within sentences-Native English speakers



Both in citation forms and within sentences the vowel length of tense vowels (represented with black bars) is longer than that of lax vowels (represented with white bars) in the two groups. For vowel duration in isolation, a repeated measures two-way Analysis of Variance (ANOVA), with the factors of vowel type and speaker, yielded highly significant effects for vowel and speaker but not for interaction of vowel and speaker [Vowel: $F=20.383>F(3,36;.05)=2.92$, $p=.000<\alpha=.05$; Speaker: $F=56.855>F(2,36;.05)=3.32$, $p=.000<\alpha=.05$; Vowel×Speaker: $F=1.266<F(6,36;.05)=2.33$, $p=.297>\alpha=.05$]. For vowel duration within sentences, the effects of vowel [$F(3,36)=7.92$, $p=.000$] and speaker [$F(2,36)=4.535$, $p=.018$] were significant, but the interaction between vowel and speaker [$F(6,36)=.147$, $p=.989$] was not significant.

Now, let us examine the vowel duration in isolation and within sentences for Korean ESL learners. Figures 3 & 4 demonstrate the vowel duration of tense and lax vowels.

Figure 3. Vowel Duration in isolation-ESL learners

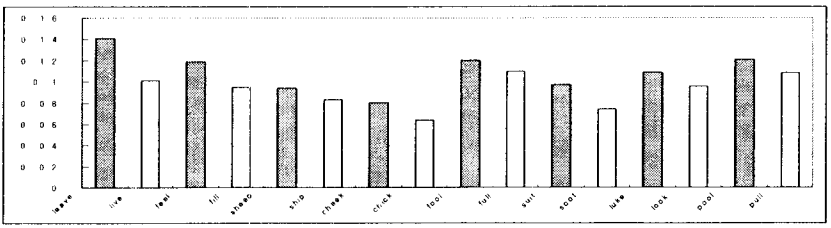
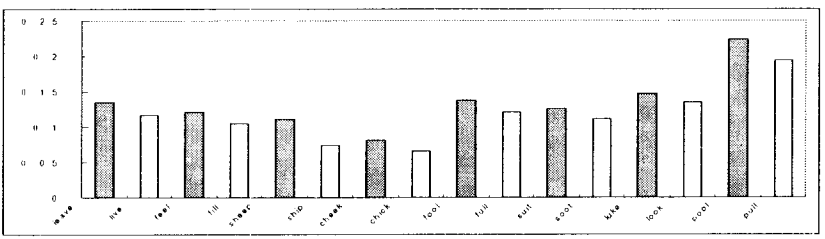


Figure 4. Vowel Duration within sentences-ESL learners



For vowel duration in isolation, a two-way ANOVA test revealed effects of vowel [F(3,36)=3.512, p=.025] and speaker [F(2,36)=33.419, p=.000] were significant, but the interaction between vowel and speaker [F(6,36)=1.017, p=.430] was not. For vowel duration within sentences a repeated two-way measurement ANOVA produced a significant main effect [Vowel: F(3,36)=7.920, p=.000]; Speaker: F(2,36)=.4.535, p=.018], but not a significant interaction effect [F(6,36)=.147, p=.989].

The durational ratios from the two groups on the 2 pairs of tense and lax vowels [i]-[ɪ] and [u]-[ʊ] were calculated and compared to see if the native and ESL speakers applied a different strategy in their productions of these vowel pairs. The tense and lax vowel duration ratios for the native vowel pairs were 1.38 for [i]-[ɪ] and 1.47 for [u]-[ʊ] respectively; for ESL speakers 1.26 for [i]-[ɪ] and 1.16 for [u]-[ʊ]. This indicates that though ESL speakers adopt temporal difference to distinguish tense and lax vowels, their duration difference is not as great as the native speakers'. This might be compatible with the trend in Korean that many young speakers are losing traditional phonemic

length distinctions (e.g., [kil] 'road' vs. [ki:l] 'long') (Magen & Blumstein, 1993)

In conclusion, the significant temporal difference in the production of tense and lax vowels by Korean ESL speakers indicates that they utilize duration cues to identify these vowels.

Finally, it should be mentioned that use of the duration cues for Korean ESL speakers to distinguish the foreign vowel contrast cannot be due to a perceptual strategy transferred from the first language based on the fact that phonemic length distinctions are almost lost in young Korean speakers.

4. Conclusion

Discriminating English tense and lax vowels is crucial in catching the meaning of sentences correctly. in particular, in the sentences differing in the two vowels only. However, it has not been easy to acquire the distinction for Koreans whose phonemic inventory does not have a tense and lax distinction. In this respect, we examined discriminability and acoustic properties of English tense and lax vowels in Korean ESL learners. Two experiments of perception and production were carried out.

The perception test was performed with 24 minimal pair words and sentences with three groups of ESL learners: college students, high school students who started learning English at earlier stages and English teachers. Results indicated that the tense and lax distinction is moderately accommodated by these ESL learners and that early learning did not seem to affect in learning the distinction.

Acoustic measurements of English tense and lax vowels produced by ESL learners showed that they do not use spectral or formant frequency differences, but duration cues to identify these vowels, in comparison with native English speakers to whom both height and length differences are found to be significant.

References

Best, C. T. (1995). A direct realist view of cross-language speech perception. *In Speech perception and linguistic experience: issues in cross-language research*, (W. Strange, editor) 171-203. Timonium, MD: York Press.

Colhoun, E. & Kim, T. W. (1976). English loanwords in Korean: Lexical and phonological problems, *Applied Linguistics 8*, 237-250.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign languages. In *Human communication and its disorders*, A Review 1988(H. Winitz, editor), 224-240. Norwood, NJ.

Flege, J. E. & Bohn. O. S. (1989). The perception of English vowels by native speakers of Spanish. *Journal of the Acoustic Society of America 85*, Suppl 1, S85.

Flege, J. E. & Munro, M. J. (1994). Auditory and categorical effects on cross-language vowel perception, *Journal of the Acoustic Society of America 95(6)*, 3623-3641.

Flege, J. E. (1995). Second language speech learning theory, findings, and problems. *In Speech perception and linguistic experience: issues in cross-language research*, (W. Strange, editor) 233-277. Timonium, MD: York Press.

Flege. J. E., Munro, M. J., & MacKay, I. R. (1995). Effects of age of second language learning on the production of English consonants *Speech Communication 16*, 1-26.

Flege, J., Bohn, O. & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics 25*, 437-460.

Ingram, J. 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, D. J. (1972). A contrastive study of English and Korean phonology on the basis of aural perception, *Language Teaching 5*, 1-36.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.

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

Tsukada, K., et al. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics 33*, 263-290.

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.

APPENDIX 1

1. seek sick
2. deep dip
3. beat bit
4. reach rich
5. I am (heating, hitting) the pan.
6. I'm taking the (lead, lid).
7. I am (heating, hitting) the pan.
8. Who found the (sheep, ship)?
9. It was a good (peach, pitch).
10. I'm taking the (lead, lid).
11. It was a good (peach, pitch).
12. Who found the (sheep, ship)?
13. fool full
14. who'd hood
15. cooed could
16. pool pull
17. Did you say "(who'd, hood)"?
18. We all cried, "(Luke, Look)!"
19. Don't (pool, pull) them right now?
20. He said they (cooed, could).
21. Did you say "(who'd, hood)"?

22. Don't (pool, pull) them right now?
23. We all cried, "(Luke, Look)!"
24. He said they (cooed, could).

Soo-Jung Kim
Tongmyong University
Yongdang-dong 535, Nam-gu
Busan, 608-711, Korea
Phone: 82-51-610-8761
Email: sjkimok@tit.ac.kr

Received: 30 Mar, 2006
Revised: 13 Jun, 2006
Accepted: 19 Jun, 2006