http://dx.doi.org/10.17959/sppm.2015.21.2.323

The perception of foreign accent in prosody-transplanted speech*

Joo-Kyeong Lee**
(University of Seoul)

Xing Liu
(Zhengzhou University of Light Industry)

Lee, Joo-Kyeong and Xing Liu. 2015. The perception of foreign accent in prosody-transplanted speech. *Studies in Phonetics, Phonology and Morphology* 21.2. 323-340.This study investigates the contribution of prosody to Korean native listeners' perception of foreign accent in Chinese-accented Korean speech. Unlike many previous studies, the current work examined synthesized speech where L2 talkers' production of prosody was transplanted onto L1 talkers' segments, and vice versa. This experimental technique allowed for observing a separate and individual role of prosody while both segmental and prosodic information was preserved as in natural utterances. This work first investigates the relative weight of segments and prosody and then narrows down the scope to prosody alone. In the experiment, the stimuli for foreign accent perception was prepared in such a way that prosodic information was cross-planted onto segmental representations between L1 and L2 Korean speech. A group of fifteen Korean native listeners rated foreign accent, and the results show that the stimuli with combined L2 prosody and L1 segments were rated as less accented than those of L1 prosody and L2 segments. This suggests that segments played a more influential role than prosody in the perception of foreign accent. In the results of individual/independent roles of prosodic parameters such as duration, pitch and intensity, duration turned out to make a more prominent contribution to Korean listeners' detection of foreign accent. That is, Korean listeners are more likely to rely on the temporal aspects of speech because the consonants/ vowels were reset in the temporal dimension once duration was transplanted. This is an extended finding of Lee and Liu (2012), Liu and Lee (2012), and Lee (2014) that the matter of which prosodic parameter contributes more to foreign accent is substantially L1-specific, reflecting the prosodic structure of the L1. **(University of Seoul and Zhenzhou University of Light Industry**)

Keywords: prosody, segments, foreign accent, TD-PSOLA, Mandarin Chinese, Korean

1. Introduction

Foreign accent in L2 speech often gives rise to unsuccessful communications with native speakers due to its phonological and phonetic deviances from native norms. Such deviances are presumably attributed to segmental errors in numerous studies; some L2 segments, though deviant to some extent, may be perceived as allophonic variations within a particular category or as actual substitutions of different phonetic segments. A dominant number of previous studies on L2 speech have focused on the segmental aspects of foreign

* This work was supported by the 2014 sabbatical year research grant of the University of Seoul.
** Corresponding author

accent, assuming that the distortion of segmental information is a primary trigger of foreign accent (Flege and Hillenbrand 1984, Flege 1987, Mack 1989, Flege 1991, Kuhl 1991, Morley 1991, Avery and Ehrlich 1992, Flege 1993, Flege et al. 1995, Flege et al. 1997, Mildner and Horga 1999). Foreign accented speech was also extensively investigated in association with intelligibility and/or comprehensibility; as the degree of foreign accent increased, L2 speech became less intelligible and/or understandable to native listeners (Munro and Derwing 1995, Derwing and Munro 1997). On the other hand, it was contented that even heavily accented speech was sometimes perfectly intelligible and/or comprehensible even though it might take more time to process than native speech (Munro and Derwing 1999). Whether L2 studies on foreign accent was affiliated with intelligibility/ comprehensibility or not, they were primarily concerned with segment as opposed to prosody. They have, however, recently paid attention to nonsegmental factors like stress, intonation, rhythm, etc. as they appreciate that prosodic deviances may contribute as much or more to native listeners' detection of foreign accent (Anderson-Hsieh et al. 1992, Flege et al. 1995, Munro 1995, Jilka 2000, Boula de Mareuiil and Vieru-Dimulescu 2006, Trofimovich and Baker 2007, Lee and Liu 2012).

A variety of techniques have been used to measure the contributions of segmental and prosodic errors to foreign accent in the empirical studies of L2, and they can be, in large, confined to three representatives; impressionistic judgment, low-pass filtering, and prosody transplanting onto a string of segments between L1 and L2. Due to technical difficulty with sorting out prosodic information alone, the L2 stimuli which many studies traditionally served was kept natural and intact. Listeners were simply asked to impressionistically make independent judgments of such prosodic features as stress and intonation with no consideration of segmental errors. (Dirven and Oakshott-Taylor 1984, Pennington and Richards 1986, Morley 1991). However, it seems to be inappropriate to ask inexperienced listeners to judge unmodified speech samples on prosodic features because speech should be almost impossible to break down into segments and prosody. In an effort to resolve such a problem, experienced or trained ESL teachers participated in judging the effect of prosody on foreign accent in some studies (van Els and de Bot 1987, Anderson-Hsieh et al. 1992), but Munro pointed out that it was not known whether inexperienced listeners and teachers were influenced in the same way by prosodic features (Munro 1995). His critical concern was that accented speech interfered most with L2 learners' communicative abilities outside of the classroom. For this reason, ratings from inexperienced listeners should be necessary to figure out the factors that cause L2 speech to be accented.

The following group working on L2 prosody attempted to overcome the difficulties discussed above and used a low-pass filtering technique (van Els and de Bot 1987, Flege et al. 1995, Munro 1995, Jilka 2000, Trofimovich and Baker 2007). When speech was filtered in this manner, high frequency

components were filtered out and segmental information was removed. Such speech became unintelligible; therefore, the prosodic information like fundamental frequency was perceptually measurable. The problem concerning naïve vs. experienced listeners seemed to be resolved to some extent, but the low-pass filtered speech still had a drawback that the synthesized speech stimuli heard like humming sound and were somewhat unnatural (Lee and Liu 2012, Liu and Lee 2012). The L2 speech which native listeners hear in their daily environments contain both segmental and prosodic information, which may incur a naturalness problem of speech. Moreover, this technique failed to compare the separate and distinctive roles between segments and prosody, but merely examined if the prosodic feature of intonation played a role or not from the speech without segmental information.

A most recent technique adopted in Boula de Mareuiil and Vieru-Dimulescu (2006), Lee and Liu (2012), Lee (2014), and Liu and Lee (2012) was TD-PSOLA (time domain pitch synchronous overlap and add) algorithm. This method manipulates speech in such a way that prosodic parameters such as pitch, duration and intensity in one talker's speech are transplanted onto segments of another talker's speech when they are identical in segments. A comparison of accentedness ratings of the speech of L1 segments and L2 prosody with the speech of L2 segments and L1 prosody may, for example, indicate which, prosody or segments, plays a more salient role in the judgment of foreign accent. Though the speech stimuli are synthesized, they may not have the naturalness problem that the low-pass filtering technique has because the stimuli do not lose any segmental or prosodic information. While speech is still manipulated to elicit the sole effect of prosodic parameters, rating results will be trustworthy in terms that naïve listeners will be able to assess separate and/or independent features of prosody in the synthesized speech of no loss of segmental or prosodic information.

In the study of Boula de Mareil and Vieru-Dimulescu (2006), recordings were made from Spanish and Italian talkers. Due to the fact that both languages were similar in phonological/phonetic structures, the sentences were almost identical in segmental and prosodic representations. The speech stimuli for the foreign accent ratings were prepared with using the TD-PSOLA algorithms; Spanish talkers' prosody such as pitch and duration were overlaid onto Italian talkers' segments and vice versa. Another group of Spanish and Italian listeners identified the synthesized stimuli as one of four options: (1) native Spanish, (2) Italian accented Spanish, (3) Spanish accented Italian, and (4) native Italian. Results showed that both Spanish and Italian listeners identified the speech combined with Italian prosody as Italian and the speech synthesized with Spanish prosody as Spanish. This suggested that the listeners relied more on prosodic information than segments in the accentedness judgment. However, it should be noticed that the segmental structures of Spanish and Italian are extremely similar and that segmental deviances, if any, might be relatively little influential. This might

mislead the conclusion that the role of prosody was more prominent in detecting foreign accent. In order to precisely measure the relative weight of segments and prosody, two languages under observation should be sufficiently distinctive both in segment and prosody.

Another research group of L2 speech benefiting from the TD-PSOLA synthesis resolved the problem of predetermined similarity between L1 and L2 (Lee and Liu 2012, Liu and Lee 2012, Lee 2014). They adopted L1 and L2 where both segmental and prosodic systems were sufficiently distinctive, examining a variety of L2 languages: Korean accented Chinese L2 speech (Lee and Liu 2012), Chinese accented English L2 speech (Liu and Lee 2012), and Korean accented English L2 speech (Lee 2014). Overall, segments played a dominant role over prosody while prosody still made a meaningful contribution to the perceived accent. When sub-components of prosody were scrutinized to determine which would be more influential, the phonological structure of native listeners' L1, especially its prosodic characteristics, was reflected straightforwardly in the rating patterns. The relative weight among prosodic features depended on the listeners' L1 background such that pitch errors were more likely to be perceived as foreign accent to Chinese native listeners (Liu and Lee 2012) while duration deviances were more strongly bonded with foreign accent in English native listeners' perception of foreign accent regardless of nonnative talkers' L1 (Lee and Liu 2012, Lee 2014).

The previous studies of Lee and Liu (2012), Lee (2014) and Liu and Lee (2012) shed light on extending L2 to Korean. The current work reports on a foreign accent experiment where Chinese accented Korean L2 speech was synthesized with Korean native L1 speech, taking advantage of the TD-PSOLA algorithm. Korean prosodic parameters produced by Chinese talkers (hereafter, L2 prosody) were transplanted onto Korean segments produced by Korean talkers (hereafter, L1 segments). Korean prosodic parameters produced by Korean talkers (hereafter, L1 prosody) were also transplanted on Korean segments produced by Chinese talkers (hereafter, L2 segments). L1 and L2 speech recordings were manipulated with attention to the multidimensional nature of prosody, assuming that pitch, duration and intensity would be all potential attributes to foreign accent. Furthermore, a variety of acoustic differences between the native and nonnative synthesized stimuli were examined in an effort to solicit comparatively potent prosodic indicators for accentedness.

2. Experiment

2.1 Talkers and Listeners

We assumed that Chinese accented Korea speech was phonetically somewhat or to a great degree deviant from L1 speech, definitely depending on L2 talker's proficiency. It would be, therefore, plausible that the accentedness ratings of the synthesized stimuli would indicate the extent to

which prosody played a role in the perception of Chinese accented Korean speech. L2 talkers' proficiency was decided in application of two different standards: length of residence (LOR) and Accentedness Task (AT). Chinese talkers were first recruited on the basis of their LOR in Korea, and their proficiency levels were confirmed by AT. We recruited Chinese college students attending a university in Seoul, Korea; 8 of them had lived in Korea longer than three years, 8 of them had lived in Korea between 1 and 1.5 years, and 8 of them barely came to Korea with the LOR of less than six months. They all started to learn Korean after they arrived in Korea. Each group was composed of four males and four females because we needed two female and two male Chinese subjects whose recordings were cross-planted with two native Korean female and male talkers. We double-numbered the Chinese talkers for each group in order to guarantee at least a half of them to participate.

The Chinese talkers were then asked to visit a speech lab at the University of Seoul and read a simple Korean passage in a sound-attenuated booth (see Appendix A). The passage was intentionally made simple and easy for the sake of the shortest LOR talkers' understanding. Their recordings were presented to three Korean native judges who did not participate as a talker. The Korean judges were asked to rate the degree of foreign accent in a 9-point Liker scale (1 = very strong foreign accent and 9 = no foreign accent). We selected two female and two male Chinese participants whose rating scores above 7 (7~9) as high proficiency and two female and two male Chinese talkers who received the scores below 3 (1~3) as low proficiency. We abandoned the remaining talkers whose scores between 4~6 to maximize the proficiency differences.

The eight Chinese talkers read five Korean sentences (see Appendix B). Two female and two male native Korean talkers were recruited and asked to record the same sentences in preparation for the speech stimuli of the perception experiment of foreign accent. Fifteen Korean natives served as listeners for judging foreign accent in it. They were all college students and did not participate in the recording process. They were all phonetically untrained listeners and had no experience to learn Chinese.

2.2 Procedure

The sentences for recording were intentionally made of 8 to 12 syllables; bugs might more possibly occur in synthesized stimuli as the sentences became longer (Lee and Liu 2012). Eight nonnative Chinese and four native Korean talkers read five Korean sentences. That is, 60 sentences were recorded in total (5 sentences * 8 non-native talkers + 5 sentences * 4 native talkers).

We carried out the recordings in a sound-attenuated booth at a sampling rate of 44.1 kHz. Prior to the transplanting synthesis, we labeled the recorded sentences in segments; the sentences whose prosody features were swapped

should be segmentized exactly the same in number and order of segments[1]. We applied the Praat script composed in the framework of the TD-PSOLA algorithm (Yoon 2006) and generated gender-matched stimuli of four different combination types: (1) non-native Chinese L2 prosody with native Korean L1 segments, (2) L1 prosody with L2 segments, (3) L2 prosody with L2 segments, and (4) L1 prosody with L1 segments. In addition to transplanting prosody within inter-language talkers (types 1 and 2), the prosody swapping was carried out in the speech between the same L1 backgrounds (intra-language talkers) in order to avoid any inconsistency between natural and synthesized stimuli (types 3 and 4).

The prosodic parameters of duration (D), pitch (F) and intensity (I) were examined. Once the script was run, an individual (D, F or I), or a combination of two or three of the parameters (DF, DI, FI, or DFI) were copied onto a string of segments in the other speech. That is, Chinese nonnative (L2) talkers' prosody was copied onto Korean native (L1) talkers' segments and vice versa. One sentence generated 14 synthesized sentences for different prosodic parameters: (1) L1 talker's duration (D) + L2 talker's segments, (2) L1 talker's pitch (F) + L2 talker's segments, (3) L1 talker's intensity (I) + L2 talker's segments, (4) L1 talker's duration & pitch (DF) + L2 talker's segments, (5) L1 talker's duration & intensity (DI) + L2 talker's segments, (6) L1 talker's pitch & intensity (FI) + L2 talker's segments, and (7) L1 talker's duration, pitch & intensity (DFI) + L2 talker's segments, and the other 7 manipulated stimuli switching between L1 and L2.

770 synthesized stimuli were generated in total. Types (1) and (2) generated 560 synthesized outputs (8 pairs of native and nonnative talkers * 5 sentences * 7 prosodic parameters * 2), Type (3) created 140 stimuli from the last method (4 pairs of non-native Chinese talkers * 5 sentences * 7 prosodic parameters) and type (4) synthesized 70 speech samples (2 pairs of Korean native talkers * 5 sentences * 7 prosodic parameters).

We assumed that listeners were not able to listen to 770 speech samples due to fatigue. We first randomized the generated stimuli and divided them into three blocks (two blocks contained 257 speech stimuli each and one contained 256 stimuli). Fifteen Korean native listeners were grouped into three; therefore, each group listened to one block of the stimuli respectively. They were asked to rate foreign accent for the speech stimuli that they had heard and press one button from 1 to 9 on the keyboard (1 = very strong foreign accent and 9 = no foreign accent) in the SuperLab software. The listeners used a headphone to listen to each stimulus only once. Native listeners' ratings were submitted to the t-test and the ANOVA analysis for statistics.

[1] See Lee and Liu (2012) and Liu and Lee (2012) for more details about the synthesis process.

2.3 Results and Discussion

Figure 1 shows the average rating scores for the stimuli manipulated between Chinese L2 high proficient talkers and Korean L1 talkers. The black bars indicate the scores for the stimuli entailing L2 prosody and L1 segments, and the gray bars represent the scores for the stimuli combining L1 prosody and L2 segments. Whether prosodic parameters are transplanted individually or in combination, the foreign accent scores are consistently higher when L2 talkers' prosody was combined with L1 talkers' segments than the other way (D: $t=12.838$, $p<0.05$; F: $t=9.337$, $p<0.05$; I: $t=11.609$, $p<0.05$; DF: $t=18.225$, $p<0.05$; DI: $t=7.560$, $p<0.05$; FI: $t=19.054$, $p<0.05$; DFI: $t=14.093$, $p<0.05$).

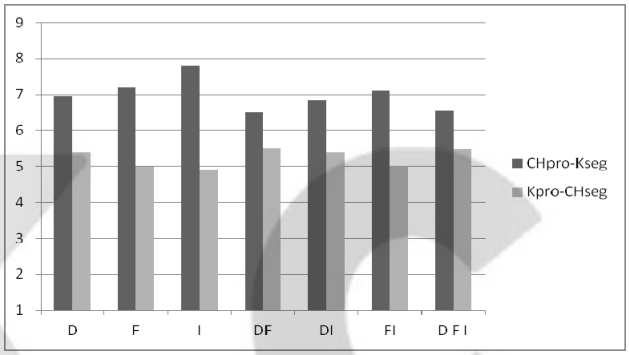

Figure 1. Foreign accent scores for the transplanted stimuli of L2 high prosody onto L1 segments (CHpro-Kseg) and those of L1 prosody onto L2 high segments (Kpro-CHseg).

Figure 2 indicates the foreign accent scores for the stimuli where Korean L1 talkers' prosody was transplanted onto Chinese L2 low proficiency talkers' segments. Similarly to Figure 1, the stimuli containing L2 talkers' prosody were evaluated significantly less accented than those of L2 talkers' segments (D: $t=7.832$, $p<0.05$; F: $t=13.509$, $p<0.05$; I: $t=24.669$, $p<0.05$; DF: $t=5.004$, $p<0.05$; DI: $t=11.476$, $p<0.05$; FI: $t=21.073$, $p<0.05$; DIF: $t=6.415$, $p<0.05$). Both figures 1 and 2 show the evidence that prosodic deviances are more tolerable than segmental ones. It seems that Korean L1 listeners rely more on segmental information than prosody when judging foreign accent of Korean L2 speech.

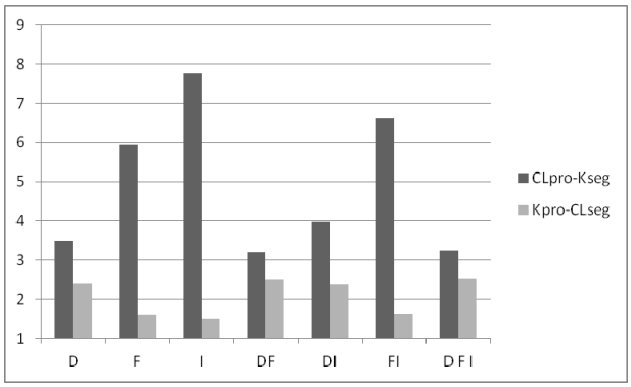

Figure 2. Foreign accent scores for the transplanted stimuli of L2 low prosody onto L1 segments (CLpro-Kseg) and those of L1 prosody onto L2 low segments (Kpro-CLseg).

Figure 3 shows a comparison of foreign accent scores between Chinese high and low proficiency talkers when Korean L1 prosody was combined with Chinese L2 segments. High proficiency talkers were evaluated consistently less accented than low proficiency talkers regardless of prosodic parameters (D: $t=12.865$, $p<0.05$; F: $t=9.007$, $p<0.05$; I: $t=15.112$, $p<0.05$; DF: $t=18.040$, $p<0.05$; DI: $t=20.109$, $p<0.05$; FI: $t=19.306$, $p<0.05$; DIF: $t=7.801$, $p<0.05$). Segments seem to provide sufficient phonological/phonetic cues to differentiate high and low proficiency talkers when prosodic information is equally ameliorated. That is, Chinese-accented Korean L2 speech is considerably distinctive in segmental representations between high and low talkers.

.

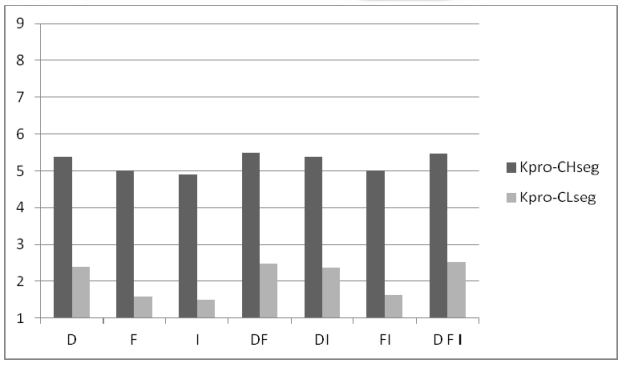

Figure 3. Foreign accent scores for the transplanted stimuli of L1 prosody with L2 high segments (Kpro-CHseg) vs. L1 prosody with L2 low segments (Kpro-CLseg)

Figure 4 presents a comparison of foreign accent scores between Chinese high and low proficiency talkers when Chinese L2 prosody was combined with Korean L1 segments. It is worth noting that accent rating scores were not consistently higher for Chinese high proficiency talkers. The scores were significantly different between high and low talkers' stimuli only when the duration parameter was entailed in the synthesized speech samples (D: $t=11.302$, $p<0.05$; F: $t=8.074$, $p=0.805$; I: $t=3.561$, $p=0.195$; DF: $t=16.104$, $p<0.05$; DI: $t=12.589$, $p<0.05$; FI: $t=10.273$, $p=0.523$; DIF: $t=1.322$, $p=0.611$). L2 talkers' production of pitch and intensity did not contribute to making recognizable differences between high and low proficiency of L2 talkers. That is, the duration parameter seemed to be a mere attribute to talkers' proficiency distinction. Due to the fact that the entire sentence was rearranged in the temporal dimension when the duration element was transplanted, the temporal lengths of consonants, vowels and even pauses were modified to those of L2 speech. High proficiency speech might be presumably more similar to native talkers in terms of temporal configurations than low speech, which resulted in the perception of better accent. This suggests that prosody in L2 speech plays a significant role when the influence of segments was suppressed. It was found that duration, i.e., temporal factors, was especially important in Korean L2 speech.

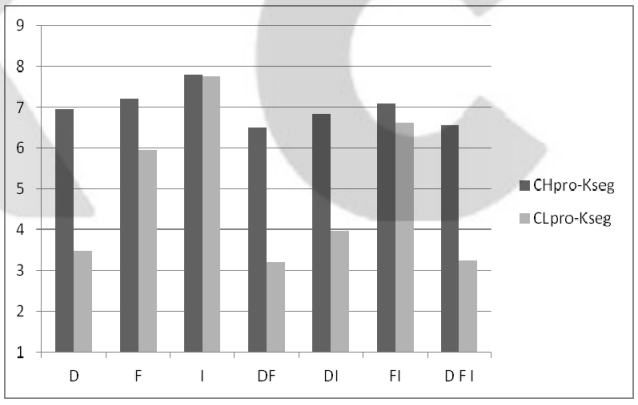

Figure 4. Foreign accent scores for the transplanted stimuli of L2 high prosody with L1 segments (CHpro-Kseg) vs. L2 low prosody with L1 segments (CLpro-Kseg)

Figure 5 displays the results when prosodic parameters were collapsed down together and compares the accent scores among four different types of combinations between L1 and L2 speech. Overall the combination of Korean L1 talkers' prosody and segments were rated as the best accent in all prosody parameters. The combinations with Korean L1 segments (Kpro+Kseg and Cpro+Kseg) were consistently better rated than those of Chinese L2 segments (Kpro+Cseg and Cpro+Cseg).

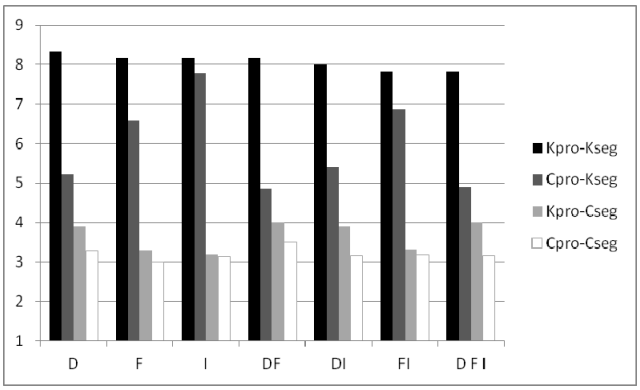

Figure 5. Foreign accent scores for the transplanted stimuli of L1 prosody and L1 segments (Kpro-Kseg), L2 prosody and L1 segments (Cpro-Kseg), L1 prosody and L2 segments (Kpro-Cseg) and L1 prosody and L2 segments (Cpro-Cseg)

The data in Figure 5 were submitted to the ANOVA analysis and its post-hoc test to see if the average scores of the four combinations were statistically different in each parameter. Table 1 presents the results of the post-hoc test. Interestingly, the stimuli of L1 prosody combined with L2 segments (Kpro-Cseg) and the stimuli of L2 prosody mixed with L2 segments (Cpro-Cseg) were not significantly different for the parameters where duration was not included (I, F and FI) as shown in Table 1. Whether the Chinese-accented segments were amalgamated with Korean talkers' prosody or Chinese talkers' prosody, Korean listeners perceived the foreign accent to a similar extent when pitch or intensity information was switched. Furthermore, the duration parameter played a critical role in differentiating Korean L1 speech from Chinese L2 speech. This is consistent with the results shown in Figures 3 and 4. Chinese high and low proficiency talkers significantly differed only when L2 talkers' duration was transplanted onto L1 segments

Table 1. Post-hoc results of the accent scores for prosodic parameters

D (F=9.388)	Cpro-Kseg	Kpro-Cseg	Cpro-Cseg
Kpro-Kseg	.000*	.000*	.000*
Cpro-Kseg		.000*	.000*
Kpro-Cseg			.000*
F (F=11.091)	Cpro-Kseg	Kpro-Cseg	Cpro-Cseg
Kpro-Kseg	.000*	.031*	.000*
Cpro-Kseg		.000*	.010*
Kpro-Cseg			.679
I (F=7.104)	Cpro-Kseg	Kpro-Cseg	Cpro-Cseg

Kpro-Kseg	.000*	.000*	.000*
Cpro-Kseg		.025*	.000*
Kpro-Cseg			.882
DF (F=5.083)	Cpro-Kseg	Kpro-Cseg	Cpro-Cseg
Kpro-Kseg	.000*	.000*	.000*
Cpro-Kseg		.000*	.000*
Kpro-Cseg			.000*
DI (F=12.971)	Cpro-Kseg	Kpro-Cseg	Cpro-Cseg
Kpro-Kseg	.000*	.000*	.000*
Cpro-Kseg		.000*	.000*
Kpro-Cseg			.000*
FI (F=9.301)	Cpro-Kseg	Kpro-Cseg	Cpro-Cseg
Kpro-Kseg	.033*	.000*	.000*
Cpro-Kseg		.025*	.000*
Kpro-Cseg			.147
DFI (F=6.584)	Cpro-Kseg	Kpro-Cseg	Cpro-Cseg
Kpro-Kseg	.000*	.000*	.000*
Cpro-Kseg		.000*	.000*
Kpro-Cseg			.002*

* $p<0.05$

Is it presumably possible to demonstrate that Chinese talkers successfully produced Korean pitch and intensity? This was not the case when Korean L1 segments were combined with Chinese L2 talkers' prosody (Cpro-Kseg); the ratings were significantly different from those for the combination of L1 segments and L1 prosody (Kpro-Kseg) for every parameter. This indicates that L1 Korean speech might be readily deteriorated by any prosody parameters produced by L2 talkers but that Chinese accented speech could be corrected by only Korean talkers' production of duration. Therefore, Korean listeners were more sensitive to duration or temporal arrangements of segments when they perceived Korean L2 speech. They might suppose that L2 speech was improved or less accented when the temporal lengths of L2 segments (including pauses) were arranged similarly to native Korean talkers. Even though it was reorganized into native Korean talkers' pitch or intensity features, Korean listeners did not judge it as improved accent. This can be an extended discussion of the previous studies (Lee and Liu 2012, Liu and Lee 2012, Lee 2014). They contended that relative weight among prosodic parameters might be different depending on the prosodic characteristics of L2. Pitch played a determining role in the perception of Chinese L2 speech while temporal deviances were more weighed in determining foreign accent of English L2 speech. This reflects that lexical and/or syntactic tone structures are dominant in Chinese as a tone language and that temporal patterns are significant in English as a stress-timed language. They consistently contented that the perception of foreign accent for prosody-transplanted L2 speech was significantly affected by the prosodic structure of

listeners' L1. Extending their discussion to different L1 and L2 as in the current experiment, Korean is a syllable-timed language. This means that temporal factors associated with syllables are perceptually more prominent than pitch or intensit. In the synthesized stimuli where the duration parameter was cross-planted between Chinese L2 and Korean L1 speech, the duration of syllables was modified in an utterance, ultimately leading to a change in speech rate. Therefore, Korean L1 listeners who are used to syllable timed utterance could more readily detect the change in syllable length than pitch or intensity.

Figure 6 shows relative significance of L2 talkers' prosody when it was overlaid onto L1 segment. The rating scores were significantly different among the 7 parameters ($F=16.668$, $p<0.05$), but the post-hoc test in Table 2 presents that the stimuli containing Chinese talkers' duration were rated as significantly lower than those with pitch or intensity. Consistently enough, L2 talkers' duration seems to trigger foreign accent significantly more than pitch and intensity.

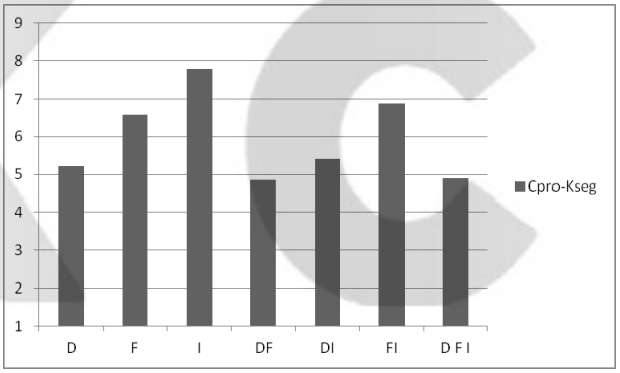

Figure 6. The foreign accent scores for the transplanted stimuli containing L2 prosody and L1 segments (Cpro-Kseg).

Table 2. Post hoc results for 7 prosodic parameters in Cpro+Kseg.

	F	I	DF	DI	FI	DFI
D	.000*	.001*	.663	.290	.004*	.067
F		.098	.000*	.036*	.839	.041*
I			.000*	.002*	.071	.043*
DF				.201	.022*	.157
DI					.000*	.192
FI						.005*

* $p<0.05$

Figure 7 compares the roles of L1 talkers' prosody in L2 speech. The ANOVA analysis demonstrates that there were no significant differences among those 7 parameters ($F=21.174$, $p=0.197$). It may be plausible that the effects of Korean talkers' prosody were suppressed because Chinese accented segments were strongly accented. As mentioned in Figures 1 and 2, segmental information was more influential than prosody in the assessment of foreign accent. Therefore, Figure 7 shows that Korean listeners mostly relied on strongly accented segments rather than Korean natives' prosody.

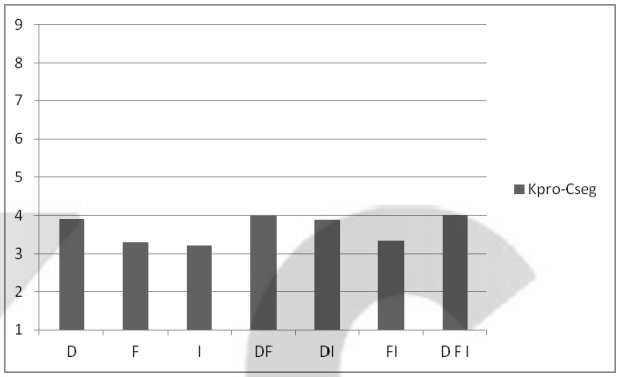

Figure 7. The foreign accent scores for the transplanted stimuli of L1 prosody and L2 segments (Kpro-Cseg).

3. Acoustic analysis

It has been shown that duration among prosodic parameters is a dominant factor in determining the foreign accent of Chinese accented Korean L2 speech. When L2 talkers' duration is copied onto the segmental representation of L1 speech, the L1 speech is reorganized in the temporal dimension. Now it seems to be necessary to look into the acoustic details associated with temporal deviance from Chinese-accented Korean speech in an attempt to find out which temporal feature best predicts the accentedness.

Many L2 studies on foreign accent have been interested in speech rate and presented the empirical finding that it is a sufficient cue for foreign accent (Anderson-Hsieh and Koehler 1988, Munro and Derwing 1994, Munro and Derwing 1995, Munro and Derwing 1998, Trofimovich and Baker 2007). Non-native speech with a slow speech rate was rated as more accented than the speech with a normal rate (Munro and Derwing 1998). L2 speech with too fast or too slow rate might easily cause native listeners to perceive it as accented (Anderson-Hsieh and Koehler 1988).

Other acoustic features reflecting temporal structures than speech rate are pause duration and frequency. These two parameters have been claimed to

directly relate with non-native talkers' proficiency (Anderson-Hsieh and Venkatagiri 1994, Munro and Derwing 1994, Flege et al. 1995, Munro and Derwing 1998, Pickering 1999, Riazantseva 2001, Kormos and Denes 2004, Iwashita et al. 2008). L2 speakers with a high proficiency produced less frequent and shorter pauses than those with an intermediate proficiency (Anderson-Hsieh and Venkatagiri 1994). L2 learners with high proficiency always produced fewer pauses than low proficient learners (Kormos and Denes 2004, Iwashita et al. 2008).

Assuming that speech rate, pause frequency and pause duration are substantially significant acoustic features relating with duration, we measured them in Chinese-accented Korean speech and observed their correlations with the scores of foreign accent in a linear regression analysis. Speech rate was calculated by dividing the total duration of an utterance (including pause time) by the total number of syllables (Kang 2010). Concerning pause-related features such as pause frequency and pause duration, we only collected the pauses that were longer than 0.1 second following Riazantseva (2001). We only measured the pauses occurring within a sentence; that is, inter-sentence pauses were not considered. Pause frequency was simply computed by counting how many pauses occurred in a given utterance, and pause duration denoted the time to be taken for each pause.

Table 3 presents the results of regression analysis. Speech rate shows the highest R^2 value of 78.5, indicating that it is the most important factor to predict foreign accent. Pause variables turned out to predict foreign accent less than speech rate while pause duration (R^2 =68.4) is a slightly better predictor than pause frequency (R^2=52.1%).

Table 3. Regression analysis for speech rate, pause frequency, and pause duration

	B	SE B	Beta	R^2	t
Speech rate					
Constant	1.77	0.211			8.378*
Speech rate	-0.101	0.075	-0.174	0.785	-1.342**
Pause duration					
Constant	-1.658	0.286			-5.791*
Pause duration	2.029	0.181	0.827	0.684	11.206*
Pause frequency					
Constant	3.132	0.353			8.796*
Pause frequency	0.299	0.223	-0.513	0.521	-3.863*

* p< .01

4. Conclusion

This study has presented the contribution of prosody to the perception of foreign accent in Chinese talkers' production of Korean. We cross-planted

the prosodic parameters between L1 and L2 speech and investigated their separate and independent roles in native listeners' assessment of foreign accent. We first found out that segmental information was a more significant factor than prosody when L1 listeners decided the overall impression of foreign accent in Chinese-accented Korean speech. Focusing on the prosody though it was less influential on accent ratings than segments, the effects of three prosodic parameters were examined individually or in combinations of two or three. Results showed that Korean listeners were more likely to put a comparatively great weight on duration. Duration, which is a temporally associated parameter, was acoustically analyzed by three different temporal measurements: speech rate, pause frequency and duration. A linear regression analysis showed that all of them were statistically correlated with foreign accent ratings, but that speech rate predicted foreign accent the most successfully, followed by pause duration and pause frequency.

Appendix A: Reading passage for accentedness task

지금 주머니 속에 동전이 몇 개쯤 들어 있나요? 자동판매기에서 음료수를 마실 때에나, 또 지하철 표를 살 때에 우리는 동전이 필요합니다. 동전은 하루에도 몇 번씩 우리의 주머니에서 들어갔다 나왔다 합니다. 이렇게 동전은 우리의 필수품이지만, 사용되는 횟수에 비해서 사람들의 관심을 끌지 못합니다. 여러분은 어떠신가요? 지금이라도 지갑 속의 동전을 꺼내서 한번 살펴보시지요.

Appendix B: Recording sentences for synthesis

① 박물관으로 갈래요?
② 불국사에 다보탑이 있습니다.
③ 한국은 쌀이 주식입니다.
④ 이 도시는 백 년 역사가 있어요.
⑤ 내가 직접 가야 돼요?

REFERENCES

ANDERSON-HSIEH, JANET, RUTH JOHNSON and KENNETH KOEHLER. 1992. The relationship between native speaker judgment of non-native production and deviance in segmental, prosody, and syllable structure. *Language Learning* 42, 529-555.

ANDERSON-HSIEH, JANET and KENNETH KOEHLER. 1988. The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning* 38, 61–613.

ANDERSON-HSIEH, JANET, and HORABAIL VENKATAGIRI. 1994. Syllable

duration and pausing in the speech of intermediate and high proficiency Chinese ESL speakers. *TESOL Quarterly* 28, 807-812.

AVERY, PETER and SUSAN EHRLICH. 1992. *Teaching American English Pronunciation*. Oxford: Oxford University Press.

BOULA DE MAREUIL, PHILLIPE and BIANCA VIERU-DIMULESCU. 2006. The contribution of prosody to the perception of foreign accent. *Phonetica* 63, 247-267.

DERWING, TRACEY and MURRAY MUNRO. 1997. Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition* 19, 1-16.

DIRVEN, RENE and JOHN OAKSHOTT-TAYLOR. 1984. Listening comprehension (part 1). *Language Teaching* 17, 89-95.

FLEGE, JAMES. 1987. The instrumental study of L2 speech production: Some methodological considerations. *Language Learning* 37.2, 285–296.

_____. 1991. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America* 89.1, 395–411.

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

FLEGE, JAMES, OCKE-SCHWEN BOHN and SUNYOUNG JANG. 1997. Effects of age of second language learning on the production of English consonants. *Speech Communication* 16, 1-26.

FLEGE, JAMES and JAMES HILLENBRAND. 1984. Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustic Society of America* 76.3, 706–721.

FLEGE, JAMES, MURRAY MUNRO and IAN MACKAY. 1995. Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustic Society of America* 97.5, 3125–3134.

IWASHITA, NORIKO, ANNIE BROWN, TIM MCNAMARA and SALLY O' HAGAN. 2008. Assessed levels of second language speaking proficiency: How difficult? *Applied Linguistics* 29, 24-49.

JILKA, MATTHIAS. 2000. *The Contribution of the Perception of Foreign Accent*. PhD Dissertation. University of Stuttgart.

KANG, OKIM. 2010. Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System* 38, 301-315.

KORMOS, JUDIT and MARIANN DENES. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32, 145-164.

KUHL, PETER. 1991. Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50, 93–107.

LEE, JOO-KYEONG. 2014. The role of prosody in the perception of foreign acent and comprehensibility: Prosody-corrected-L2 speech vs. prosody-

distorted-L1 speech. *Korean Journal of Linguistics* 39.1, 155-179.
LEE, JOO-KYEONG and XING LIU. 2012. The relative weight of prosody and segment in the perception of Korean accented Chinese speech. *Language and Linguistics* 54, 263-293.
LIU, XING and JOO-KYEONG LEE. 2012. The contribution of prosody to the foreign accent and Chinese talkers' English speech. *Journal of the Korean Society of Speech Sciences* 4.3, 59-73.
MACK, MOLLY. 1989. Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception and Psychophysics* 46.2, 187-200.
MILDNER, VESNA and DIANA HORGA. 1999. Relations between second language proficiency and formant defined vowel space. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 1455–1458.
MORLEY, JOAN. 1991. The pronunciation component of teaching English to speakers of other languages. *TESOL Quarterly* 25, 481-520.
MUNRO, MURRAY. 1995. Non-segmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition* 17, 17-34.
MUNRO, MURRAY and TRACEY DERWING. 1994. Evaluations of foreign accent in extemporaneous and read material. *Language Testing* 11.3, 253-266.
MUNRO, MURRAY and TRACEY DERWING. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 49, 285-310.
_____. 1998. The effect of speaking rate on listener evaluations of native and foreign accented speech. *Language Leaning* 48, 159-182.
_____. 1999. Foreign accent, comprehensibility and intelligibility in the speech language of learners. *Language Learning* 49, 73-97.
PENNINGTON, MARTHA and JACK RICHARDS. 1986. Pronunciation revisited. *TESOL Quarterly* 20, 207-225.
PICKERING, LUCY. 1999. *An Analysis of Prosodic Systems in the Classroom Discourse of Native Speaker and Nonnative Speaker Teaching Assistants.* Unpublished PhD Dissertation. University of Florida.
RIAZANTSEVA, ANASTASIA. 2001. Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language and Acquisition* 23, 497-526.
TROFIMOVICH, PAVEL and WENDY BAKER. 2007. Learning prosody and influency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics* 28, 251-276.
VAN ELS, THEO and KEES DE BOT. 1987. The role of intonation in foreign accent. *Modern Language Journal* 71, 147-155.
YOON, KYU-CHEOL. 2006. Imposing native speaker's prosody onto non-native speaker's utterance. Paper presented at WESPAC IX.

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

Xing Liu
School of Foreign Languages
Zhengzhou University of Light Industry
No. 5. Dongfeng Road, Zhengzhou, Henan Province
450002, China
e-mail: fayemeteor@naver.con

received: July 19, 2015
revised: August 18, 2015
accepted: August 20, 2015