Linguistic Research 33(3), 487-517
DOI: 10.17250/khisli.33.3.201612.006

Optimal prosody for different proficiency Chinese talkers of English: Pitch and speech rate*

Joo-Kyeong Lee
(University of Seoul)

Lee, Joo-Kyeong. 2016. Optimal prosody for different proficiency Chinese talkers of English. *Linguistic Research* 33(3), 487-517. This paper investigates native listeners' perceived accentedness for pitch- and duration-manipulated speech of Chinese L2 talkers and compared it with that of nonnative Chinese and Korean listeners. It attempts to find out actual numeric values of pitch and speech rate where low proficiency talkers are perceived as optimally improved proficiency and where high talkers are judged as the least accented. Results showed that Chinese low talkers were judged as a higher level of proficiency, intermediate, when the pitch ranges were synthetically expanded twice. High talkers were, on the other hand, degraded into intermediate proficiency when their pitch ranges was manipulated to be almost zero, but they were rated as optimally accented, close to the score of 1, when the pitch ranges were extensively widened two or three times. Speech rate did not, however, affect native listeners' perception of low talkers' stimuli, but the rating scores showed a curvelinear shape along with the changes of speech rate for high talkers. The optimal speech rate, where high talkers were perceived as the least accented, emerged as 6~6.5 s/s. It was 1.2 to 1.5 times faster than high talkers' original speech. These results cast a few pedagogical implications which can readily apply to teaching English. Prosodic factors such as pitch range or speech rate are sufficient attributes to altering the perceptual categories of L2 talkers' proficiency (except for speech rate for low talkers). The optimal numeric values of pitch range and speech rate will be solicited practically for low and high proficiency talkers in the pedagogical environment. In addition, nonnative Chinese and Korean listeners did not show statistically significant correlations with native listeners in the perception of foreign accent regardless of their L1 backgrounds or L2 proficiency. This might be plausibly due to the fact that prosodic features of nonnatives' interlanguages were considerably dissimilar to those native listeners. This might induce inconsistent ratings between native and nonnative listeners. **(University of Seoul)**

Keywords pitch range, speech rate, proficiency, foreign accent, L2 speech

---

* This work was supported by the National Research Foundation of Korea Grant funded by the Korean National Government (NRF-2014S1A5A2A01011617). I would like to thank the two anonymous reviewers of this journal for their helpful comments. All errors are, of course, mine.

1. Introduction

It has been widely claimed in L2 studies that foreign accent very often detected in nonnative speech is induced by phonological and phonetic deviances from native norms, giving rise to unsuccessful communications with native speakers. Such deviances may be attributed to both segmental and suprasegmental inaccuracy and/or errors, but numerous studies have focused on segmental features of L2 speech (Kuhl 1991; Flege 1991; Flege, Munro, and MacKay 1995; Flege, Takagi, and Mann 1995, 1996; Flege, Bohn, and Jang 1997; Munro, Flege, and MacKay 1996; Pallier *et al*. 1997; Takagi and Mann 1995; among others). They assumed that distortion of segmental information is primarily ascribed to relative accentedness of L2 speech; for instance, foreign accent was strongly related to segmental errors such as substitution, deletion and insertion (Anderson-Hsieh, Johnson, and Kohler 1992; Lee and Xue 2013).

Comparatively little attention has been paid to suprasegmental attributes such as intonation, speech rate, rhythm, pause, etc. Some studies on L2 speech have recently appreciated that prosodic deviances may contribute as much or more to native listeners' detection of foreign accent (Flege, Munro, and MacKay 1995; Munro 1995; Jilka 2000; Boula de Mareuiil and Vieru-Dimulescu 2006; Trofimovich and Baker 2007; Lee and Liu 2012, 2015). They sometimes executed a straightforward comparison between segmental and suprasegmental features in the judgment of foreign accent, demonstrating that suprasegmentals played a relatively dominant role over segmentals in L2 speech (Boula de MareuiilandVieru-Dimulescu 2006; Lee and Liu 2012; Lee and Liu 2015). On the other hand, suprasegmental factors were solely examined in segmentally nullified (low-pass filtered) L2 speech; intonation alone served as a sufficient cue to the perception of foreign accent (Munro 1995), and speech rate was the most prominent attribute to foreign accent (Trofimovich and Baker 2007). It seems to be inconclusive to determine a relative weight of suprasegmental parameters for perceived accentedness of L2 speech, but those studies all agreed that foreign accent was sufficiently detected from suprasegmentals, and provided an ample amount of empirical evidence taking advantage of advanced technical experiments.

The L2 studies mentioned above successfully presented a significant role of suprasegmentals in the perception of L2 accent, which otherwise would have been ignored due to comparatively more focus on segmental factors. However, they did not attempt to associate their results to pedagogical applications. Suprasegmentals have been claimed to be a guiding principle for pronunciation teaching in numerous studies of ESL/EFL pedagogy (Morley 1991; Nida 1957; McNerney and Mendelson 1992; Brazil, Coulthard, and Johns 1980; Pennington and Richards 1986; Brown 1995; Clennell 1996; Chela-Flores 1998; etc.). They have commonly asserted from either theoretical understanding or experimental evidence that current practice should redirect priorities within the sound system to a focus on the critical importance of suprasegmentals (Hahn 2004). Such assertions arise from the researchers' general appreciation that teaching suprasegmentals crucially improved the intelligibility of nonnative talkers' L2 speech in communication, more specifically grasping meaning in the context of discourse. These studies in pronunciation pedagogy shed light on the ideas about how the results of L2 speech studies will necessarily serve as substantial resources for teaching pronunciation.

It may be, then, plausibly true that instructors in the sites need practical values of suprasegmental features to apply to their teaching of second/foreign language pronunciation. What the instructors may actually be in need are the numeric measures of prosody such as pitch or speech rate to substantially minimize nonnative talkers' foreign accent and/or maximize speech intelligibility, that is, the prosodic values optimized for different L2 talkers.[1] Previous studies on L2 speech have merely investigated separate and independent roles of segmental and suprasegmental parameters and compared their relative contribution to perceived accentedness, seeking for its phonetic and/or phonological source. They haven't provided further resources applicable to L2 pronunciation pedagogy. The current study, however, concentrates on the practical values of prosody where English native listeners are perceptually sensitive enough to detect a categorical change of foreign accent and endeavors to extend the results of pitch and speech rate to pedagogical applications. In the experiment, Chinese talkers' production of English sentences were manipulated in pitch and duration in an attempt to seek for the numeric measures of pitch and speech rate to optimally decrease foreign accent for high and low proficiency Chinese L2 talkers.

Among the studies advocating that the role of prosody is equal to or prevalent over that of segments are two distinctive contentions; (1) pitch contour/intonation was found to be the most crucial factor attributable to the judgment of foreign accent, and (2) speech rate was the most sensible parameter adopted for natives'

---

1 Optimal values of prosody in L2 speech may be different according to L2 talkers' proficiency or L1 backgrounds.

perception of foreign accent. For example, Trofimovich and Baker (2007) examined low-pass filtered speech of Korean learners of English to determine the relative role of five prosodic elements such as stress timing, tonal peak alignment, speech rate, pause frequency, and pause duration. They demonstrated that speech rate along with the pause-related parameters outweighed stress timing and tonal peak. Furthermore, Liu and Lee (2012) transplanted the pitch and duration between English L1 and Chinese L2 talkers and created prosodically synthesized stimuli. Since the manipulations of pitch and duration were carried out over entire sentences, both L1 and L2 stimuli were altered in intonation and speech rate. Native listeners of English identified the L2 stimuli with L1 duration as less accented than those with L1 pitch and L1 stimuli with L2 duration as more accented than those with L2 pitch. Liu and Lee's work was extended in Lee (2014) where pitch and duration were transplanted between English L1 and Korean L2 talkers. Consistently enough, speech rate was found to be the more contributing feature to judging foreign accent than intonation.

Anderson-Hsieh, Johnson and Koehler (1992), on the other hand, showed that pitch, more precisely intonation, had a significant correlation with the perception of foreign accent. They collected 60 oral reading passages produced by male speakers of 11 different language groups and asked three experienced ESL teachers for foreign accent ratings. The speech stimuli were analyzed by trained phoneticians in terms of three errors: intonation, segments and syllable structure. The degree of deviances in each error was correlated with the accent ratings in such a way that foreign accent was statistically more correlated with intonation than segments or syllable structure. A more advanced technology was employed in Munro (1995). He compared low-pass filtered L1 and L2 speech with corresponding unfiltered speech and showed that native listeners correctly judged that L2 speech was more strongly accented than L1 even when intonation information was only available. Munro suggested that intonation alone was sufficient enough for the identification of foreign accent. His work was extended to English and German by Jilka (2000). The low-pass filtered speech of German and American English was presented to both native German and American English listeners for the identification of language and accent. The listeners successfully differentiated German from English when listening to the stimuli with pitch information alone.

Similarly Kang (2010) showed that a pitch range was the best predictor of foreign accent. She investigated the effects of the individual suprasegmentals such as speech rate, pauses, stress, and pitch range on native listeners' perception of foreign accent. The results demonstrated that the overall pitch range best predicted the nonnative talkers' accentedness, i.e., listeners judged the English L2 speech with less pitch variations as more accented.

A most recent technology was used to measure suprasegmental features of L2 with keeping segmental information intact in Lee and Liu (2012). They adopted the TD-PSOLA algorithm and crossed prosodic and segmental elements between Chinese L1 and Korean L2 speech. Chinese listeners identified Korean L2 speech with Chinese pitch as less accented than that of Chinese duration. They concluded that pitch contour or intonation played a more important role than speech rate in the perception of foreign accent of Chinese L2 speech. Lee and Liu integrated their results with those of Liu and Lee (2012) and asserted which prosodic factor is more weighed might depend on native listeners' L1. English native listeners put more importance on speech rate for the accent judgment of Chinese L2 talkers' production of English (Liu and Lee 2012) while Chinese native listeners were more sensitive to pitch information when listening to Korean L2 talkers' speech of Chinese (Lee and Liu 2012).

It seems to be inconclusive in the previous work to determine which of pitch and speech rate is more contributable to perceived foreign accent, but it may probably rather depend on complicated relations between L1 and L2 phonological and/or phonetic sound structures and systems. For example, in the experiments with similar technical methodology, duration/speech rate was born out to have more weight in the assessment of foreign accent for Chinese L2 talkers' English speech (Liu and Lee 2012), Korean L2 talkers' English speech (Lee 2014) and Chinese L2 talkers' Korean speech (Lee and Liu 2015) while pitch/intonation played a more crucial role in Chinese listeners' judgment of Korean talkers' Chinese L2 speech (Lee and Liu 2012). The current study, therefore, assumes that both speech rate and pitch are equally responsible for the detection of foreign accent and examines the extent to which perceived foreign accent changes across the categories of talkers' proficiency when each factor is manipulated to increase and decrease.

It has been broadly accepted that foreign accent of L2 speech is evaluated by native listeners; therefore, there have been relatively a small number of studies where nonnative listeners took part in judging foreign accent (Chen 2011; Flege 1988; Munro, Derwing, and Morton 2006; Lee and Xue 2013; Wester and Mayo 2014). Foreign accent ratings were correlated between native and nonnative listeners when

L2 listeners' proficiency was high and the listeners' L2 experience was long enough (Flege 1988; Lee and Xue 2013). Moreover, nonnative listeners rated nonnative talkers as less accented than native listeners (Lee and Xue 2013; Western and Mayo 2014). It was interpreted as stating that high proficiency listeners have a more similar interlanguage to native speakers than low proficiency listeners and that it enabled the high listeners to evaluate foreign accent more accurately and closely to native listeners (Lee and Xue 2013). Lee and Xue further asserted that nonnative listeners and talkers with the same language background share their interlanguage which facilitates sound familiarity between them, resulting in less harsh ratings. Along with the same line, the present study investigates both native and nonnative listeners' judgments of foreign accent to see how similar or different they will be in association with nonnatives' L1 backgrounds and L2 proficiency.

2. Experiment

2.1 Talkers and listeners

Six Chinese female talkers of English participated in recording the stimuli for the perception experiment of foreign accent: three of them were high proficiency talkers and three of them were low proficiency talkers.[2] The English L2 proficiency of Chinese speakers was determined in an Accentedness Rating (AR) task (Munro, 1998). More specifically, prior to their recording, fifteen female Chinese speakers were recruited at a university in Seoul to collect three high and three low proficiency talkers. They were asked to read a passage consisting of five sentences with various structures and length (See Appendix I).[3] Their recordings were presented to three native English listeners for accentedness ratings on a 9-point Liker scale (1=native like; 9=strong foreign accent). The participants who were rated between 7 and 9 were categorized into low proficiency, those who were rated between 4 and 6 were intermediate proficiency, and those who were evaluated as 1 through 3 were high proficiency. Three were randomly chosen from those who judged as high proficiency, and three were randomly selected from those who perceived as low proficiency. The participants who turned out to be

2 Since pitch is peculiar to gender, male and female pitch measures can't be mixed to generalize nonnatives' optimal pitch values. Therefore, the talkers were intentionally confined to females.

3 We used the same passage as in Xue and Lee (2014).

intermediate did not take part in the experiment as talkers. The three native raters did not participate in the later perception experiment as listeners.

As for the listeners, 10 English listeners served as native listeners, and 40 Chinese and Korean listeners served as nonnative listeners; 10 high and 10 low proficiency Chinese listeners and 10 high and 10 low proficiency Korean listeners. All of them reported that they had no hearing impairments. The L2 English proficiency of nonnative Chinese and Korean listeners was determined in another AR task. Thirty five Chinese speakers were recruited for the AR task; 11 of them were exchange students at a university in Korea, and 24 of them were college students at a university in China. They recorded the same passage that Chinese listeners used in their AR task, and the recordings of 35 potential listeners were presented to three native English speakers for accentedness ratings. Among 35 Chinese participants were 13 rated as scores 1, 2 and 3, and 10 of them were randomly selected for high proficiency listeners. Eleven participants were rated as scores 7, 8 and 9, and one of them was simply expelled to make 10 low proficiency listeners. It turned out that 3 of 10 high proficiency and 8 of 10 low proficiency Chinese listeners were from a college in China. The remaining 11 Chinese speakers who were presumably categorized into intermediate proficiency (scores 4, 5 and 6) did not participate as listeners because the current experiment only focuses on high and low listeners to maximize the difference in proficiency.

To organize 20 high and low proficiency Korean listeners, 40 Korean college students were recruited and asked to read the same passage that Chinese listeners used in their AR task. Koreans' recordings were rated for accentedness by the same native English speakers who participated in the Chinese talkers' and listeners' AR tasks. Thirteen of the Korean participants were judged as between scores 1 and 3, and ten of them were randomly selected for Korean high proficiency listeners. Sixteen were rated as between scores 7 and 9, and ten of them were chosen for Korean low proficiency listeners. The remaining eleven Korean participants were assessed to be potentially intermediate, but they didn't serve as listeners in the experiment.

2.2 Stimuli

To prepare the stimuli for the main perception test, three high proficiency

Chinese talkers and three low proficiency Chinese talkers were asked to read three English statements and three yes-no questions (see Appendix II).[4] There was only one pitch accent in each sentence with no doubt because the high pitch accent (H*) in statements and the low pitch accent (L*) in yes-no questions were synthesized to increase and decrease in their height. All the recordings were carried out in a sound-attenuated booth using the Praat program. Thirty six natural sentences (6 talkers * 6 sentences) were submitted for pitch and duration manipulations.

For pitch manipulations, 36 sentences were synthesized in Praat; the pitch value of a high pitch accent (H*) in each statement was raised by 20Hz up to +100Hz, and lowered by 20Hz down to -100Hz (see more details in Xue and Lee, 2014). There were ten synthesized statement stimuli generated and one original natural statement audio stimulus from one statement; therefore, 198 pitch-manipulated statement stimuli (11 stimuli * 3 statements * 6 talkers) were provided to one listener. Concerning the pitch manipulations of yes-no questions, the f0 of the low pitch accent (L*) in each question was synthesized to increase up to +100Hz and decrease down to -100Hz in 20Hz intervals. Consequently, there were ten synthesized question stimuli produced and one original recording from one yes-no question; therefore, 198 pitch-manipulated yes-no question stimuli (11 stimuli * 3 questions * 6 talkers) were served to one listener. In total, 396 pitch-manipulated audio stimuli were presented to the listeners for the judgment of foreign accent.

For duration manipulations, 36 sentences (3 statements * 6 talkers + 3 yes-no questions * 6 talkers) were synthesized in the Praat; the duration of each sentence was synthesized to expand twice and compress a half, and there were five steps implemented in both expansion and compression processes. That is, sentence duration was manipulated to increase to 1.2, 1.4, 1.6, 1.8 and 2.0 times and decrease to 0.9, 0.8, 0.7, 0.6 and 0.5 times; therefore, 11 stimuli (ten synthesized and one natural) were generated from each sentence (see more details in Xue and Lee 2015). Since the duration manipulation was carried out over the entire sentence, the expansion of the duration resulted in slower speech, and the compression of it resulted in faster speech. In other words, duration manipulations are straightforwardly associated with speech rate alterations. There were, in total, 398 duration-manipulated audio stimuli (11 stimuli * 3 statements * 6 talkers + 11 stimuli * 3 yes-no questions * 6 talkers)

---

4 We adopted the sentences from Xue and Lee (2014).

generated for the perception experiment.

2.3 Procedure

The listeners made two visits to the speech lab located in the schools of Korea or China; one for the perception of pitch-manipulated L2 speech and the other for the perception of duration-manipulated L2 speech. The 11 Chinese listeners (3 high and 8 low proficiency) who were recruited in a college of China participated in their school. They had as many breaks as they wanted during the experiment to avoid any fatigue. They were all paid upon completion.

Native listeners' responses were analyzed to seek for the pitch and speech rate values where high proficiency talkers who were rated as scores 1 to 3 in the AR task were perceived to be intermediate in foreign accent ratings, obtaining scores 4 to 6. In the case of low proficiency talkers who were rated as scores 7 to 9 in the AR task, the pitch and speech rate values were explored where their accent ratings improved to a higher category of intermediate (scores 4 to 6). Nonnative listeners' responses were further submitted to a statistical analysis of correlation with those of native listeners to see how similar or different nonnative listeners' responses would be in accordance with nonnonnative listeners' L2 proficiency (high vs. low) or L1 backgrounds (Chinese vs. Korean).

2.4 Results

Among the five listener groups was the native group first analyzed in order to see how native listeners' responses of foreign accent would change as the pitch increased or decreased on the high and low pitch accents (H* and L*). Figure 1 shows native listeners' (NL's) average rating scores for the manipulated H*s of Chinese high proficiency talkers' (CHT) statements. The average pitch of H*s of the CHTs in their original stimuli, which was represented by the dotted vertical line in the graph, was 219Hz, and native listeners' average rating was 2.3. Overall, the accent scores continuously decreased when the H* synthetically increased, and the scores increased when H* decreased. Therefore, it was seen that native listeners preferred higher pitch values of H*s from CHTs' production of English statements. Interestingly enough, the mean rating 2.3 on 219Hz dropped down to 1's when the peak of H* gradually increased to 270Hz ~ 350Hz. The score of 1 means the most native-like and favorable accent, and native talkers are even rated as scores between 1 and 2 in average (Lee, 2014). Therefore, the scores of 1's (or between 1 and 2) are assumed to be optimal for CHTs' high pitch accent (H*). See the right ward frequencies from the solid vertical line at 270Hz where the rating scores were all under 2.

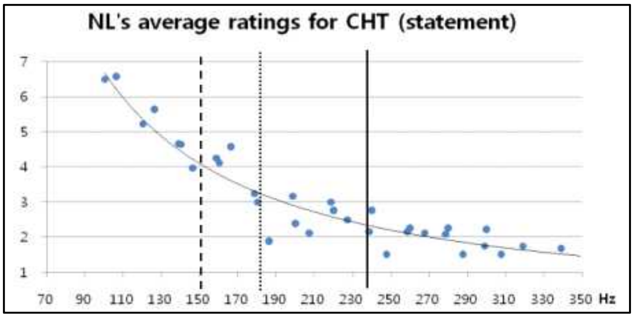

Figure 1. Native listeners' average rating scores for CHT (pitch-manipulated statements)

The accent scores, however, increased as high as 6.5 when the H* decreased down to around 100Hz. What is interesting to note in Figure 1, the prediction line shows that CHTs, though they were evaluated as lower than 3 in the AR (Accentedness Rating) task, were perceived as intermediate proficiency when the peak of H* dropped as low as 150Hz.[5] See the left ward frequencies from the broken line at 150Hz where the rating scores were all above 4. That is, such a drop of H* alone will sufficiently change the perceptual category of foreign accent from high to intermediate proficiency.

Figure 2 shows average rating scores for manipulated H*s in three low proficiency talkers' (CLT) production of English statements. The mean value of H*s in the unsynthesized original stimuli was measured to be 210Hz, and its accent score was 7.1 in average. The accentedness (scores) tended to be stronger (higher) as the peaks of H*s decreased and vice versa. As seen in the prediction line, the accent scores went higher,

---

5 Recall the 9-point Liker scale on which native English evaluators accessed the accentedness of nonnative participants in the AR task. Their proficiency was determined by three criteria: those who obtained the scores 1 to 3 were categorized to be high, those who belonged to the scores 4, 5 and 6 were intermediate, and those who were evaluated between 7 and 9 were low.

gradually increasing close to 9 when the peak of H* decreased from the original value (see the scores on the leftward pitch values from 210Hz along the prediction line). When the high pitch accents (H*) were, however, higher than 210Hz (in the rightward direction from 210Hz), Chinese low talkers will be consistently perceived to have better accent. When the peak was as high as 295Hz as shown by the solid vertical line, their proficiency will categorically change to intermediate because it was the point where the foreign accent score starts from 6 to lower. This indicates that low proficiency talkers can satisfactorily improve their accent when correcting their H* to be higher than 295Hz and that this will be an optimal value of H* for low proficiency talkers.

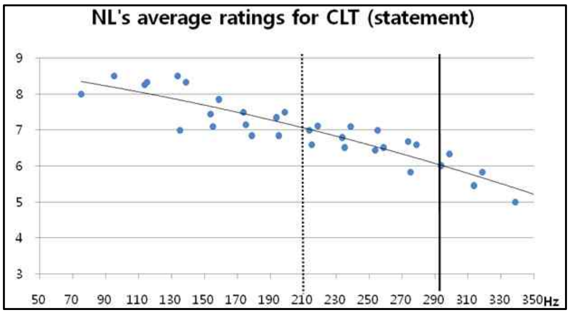

Figure 2. Native listeners' average rating scores for CLT (pitch-manipulated statements)

Turning to the results of yes-no questions, native listeners' responses for CHT are shown in Figure 3. The average score for high talkers' original production was 1.9 when the low pitch accent (L*) was 190Hz as indicated by the dotted line. Note that CHTs were perceived as the lower/intermediate proficiency (that is, scores between 4 and 6) when the pitch of L* went up to 275Hz and higher in their yes-no question productions. See that the scores are higher than 4 on the rightward pitch values from the broken line at 275Hz. On the other hand, the ratings improved to be 1's when the valley of L* was more deepened between 110Hz and 190Hz. Therefore, this indicates that native listeners preferred L* was as low as 110Hz ~ 190Hz in CHTs' production of yes-no questions.

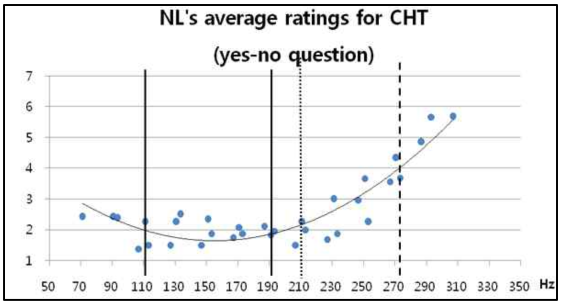

Figure 3. Native listeners' average rating scores for CHT (pitch-manipulated yes-no questions)

Native listeners' evaluation of pitch-manipulated stimuli excerpted from Chinese low talkers' (CLT) yes-no questions as shown in Figure 4. The pitch of L* was 195Hz in their original production, which was represented by the dotted vertical line. The scores on the left of the vertical line were mostly distributed below 7, but the scores on the right of the vertical line were concentrated above 8. As shown in the prediction line, accent scores dropped down to 6 and lower when the L* was below 110Hz (see the leftward direction from the solid vertical line). In other words, the low talkers were perceived as intermediate proficiency when they produced low pitch accents (L*) at 110Hz or lower. Therefore, these would be optimal pitch values of L* for low proficiency talkers to be perceived as better accent. When the pitch of L* increased higher than 195Hz, the original pitch value, the scores changed to 8 and 9. Since the original speech was already very strongly accented, i.e., 8.7, the increase of L* did not seem to have room to worsen the perceived accentedness.

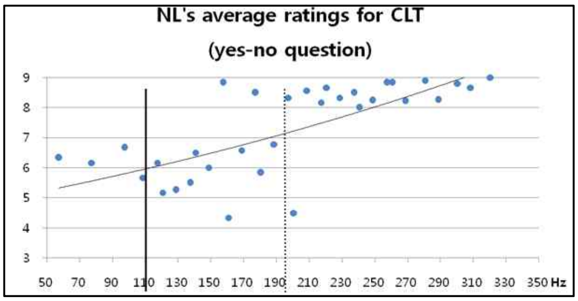

Figure 4. Native listeners' average rating scores for CLT (pitch-manipulated yes-no questions)

We have seen native listeners' responses to Chinese nonnative talkers with high and low proficiency for their foreign accent. Now, nonnative Chinese and Korean listeners' responses are presented along with those of natives, taking both nonnatives' L1 background and L2 proficiency into consideration. The average scores obtained from five listener groups were scatter-plotted in Figure 5. The regression line of the natives' scores seems to be almost inversely proportional between the x-axis (pitch) and the y-axis (score). In the cases of CHL and CLL, pitch and accent scores seem to show inverse proportion although the slopes of the regression lines are much moderate. The cases of KHL and KLL do not, however, look to have such a relation; they are almost parallel to the x-axis with little change in accent scores. According to the Pearson-r correlation analyses, the rating scores of CHL and CLL were significantly correlated with those of NL ($r=0.716$ for CHL, $r=0.631$ for CLL, $p<0.005$), but KHL and KLL did not show significant correlations with NL ($r=0.171$, $p=0.340$ for KHL, $r=0.133$ $p=0.461$ for KLL). Therefore, Chinese listeners rated foreign accent for Chinese high talkers in a similar way to native listeners.

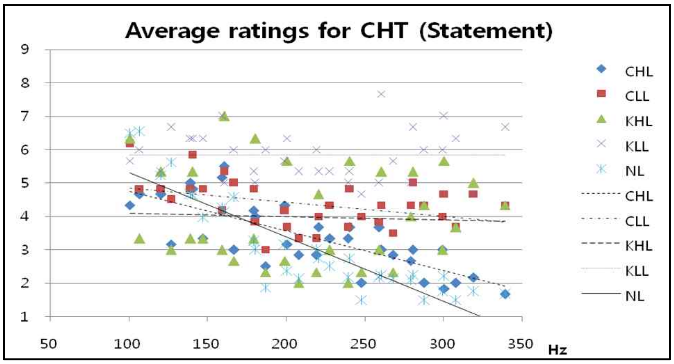

Figure 5. Average rating scores of five listener groups for Chinese high proficiency talkers' statements: Chinese high proficiency listeners (CHL), Chinese low proficiency listeners (CLL), Korean high proficiency listeners (KHL), Korean low proficiency listeners (KLL), and native listeners (NL)

Figure 6 shows the accent scores of all listener groups for Chinese low proficiency talkers' (CLT) statements. NL's scores are undoubtedly reversely proportional to the pitch values of H* as shown in the regression line. Similar reverse patterns are found in CHL and CLL, though the slopes do not seem to be as stiff as that of NL. On the other hand, KHL and KLL do not seem to show such a reverse relation between foreign accent and pitch; their regression lines are rather flat. This is confirmed by the statistics of Pearson -r correlations. The accent scores of CHL and CLL were significantly correlated with those of NL ($r = 0.542$, $p = 0.001$ for CHL and $r = 0.638$, $p < 0.005$), but the correlation coefficients were not statistically significant between KHL and NL ($r = 0.125$, $p = 0.489$) or between KLL and NL ($r = -0.089$, $p = 0.623$).

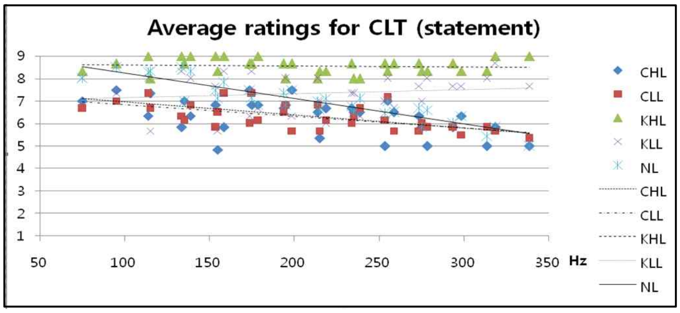

Figure 6. Average rating scores of five listener groups for Chinese low proficiency talkers' statements

Concerning the foreign accent responses to yes-no questions of CHT, the correlations between native listeners and nonnative listeners are graphically shown in Figure 7. NL's scores are directly proportional to the pitch, which means that foreign accent improves as the bottom value of L* deepens. Among the four groups of nonnative listeners does the KHL group look similar to NL, but there was no significant correlation between them ($r = 0.362$, $p = 0.041$).

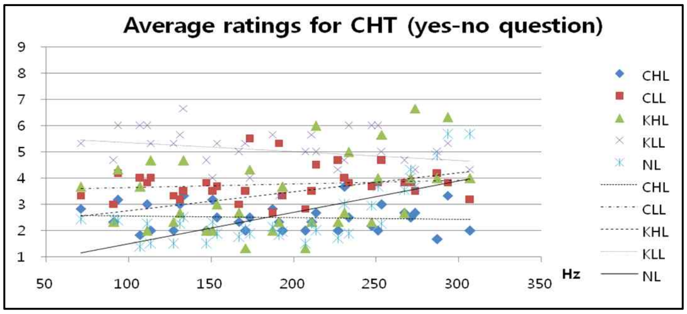

Figure 7. Average rating scores of five listener groups for Chinese high proficiency talkers' yes-no questions

The other nonnative listeners seem to show somewhat dissimilar shapes of regression lines from NL, and this was confirmed by the statistics. The Pearson-r correlation analyses demonstrated that their rating scores did not significantly agree with those of NL (CHL vs. NL, r = 0.131, p = 0.466, CLL vs. NL, r = -0.074, p = 0.683, and KLL vs. NL, r = -0.091, p = 0.613

Figure 8 presents native and nonnative listeners' average scores for CLT's yes-no questions. NL's scores consistently show a direct proportion with the pitch. Only CHL's scores are merely proportional in a gentle slope, but their correlation was not statistically correlated with those of NL (r = 0.312, p = 0.048). The other nonnative listeners do not even seem to be similar to NL in the shapes of regression lines. This was consistently supported by the Pearson-r correlation analyses; CLL, KHL, and KLL do not show significant correlations with NL (r = 0.062, p = 0.730 for CLL, r = 0.058, p = 0.747 for KHL, and r = -0.073, p = 0.688 for KLL).

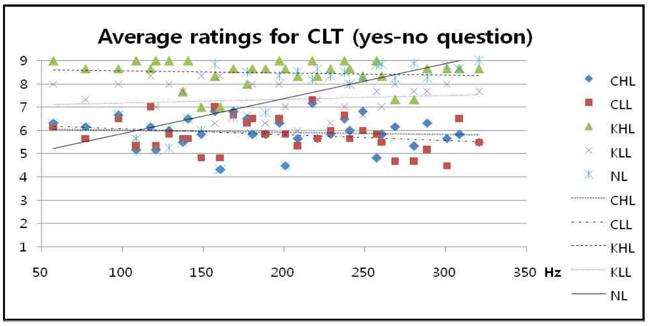

Figure 8. Average rating scores of five listener groups for Chinese low proficiency talkers' yes-no questions (pitch-manipulated)

Moving onto the results of duration-manipulated stimuli, Figure 9 shows NL's responses to duration-manipulated CHT's statement stimuli. The average accent score, which was measured from the unsynthesized original stimuli, was 2.3 at 5.2 syllables per second (s/s) as indicated by a dotted line. The scores were in a curvelinear relation with speech rate; that is, native listeners judged it as stronger when nonnative talkers' speech was too fast or too slow. The curvelinear prediction line shows that high talkers were perceived as lower or intermediate proficiency in their foreign accent when the speech rates were slow enough at 4 s/s and fast enough at 8.8 s/s as shown by the broken lines. However, an optimal speech rate was found to be 6.5 s/s when the rating score reached the bottom (the lowest). It was represented by a solid line. This suggests that CHT may need to speak English statements slightly faster (6.5 s/s) than their normal speech (5.2 s/s) to minimize their perceived accentedness.

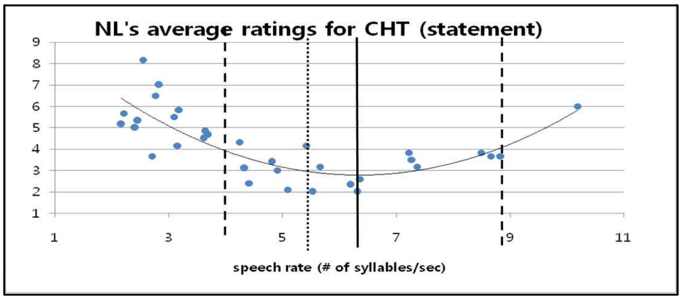

Figure 9. Native listeners' average rating scores for CHT (duration-manipulated statements)

Duration-manipulated stimuli of CLTs' statements were assessed by native listeners as shown in Figure 10. The average score of the unsynthesized speech was 7.1 at the speed of 4.4 s/s as indicated by a dotted line. What is interesting to note here is that a clear curvelinear relation is not observed unlike CHTs; the scores were distributed rather flat along with all speech rates. Moreover, no stimuli were judged below 6, indicating that a change in speech rate may not facilitate the foreign accent of low proficiency talkers.

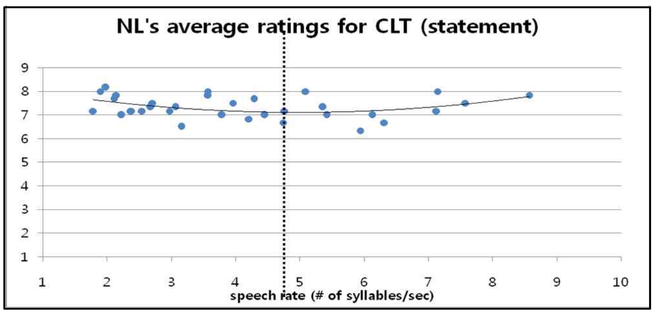

Figure 10. Native listeners' average rating scores for CLT (duration-manipulated statements)

Native listeners' accent scores for CHT's yes-no questions were scatter-plotted in Figure 11. Their distribution generates a curvelinear shape as in the case of statements (Figure 10), indicating that native listeners judged nonnative speech as strongly accented if it was too fast or too slow. The mean score of unsynthesized original yes-no questions was 1.9 at the speed of 4.2 s/s (see the dotted line), but the scores went higher than 4 when the speech rate was as slow as 3.0 s/s and as fast as 8.1 s/s (see the broken lines). Therefore, CHTs were downward perceived as intermediate proficiency when their speech was too slow or too fast. An optimal speech rate was found to be 5.8 s/s/ for high proficiency talkers when the score reached the lowest (see the solid line). Since the speed of their original speech was 4.2 s/s, CHT may need to speak a bit faster to be perceived as optimal/best accent.

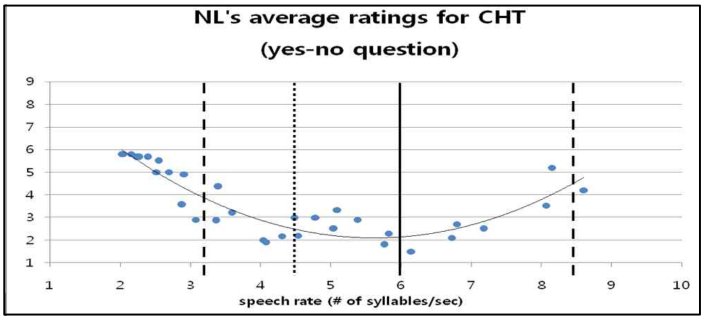

Figure 11. Native listeners' average rating scores for CHT (duration-manipulated yes-no questions)

Figure 12 illustrates the results for CLTs' yes-no questions. The mean speech rate was 4.2 s/s as indicated by a dotted line. Similarly to CLTs' statements in Figure 10, the curve-linear relation was not shown as clearly as in the case of CHTs' statements. Moreover, a category change in proficiency perception did not occur because the scores were all concentrated between 7 and 9 as speech rates changed. That is, speech rate did not seem to be an influencing contributor to improving the perception of foreign accent in the case of low proficiency talkers.

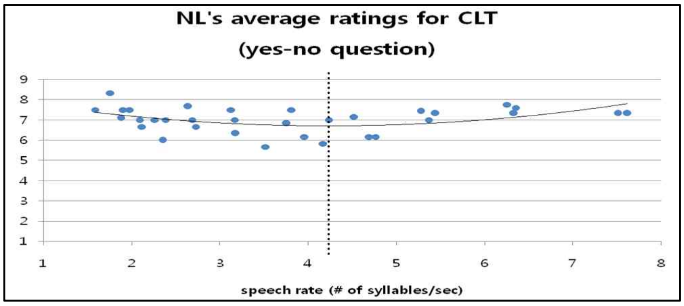

Figure 12. Native listeners' average rating scores for CHT (duration-manipulated yes-no questions)

Now let's look at nonnative listeners' scores for Chinese talkers and see if nonnative ratings will pattern together with those of natives in the perception of duration-manipulated L2 speech. Figure 13 presents native and nonnative listeners' perception results for CHT statements. The Pearson-r correlation analysis demonstrated that NL had no significant correlations with any groups of nonnantive listeners (r = 0.266, p = 0.97 for CHL, r = 0.273, p = 0.088 for CLL, r = 0.299, p = 0.061 for KHL, and r = 0.052, p = 0.749 for KLL).

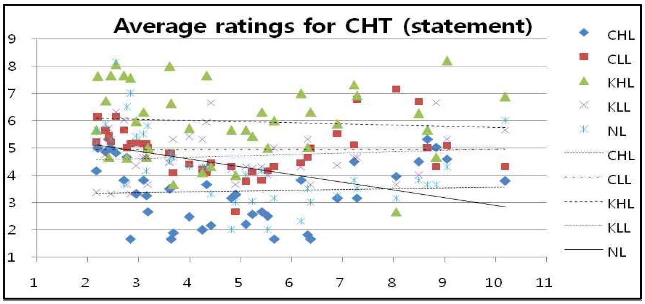

Figure 13. Average rating scores of five listener groups for CHT (duration-manipulated statements)

Figure 14 shows average accent scores of five native and nonnative listener groups for CLT statements. Similarly to the results for CHT statements above, there were no significant correlations between NL and any nonnative listener groups (r = 0.229, p = 0.167 for CHL, r = 0.196, p = 0.238 for CLL, r = 0.169, p = 0.312 for KHL, and r = 0.014, p = 0.935 for KLL).

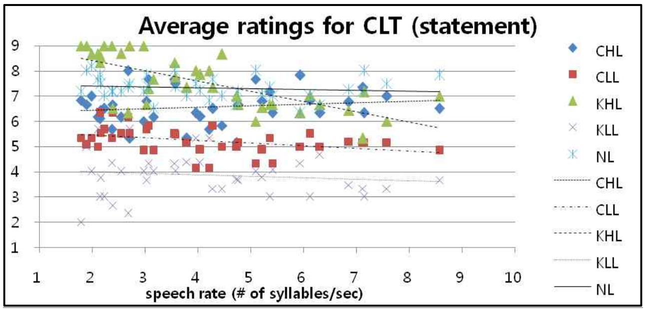

Figure 14. Average rating scores of five listener groups for CLT (duration-manipulated statements)

Next, Figure 15 shows native and nonnative listeners' assessments for duration-manipulated yes-no question stimuli of CHT. Concerning the Pearson-r correlations, the stimuli which were temporally synthesized did not show statistically significant correlations between native and nonnative listeners ($r = 0.159$, $p = 0.287$ for CHL, $r = 0.049$, $p = 0.765$ for CLL, $r = 0.143$, $p = 0.386$ for KHL, and $r = -0.387$, $p = 0.067$ for KLL).

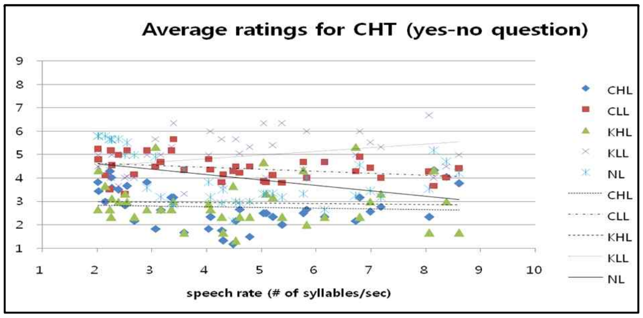

Figure 15. Average rating scores of five listener groups for CHT (duration-manipulated yes-no questions)

Both natives' and nonnatives' responses to different speech rates of CLTs' yes-no questions are illustrated in Figure 16. As far as the correlations of NL with nonnative listeners are concerned, native scores were not significantly correlated with those of any groups of nonnative listeners ($r = 0.258$, $p = 0.112$ for CHL, $r = 0.028$, $p = 0.868$ for CLL, $r = 0.034$, $p = 0.838$ for KHL, and $r = -0.073$, $p = 0.661$). As shown in the figure, there was little change in native judgments as the speech rate changed. However, the scores of nonnative listeners showed inconsistent patterns with those of NL.

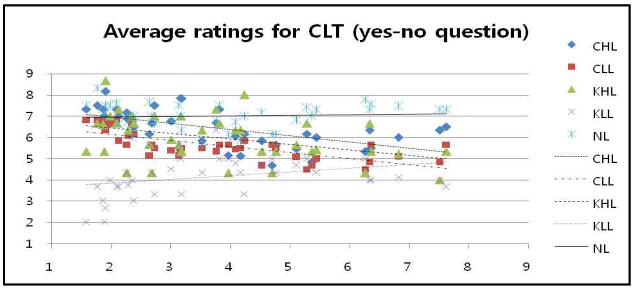

Figure 16. Average rating scores of five listener groups for CLT (duration-manipulated yes-no questions)

2.5 Discussion

We have investigated native listeners' perception of foreign accent for prosodically manipulated L2 speech and compared their rating patterns with those of nonnative Chinese and Korean listeners. Unlike many previous studies on natives' perception of foreign accent, the current work attempts to seek for optimal values of pitch and speech rate for different proficiency Chinese talkers, assuming that segmental deviances derived from different L2 proficiencies may be associated with particular prosodic features in order to decrease, ultimately minimize, foreign accent. Such substantial resources can be, in various ways, applied to implement such practical tools as computer software for teaching English prosody, more specifically, teaching how to decrease foreign accent by correcting prosody.

For the perception experiments, we synthesized the stimuli in pitch and duration respectively, keeping all the segmental information intact. Therefore, if there is any change in accent ratings, it will undoubtedly arise from the manipulated prosody. Concerning the perception of pitch-manipulated speech, Figures 1 to 4 reported the pitch values of H* and L* where native listeners changed their judgments across proficiency categories for high and low Chinese talkers. That is, the results were analyzed to find out the pitch values where low proficiency talkers perceptually improved to an intermediate level and where high proficiency talkers worsened to an intermediate category. However, it has been widely acknowledged that local pitch changes and global pitch ranges are perceptually more important than absolute pitch values (Pierrehumbert and Hirschberg, 1990). This might be due to the fact that pitch values are not invariable in both inter- and intra-speakers' speech. For instance, H* might not be sufficiently prominent to listeners when a lowest pitch value is relatively high within an intonational phrase because this constitutes a smaller pitch range.

Along with this line, the results in Figures 1 to 4 are redeemed in consideration of pitch ranges in Tables 1 and 2. The pitch values of H* are very similar between high and low talkers: 219Hz vs. 210Hz. The mere difference of 9Hz might not give a drastic and/or sufficient difference with which native listeners made a distinction between high and low proficiency in foreign accent, but other factors than the absolute height of H* should determine talkers' proficiency, for example, phonological and phonetic features of segments and/or other prosodic features like speech rate, rhythm, etc. Another possible factor relevant to pitch, i.e., pitch range, is examined at this point. As presented in Tables 1 and 2, CHT's pitch ranges are higher than those of CLT in both statements (69Hz vs. 52Hz) and yes-no questions (80Hz vs. 50Hz), and their differences were statistically significant ($t=2.59$, $p<0.05$ for statements and $t=11.96$, $p<0.005$ for yes-no questions).

Table 1. Pitch ranges of original and synthesized speech (statements)

	Original speech			Synthesized speech	
	bottom	H*	pitch range	H* at proficiency change	pitch range
CHT	150Hz	219Hz	69Hz	150Hz (High to Intermediate)	0
CLT	158Hz	210Hz	52Hz	295Hz (Low to Intermediate)	137Hz

Table 2. Pitch ranges of original and synthesized speech (yes-no questions)

	Original speech			Synthesized speech	
	top	L*	pitch range	L* at proficiency change	pitch range
CHT	270Hz	190Hz	80Hz	270Hz (High to Intermediate)	0
CLT	235Hz	195Hz	50Hz	110Hz (Low to Intermediate)	105Hz

When CHTs were perceptually downgraded to intermediate proficiency, the pitch range turned out to be zero. On the other hand, CLTs were judged as improved accent (intermediate proficiency), when the pitch range increased to be 137Hz as the peak of H* was manipulated as high as 295Hz. Furthermore, in yes-no questions, the accent ratings of CHTs worsened to be intermediate proficiency when the valley of L* synthetically changed to as high as 270Hz and the pitch range was zero. CLTs were also perceived as better accent (intermediate) when the L* dropped down to 110Hz and the pitch range was as wide as 105Hz.

The absolute values of H* which incurred a categorical change of proficiency might as well be reanalyzed from the perspective of pitch ranges in Tables 1 and 2. The foreign accent of CHTs was readily stronger, being judged as the intermediate level when their H* dropped to 150Hz. This was actually the same pitch as the bottom line of the tonal contour; therefore, it should be restated in such a way that high talkers can be possibly under estimated as intermediate when the pitch range is almost 0Hz. Similarly, CHTs' accentedness changed to intermediate proficiency when the pitch range was 0Hz in the case of yes-no questions. This indicates that high proficiency talkers will be categorically degraded to intermediate when their L2 speech is almost flat and monotonous. This was shown in Figures 1 and 3 that Chinese high talkers' accent became stronger, i.e., the scores went higher, as the peak of H* decreased and the valley of L* increased, i.e., pitch ranges became smaller. The pitch range phonologically salient enough to provoke a categorical change of native listeners' perception in proficiency turned out to be almost flat. Such a zero pitch range does not seem to carry a prosodic function, but the reason why they are still assessed as intermediate proficiency might be ascribed to high talkers' interlanguage which has relatively similar segmental and other prosodic features to native L1 speakers

Low proficiency talkers, on the other hand, made their accent categorically better, intermediate, when the pitch range increased from 52Hz to 137Hz in the production of English statement and from 80Hz to 105Hz in yes-no questions. Recall that the pitch range was almost zero when high frequency talkers were degraded to intermediate, but low talkers were required to have 137Hz and 105Hz when they were upgraded to intermediate. High and low talkers were perceived as the same intermediate proficiency when their pitch ranges were modified, but the huge ranges, in the case of low talkers, plausibly had to compensate for considerably deviant segments and other prosodies like speech rate, rhythm, etc. This suggests that pitch alone can successfully contribute to altering the perception of foreign accent.

Turing to the results of duration-manipulated speech, recall Figures 9 to 12. The rating scores of CHTs showed a curvelinear relation with the speech rates in both statements and yes-no questions, but in the case of CLTs, speech rate did not seem to affect the perception of foreign accent. The details are summarized in Tables 3 and 4.

Table 3. A summary of the results of duration-manipulated speech (statements)

	Original (s/s)	Synthesized (s/s)	Optimal (s/s)
CHT	5.2	4 > x or x < 7.8 (x=intermediate)	6.5
CLT	4.4	No change	

Table 4. A summary of the results of duration-manipulated speech (yes-no questions)

	Original (s/s)	Synthesized (s/s)	Optimal (s/s)
CHT	4.2	3 > x or x < 8.5 (x=intermediate)	6.0
CLT	4.4	No change	

As presented in Table 3, the average speech rate was 5.2 s/s in the case of CHTs' statements, but the curvelinear prediction line showed that their foreign accent was perceptually deteriorated to be higher than 4 when they speak slower than 4 syllables per second (s/s) or faster than 7.8 s/s. Even high talkers' accent improved, reaching the lowest accent score, when the speed was manipulated to be 6.5 s/s. We call it optimal speech rate for CHTs, and it was slightly faster than their original speed. However, CLTs showed little change of scores along with the increase/decrease of speech rates. Presumably the changes of speech rate might be perceptually masked by CLTs' strong accent in segments as well as other prosodic features such as pitch. The same thing happened to yes-no questions as shown in Table 4. High talkers, on the other hand, showed their proficiency change to intermediate when the speech rates were lower than 3 s/s and higher than 8.5 s/s, and the lowest accent score occurred at 6.0 s/s. Again this is quite faster than their original speech rate 4.2 s/s. This is, therefore, interpreted as indicating that native listeners prefer a little faster speech for high proficiency L2 talkers. On the other hand, they did not respond to the changes of speech rate when they listened to low talkers. Strong accent, that is, considerably dissimilar phonological/phonetic features of low talkers might conceal natives' perceptual sensitivity to the changes in speech rate.

The correlations of native listeners' ratings were solely significant with Chinese listeners (both HLs and CLLs) in pitch-manipulated statement stimuli. Korean listeners, whether high or low proficient, did not show significant correlations with native listeners. However, in the perception of pitch-manipulated stimuli of Chinese talkers' yes-no questions, none of Chinese or Korean listener groups showed significant correlations with natives. Chinese listeners rated Chinese talkers' speech in a similar manner to native listeners for pitch-manipulated statement stimuli, but such a pattern was not found in the perception of yes-no question stimuli. This indicates that listeners' L1 background, rather than L2 proficiency, played a role in the statement stimuli, but not in the yes-no question stimuli.

In consideration of the effect of listeners' L1 on the rating agreement between native and nonnative Chinese listeners, we may need to compare the pitch-related prosodic systems among the listeners' L1, English, Chinese and Korean. Both Chinese and English listeners have their L1 where pitch plays an important role (Xue and Lee, 2014). While Chinese has lexical tones which are phonologically distinctive and faithfully realized in the surface intonation, English has lexical stress which is associated with pitch prominence in the surface intonation. Due to the similar features of pitch-related prosody between Chinese and English, both listener groups might be equally sensitive to the pitch modifications of statements. On the other hand, Korean listeners' L1 prosodic system does not involve a critical role of pitch equivalent to stress-associated pitch accents in English or lexical tone-associated pitch contours in Chinese, but merely concerns the surface intonational configurations which are mostly related with semantic and/or syntactic structures (Xue and Lee 2014). Therefore, Korean listeners did not seem to be as sensitive to the changes in pitch as native or Chinese listeners.

What might, then, cause no correlations between native and Chinese listeners in the perception of pitch-manipulated yes-no question stimuli. Chinese has four distinctive lexical tones which are interactive in the end of questions in complex manners. The rising pitch structurally induced by questions may not be successfully or faithfully realized when the lexical tone of the last word is sustained (tone 1) or falling (tone 4). The yes-no questions, whether the height of L* was synthesized to go up and down, all had a rising edge tone in the current experiment. Chinese listeners might not, therefore, correctly respond to the pitch changes of L* followed by the consistently rising pitch at the end. According to Xue and Lee (2014), correlations of listeners' ratings seems to be closely affiliated with how effectively the particular prosodic parameter under control functions in the listeners' L1.

The correlation results should be further discussed in terms of nonnative listeners' L2 proficiency. Flege (1988) and Lee and Xue (2013) asserted that nonnative listeners with much L2 experience or high L2 proficiency showed a significant agreement with native listeners in rating foreign accent of L2. This was, however, not found in the current experiment; Korean high proficiency listeners did not consistently agree with Chinese high listeners in the perception of pitch-manipulated English speech. It could be speculated that the high proficiency Chinese and Korean listeners who participated in the current study might not be as proficient as the experienced nonnative listeners in Flege (1988) and the high proficient listeners in Lee and Xue (2013). The methodology that was used to determine talkers' proficiency was basically the same as in Lee and Xue (2013), but the academic institution where Korean listeners and some of Chinese listeners were recruited was different from the one in Lee and Xue. Universities in Korea are ranked according to students' scholastic ability/scores upon their entrance. Scholastic ability, though it may not be consistently and/or directly correlated with English proficiency, is assumed to influence their English proficiency to some degree. The institution where Korean listeners were recruited in the current work has been ranked a bit lower than that of Lee and Xue.

Moreover, a majority of Chinese potential listeners (24 out of 35) were recruited at a university in China rather than in Korea, but only 3 of them were rated as high proficiency. This indicates that they were considerably lower in English proficiency than the Chinese exchange students recruited in a university in Korea. Overall the listener pool in this study might be presumably less proficient. Therefore, the Korean listeners who were rated as high proficiency might be slightly lower because natives' ratings were executed within the pool.[6] In other words, the high proficient listeners in this experiment might not be as high as those in Lee and Xue and did not have their interlanguages similar enough to natives. If they had been sufficiently proficient to access phonological/phonetic characteristics similar to native listeners, they would have agreed. There seem to be various and/or different backgrounds of pools to choose from when determining

6 This kind of sampling issue has been raised in the studies of L2 speech, more specifically in those particularly sensitive to L2 proficiency. Scores of standardized English tests like TOEFL or TOEIC were once popularly used in L2 speech, but has been substituted by an Accentedness Rating task due to the fact that those tests do not directly reflect English proficiency associated with speaking. However, proficiency ratings are carried out within a sample in the AR task; therefore, overall levels of samples may vary depending on where and how they are collected.

nonnatives' proficiency; therefore, proficiency levels are not comparable between studies.

Continuing to discuss the correlations between native and nonnative listeners' ratings for the duration-manipulated stimuli, no groups of nonnative listeners showed their ratings significantly correlated with those of natives. Native listeners were sensitive enough to judge too fast or too slow speech as unfavorable and a little faster than their original speech as optimal, but such dynamic responses were not found in either Chinese or Korean ratings. They might not have their own favorable L2 speech rates of English that they judge as least accented. They are not as sensitive as native listeners to the changes in speech rate just because L2 is not as familiar.

The results in the current study can be extensively interpreted in the pedagogical perspectives. Low talkers are able to successfully amend their proficiency better with increasing the pitch of H* and decreasing L*, more specifically expanding the pitch ranges drastically from about 50Hz to 105 ~ 140Hz. If they increase their pitch ranges two to three times even with no attempt to correct other segmental and suprasegmental factors, their overall accent will readily improve to be judged as intermediate. This will shed light on the importance of explicit instruction of intonation when teaching English speaking. It seems to be necessary to design a promising methodology about how to teach low proficient learners to reduce foreign accent through expanding their pitch ranges two to three times.

High talkers who were rated as scores between 1 and 3 in the Accentedness Task could be perceptually degraded to intermediate when their pitch contours were almost flat or monotonous. Therefore, it seems to be very important to maintain their pitch ranges. In addition, they can possibly hear almost native-like if they expand their pitch ranges to about 220Hz and higher in both statement and yes-no question (see Figures 1 and 3). They were rated as the scores of 1's (under 2) when the pitch ranges were wider than 220Hz. Since their original pitch ranges were 70 ~ 80Hz, they should increase the overall range of intonation almost three times to satisfy native listeners' expectation of native-like accent if everything else is kept intact.

Concerning the results of speech rate-manipulated stimuli, the low proficient talkers unfortunately did not acquire better ratings even if speech rate increased or decreased. Native listeners showed no response to the temporal changes of low talkers' L2 speech, which might be ascribed to the perceptual mask of strongly accented segments over prosodic ameliorations of speech rate. Native listeners of English seem to be more sensitive to the changes in pitch than speech rate in the assessment of the foreign accent of low proficiency talkers if and only if the same segmental information is given. Therefore, pitch ranges should take priority over segments, and segments should take priority over speech rate if low proficiency talkers are trained to reduce foreign accent in the pedagogical environment.

High proficiency talkers were, on the other hand, rated as the least accented or optimal when the speech rate was 6.0 ~ 6.5 s/s. Since their average rates were 5.2 and 4.2 s/s respectively in statements and yes-no questions, native listeners preferred high talkers' speech 1.2 ~ 1.5 times faster. Unlike low proficiency talkers, both pitch range and speech rate turned out to be contributing factors to better accent.

3. Conclusion

The current study has investigated native English listeners' perception of pitch- and duration-manipulated speech of Chinese L2 talkers and compared it with that of nonnative Chinese and Korean listeners. It has attempted to find out actual numeric values of pitch and speech rate where Chinese high and low proficiency talkers change in the category of English proficiency in the experiment. Low talkers improved to enter the category of intermediate in native listeners' perception of foreign accent when the pitch range expanded two to three times, i.e., from about 50Hz to 105 ~ 140Hz. High talkers were perceived as having almost native-like accent when their pitch range expanded almost most than three times from 70~80Hz to 220Hz and higher. Speech rate, on the other hand, made a different aspect of contribution to foreign accent. Low talkers' speech, when synthesized in speech rate, did not improve or worsen in perceived accent, but high talkers' speech showed a curvelinear pattern of ratings. That is, the rating scores went up (or foreign accent became stronger) when the speech was too fast or slow. The optimal speech rate, where high talkers were perceived as the least accented, emerged 6 ~ 6.5 s/s. It was 1.2 to 1.5 times faster than high talkers' original speech.

The results suggest a few pedagogical implications which can readily apply to teaching English. Prosodic factors such as pitch range or speech rate are sufficient attributes to change the perceptual category of L2 talkers' proficiency (except for speech rate for low talkers); therefore, it would be more effective to focus on pitch range or speech rate to reduce foreign accent of L2 talkers. Chinese L2 talkers need to expand their pitch ranges twice or three times for their optimal accentedness while their speech rate should accelerate 1.2 to 1.5 times faster in the case of high talkers.

References

Anderson-Hsieh, Janet, Ruth Johnson, and Kenneth Koehler. 1992. The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody and syllable structure. *Language Learning* 42(4): 529-555.

Boula de Mareuil, Phillipe and Bianca Vieru-Dimulescu. 2006. The contribution of prosody to the perception of foreign accent. *Phonetica* 63: 247-267.

Brazil, David, Malcom Coulthard, and Catherine Johns. 1980. *Discourse intonation and language teaching*. London: Longman Group.

Brown, Adam. 1995. Minimal pairs: Minimal importance? *ELT Journal* 49(2): 169-175.

Chen, Hsueh Chu. 2011. Judgments of intelligibility and foreign accent by listeners of different language backgrounds. *The Journal of Asia Tefl* 8: 61-83.

Chela-Flores, Bertha. 1998. *Teaching English rhythm: From theory to practice*. Caracas, Venezuela: Fondo Editorial Tropykos.

Clennell, Charles. 1996. Promoting the role of English prosody in a discourse-based approach to oral interaction. *Prospect* 11(3): 17-28.

Flege, James E.1988. Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America* 84(1): 70-79.

Flege, James E. 1991. Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in second language. *Journal of the Acoustical Society of America* 89(1): 395-411.

Flege, James E., Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of experience of non-native speakers' production and perception of English vowels, *Journal of Phonetics* 25(4): 437-470.

Flege, James E, Murray Munro, and Ian MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97: 3125-3134.

Flege, James E., Naoyuki Takagi, and Virginia Mann. 1995. Japanese adults can learn to produce English /r/ and /l/ and accurately. L*anguage and Speech* 38(1): 25-55.

Flege, James E., Naoyuki Takagi, and Virginia Mann. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of the Acoustical Society of America* 99(2): 1161-1173.

Hahn, Laura. 2004. Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly* 38: 201-223.

Jacobs, Roderick A. and Peter S. Rosenbaum (eds.). 1970. *Readings in English Transformnational Grammar*. Waltham, MA: Ginn and Company.

Jilka, Matthias. 2000. *The contribution of intonation to 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.

Kuhl, Patricia K. 1991. Human adults, and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics* 50(2): 93-107.

Lee, Joo-Kyeong. 2014. The role of prosody in the perception of foreign accent and comprehensibility: Prosody-corrected-L2 speech vs. prosody-distorte-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.

Lee, Joo-Kyeong and Xing Liu. 2015. The perception of foreign accent in prosody-transplanted speech. *Studies in Phonetics, Morphology, and Phonology* 21(2): 323-340.

Lee. Joo-Kyeong and Xiaojiao Xue. 2013. The effects of non-native listeners’ L1 backgrounds and L2 proficiency on the evaluation of foreign accent. *Korean Journal of Linguistics* 38(4): 971-996.

Liu, Xing and Joo-Kyeong Lee. 2012. The contribution of prosody to the foreign accent of Chinese talkers’ English speech. *Journal of the Korean Society of Speech Sciences* 4(3): 59-73.

McNerney, Mellisa and David Mendelsohn. 1992. Suprasegmentals in the pronunciation class: Setting priorities. In Peter Avery and Susan Ehrlich (eds.), *Teaching American English pronunciation*, 185-196. Oxford: Oxford University Press.

Morley, Joan. 1991. The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly* 25(3): 481-520.

Munro, Murray. 1995. Non-segmental factors in foreign accent: ratings of filtered speech. *Studies in Second Language Acquisition* 17(1): 17-34.

Munro, Murray, Tracey Derwing, and Susan Morton. 2006. The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition* 28(1): 111-131.

Munro, Murray, James E. Flege, and Ian MacKay. 1996. The effects of age of second-language learning on the production of English vowels. *Applied Psycholinguistics* 17: 313-334.

Nida, Eugene. 1957. *Learning a foreign language*. Ann Arbor: Friendship

Pallier, Christophe, Laura Bosch, and Núria Sebastián-Gallés. 1997. A limit on behavioral plasticity in speech perception. *Cognition* 64(3): 9-17.

Pennington, Martha and Jack Richards. 1986. Pronunciation revisited. *TESOL Quarterly* 20(2): 207-225.

Pierrehumbert, Janet and Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip Cohen, Jerry Morgan, and Martha Pollack (eds.), *Intentions in Communication*, 271-311. Cambridge, MA: The MIT Press.

Takagi, Naoyuki and Virginia Mann. 1995. The limits of extended naturalistic exposure on the perceptual mastery of English /r/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics* 16: 379-405.

Trofimovich, Pavel and Wendy Baker. 2007. Learning prosody and fluency characteristics of second language speech: the effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics* 28(2): 251-276.

Xue, Xiaojiao and Joo-Kyeong Lee. 2014. The effects of non-native listeners' L1 background and L2 proficiency on the perception of foreign accent for pitch-manipulated native and Chinese accented English speech. *Linguistic Research* 31(2): 275-303.

Xue, Xiaojiao and Joo-Kyeong Lee. 2015. Native and non-native listeners' perception of foreign accent for duration-manipuated English L2 speech. *Language and Linguistics* 66: 509-537.

Wester, Mirjam and Cassie Mayo. 2014. Accent rating by native and nonnative listeners. In Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 7749-7753

Appendix I: A passage for the Foreign Accented (FA) Task

When Frank was young, his job was to repair bicycles and at that time he used to work fourteen hours a day. He saved money for year. And in 1958 he bought a small workshop of his own. In a few years the small workshop had become a large factory which employed seven hundred and twenty-eight people. Frank smile when he remembered his hard early years and the long road to success.

Appendix II: Sentences for pitch- and speech rate-manipulations
(Boldfaced sentences were the target sentences)

(A) Statements

Q: How many apples did you buy?

A: **I bought eleven apples.**

Because I don't like coffee, **I ordered lemonade.**

My uncle isn't a teacher, but **he's a lawyer.**

(B) Yes-no questions

Do you need an orange?

Are you married?

Is it raining?

Joo-Kyeong Lee
Department of English Language and Literature
University of Seoul
163 Seoulsiripdaero, Dongdaemun-gu, Seoul 02504, Korea
E-mail: jookyeong@uos.ac.kr

Received: 2016. 07. 25.
Revised: 2016. 11. 01.
Accepted: 2016. 11. 01.