*Linguistic Research* 35(1), 203-232
DOI: 10.17250/khisli.35.1.201803.008

# Prosodic remedies of Korean talkers' English L2 speech: Optimal pitch and speech rate*

**Joo-Kyeong Lee**
**(University of Seoul)**

**Lee, Joo-Kyeong. 2018. Prosodic remedies of Korean L2 speech: Optimal pitch and speech rate.** *Linguistic Research* 35(1): 203-232. The current study investigates native English listeners' perceived foreign accent for pitch- and duration-manipulated speech of Korean EFL talkers with high and low proficiency. This work contents that low proficiency talkers' L2 speech can be perceptually better accented when a single prosodic parameter such as pitch or speech rate is merely corrected. As nonnative talkers' proficiency was divided into three categories, high, intermediate and low in the Accentedless Rating task, low talkers were categorically promoted to the intermediate when H* and L* were remedied more native-like and their L2 speech was synthesized to be faster. The corrected prosodic features seemed to be readily detected by native listeners because such corrections might be comparatively salient over the segmental features seriously deviant from native norms in their interlanguage. The results of the current experiment suggested actual numeric values of optimal pitch and speech rate for upgrading low talkers' proficiency; H* should increase roughly by 30% to 45% higher than the ones that they usually produce in statements, and L* should deepen approximately by 20% to 40% lower than in their normal production of yes-no questions. Speech rate should be 1.2 to 2 times faster for low talkers to be judged as intermediate. On the other hand, Korean high proficiency talkers didn't show a categorical decline to intermediate when pitch or speech rate was synthetically deteriorated. Due to their little accented L2 speech, phonological/phonetic features, which are very similar to those of native speakers, seemed to firmly tolerate the degrading portion of prosody. In addition, the actual numeric values of pitch and speech rate obtained in the results should be applied to the pedagogical environment and used as references to facilitate improving low talkers' proficiency. **(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

Many studies on L2 speech have attempted to technically modify segmental and/or prosodic features in nonnative talkers' production of second or foreign language to see how native listeners' perceived accentedness changes (de Mareuiil and Vieru-Dimulescu 2006; Lee 2014, 2016; Lee and Liu 2012, 2015; Magen 1998; Rekart et al. 1985; Winters and O'Brien, 2013 among others). Most recent and advanced methodologies have enabled a particular parameter of L2 speech to be altered so that listeners' identification can be interpreted as the outputs in association with using a unidimensional change. They have technical supremacy over low-pass filtered or flat intonation skills which lose critical linguistic information; segmental cues are cut-off in the low-pass filtered speech, and the overall tone is completely monotonous at a certain pitch in the flat intonation speech (Jilka 2000; Munro 1995; Trofimovich and Baker 2007). The output stimuli should, therefore, sound like unnatural or unhuman speech, which has been criticized for inducing inaccurate perception of naive listeners.

What seems to be important in manipulating L1/L2 stimuli is speech naturalness; that is, the synthesized stimuli should be as close as original natural speech where individual phonetic characteristics are kept intact with the target feature(s) only modified. Recently much attention have been paid to prosodic features such as intonation or speech rate in L2 speech studies as they have been appreciated as much contribution to native listeners' detection of foreign accent as segments. Recently, an algorithm called 'TD-PSOLA (time domain pitch synchronous overlap and add)' has drawn much attention. A prosodic parameter such as pitch or duration is translpanted between two different talkers' speech while other linguistic information is maintained (de Mareuiil and Vieru-Dimulescu 2006; Lee 2014; Lee and Liu 2012, 2015; Major 1987; Magen 1998; Moulines and Charpentier, 1990; Rekart 1985; Winters and O'Brien, 2013).

Prosodic features were also manipulated with their values increasing or decreasing while other segmental or prosodic information entirely remains the same (Xue and Lee 2014, 2015; Lee 2016). For instance, the pitch value of a particular accented syllable is synthesized to rise or fall, and speech rate is also holistically synthesized to accelerate or slow down to see the degree of listeners' sensitivity to the changes of the target prosody with invoking no naturalness problem. Both

transplantation between L1 and L2 speech or single feature alteration of L2 speech are highly advanced technologies, for which a separate and independent role of a segmental or prosodic parameter has been successfully investigated in the perception of foreign accent.

Prosody has been shown to play a prevalent or at least an equal role in an ample amount of empirical work of L2 speech (Anderson-Hsieh, Johnson and Kohler 1992; Flege, Munro and MacKay 1995; Munro 1995; Magen 1998; Jilka 2000; de Mareuiil and Vieru-Dimulescu 2006; Trofimovich and Baker 2007; Lee 2014, 2016; Polyanskaya, Ordin and Crazia, 2016). Among those are two different claims, pitch or intonation is a more important attribution to the perception of foreign accent than other prosodic parameters (Van Els and DeBot 1987; Munro 1995; Magen 1998; Jilka 2000; de Mareuiil and Vieru-Dimulescu 2006; Lee and Liu 2012) and duration or speech rate outweighs intonation in the judgment of foreign accent (Trofimovich and Baker 2007; Liu and Lee 2012; Lee 2014; Lee and Liu, 2015). Due to lack of consistency between intonation and speech rate, the current study manipulated both pitch and duration separately in Korean L2 speech, investigating its optimal values for high and low proficiency Korean talkers.

The transplantation skill of segments and prosody between two talkers is an application of the TD-PSOLA algorithm and switches the temporal and tonal components of prosody between two talkers as mentioned earlier. De Mareuiil and Vieru-Dimulescu (2006) was one of the earliest studies that adopted the transplantation methodology in speech. They recorded almost identical sentences from Spanish and Italian L1 talkers, taking advantage of the phonological/phonetic similarities between the two languages. They swapped pitch and duration separately with segments between Italian and Spanish talkers. The synthesized speech was mostly identified as the language which was the source of the prosody. De Mareuiil and Vieru-Dimulescu, therefore, asserted that prosody outweighed segments in language identification, but this might be attributed to the mostly indistinguishable segmental patterns between Italian and Spanish as already criticized in Lee and Liu (2012). The role of segments might be comparatively insignificant, which resulted in more contribution of prosody. Although their work was not a study of L2 speech, it shed light on a new and renovative methodology in relevant areas, enabling L2 stimuli to be synthesized with no loss of any linguistic elements.

The TD-PSOLA algorithm was applied to L2 speech and synthesis in Lee and

Liu (2012). They recorded the same Chinese sentences from Chinese L1 and Korean L2 talkers and transplanted Chinese talkers' prosody onto Korean talkers' segment and Korean talkers' prosody onto Chinese talkers segment. Chinese L1 listeners participated in judging foreign accent of the stimuli; they perceived the stimuli of Korean talkers' segment and Chinese talker's prosody as less accented. This was interpreted as the fact that segments played a more influential role in the perception of Korean accented Chinese L2 speech. Moreover, when the prosodic parameters, duration, pitch, and intensity, were individually transplanted onto each other, pitch was the most significant factor. Korean talkers' deviant production of intonation rendered Chinese listeners' detection of foreign accent.

Liu and Lee (2012) extended Lee and Liu (2012)'s study to English L1 and Chinese L2 talkers' speech. Native English listeners rated the combination of Chinese talkers' segments and English talkers prosody as more accented than that of English talkers' segments and Chinese talkers' prosody. This was similar to Lee and Liu (2012) in that segments contributed more to the native listeners' perception of foreign accent. However, in the analysis of a separate role of each prosodic parameter, duration and pitch, duration was comparatively more attributable to perceived accentedness. Due to the fact that the duration of an entire sentence was swapped, the duration-manipulated stimuli were reorganized in speech rate. That is, native English listeners were more sensitive to foreign deviances of speech rate than those of intonation.

Similarly to Lee and Liu (2012) and Liu and Lee (2012), Lee (2014) also adopted the TD-PSOLA algorithm in her study, comparing prosody-corrected L2 Korean talkers' production of English (Korean talker's segments + English talkers' prosody) with that of prosody-distorted L1 English talker's production of English (English talkers' segments + Korean talkers' prosody). The prosodically synthesized L1 and L2 speech did not show statistical differences from original L1 and L2 speech in accent ratings. This means that segments played a dominant role over prosody, consistently enough to the previous studies (Lee and Liu 2012; Liu and Lee 2012). Furthermore, foreign accent was rated statistically stronger when L1 speech was distorted by L2 duration, speech rate per se, but intonation did not make any significant differences. Speech rate as opposed to intonation was found to contribute more to English L1 listener's perception of Korean accented English speech than intonation, which is similar to Liu and Lee (2012).

Lee and Liu (2015) extensively investigated transplanted speech between L1 and L2 talkers, yet synthesizing Korean sentences produced by Korean L1 and Chinese L2 talkers. In the results of individual/independent roles of prosodic parameters such as duration, pitch and intensity, duration was found to make a more prominent contribution to Korea listeners' detection of foreign accent. That is, Korean listeners were more detective to talkers' speech rate, showing disfavor to slower speech with foreign-accented segments. It may be plausible that L1 listeners were sensitive to a different prosodic factor of L2 accented speech depending on their L1 (Lee 2014). As discussed above, pitch information, that is, intonation was more influential to Chinese L1 listeners (Lee and Liu 2012). On the other hand, duration or speech rate was more detectable to English or Korean L1 listeners (Liu and Lee 2012; Lee 2014; Lee and Liu 2015). Lee and Liu (2012) examined Korean talkers' speech of Chinese whose prosody and segments were transplanted with those of native Chinese talkers' speech. In their study, L2 was Chinese, and it is a tone language where lexical tones not only determine the overall prosodic pattern at the lexical level but also at the phrasal level (Pennington and Ellis 2000). Once lexical or phrasal tones are distorted in talkers' speech of Chinese, it seems to be readily detected as critical deviances, which results in strong foreign accent. A temporal aspect is more likely to affect native listeners' perceived accentendness than intonation in the languages other than tone-rich one, although the number of languages examined so far does not seem to be sufficient to make such a solid conclusion.

The numerous empirical studies mentioned so far show that both intonation and speech rate are equally important and deserve to be investigated in L2 speech studies. However, very few studies attempted to associate their results to pedagogical methodology or resources. Most of them merely stopped at reporting experimental results even though the results should be sufficiently useful and/or effectively applied to teaching foreign or second language. This sheds light on the emergency of the current work where the optimal measures of both pitch and speech rate are discovered for different levels of talkers' proficiency and will readily serve as resources and references in pedagogical environments.

Another technology which has been long utilized in speech synthesis is to straightforwardly modify a particular acoustic cue and interpret listeners' responses as perceptual sensitivity to the changes of the cue. When applied to the studies of L2, this methodology has been recently selected in Xue and Lee (2014, 2015) to see

what might happen to the degree of foreign accent if pitch or duration of talkers' speech is manipulated to increase or decrease while other individual characteristics of the speech are kept constant? Recently Lee (2016) has presented the optimal values of prosody for Chinese talkers of English with high and low proficiency. She manipulated pitch and speech rate of Chinese L2 speech of English to seek for the optimal measures where native listeners judged it as the least accented for Chinese talkers with different proficiency. More specifically, when low proficiency talkers were perceptually identified as a better accent category of 'intermediate' at a certain synthesized pitch or speech rate, such prosodic measures were assumed to optimally promote talkers' proficiency. Chinese talkers could improve their L2 accent with a sole change of either pitch or speech rate, which suggests that prosody made a significant contribution to perceived foreign accent. The current study attempts to extend Lee (2016)'s study to Korean L2 talkers of English, exploring their optimal prosodic values for low and high proficiency talkers. It further compares them with the actual values of prosody optimized for Chinese talkers in an attempt to provide a substantial explanation about the differences/similarities based on prosodic systems and structures between Korean and Chinese.

## 2. Experiment

### 2.1 Talkers and listeners

For the speech stimuli which served as inputs to the perception of foreign accent, three low and three high proficiency Korean talkers participated. Five different listener groups (10 native English, 10 Chinese low & 10 high proficiency listeners and 10 Korean low proficiency and 10 high proficiency listeners) listened to pitch- or duration-synthesized speech of Korean talkers and judged a degree of foreign accent. Prior to the perception experiment, an Accentedness Rating (AR) task was carried out to determine Chinese and Korean participants' proficiency as specified in Munro (1998).

To organize 13 high and 13 low proficiency Korean participants (6 talkers and 20 listeners), 85 college students attending a Korean university in Seoul were recruited and asked to read the English passage used in Xue and Lee (2014).[1] Their

recordings were submitted to three native English listeners for accentedness ratings. The native listeners judged it on a 9-point Likert 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. Thirty four participants out of 85 were turned out to belong to low proficiency, 27 were judged as high proficiency, and the remaining 24 participants were categorized into intermediate proficiency. Thirteen were randomly chosen from the 27 high proficiency participants and served as high proficiency talkers and listeners, and thirteen were randomly selected from the 34 low proficiency participants and took part in the experiment as low proficiency talkers and listeners. The remaining 24 Korean speakers who were presumably categorized into intermediate proficiency (scores 4, 5 and 6) did not participate because the current experiment only focuses on high and low listeners to maximize the difference in proficiency.

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 Korean participants read 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 participants who were assessed as intermediate did not take part in the experiment as listeners.

As for the native English listeners, they were either instructors or exchange students in a college in Seoul, Korea. They were either American or Canadian, and their length of residence in Korea varied. The three native raters of the AR task did not participate in the later perception experiment as listeners.

## 2.2 Stimuli

The speech stimuli were prepared from the recordings of three high proficiency

1 See Appendix I.

Korean talkers and three low proficiency Korean talkers. They were asked to read three English statements and three yes-no questions.[2] The statements contained only one high pitch accent (H*) because a single narrow focus was intentionally induced in dialogues. The yes-no questions were also designed to have only one low pitch accent (L*) assigned to the last content word. This was due to the synthesis process that the peaks of high pitch accents and the valleys of low pitch accents were manipulated to increase or decrease. 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, 18 statement sentences were synthesized in Praat; the pitch value of a high pitch accent (H*) in each statement were increased in 5 steps and decreased in 5 steps respectively. The synthesis intervals were 20Hz; therefore, the peaks of high pitch accents incremented up to +100Hz, and lowered down to -100Hz. Ten sentences with different slopes of a high pitch accent were generated from one statement. Eleven audio stimuli including one original natural statement were produced from one statement. Therefore, 198 pitch-manipulated statement stimuli (11 stimuli * 3 statements * 6 talkers) were provided to one listener for foreign accent judgment. The pitch manipulation of yes-no questions were exactly the same as that of statements. The valley of the low pitch accent (L*) in each question was synthesized to increase up to +100Hz and decrease down to -100Hz in 20Hz intervals. Ten synthesized stimuli with different slopes of a low pitch accent and one original recording speech were produced from one yes-no question. Consequently, 198 pitch-manipulated yes-no question stimuli (11 stimuli * 3 questions * 6 talkers) were served to one listener. In total, 396 pitch-manipulated statement and yes-no questions audio stimuli were randomized and presented to the listeners for accentedness ratings.

For duration manipulations, 36 sentences (3 statements * 6 talkers + 3 yes-no questions * 6 talkers) were synthesized to expand twice and compress a half in the temporal dimension of sentences. 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. Duration expansion or compression operated over the entire length of

2 We adopted the same sentences as in Lee (2016). See Appendix II.

sentences, and consequently altered speech rate; the expansion of sentence duration resulted in slower speech, and its compression gave rise to faster speech. In total, 398 duration-manipulated audio stimuli (11 stimuli * 3 statements * 6 talkers + 11 stimuli * 3 yes-no questions * 6 talkers) were randomized and presented to listeners for the perception of foreign accent.

## 2.3 Procedure

The listeners visited the speech lab twice in the schools of Korea or China; they carried out the experiment of pitch-manipulated stimuli in the first visit and that of duration-manipulated stimuli in the other visit. The 11 Chinese listeners (3 high and 8 low proficiency) who were recruited in a college of China participated in their school. They were paid for participation.

Native listeners' responses were averaged at eleven different pitch locations and speech rates (the number of syllables per second, syl/sec) to observe where perceived accentedness changed to another category of L2 proficiency. Native listeners' responses were further submitted to a statistical analysis of Correlation with those of nonnative listeners to see if native and nonnative responses would be consistent and if nonnative listeners' L2 proficiency (high vs. low) and L1 backgrounds (Chinese vs. Korean) have an effect on the correlation results.

## 2.4 Results

### 2.4.1 Pitch and duration manipulations

Figure 1 and Table 1 show native listeners' average ratings corresponding to eleven different pitch values of H* in statement stimuli. The pitch value 246Hz in the center (grey-colored box) is the average pitch from original productions, and the foreign accent score is 1.89. Due to the fact that the talkers were high proficiency, it precisely belongs to the range between 1 to 3. The pitch 246Hz decreases by 20H from the center to the left and increases to the right. Rating scores drastically increase as the pitch of H* drops down. The scores once decrease to 1.66 when the pitch rises up by 20Hz to 266Hz, but they seem to stay steady around 1.56 ~ 1.62 when the pitch goes further higher than 266Hz.[3] That is, Korean talkers of high

proficiency were perceived as stronger foreign accent as the peak of H* becomes lower. Native listeners' perceived accentedness goes beyond the rating score 3 (precisely, 3.11) at 166Hz and reaches 3.5 at 146Hz.[4] This indicates that high talkers started to be out of the high proficiency category when the peak of H* was as low as 166Hz. They were not judged to be necessarily high proficiency any more at 166Hz and 146Hz.

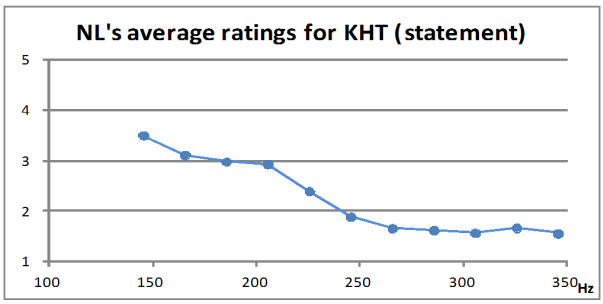


Figure 1. Native listeners' (NL) average ratings of foreign accent (FA) for Korean high talkers' (KHT) statements (pitch-manipulated stimuli)

Table 1. NL's average FA scores for KHT (pitch-manipulated statements)

| pitch (Hz) | 146 | 166 | 186 | 206 | 226 | 246 | 266 | 286 | 306 | 326 | 346 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FA | 3.50 | 3.11 | 2.98 | 2.93 | 2.39 | 1.89 | 1.66 | 1.62 | 1.57 | 1.67 | 1.56 |

Figure 2 shows native listeners' average scores of foreign accent at eleven different pitch values of H* in English statement sentences produced by Korean talkers with low proficiency. The score of unsynthesized original stimuli were 7.68 at 217Hz. As compared with high proficiency talkers, the peak of H* was considerably lower. As the pitch increases rightward, rating scores decrease. That is,

3 It is not necessarily important to see which step of pitch increment or decrement shows a statistically significant change in FA. This study focuses on where native listeners' perception shows a categorical change of nonnative talkers' proficiency, seeking for actual numeric values of pitch.

4 Remember that Korean or Chinese talkers' and listeners' proficiency was determined by the AR tasks, and those who were perceived as 1, 2 and 3 were high proficiency, and those who were perceived as 7, 8, and 9 were low proficiency.

native listeners detected a lesser degree of foreign accent when H* was higher even for low proficiency talkers. The accent score dropped lower than 7 and escaped from the category of low proficiency exactly when the H* peak was raised in one single step by 20Hz. Note that it was rated as 6.68 at 277Hz. When the pitch value incremented to 297Hz and 317Hz, the rating scores safely stayed in the intermediate proficiency category (5.68 & 5.94). This suggests that foreign accent of low proficiency talkers can improve when the height of H* is satisfactorily corrected. The strong accent derived from low talkers' sound information other than pitch contour can be possibly masked by prosody correction, which results in better accent. On the other hand, low talkers' accent seemed to be slightly stronger from 7.68 to 8.16 as the peak of H* became gradually lowered in 5 steps by 20Hz to 100Hz. Low proficiency talkers already had strongly deviant segments, making ceiling effects *per se* on the perceived accentedness. That is, prosody deterioration did not worsen their foreign accent as much.

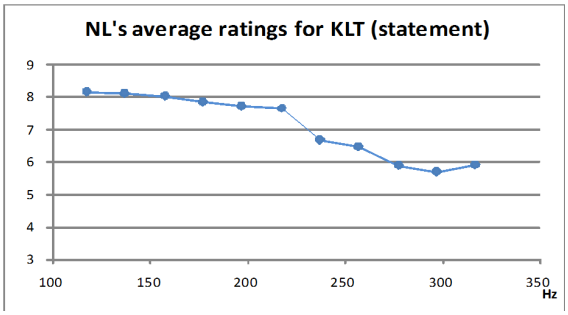


Figure 2. NL's average FA ratings for KLT's statements (pitch-manipulated stimuli)

Table 2. NL's average FA scores for KLT (pitch-manipulated statements)

| pitch (Hz) | 117 | 137 | 157 | 177 | 197 | 217 | 237 | 257 | 277 | 297 | 317 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FA | 8.16 | 8.14 | 8.02 | 7.87 | 7.74 | 7.68 | 6.68 | 6.49 | 5.89 | 5.68 | 5.94 |

Turning our attention to the results of yes-no questions stimuli, Figure 3 and Table 3 exhibit native listeners' average rating scores for high proficiency Korean talkers' stimuli of yes-no questions. Note that the median score, 2.56 at 184Hz, is

from original and unsynthesized stimuli. The valley-like low pitch accent (L*) became flatter as its bottom pitch increased. This incurred listeners' perception of stronger accent as seen in Figure 3: foreign accent increases as pitch increases rightward. When the bottom pitch of L* increased to 244Hz ~ 284Hz, the accent scores went beyond 3 (3.45 ~ 3.72), and high talkers who were rated as 1 ~ 3 did not seem to maintain their own category of high proficiency. On the other hand, the scores did not appear to change much (from 2.56 to 2.34) even though the pitch of L* decreased by 20Hz to 100Hz. The low pitch accents produced by high talkers were presumably low enough, and less deviant (accented) segments seemed to back up native listeners' perception of foreign accent. Similar to the floor effects of low talkers, high talkers' L* lowering did not result in a drastic drop of FA ratings due to floor effects.

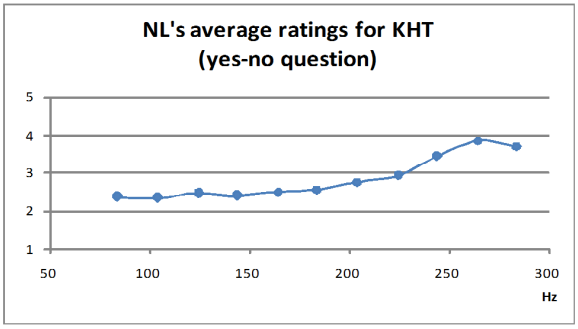


Figure 3. NL's average FA ratings for KHT's yes-no questions (pitch-manipulated stimuli)

Table 3. NL's average FA scores for KHT (pitch-manipulated yes-no questions)

| pitch (Hz) | 84 | 104 | 124 | 144 | 164 | 184 | 204 | 224 | 244 | 264 | 284 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FA | 2.37 | 2.34 | 2.45 | 2.39 | 2.5 | 2.56 | 2.76 | 2.93 | 3.45 | 3.87 | 3.72 |

Figure 4 and Table 4 present native listeners' perception of foreign accent for synthesized stimuli of Korean low talkers' yes-no questions. The median value of pitch is 203Hz, which is L* of original stimuli. The accent score at this pitch point

is 7.23. As the pitch of L* decreases by 20Hz, the accent scores consistently decrease, and finally drop down to the category of intermediate proficiency (5.85) at 143Hz. The scores go even lower to 5.76 and 5.65 at 123Hz and 103Hz and safely belong to the intermediate category. This shows that even low proficiency talkers can be judged as higher proficiency when the valley of L* is synthesized to go lower. That is, strong accent might be perceptually masked by improved tonal contour of L*.

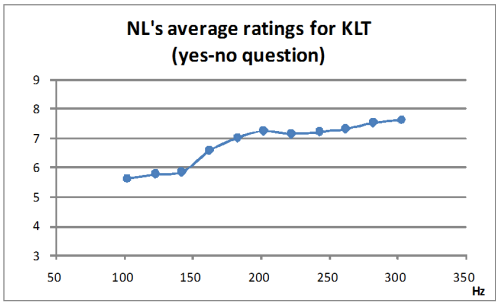


Figure 4. NL's average FA ratings for KLT's yes-no questions (pitch-manipulated stimuli)

Table 4. NL's average FA scores for KLT (pitch-manipulated yes-no questions)

| pitch (Hz) | 103 | 123 | 143 | 163 | 183 | 203 | 223 | 243 | 263 | 283 | 303 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FA | 5.65 | 5.76 | 5.85 | 6.57 | 7.00 | 7.23 | 7.13 | 7.21 | 7.30 | 7.51 | 7.64 |

Next, we move onto the results of duration-manipulated stimuli and their accent ratings from native listeners. Because types of sentences like statement and yes-no question do not matter in terms of duration, accent scores of two types of sentences were integrated and demonstrated together. Results are, therefore, shown only in accordance to talkers' proficiency. As seen in Figure 5 and Table 5, the original stimuli of high proficiency talkers were rated 1.76 at 5.08 syl/sec. Foreign accent was perceived to be stronger as speech rate was slower as the scores increased from the center (original) point to the leftward direction. When speech rate decreased to 1/2 of the original stimuli (i.e., 2.54 syl/sec), the rating score was as high as 3.28.

When the synthesized stimuli was 0.6 times as slow as the orignal stimuli, the score was already higher than 3 and began to be out of the high proficiency category.

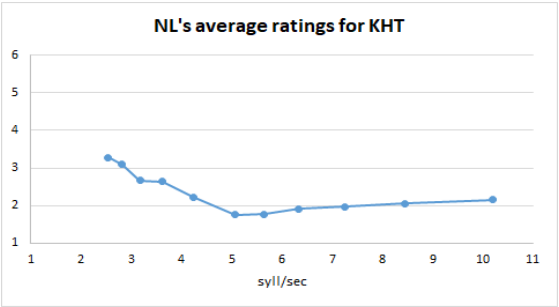


Figure 5. NL's average FA ratings for KHT (duration-manipulated stimuli)

Table 5. NL's average FA scores for KHT (duration-manipulated speech)

| speech rate (syl/sec) | 2.54 | 2.82 | 3.17 | 3.63 | 4.23 | 5.08 | 5.64 | 6.35 | 7.25 | 8.46 | 10.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FA | 3.28 | 3.09 | 2.66 | 2.64 | 2.22 | 1.76 | 1.77 | 1.91 | 1.97 | 2.06 | 2.16 |

What is interesting in Figure 5 and Table 5 is that foreign accent also became stronger as speech rate was faster and reached 2.16 when speech rate was twice as fast (10.02 syl/sec). High proficiency talkers were judged to have stronger accent when they spoke either too slow or too fast even though accent scores of faster speech were not as high as those of slower speech. This is consistent with the results of Chinese talkers (Lee 2016). Both L2 talkers of Korean and Chinese showed a curve-linear shape between speech rate and accentedness scores. Faster speech of Korean high talkers seems to perceptually stay within the scope of high proficiency though accent scores go higher, but the high talkers may start to be out of the boundary of high proficiency when their speech is twice as slow. That is, faster speech did not make as strong accent (high scores) as slower speech. Note that the score 2.16 at 10.2 syl/sec is still within the category of high proficiency. It seems that native listeners expect a certain degree of phonological/phonetic deviances to be associated with an optimal speech rate. If accented speech is too fast or too slow, sound deviances are not consistent with the predicted speech rate, which results in stronger foreign accent. Such a mismatch brings about more serious deterioration of

foreign accent (category change of L2 proficiency to intermediate) when speech rate decreases.

Korean low talkers showed quite different results from high talkers as displayed in Figure 6 and Table 6. Their scores did not show a curvelinear shape unlike high talkers but a inverse relation with speech rate. Low talkers' foreign accent decreased (became weaker) as speech rate increased. The original stimuli were judged as 7.20 at the speed of 3.93 syl/sec. The accent scores went higher as speech rate decreased to the left from the mid-point (7.20 at 3.93 syl/sec). The scores also went lower as speech rate increased to the right, though they sustained at the points where the original stimuli were 0.7, 0.6 and 0.5 times as slow. In addition, the scores started to drop below 7 (6.74) when the speech rate was 1.4 times as fast as the original stimuli (4.87 syl/sec).

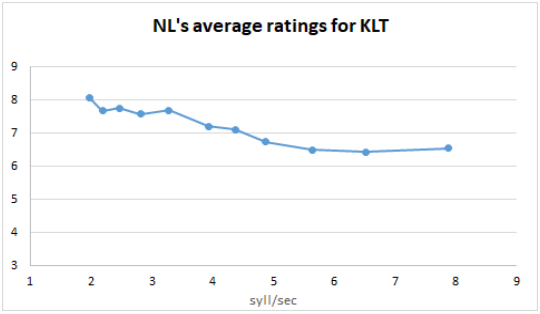


Figure 6. NL's average FA ratings for KLT (duration-manipulated stimuli).

Table 6. NL's average FA scores for KLT

| speech rate (syl/sec) | 1.97 | 2.19 | 2.46 | 2.81 | 3.28 | 3.93 | 4.37 | 4.87 | 5.63 | 6.51 | 7.87 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| FA | 8.07 | 7.67 | 7.75 | 7.57 | 7.69 | 7.20 | 7.11 | 6.74 | 6.50 | 6.44 | 6.55 |

We have seen that native listeners were sufficiently sensitive to changes in pitch and speech rate of Korean L2 speech. Low talkers' English was improved and perceptually escaped out of the low category when the pitch values of high or low pitch accents were merely corrected. Low talkers all promoted to the category of intermediate proficiency when the slope of H* was stiffer with its peak increasing by 40Hz and the slope of L* was stiffer with its valley decreasing by 40Hz. In case of

high talkers, pitch-manipulated speech was perceived as stronger accent than their high proficiency level, but not as strong as the completely intermediate category (scores 4 to 6). Presumably, their interlanguage containing similar segmental features to those of native talkers compensated for the aggravated portion of pitch. Their favorable segments seemed to prevent the pitch-deteriorated speech of high proficiency talkers from going through a categorical change to intermediate or low proficiency.

Duration-manipulated speech was perceived as better accent as speech rate was faster for both high and low proficiency talkers as shown in Figures 5 and 6. High talkers' speech was rated higher than 3.0 when speech rate was 3/5 and 4/5 times as slow as the original speech. It was not, however, perceived as completely intermediate proficiency. Similarly, low talkers' speech was better scored below 7.0 when speech rate was 9/5 and 2 times as fast as the original speech, but again its accent scores did not show a categorical promotion to the intermediate proficiency.

### 2.4.2 Correlations of native and nonnative listeners

Native listeners' ratings of foreign accent were examined in comparison with those of Korean and Chinese listeners. As s get more frequently exposed to communications with other s in English, the accuracy of s' perception of foreign accent seems to be important. Therefore, the current experiment investigated how similarly listeners' judgments of foreign accent patterned with those of native listeners. Different L1 backgrounds between talkers and listeners may have an effect on the judgment of foreign accent; therefore, perception was extended to Chinese listeners. Korean and Chinese listeners' accent scores are presented in terms of how they are correlated with native listeners' scores.

Figure 7 shows the correlation of native listeners' scores with those of four listener groups, Korean high listeners (KHL), Korean low listeners (KLL), Chinese high listeners (CHL), and Chinese low listeners (CLL) for Korean high talkers' (KHT) statement stimuli. According to Pearson-r Correlation statistics, CHL showed a marginal significance with NL ($r = 0.521$, $p = 0.017$). The other three groups of KHL, KLL, and CLL did not show a significant correlation with NL (NL vs. KHL: $r = 0.211$, $p = 0.238$, KLL vs. NL: $r = 0.278$, $p = 0.117$, CLL vs. NL: $r = 0.218$, $p = 0.220$).

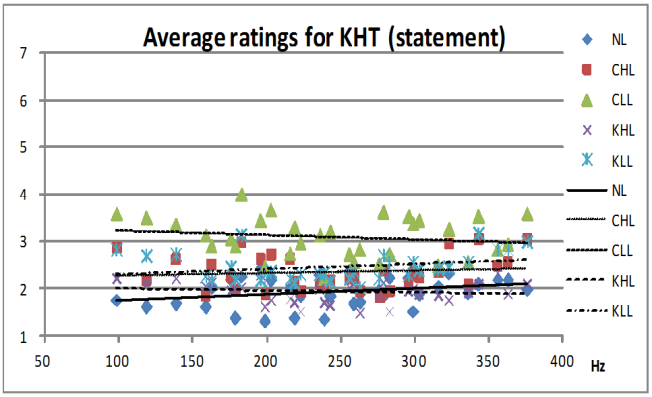


Figure 7. Average rating scores of five listener groups (CHL, CLL, KHL, KLL, and NL) for KHT's pitch-manipulated statements

Figure 8 displays the correlation of NL's scores with those of four groups of listeners for KLT pitch-manipulated statements. Similar to the results of KHT, only CHL showed a marginally significant correlation with NL ($r = 0.430$, $p = 0.012$). The other groups of CLL, KHL and KLL did not show statistically significant correlations with NL (NL vs. CLL: $r = 0.311$, $p = 0.077$, NL vs. KHL: $r = 0.301$, $p = 0.89$, NL vs. KLL: $r = 0.378$, $p = 0.030$).

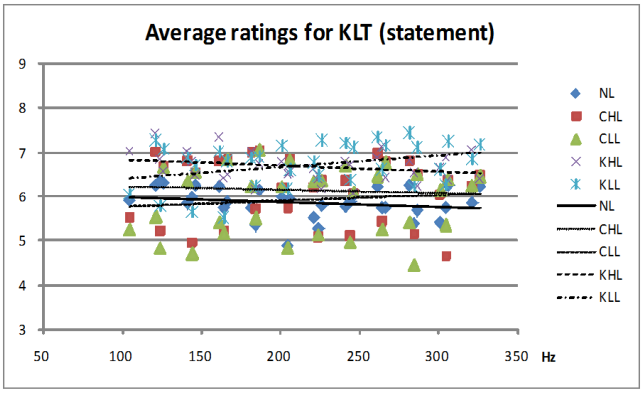


Figure 8. Average rating scores of five listener groups (CHL, CLL, KHL, KLL, and NL) for KLT's pitch-manipulated statements

Figure 9 presents five listener groups' FA responses to the increase/decrease of low pitch accents (L*) in KHT's yes-no questions. Their scores were submitted for the Pearson-r Correlation analysis between native listeners and four listeners each. Results showed that both Chinese high and low listeners' scores were statistically correlated with those of native listeners (NL vs. CHL: $r = 0.774$, NL vs. CLL: 0.692, $p < 0.005$). Korean listeners, regardless of their proficiency, did not show significant correlations with native listeners (NL vs. KHL: $r = 0.264$, $p = 0.0886$, NL vs. KLL: 0.198, $p = 0.319$).

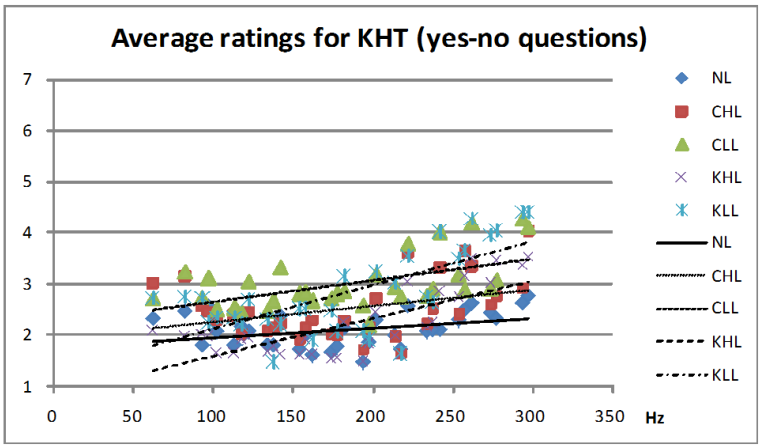


Figure 9. Average rating scores of five listener groups (CHL, CLL, KHL, KLL, and NL) for KHT's pitch-manipulated yes-no questions

Figure 10 displays listeners' ratings scores for KLT's yes-no question stimuli. According the Pearson-r Correlation statistics, Chinese high listeners were the only group that showed a significant correlation with native listeners (NL vs. CHL: $r = 0.629$, $p < 0.005$). The other three groups of listeners (CLL, KHL, and KLL) did not show a statistically similar pattern to native listeners. That is, Chinese listeners with high proficiency of English were sensitive enough to detect the pitch change of L* while the others were not sufficiently accurate to identifying foreign accent associated with pitch.

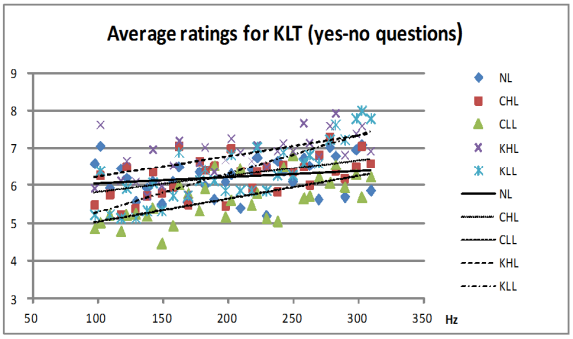


Figure 10. Average rating scores of five listener groups (CHL, CLL, KHL, KLL, and NL) for KLT's pitch-manipulated yes-no questions

Listeners' responses to duration-manipulated stimuli are presented in Figures 11 and 12. Both native and listeners all showed inverse proportion; faster speech was judged as lower foreign accent. As far as the Correlation statistics is concerned, four groups of listeners were all significantly correlated with native listeners when they judged foreign accent (NL vs. CHL: $r = 0.851$, NL vs. CLL: $r = 0.753$, NL vs. KHL: 0.869, NL vs. KLL: 0.829, $p < 0.005$). Similar to the results for KHT, listeners, irrespective of L1 background or L2 proficiency, showed significant correlations with native listeners for KLT (NL vs. CHL: $r = 0.773$, NL vs. CLL: $r = 0.535$, NL vs. KHL: 0.886, NL vs. KLL: 0.875, $p < 0.005$).

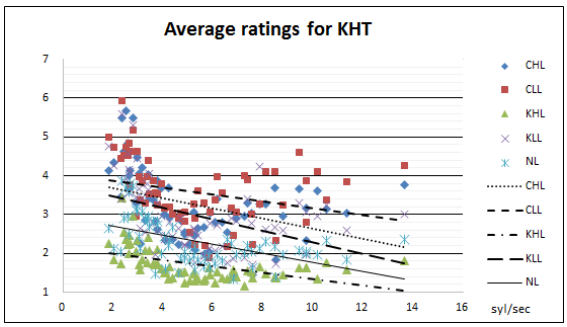


Figure 11. Average rating scores of five listener groups (CHL, CLL, KHL, KLL, and NL) for KHT's duration-manipulated stimuli

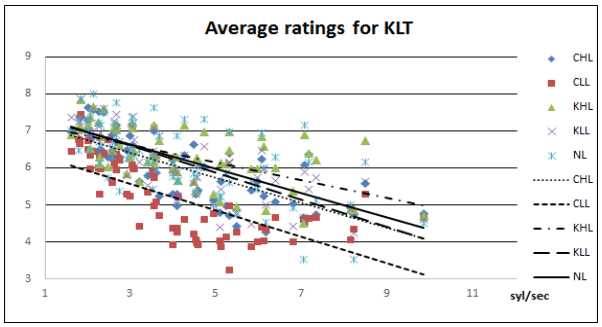


Figure 12. Average rating scores of five listener groups (CHL, CLL, KHL, KLL, and NL) for KLT's duration-manipulated stimuli

## 2.5 Discussion

We have explored native listeners' perception of optimal prosodic values for Korean L2 talkers' English speech. Prosodic parameters like pitch and speech rate were merely manipulated with all the segmental cues kept intact, so perceived accentedness was considered as reflecting native listeners' responses only to such prosodic changes. Korean talkers were varied to be high and low in proficiency in order to seek for L2 English prosody optimal for each proficiency. We assumed that segmental imprecision originated from different L2 proficiencies would be compensated for by different levels of prosodic promotion, minimizing overall foreign accent. According to many studies on the role of prosody (Munro 1995; Jilka 2000; Boula de Mareuiil and Vieru-Dimulescu 2006; Trofimovich and Baker 2007; Lee and Liu 2012, 2015), prosodic deviances contributed as much or more to native listeners' detection of foreign accent in L2 speech. This sheds light on synthetic modifications of prosodic parameters to examine how the preception of overall accentedness would change for high and low proficiency talkers, respectively.

Figures 1 to 4 showed native listeners' responses to eleven different pitch levels of H* and L*. The speech stimuli were synthesized in terms of pitch from English statements and yes-no questions produced by Korean high and low proficiency talkers. Overall, accent scores increased (became stronger) as the peak of H* decreased, but the scores decreased (became weaker) as the peak of H* increased. This pattern was observed consistently for both high and low proficiency talkers.

High talkers who were rated between 1 and 3 in the Accentedness Rating (AR) task started to be perceived as stronger accent than the score 3, i.e., 3.11, when the peak of H* was as low as 166Hz. It became worsened to 3.50 at 146Hz, but native listeners did not judge Korean high talkers as a lower category of L2 proficiency (intermediate) at any synthetic point of lowering H*. In other words, their scores did not go up higher than 4 even when the peak of H* was manipulated to decrease in five steps by 100Hz (20Hz per step). Similarly in yes-no questions, perceived accentedness was barely higher than the score 3 (3.45) when the valley of L* went up to 244Hz from the original pitch 184Hz. It reached 3.87 and 3.72 when the valley of L* increased to 264Hz and 284Hz. High talkers were perceived merely out of the boundary of high proficiency, but they were not categorically affiliated to a lower level of intermediate proficiency (scores 4 to 7) even when L* increased in five steps by 100Hz.

Low proficiency talkers, on the other hand, showed a category change of perceived accentedness as shown in Figures 2 and 4. The score started to drop out of 7, i.e., 6.68 when the height of H* increased to 237Hz from 217Hz of original speech. The scores safely stayed in the boundary of intermediate proficiency when H* was 277Hz to 317Hz. This indicates that low talkers' accent can be perceptually ameliorated to be a higher proficiency category if pitch is corrected. That is, H* should be at least 277Hz or higher when they produce statement sentences. Similarly, low talkers were rated as 7.23 when L* was 203Hz in original yes-no question stimuli and started to be judged out of the low proficiency category (6.57) at 163Hz. They were finally rated below 6 when the valley of L* decreased to 143Hz or lower; their scores were consistently within the range of intermediate proficiency, 5.85, 5.76, and 5.65 at 143Hz, 123Hz, and 103Hz, respectively. This also shows that low proficiency L2 speech successfully improves when high and low pitch accent values are corrected more similarly to native talkers. More specifically, H* should be increased and L* should be decreased by the amount of 40Hz ~ 80Hz in order to induce low talkers' categorical enhancement.

Similarly, Lee (2016) reported that Chinese low talkers were perceptually improved to the category of intermediate proficiency when H* was manipulated to increase to 295Hz from 210Hz and when L* was synthesized to decrease to 110Hz from 195Hz. Unlike Korean high talkers who did not show an explicit categorical downgrade to intermediate, Chinese high talkers were perceptually deteriorated to the

intermediate category when H* synthetically became smooth to 150Hz (from the original average 219Hz) and when L* changed from 190Hz to 270Hz. There seems to be differences between Chinese and Korea high talkers. As mentioned in Lee (2016), the recruiting pool of Chinese participants showed relatively low proficiency, and the Chinese talkers who were rated as high within the pool might be lower in proficiency. This seems to hold true because their average pitch of H* was considerably lower than that of Korean high talkers (219z vs. 246Hz), and that L* was also higher than that of Korean high talkers (190Hz vs. 184Hz). Therefore, their interlanguage entails less similar phonological/phonetic features to English natives than Korean high talkers, and those sound features must have been more readily worsened and perceived as intermediate. On the other hand, Chinese and Korean lower proficiency talkers had very similar pitch values of H* and L* (210Hz vs. 217Hz; 195Hz vs. 203Hz). The differences are merely 7 to 8 Hz, so their proficiency was assumed to be almost identical at least in terms of prosody. Consequently, both Chinese and Korean low talkers were perceptually upgraded to intermediate when their intonation was corrected.

Lee (2016) suggested that a pitch range rather than absolute pitch values of H* and L* should be important. Since the sentences were intentionally made to have one pitch accent, pitch ranges were measured between the peak of H* and the bottom at the end of a statement and between the valley of L* and the ceiling at the end of a yes-no question. Tables 7 and 8 show the average pitch ranges of original and synthesized speech of statements where optimal pitch ranges for proficiency enhancement were also specified in case of Korean low proficiency talkers (KLT). What should be worth to note is that low talkers' exertion to make a pitch range as wide as or wider than high proficiency talkers ended up with intermediate proficiency. This is due to much dissimilar sound features derived from stronger accented segments and prosodic factors other than pitch.

Table 7. Pitch ranges of original and synthesized speech of statements

| | original speech | | | synthesized speech | |
|---|---|---|---|---|---|
| | bottom | H* | pitch range | optimal H* | pitch range |
| KHT | 163Hz | 246Hz | 83Hz | -- | -- |
| KLT | 173Hz | 217Hz | 44Hz | 277~317Hz | 104~144Hz |

Table 8. Pitch ranges of original and synthesized speech of yes-no questions

| | original speech | | | synthesized speech | |
|---|---|---|---|---|---|
| | ceiling | L* | pitch range | optimal L* | pitch range |
| KHT | 326Hz | 184Hz | 142Hz | -- | -- |
| KLT | 282Hz | 203Hz | 79Hz | 163~123Hz | 119~159Hz |

The results of pitch-manipulated speech should be applied to the pedagogy of English intonation, especially for low proficiency Korean learners. Since actual pitch values are not easy to teach precisely, instructors should take advantage of the percentile from learners' own pitch values. Based on the results shown in Tables 7 and 8, the H* values which induced enhancing perception of foreign accent to the intermediate proficiency are 277Hz ~ 317Hz in statements and 163Hz ~ 123Hz in yes-no questions. When compared with those of unsynthesized speech, (217Hz for H* and 203Hz for L*), the optimal H* values are 28% ~ 46% higher than the original one, and the optimal L* values are 20% ~ 39% lower than the original one. Therefore, low Korean talkers should be taught to produce H* roughly 30% to 45% higher than the ones that they usually produce in statements. In case of yes-no questions, they should be taught to produce L* approximately 20% to 40% lower than in their normal utterance.

Turning our attention to the results of duration-manipulated speech, Korean high talkers did not go through a categorical drop to intermediate when their speech was synthesized to be elongated twice or shrunk to a half. Native listeners seemed to disfavor slower speech over faster speech although their actual rating scores consistently stayed in the high proficiency category for either faster or slower speech. On the other hand, low proficiency talkers were rated drastically lower scores and they were successfully judged as intermediate proficiency as shown in Figure 6. Table 9 summarizes average speech rates and optimal speech rates for low talkers to be perceptually enhanced to higher proficiency. The optimal speech rates to KLT are 4.87 to 7.87 where native listeners rated them as intermediate proficiency. They are comparable with KHT's original speech rate 5.08, but their ratings were at best intermediate. This suggests that segmental and/or other prosodic features than speech rate were strongly accented and that they were not presumably compensated as much by the corrected speech rate. In the pedagogical perspectives,

low proficiency talkers should be instructed to produce English sentences 1.2 to 2 times faster according to Table 9.

Table 9. A summary of the results of duration-manipulated speech

| | original (syl/sec) | optimal (syl/sec) |
|---|---|---|
| KHT | 5.08 | -- |
| KLT | 3.93 | 4.87 ~ 7.87 |

According to Lee (2016), Chinese high talkers showed a categorical deterioration of perceived accentedness when speech rate was too fast (higher than 7.8 syl/sec) or too slow (lower than 4 syl/sec). On the other hand, native listeners' rating scores were invariable as Chinese low talkers' speech was manipulated to be faster or slower. As mentioned in Lee, the recruiting pool of Chinese participants was overall lower proficient in English; therefore, proficiency decision in the AR task was executed within the pool, which resulted in overall downgrading in L2 proficiency of Chinese talkers. Chinese high talkers who were somewhat less proficient (than Korean high talkers in the current work) were easily degraded to intermediate when speech rate increased or decreased. On the other hand, Korean high talkers who were proficient enough seemed to have their interlanguage very similar to native speakers, and their foreign accent was little affected by speech rate changes. Native-like phonetic information appeared to be sufficiently solid, which resulted in the maintenance of high proficiency.

Figures 7 to 12 showed whether nonnative listeners' responses to Korean talkers agreed with those of native listeners. When Korean talkers' pitch-manipulated stimuli were assessed by native and four nonnative listeners (Korean high and low listeners and Chinese high and low listeners), results are summarized in Table 10. Chinese high listers seem to have taken advantage of their L1 Chinese where pitch playes an important role (Xue and Lee 2014). As stated in Lee (2016), Chinese has lexical tones which are phonologically distinctive and faithfully realized in the surface intonation, and English has lexical stress which is associated with pitch prominence in the surface intonation (p. 510). Chinese high listeners who have a similar interlanguage to L2 seemed to be sufficiently responsive to pitch changes in English intonation. Korean higher listeners, on the other hand, were not sensitive enough to pitch manipulations of English sentences even though they have a similar

interlanguage to L2. They might be affected by their L1 because Korean does not have a critical role of pitch equivalent to stress-associated pitch accents in English or lexical tone-associated pitch contours in Chinese (Lee, 2016). Chinese low listeners, who have a dissimilar interlanguage from L2, mostly showed disagreement with native listeners. This shows that they didn't take advantage of their L1 Chinese. Since they have very deviant English prosodic features in their interlanguage, pitch changes of H* and L* in English sentences did not seem to be comprehensibly captured within their intelanguage. The function of L1 pitch failed to play a role in the perception of English pitch change due to their interlanguage. Moreover, Korean low listeners have their L1 Korean like Korean high listeners and a dissimilar interlanguage to L2 like Chinese low listeners. For such coupled reasons, Korean low listeners showed significantly distinctive perceptual responses from native listeners.

Table 10. A summary of Correlations between native and nonnative listeners (pitch-manipulated speech)

| | | CHL | CLL | KHL | KLL |
|---|---|---|---|---|---|
| KHT-state | NL | * | ≠ | ≠ | ≠ |
| KLT-state | NL | * | ≠ | ≠ | ≠ |
| KHT-yes-no | NL | ** | ** | ≠ | ≠ |
| KHT-yes-no | NL | ** | ≠ | ≠ | ≠ |

In case of Chinese talkers' pitch-manipulated stimuli as reported in Lee (2016), nonnative listeners showed similar patterns. both Chinese high and low listeners statistically significant agreement with native listeners. Unlike the experiment of the current study where Chinese low listeners did not agree with native listeners, Chinese low listeners seemed to take benefits of L1 familiarity with Chinese talkers. That is, Chinese accented English speech, though its pitch was synthesized, entailed similar patterns of segmental and prosodic deviances to Chinese listeners due to their shared interlanguages. Despite differences in proficiency between Chinese talkers and listeners, the shared L2 phonology contributed to favorable ratings of foreign accent for the same L1 talkers (Munro 2006). This might extend to Chinese low listeners' similar sensitivity toward the pitch chances of Chinese talkers' L2 speech.

Recall Figures 11 and 12 here. When duration-manipulated Korean L2 speech

was evaluated by native and nonnative Chinese and Korean listeners, both Korean and Chinese listeners, irrespective of their proficiency, showed statistically significant correlations with native listeners. Native listeners detected stronger foreign accent as Korean L2 speech was synthesized to be slower, and nonnative listeners similarly preferred faster L2 speech. What is worthy noting is that nonnative Chinese and Korean listeners did not show significant correlations with native listeners for duration-manipulated Chinese speech as demonstrated in Lee (2016). Native listeners showed a clear curvelinear pattern of accent ratings for high talkers, but steady rating scores for low talkers. However, nonnative listeners all consistently rated faster speech as lower scores (weaker accent). There were distinctive mismatches of rating patterns between native and nonnative listeners, which resulted in no correlations between them. This does not support Munro and Derwing (1998)'s claim that native listeners' detection of foreign accent is curvelinear along with changes in speech rate. The relation between accent scores and speech rate is not necessarily curvelinear, but it may differ depending on L1 or L2 talkers' proficiency, etc.

## 3. Conclusion

We have investigated native English listeners' perception of foreign accent for pitch and duration manipulated speech of Korean high and low talkers and contented that low proficiency talkers could be perceptually better accented when a single prosodic parameter such as pitch or speech rate was merely corrected. They categorically promoted to the intermediate when H* and L* were remedied to be less accented and their L2 speech was synthesized to be faster. The corrected prosodic features seemed to be readily detected by native listeners because such corrections might be comparatively salient over segmental features seriously deviant from native norms in their interlanguage. The results of the current experiment suggested actual numeric values of optimal pitch and speech rate for upgrading low talkers' proficiency; H* should increase roughly 30% to 45% higher than the ones that they usually produce in statements, and L* should deepen approximately 20% to 40% lower than in their normal utterance. Speech rate should be 1.2 to 2 times faster for low talkers to be perceived as intermediate. On the other hand, Korean high talkers didn't show a perceptually categorical decline to intermediate when pitch or speech

rate was synthetically deteriorated. Due to their little accented L2 speech, phonological/phonetic features, which are very similar to those of native speakers, seemed to firmly tolerate the degrading portion of prosody.

Concerning the correlations between native and nonnative listeners' ratings, Chinese high listeners merely showed a concord with native listeners in pitch-manipulated Korean speech, but all of nonnative listeners agreed with native listeners in duration-manipulated speech, preferring faster speech,

## 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.

Flege, James and James Hillenbrand. 1984. Limits on pronunciation accuracy in adult foreign language speech production. *Journal of the Acoustical Society of America* 76: 708-21.

Flege, James, 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.

Jilka, Matthias. 2000. *The contribution of intonation to the perception of foreign accent*. PhD Dissertation, University of Stuttgart.

Lee, Joo-Kyeong. 2014. The role of prosody in the perception of foreign accent and comprehensibility: Prosody-corrected-L2 speech vs. prosody-distorted-L1 speech. *Korean Journal of Linguistics* 39(1): 155-179.

Lee, Joo-Kyeong. 2016. Optimal prosody for different proficiency Chinese talkers of English: Pitch and speech rate. *Linguistic Research* 33(3): 487-517.

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.

Levi, V. Susannah, Stephen J. Winters and David B. Pisoni. 2007. Speaker-independent fac-

tors affecting the perception of foreign accent in a second language. *Journal of the Acoustical Society of America* 121(4): 2327-38.

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.

Magen, Harriet. 1998. The perception of foreign-accented speech. *Journal of Phonetics* 26: 381-400.

Major, Roy C. 1987. Foreign accent: Recent research and theory. *International Review of Applied Linguistics in Language Teaching* 25: 185-202.

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

Moulines, Eric and Francis Charpentier. 1990. Pitch synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication* 9(5-6): 454-467.

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

Munro, Murray and Tracey Derwing. 1998. The effects of speech rate on the comprehensibility of native and foreign accented speech. *Language Learning* 48(2): 159-182.

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

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.

Polyanskaya, Leona, Mikhail Ordin and Maria Crazia, 2016 Relative salience of speech rhythm and speech rate on perceived foreign accent in L2. *Language and Speech* 60(3): 333-355.

Rekart, Deborah, Paul Hoffman and Raymond Daniloff. 1985. Evaluation of foreign accent using synthetic speech. *Journal of the Acoustical Society of America* 77. [https://doi.org/10.1121/1.2022635].

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.

Van Els, Theo & Kees de Bot. 1987. The role of intonation in foreign accent. T*he Modern Language Journal* 71(2): 147-155

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-manipulated English L2 speech. *Language and Linguistics* 66: 509-537.

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

Winters, Stephen and Mary Grantham O'Brien. 2013. Perceived accentedness and intelligibility: The relative contribution of F0 and duration. *Journal of Speech Communication* 55(3): 486-507.

## Appendix I: A passage for the Accented Rating (AR) 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 years. 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 smiled when he remembered his hard early years and the long road to success.

## Appendix II: Sentences for pitch- and speech rate-manipulations (Target sentences are boldfaced.)

(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: 2017. 12. 08.
Revised: 2018. 02. 13.
Accepted: 2018. 02. 13.