Intonational Variations in Korean Learners' English

Seokhan Kang*
Konkuk University

Hyunkee Ahn**
Seoul National University

Abstract

This study investigated the effects of English proficiency level of Korean L2 learners on the production of intonational features in L2 English. It was hypothesized that the Korean group who had high proficiency level in the English language would have intonation patterns more similar to those of the native English speakers than a low level group of Korean speakers. Forty-nine subjects in three groups – 21 Korean adults in the high-level group, 19 Korean adults in the low-level group, and 9 native English speakers in a control group – took part in the experiment. The results showed that the high group was more native-like by having a steeper F0 declination tilt (a wider F0 range and fast speech rate). However, even this high-level group exhibited difficulty in controlling phonetic cues at the boundary. These results suggest that the acquisition of second language intonation is affected by English proficiency level, but the degree of intonational acquisition varies by sub-areas such as acoustic cues at the boundary which proved to be hard to acquire.

Key words: intonation, F0, second language acquisition, English proficiency, speech rate, declination tilt, boundary cues

* First author (kangs45@kku.ac.kr)
** Corresponding author (ahnhk@snu.ac.kr)

Ⅰ. Introduction

Intonation plays a crucial role in deciding the quality of L2 speech. Intonation uses suprasegmental phonetic features to convey sentence-level pragmatic meanings in a linguistically structured way (Ladd, 1996; Kang, 2014). The structure is affected by the subject variables such as the starting age of L2 acquisition, L1 background, L2 experience and training. The aim of the study is to investigate how the effect of L2 proficiency level relates to acquiring L2 intonation. For this investigation, intonational phonetic features will be analyzed in detail.

Intonation has been considered to be an important aspect of prosody affecting the quality of L2 speech (Laures & Weismer, 1999; Mennen, 2006). Even though its quality is related to both segmental (Jenkins, 2000; Munro & Derwing, 1995) and prosodic features (Anderson-Hsieh & Koehler, 1988; Tajima, Port, & Dalby, 1994), some researchers have reported that prosodic factors in intonation are more important than segmental factors in deciding the L2 intelligible quality (Bradlow, Torretta, & Pisoni, 1996; Tajima, Port, & Dalby, 1994). Since intonational features lead listeners to interpret the speech elements in a sentence as new versus old, salient versus weak, or foreground versus background, its function is crucial in verbal communication among humans (Morley, 1992; Wennerstrom, 1991, 1994).

L2 Intonational features are affected by various variables: the starting age of L2 acquisition (Guion, et al., 2000; Tahta, et al., 1981), L2 language experience (Mennen, 2004; Trofimovich & Baker, 2006), L1 background (Archibald, 1998; Davis & Kelly, 1997; Guion, et al., 2004), or L2 motivation (Conrad, 1991; Moyer, 1999). Among these factors, the L1 transfer factor has been suggested to play a key role in reforming L2 intonational production. Delattre (1963) claimed that L2 speakers tended to impose their L1 intonation patterns on producing their second language. In his study, a French learner of English produced various L1 like English intonations. Aoyama and Guion (2007) reported that there were considerable differences in duration of linguistic unit and overall F0 range between English spoken by

Japanese speakers and that by native English speakers, and suggested that the prosodic differences come from their L1 background. These studies reach the same conclusion that L2 experience is the most influential factor in acquisition of L2 intonation.

To date, however, a question of what influence L2 learners' English proficiency in non-immersed EFL environment has on L2 intonational acquisition was rarely studied, even though a lot of research was done in an ESL environment, mainly in the north America. This limitation resulted in poor implications to the current English pronunciation teaching in Korea. Thus, hopefully, this study would supplement this weak point for the L2 studies in Korea.

Ⅱ. L1 intonational structure

Phonologically, intonation consists of prosodic words with strong and weak rhythms. This rhythm is a basic structure of the intonational phrase. More specifically, syllables are grouped into feet in such a way that every unstressed syllable is grouped with a stressed syllable that precedes it. Foot is a unit of recurrent form: being a phonological unit, they should be represented by nodes on the foot level. These feet are also subject to a higher-level contouring, governed by the Phrase Prominence Rule: In a pair of sister nodes [N1 N2]p, where p is a phrasal category, N2 is strong (Giegerich, 2005). Thus, a stressed syllable in the final foot can obtain the phrasal accent. This phonological structure shapes a typical intonational pattern in English.

Phonetically, there is a general tendency for F0 to begin with a moderate frequency, moving to a higher frequency, and then being lowered across the sentence (Pike, 1945). The intonation in most languages can be characterized by the declination theory which asserts that F0 gradually fall throughout the course of a sentence, which is one of the linguistically salient F0 aspects well-documented in linguistics.

The degree and type of the F0 movement, however, may vary by intonational structures which each language possesses. For example, Korean has a different intonational structure from English. It has two different prosodic units above the prosodic word: the intonational phrase (IP) and the accentual phrase (AP) (Jun, 2005). An IP is characterized by phrase-final lengthening as a form of boundary tone and is also the highest prosodic unit including one or more APs. APs in Korean do not have any pitch accents associated with a stressed syllable in its domain and lack a unit like a phrase accent which occurs at the end of the intermediate phrase in English. However, it is difficult to find a distinctive declination tilt in Korean because an IP-final part is too short to represent the distinctive declination tilt. There is usually only one accented syllable in IP-final part so that it is hard to maintain in prosody of Korean.

English, unlike Korean, is a stress-timed language in which one syllable is stressed within a prosodic foot. The stressed syllable tends to have a greater duration, higher pitch, and more complicated contour of F0 than neighboring unstressed syllables. In general, English has three prosodic units above the prosodic foot: intonation phrase (IP), intermediate phrase (iP), and accentual group (AG) (Wells et al., 2004). An IP is the highest prosodic unit characterized by intonation and may contain one or more iPs. The IP has a pattern of final lengthening with F0 falling finally in declarative sentences. An iP contains at least one pitch accent with either falling or rising F0. Each iP consists of one or more AGs, functioning as a domain for a pitch accent configuration. In addition, AGs contain one or more Feet, each of which comprises a strong initial syllable and weak syllables that follow it.

It is clear that L1 intonational structure has some influences on acquiring L2 intonation, which forms a major part for the intelligible L2 speaking. Jilka (2000) reported that German/American bilinguals used a wider pitch range in their American English than pitch range in German. Such difference is assumed to be due to the intonational differences between L1 and L2.

Likewise, it is hypothesized in this study that intonational difference is affected by the English proficiency level of Korean speakers in this EFL environment. The more specific aims of the study are to investigate to what extent Korean learners of English produce native-like intonation and whether English proficiency level has significant effect on the acquisition of L2 intonation. In this pursuit, we will employ F0 contour as an estimate of F0 declination, and duration and mean F0 at the boundary will also be treated in experiments.

Ⅲ. Methodology

A. Participants

The data were collected from a total of 49 adult participants. None reported being diagnosed with a language or speech disorder. The participants were divided into three groups: HEP (High English Proficiency Speakers) of Koreans; LEP (Low English Proficiency Speakers) of Koreans; and NE (Native English Speakers). Most of the NE participants were the visiting students at a university in Seoul. They did not speak any language other than American English when they took part in the experiment.

The Korean subjects have learned English for more than ten years in Korea. There are no subjects who have lived in English-speaking countries. TEPS (Test of English Proficiency developed by Seoul National University) scores were reported by all Korean participants and they were divided into two groups based on the scores. TEPS score was chosen as a criterion because it was designed to decide English proficiency at adult level and it is also known to be a reliable tool for the measurement of English proficiency (see www.teps.or.kr). The LEP group performed at 64% correctness (642/1000) with a standard deviation of 82.91. On the contrary, the HEP group performed at 86% correctness (863/1000) with a standard deviation of 64.72. The one way analysis of variance confirmed

that there was a significant effect on the group variable ($F(3,29) = 8.276$, $p < .001$).

B. Materials and Procedures

All the participants were recorded individually in a quiet room using a portable digital recorder (Marantz PMD 650). Subjects heard the recorded stimuli before recording. The elicitation procedure helped them to produce sentences fluently, while minimizing the likelihood of mimicry.

We used a picture-story telling experiment (see Appendix), in which the pictures were presented to the subjects and they have to make sentences for 1 minute with given target words. We confirmed that they knew how to pronounce all the given words. The sounds were recorded with a Marantz PMD 650 using a Shure SM 10A microphone, and digitized at 44.05 kHz and 16 bit resolution.

C. Intonational Measurements

Several acoustic measurements such as fundamental frequency (in Herz) and duration (in milliseconds) were made using a waveform display with a time-locked wideband spectrogram with software PRAAT (5.1.17). All acoustic cues were measured from the initial acoustic signals in both the waveform and the spectrogram to the final acoustic cues of the boundary such as burst or spectral cues (Kent and Read, 2003). Measures for F0 declination tilt were also made in each sentence as specified in detail below.

The F0's were measured at the onset of the phrase, at the absolute maximum point of the F0 peak, at the absolute minimum point of the F0 valley, at the local maximum point of the final F0 peak, at the local minimum point of the final F0 valley, and at the phrase offset. Thus, the F0's and times in six points of a sentence were collected and calculated to form a slope line, in which F0 range forms on the x-axis and time duration (otherwise speech rate) shapes on the y-axis. The

declination tilt was computed as follows.

$\gamma = \triangle f / \triangle t$
where $\triangle f$ is the range difference of f0 declination and $\triangle t$ is the durational time.

Based on the formula, the intonation contours for the upper-line and lower-line were determined in mixed fashions as in Cohen et al. (1982), Thosen (1985), and Lieberman et al., (1985). In this study, the adopted method provides information to compare the size of F0 slope and duration across the three groups. The upper-line connects the first maximum peak of F0 appearing in the initial part of the sentence to the final peak F0 of the utterance, while the lower-line connects the initial minimum point of F0 to the final valley point of F0 in the sentences. The formulas are as follows:

Upper-line = the initial peak of F0 – the boundary peak of F0 / duration * 100

Lower-line = the initial valley of F0 – the boundary valley of F0 / duration * 100

If the slope approximates 0, a level intonation between the two measured points of F0 is indicated. If the slope has a negative value approaching -1, it means that the initial peak point of F0 is higher than the final peak point of F0. On the contrary, in the case of a positive value approaching +1, it means that the final peak point of F0 is higher than the initial peak point of F0 (cf. Positive values usually appear in the interrogative sentences.).

Several acoustic features were also measured as follows:

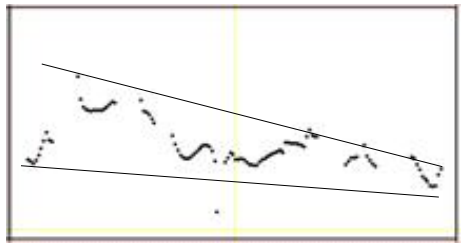
Y-features: the declination tilt was measured using F0 range as a Y-axe, and speech rate as an X-axe. The F0 range, one of the Y-features, is known to be an indicator of English proficiency (e.g., Backman, 1979; Willems, 1982). Generally, low English

proficiency is related to a narrower F0 range. In this study, the range was measured from the highest point to the lowest point of the F0 and we used the F0 tracing function in Praat to determine peaks and troughs. F0 was also calculated from the duration measurement of an individual cycle in the waveform when it is needed to check on the accuracy of the F0 tracker.

X-feature: As an X-axis feature, the speech rate was proved to be a good indicator of the second language proficiency (e.g., Derwing & Munro, 1997; Guion et al., 2000). In this study, the speech rate is operationalized as a duration from the initial acoustic signal of the phrase in both the waveform and the spectrogram to the final acoustic or spectral cues of the phrase boundary.

Boundary features: Final strengthening at boundary is realized in prosodic domains (e.g., mora, syllable, foot, or prosodic word) at the end of the phrase in forms of longer duration (Beckman & Edwards, 1991; de Pijper & Sanderman, 1994; Wightman et al., 1992), strengthening (Fourgeron & Keating, 1997), and alternation of overlap with adjacent segments (Chun, 2002). This study measured duration and mean F0 in the phrase-final syllable because the final lengthening and strengthening were shown as a form of local peak F0 or longer duration.

See the following picture where the upper and lower lines are described covering the sentence "My brother is coming on Friday."



Picture 1. The upper and lower lines of intonation over "My brother is coming on Friday."

These measures were analyzed with Repeated Measures of Analyses of Variance (RM ANOVAs) for the statistical evaluation of the groups. Such dependent variables as F0, speech rate (duration), boundary cues, and the declination tilt measures were examined by Group (three levels: NE, HEP, LEP). The RM ANOVA method was employed in order to factor out some of the variation that occurs within individuals.

D. Data analysis

Three analyses were performed. The first analysis was to examine the extent to which the learners were able to produce L2 intonation intelligibly. In this analysis, the sentences spoken by the subjects were presented to nine native English raters for evaluation. They judged the sentences produced by both groups of L2 Korean learners. The second analysis was to examine the extent to which the learners were able to accurately produce specific intonations. The results of the acoustic measurements obtained were compared. The purpose of the experiment was to analyze to what extent each measurement has been affected by English proficiency. The final analysis was to extend the findings of the first two analyses by using a multiple regression procedure and to investigate how specific intonational features of L2 speakers relate to the native listeners' intelligibility judgment on the L2 speech.

Ⅳ. Ratings of intelligibility

A. Ratings and Raters

The recorded samples were 40 in total (twenty-one for 21 HEP speakers and nineteen for 19 LEP speakers) and they were randomized and presented to the native English raters. The total of nine native-speaking English listeners (five males and four females; age range 20-26 years, mean=23) were recruited to evaluate the intelligibility of L2 speakers' production using a 9-point Likert scales (from 1 = definitely unintelligible speaking to 9 = native-like intelligible speaking). All the raters were native English speakers who had some teaching experiences in Korea. All the raters reported normal hearing. For the judgment, the raters listened to the sample paragraph as a warm-up before rating each subject. The raters were encouraged to use the entire scale and to give a score even when they were unsure.

B. The Results

The first purpose of the present study was to test the hypothesis that the HEP learners were to produce L2 intonation more intelligibly than the LEP counterparts. Most of the raters showed the results as expected. The dependent variable in this analysis was the mean score of the nine English listeners' rating of the forty Korean subjects. The intra-class correlation coefficient was used to measure the degree of inter-rater reliability. The coefficient was highly correlated ($r(9) = 0.96$, $p < .0001$). These results indicated the high level of agreement among all the raters.

Figure 1 presents the mean scores of fluency rating for both Korean groups. The scores for both Korean groups are quite different, ranging from 3.0 to 6.0 out of 9. In Figure 1, higher mean score was obtained for the HEP group. The mean scores in the intelligibility rating were like this: LEP = 4.32 and HEP = 5.45. The two groups were tested by an independent t-test and the analysis revealed a significant group difference ($t = 25.352$,

df = 39. P < .0001). This indicates that English proficiency level has a significant relation to L2 intonation intelligibility.

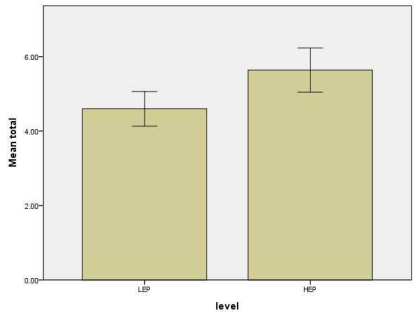


Figure 1. Mean scores of intelligibility rating (±1 SE) for both Korean groups.

Ⅴ. Production experiment

The English intonational phrases produced by the subjects were analyzed to find differences among the three groups. Table 1 presents the mean values and standard deviations of the F0 range, speech rate, mean F0 and duration at boundary, and slope patterns of upper- and lower-lines. As seen in descriptive statistics in Table 1, there were differences in every parameter among the three groups.

Table 1. Mean and standard deviation of parameters used in the experiment

Measure		NE	HEP	LEP
Slope pattern	Upper-line	-0.71(1.08)	-0.57 (0.43)	-0.39 (0.34)
	Lower-line	-0.37(0.34)	-0.31 (0.28)	-0.22 (0.21)
F0 range (Hz)	Overall Range	68 (32)	61 (22)	56 (33)
Speech rate (S)		1.71 (9)	2.42 (28)	2.91 (18)
Boundary cues	Mean F0 (Hz)	115 (35)	128 (31)	130 (40)
	Duration (S)	0.40 (0.10)	0.56 (0.11)	0.60 (0.11)

A. Declination tilt

For the intonational slope, the correlation between upper and lower-line was statistically significant for the three groups (r = 0.417, p < .01). This implies that the two lines were closely related. However, the groups diverged markedly in terms of the upper-line; the NE group had the steepest slope, followed by the HEP and then by the LEP groups.

The RM ANOVA confirmed that there was a significant group effect for the upper-lines, (F(2,799) = 27.326, p < .001). Tukey's tests (p < .05) revealed that the upper-lines slope was the steepest for the NE group, intermediate for the HEP group, and the least steep for the LEP group. The RM ANOVA also showed a significant effect for the lower-lines, (F(2,799) = 24.722, p < .001). Tukey's tests (p < .05) revealed that the lower-line is steepest for the NE group, intermediate for the HEP group, and then for the LEP group.

The NE group showed the following features: comparatively steep slope for the upper-line and lower-line (-0.71 and -0.37), wider F0 range of 68 HZ for the entire sentence, 82 Hz in the initial part and 31 Hz in the final foot, and 1.71 seconds of sentence duration. The HEP group is generally located between

the NE and LEP groups. They showed the followings: intermediate slope of the upper-line and lower-line (-0.59 and -0.31), intermediate F0 range of 61 Hz in the entire sentence, 73 Hz in the initial part and 35 Hz in the final foot, and 2.42 seconds of sentence duration. Finally the LEP group showed the followings: gentle slope of the upper-line and lower-line (-0.36 and -0.22), narrow F0 range of 56 Hz in the entire sentence, 64 Hz in the initial part and 37 Hz in the final foot, and 2.91 seconds of sentence duration.

B. Y-axis (F0)

For the F0 range on an y-axis, the RM ANOVA confirmed that there was a significant group effect on F0 range, ($F(2,899) = 21.967$, $p < .001$). Tukey's tests ($p < .05$) revealed that the F0 range of the LEP group was smaller than those of the NE and HEP groups. Figure 2 presents the F0 range produced by the three groups. The results showed that members of the HEP group produced the F0 range which is similar to that of the native speakers. However, members of the LEP group produced smaller range in F0 when speaking. Over all, this result supports our hypothesis that more fluent L2 learners of English have a wider F0 range than less fluent L2 learners (cf, Bradlow et. al., 1996; Mennen, 2006).

The narrow F0 range could be realized by the influence of the native language (Scherer, 2000; van Benzooijen, 1995), or by a lack of L2 proficiency (Backman, 1979; Willems, 1982). At this moment it is not clear whether the narrow F0 range in this experiment results from the L1 influence or the low proficiency level of the LEP group. However, it is noteworthy that the F0 range in the LEP group is only 70% of the F0 range in the NE group. The F0 range in the HEP group approximated the FO range in the NE group, which shows less influence from the native language that is Korean.

An interesting observation about the F0 range is that, although all groups have a wider range in the initial part of the sentence and a narrower range in the final syllable of the

sentence, the degree is different among the three. In NE, the initial F0 range is almost two times larger than that of the final word, but the LEPs' F0 range is almost same between the two areas, which naturally leads to a relatively more monotonous intonation.

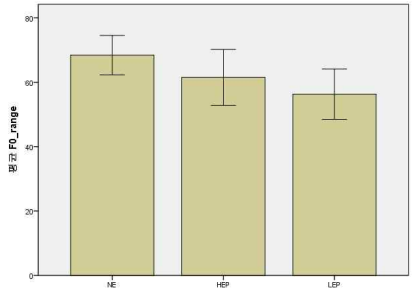


Figure 2. Mean values and standard errors for the F0 range.

C. X-axe (Speech rate)

For the speech rate on an x-axe, the results of the RM ANOVA confirmed a significant group effect on the speech rate, ($F(2,799) = 348.764$, $p < .001$). Tukey's tests ($p < .05$) revealed that the phrase duration was shortest for the NE group, intermediate for the HEP group, and longest for the LEP group. Figure 3 presents the mean duration values for the three groups. These results were in agreement with the previous works in which more native-like speech was produced with fast speech rate (Adams & Munro, 1978; Guion et al., 2000; Munro & Derwing, 1995; Sluijter & Van Heuven, 1996). Likewise the results showed that the HEP group produced sentences with intermediate duration which lies between those of the LEP and NE groups, indicating that immersion has an influence on the speech rate.

The reason why the LEP group produces sentences more slowly may come from not being able to control various

components of an utterance. For example, they tend to pronounce function words and unstressed syllables with higher pitch and longer duration than the NE group. However, the HEP group comes close to the NE group, distinguishing stress and unstressed syllables, strengthening content words, and lengthening the pitch contour at boundary.

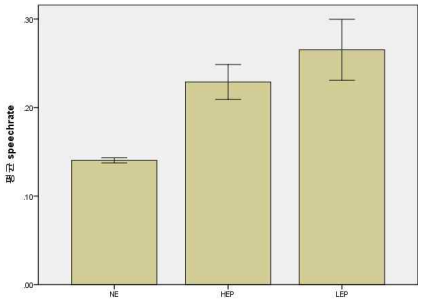
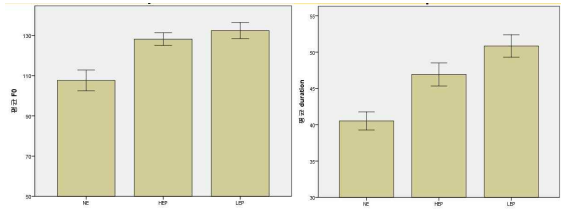


Figure 3. Mean values and standard errors for the speech rate

D. Mean F0 and duration of Phrase-final foot

As for the mean value of F0 in the phrase-final foot, the result of RM ANOVA returned a significant group effect, ($F(2,799) = 9.653$, $p > .05$). Tukey's tests ($p < .05$) revealed that the mean value of F0 was meaningless cue among three groups. The results are summarized in Figure 4.

Also the RM ANOVA revealed that there was a significant group effect on duration of the phrase-final foot, ($F(2,799) = 33.876$, $p < .001$). Tukey's tests ($p < .05$) revealed that the duration in the LEP group was longer than those in the NE and HEP groups. See Figure 4. In a short summary, the results suggest that the final foot in the terminal phrases produced by the NE group showed low F0 and shorter duration, while the LEP group showed high F0 and long duration.



(a) Mean F0 at boundary (Hz) (b) Foot Duration at boundary (S)

Figure 4. Mean values and standard errors of F0 and duration at the boundary foot.

VI. Relationship between production and perception

The analysis of the production data showed that the level of English proficiency have a relation to the degree of the L2 learners' acquisition of English intonation. One remaining task was to investigate to what extent their production could correlate to the intelligibility judgment. For the analysis, both rating scores of the Korean subjects and their acoustic values of various intonational features employed in this study were submitted to the measures of correlation and regression. Zero-order correlations were computed between the learners' intelligibility rating scores (n=20) and their intonation-related acoustic values. See Table 2 below.

Table 2. Summary of correlation analyses between intelligibility rating scores and acoustic measurement values.

	F0 range	Speech rate	F0 at boundary	Duration at boundary
intelligibility rating scores	-.379*	-.711**	-.089	-.248*

**: $p < .001$, *: $p < .05$

The result clearly indicates that some acoustic values are significantly correlated with the intelligibility rating scores. This leads us to the conclusion that the speech rate and F0 range

play a role as strong predictors in deciding intelligibility in English. The results may reflect some universal features of intelligibility judgment, which is that whether what is heard is native-like is connected to some temporal cues such as speech rate and duration.

Ⅶ. Discussion and Conclusion

In this study we found that the English learners in the HEP group exhibited patterns more similar to those of the native English speakers, while the English learners in the LEP group did not. Specifically, compared to the LEP group, the HEP subjects showd steeper declination tilt including wider F0 range, much lower F0 and shorter duration at phrase-final boundaries, faster speech rate, and shorter duration in pauses.

The fact that the English proficiency level has a significant relation to the L2 intonation acquisition is clear, but it is shown that the acoustic cues at boundary maynot be affected by the English proficiency level. At F0 in boundary, there is no statistical difference in acoustic cues between both groups. It is interesting that the HEP group failed to produce native-like spectral cues at boundary, even though the duration of the boundary foot for the HEP group came close to that for the NE group. This indicates that English learning in an EFL setting does not guarantee a native-like production of intonation. More specifically, L2 learners with high English proficiency level appear to have native-like speech rate and F0 patterns such as declination tilt and F0 range. On the other hand, more local cues at boundary were found to be less native-like. It means that the advanced EFL learners have still difficulty in controlling the cues at boundary. It may cause confusion about whether or not the sentence has ended. Since this is a question not explored fully at the current study, we will save it for the future study.

In conclusion, by the experiment the HEP group was found to produce more native-like intonation than the LEP learners who had similar length of English instruction. From this result

we can infer that general proficiency level in English has a facilitating effect on the native-like production of English intonation: steeper declination with wide F0 range and faster speech rate than Korean. However, some acoustic cues at boundary were found to be hard to change regardless of proficiency levels. It seems likely that these characteristics come from the interference of L1 prosody. Future research is needed to investigate this question fully which will hopefully contribute to the more complete comprehension of foreign accent in English prosody.

References

Adams, C., & Munro, R. (1978). In search of the acoustic correlates of stress: Fundamental frequency, amplitude, and duration in the connected utterances of some native and nonnative speakers of English. *Phonetica, 35*, 125-156.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language learning, 38*, 561-613.

Aoyama, K., & Guion, S. G. (2007). Prosody in second language acquisition: An acoustic analysis on duration and F0 range. In O-S. Bohn & M.J. Munro (Eds.), *The role of language experience in second-language speech learning: In honor of James Emil Flege* (pp. 281-297). Amsterdam: John Benjamins.

Archibald, J. (1998). *Second language phonology.* Amsterdam: John Benjamins.

Backman, N. E. (1979). Intonation errors in second language pronunciation of eight Spanish speaking adults learning English. *Interlanguage Studies Bulletin, 4*(2), 239-266.

Beckman, M.E., & Edwards, J. (1991). Prosodic categories and duration control. *Journal of the Acoustical Society of America, 89,* (Suppl.1), S65.

Bradlow, A.R., Torretta, G.M., & Pisoni, D. B. (1996). Intelligibility of normal speech: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication, 20,* 255-272.

Chun, D. M. (2002). *Discourse intonation in L2. From theory and research to practice*. Amsterdam, PA: John Benjamins Publishing Company.

Cohen, A., Collier, R., & t'Hart, J. (1982). Declination: Construct or intrinsic feature of speech pitch? *Phonetica, 39,* 254–273.

Conrad, B. K. (1991). The relationship between empathy and pronunciation ability: A study of elementary-level college students of German. Unpublished doctoral dissertation, University of Texas at Austin.

Davis, S. M., & Kelly, M. H. (1997). Knowledge of the English noun–verb stress difference by native and nonnative speakers. *Journal of Memory & Language, 36,* 445-460.

de Pijper, J.R., & Sanderman, A.A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of Acoustical Society of America, 96,* 2037-2047.

Delattre, P. (1963). Comparing the prosodic features of English, German, Spanish, and French. *Interrnational Review of Applied Linguistics, 1,* 193-210.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in second language acquisition, 19*(01), 1-16.

Fourgeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the acoustical society of America, 89,* 369-82.

Giegerich, H. J. (2005). *English Phonology.* Boston: Cambridge Press.

Guion, S. G., Flege, J.E., & Loftin, J.D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics, 28,* 27-42.

Guion, S.G., Harada, T., & Clark, J. J. (2004). Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition, 7,* 207-228.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent.* Unpubished doctoral dissertation, University of Stutgart.

Jun, S. (2005). *Prosodic typology: The phonology of intonation and phrasing.* London: Oxford University Press.

Kang. S-H. (2014). The production of the English sentence-final foot by the effect of immersion. *Studies in Linguistics 32,* 1-20.

Kent R., & Read, C. (2003). *Acoustic analysis of English.* Boston: Singular Thomson.

Ladd, D.R. (1996). *The structure of intonational meaning*. Bloomington, IN: Indiana University Press.

Laures, J.S., & Weismer, G. (1999). The effect of flattened F0 on intelligibility at the sentence level. *Journal of Speech, Language, and Hearing Research, 42,* 1148 – 1156.

Lieberman, P., Katz, W., Jongman, A., Zimmerman, R., & Miller, M. (1985). Measures of the sentence intonation of read and spontaneous speech in American English. *Journal of the Acoustical Society of America, 77,* 649-657.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics, 32,* 543-563.

Mennen, I.(2006).Phonetic and phonological influences in non-native intonation: An overview for language teachers. *QMUC speech science research center working paper WP, 9,* 1-18.

Morley, J. (1992). *Rapid review of vowel and prosodic contexts*. Ann Arbor: University of Michigan Press.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21*(1), 81-108.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Pike, K.L. (1945). *The intonation of American English*. AnnArbor: University of Michigan Press.

Scherer, K. R. (2000). A cross-cultural investigation of emotion inferences from voice and speech: Implications for speech technology. In B. Yuan, T. Huang, & X. Tang (Eds.), *Proceedings of the sixth international conference on spoken language processing 2* (pp. 379-382). Beijing: China Military Friendship Publish.

Sluijter, A.M.C., & Heuven, V.J.V. (1996). Acoustic correlates of linguistic stress and accent in Dutch and American English. *Internat. Conf. on Spoken Language Process, 2,* 630-633.

Tahta, S., Wood, M., & Loewenthal, K. 1981. Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech* 24(3), 265-272.

Tajima, K., Port, R., & Dalby, J. (1994). Influence of timing on intelligibility of foreign accented English. *The Journal of the Acoustical Society of America, 95*(5), 3009-3009.

Thorsen, N. G. (1985). Intonation and text in standard Danish. *JASA, 77*(3), 1205-1216.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*(1), 1-30.

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech, 38*(3), 253-265.

Wells, B., Peppé, S., Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31,* 749-778.

Wennerstrom, A. (1991). *Techniques for teachers: A guide for nonnative speakers of English*. Ann Arbor: University of Michigan Press.

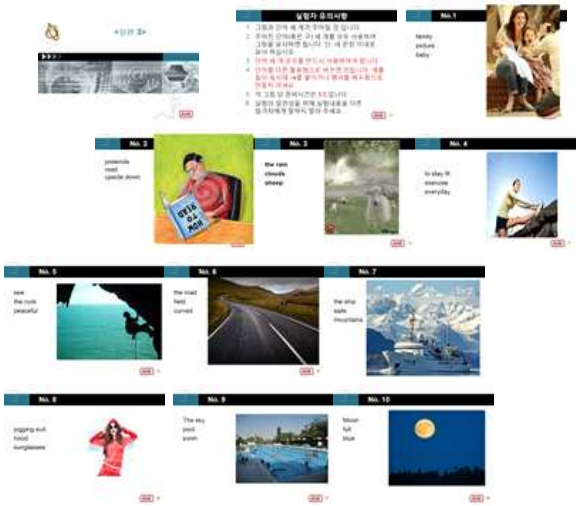
Wennerstrom, A. (1994). Intonational meaning in English discourse: A study of non-native speakers. *Applied Linguistics, 15*(4), 399-420.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendof, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries, *JASA, 91*(3), 1707-1717.

Willems, N.J. (1982). *English intonation from a Dutch point of view*. Dordrecht: Foris Publication.

Appendix

Experiment: Picture story telling



Received: 15th Octover, 2014
Reviewed: 14th November, 2014
Revised version accepted: 30th November, 2014