Check for updates

Post-focus compression is not automatically transferred from Korean to L2 English*

Jun Liu[1,2] · Yi Xu[3] · Yong-cheol Lee[1]**

[1]*Department of English Language and Literature, Cheongju University, Cheongju, Korea*
[2]*School of Foreign Language Education, Liaocheng University, Liaocheng, China*
[3]*Department of Speech, Hearing and Phonetic Sciences, University College London, London, United Kingdom*

Abstract

Korean and English are both known to show on-focus pitch range expansion and post-focus pitch range compression (PFC). But it is not clear if this prosodic similarity would make it easy for Korean speakers to learn English focus prosody. In the present study, we conducted a production experiment using phone number strings to examine whether Korean learners of English produce a native-like focus prosody. Korean learners of English were classified into three groups (advanced, intermediate and low) according to their English proficiency and were compared to native speakers. Results show that intermediate and low groups of speakers did not increase duration, intensity, and pitch in the focus positions, nor did they compress those cues in the post-focus positions. Advanced speakers noticeably increased the acoustic cues in the focus positions to a similar extent as native speakers. However, their performance in post-focus positions was quite far from that of native speakers in terms of pitch and excursion size. These results thus demonstrate a lack of positive transfer of focus prosody from Korean to English in L2 learning, and learners may have to relearn it from scratch, which is consistent with a previous finding. More importantly, the results provide further support for the view proposed in other works that acoustic properties of PFC were not easily transferred from one language to another.

Keywords: prosodic focus, post-focus compression, L2 prosody, Korean learners of English

1. Introduction

In many languages, a focused element is realized with increased duration, intensity, and pitch, as has been found for English (Cooper et al., 1985; Lee et al., 2015; Xu & Xu, 2005), German (Féry & Kügler, 2008), Arabic (Alzaidi et al., 2019), Turkish (Ipek, 2011), Japanese (Lee & Xu, 2012), and Korean (Lee & Xu, 2010). Also in these languages, post-focus elements are realized with lower pitch and weaker intensity relative to the same elements in a neutral focus condition, a phenomenon known as post-focus compression (PFC). Neither of these two features are universal, but PFC has been found to be the most consistent across these languages (Ipek, 2011; Xu,

2019). However, in many other languages, including Taiwanese (Southern Min) and Cantonese, PFC is absent while on-focus expansion of pitch, intensity or duration is sometimes detectable (see Xu et al. (2012) for more details). The uneven cross-linguistic distribution of focus prosody has led to the hypothesis that PFC is shared only by languages that are historically related (Xu, 2011; Xu et al., 2012). An important basis of this hypothesis is that PFC is a rarely occurring prosodic feature that does not easily transfer across languages. This is based on evidence that a) it is never transferred from a PFC language to a non-PFC language even by the same bilingual speakers (Wang et al., 2011; Wu & Chung, 2011), b) it is not easily acquired by L2 learners whose L1 has no PFC (Chen et al., 2012), and c) it is not easily acquired even by L2 learners whose L1 already has PFC (Chen, 2015). The third kind of evidence is particularly interesting as it is a case of lack of positive transfer common in L2 learning. Given, however, that it was only from a single study, there is a need to test how general it is the case. The present study will therefore examine whether there is positive transfer of PFC in Korean L2 learners of English.

Although both English and Korean have on-focus expansion and post-focus compression, the two languages have also been said to differ in how they realize prosodic focus. Because English has lexical stress, it is sometimes described as a stress-accent language, and so its prosodic focus is realized by a nuclear pitch accent on the primary stressed syllable in the focused word (de Jong, 2004; Ueyama & Jun, 1998). Korean has neither lexical stress nor lexical pitch accent (Jun, 1998), and so the prominence of a focused word is said to be marked by inserting a phrase boundary at the beginning of the focused word (Jun, 2011; Lee, 2012, 2017). As can be seen in Figure 1, however, although the focus-cueing difference is larger in magnitude in English than in Korean, there is no sign that a phrase boundary is inserted in Korean to mark focus. The pitch contour of the focus item is just an enlarged shape of what is already there in the neutral focus condition. On the other hand, when measured in terms of the maximum pitch of the target words between the two focus conditions (neutral vs. discourse-new focus), the difference was about 2.5 semitones in English and just about 1.2 semitones in Korean. Thus the size of prominence of a focused word shown here is more than twice larger in English than in Korean. Furthermore, in the production of Korean phone number strings, the difference between the two focus types was found to be about 1.0 semitones (Lee, 2015)—quite similar to the difference in Korean regular sentences in Figure 1.
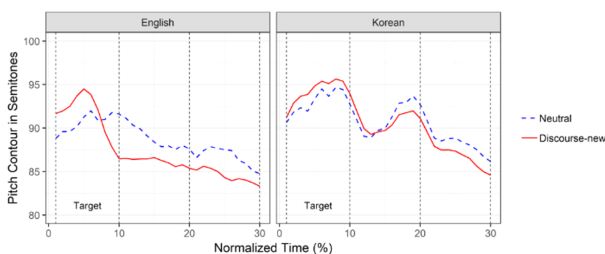

Figure 1. Time-normalized mean pitch contours of the sentence *Jonathan remembered Jessica* (left panel) and *Minsuga manduɾil mʌkninda* ('Minswu is eating dumplings') (right panel) produced by six speakers. The first area of each panel contains a target word (Modified from Lee (2009) for the left panel and from Lee & Xu (2010) for the right one).

Given the similarities and differences in focus prosody between Korean and English, this study aims to find out what happens to Korean learners of English as a second language. We will examine two alternative hypotheses. The first is that the learners will show positive transfer by producing focus with a Korean-like pattern of small on-focus expansion and post-focus compression, and then fully acquire English-like patterns when their L2 proficiency improves. This would be consistent with Lado's Contrastive Analysis Hypothesis (1957) that properties similar to first language (L1) are acquired relatively easily but those different from L1 are difficult to learn. Alternatively, there is no positive transfer, so that learners show neither on-focus expansion nor post-focus compression, and English-like focus prosody only starts to emerge when their L2 becomes proficient. This would be consistent with previous findings that learners' focus prosody improves with proficiency and level of usage of L2 (Chen, 2015; Chen et al., 2012; Swerts & Zerbian, 2010). Note here that, by transfer, we mean that L2 learners automatically reproduce a prosodic pattern in their L1. In order to examine the two hypotheses, we tested four language groups (native, advanced, intermediate, low) and conducted a production experiment with phone number strings. Among several focus types, corrective focus was elicited by correcting a wrong digit in a previous question and neutral focus was also used for comparison.

2. Method

2.1. Stimuli

The stimuli were 100 10-digit phone number strings, arranged in the form of (NNN)-(NNN)-(NNNN). The strings were generated randomly through a Python script by applying the following two criteria (Lee, 2015): (i) each digit (0–9) appears equally often (i.e., ten times) in each position of each digit string; and (ii) a combination of every two digits occurs equally often in each digit string. The phone number strings as target stimuli for focus prosody include the following advantages over regular sentences. Morphological and syntactic modifications are completely ruled out and more importantly, prosodic focus can be placed equally in every position within a digit string.

The 100 phone number strings were embedded in two focus conditions: neutral and corrective focus. The stimuli in the neutral-focus condition were produced in isolation as a background reading (i.e., Mary's number is 787-412-4699). The same sequences in the corrective-focus condition were presented in the form of question and answer (Q&A). A prerecorded question asks whether the phone number is correct and then each speaker responds by correcting one incorrect digit in the question (i.e., A: Mary's number is 887-412-4699. Right? B: No, Mary's number is 7̲87-412-4699).

2.2. Participants

Three groups of Seoul Korean learners of English, classified according to their scores on the Test of English for International Communication (TOEIC), participated in the experiment. The three groups were: an advanced group with TOEIC scores in the 900s; an intermediate group with TOEIC scores in the 700s; and a low group with TOEIC scores below 500. Each group included two females and three males, and all were undergraduates at Cheongju University at the time of recruitment, except for one advanced-level female speaker who came from Kyung Hee University in Seoul. None of
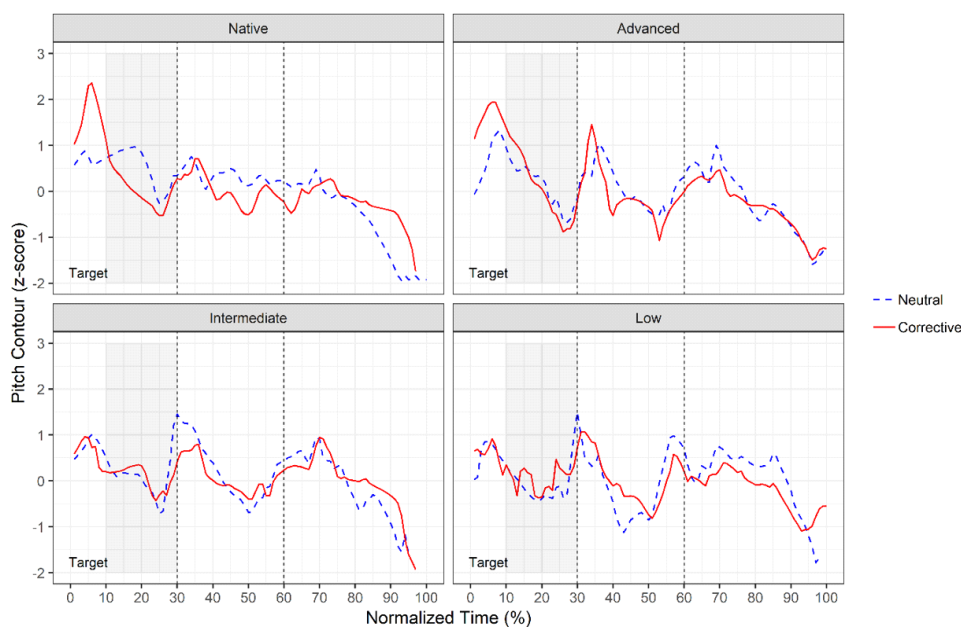
Figure 2. Time-normalized mean pitch contours of the digit string (637-686-7664) produced by five speakers in the two focus conditions, separated by language group.

the speakers reported a history of speaking or hearing problems. Each speaker received ₩10,000 (approximately USD10) after the experiment, as compensation for their participation. As control data, recordings of native English speakers (three females and two males)—produced in the same format as the current experiment's recordings—were borrowed from Lee (2015).

2.3. Recording Procedure

Recordings were conducted in a sound-attenuated booth at Cheongju University. Participants were seated in front of a laptop, wearing a head-mounted microphone, and were given a bottle of water. Stimuli were presented through PowerPoint slides in the middle of the laptop screen. To familiarize themselves with the experimental procedure and speech materials, speakers had a practice session with three trial phone number strings for each focus condition. After the practice session, neutral-focus recordings took place first, followed by the corrective-focus ones. In this experiment, participants were instructed to produce the target stimuli as naturally as possible and had a five-minute intermission between the two focus conditions. When a mistake was detected in the production of stimuli, speakers were instructed to repeat the digit string. We saved the recordings as *.wav* files on a laptop through Praat. The present study consisted of 4,000 digit strings in total: 3,000 digit strings (100 digit strings×5 speakers×2 focus types×3 groups) from the current experiment and 1,000 digit strings borrowed from Lee (2015).

2.4. A Sketch of Pitch Contours

We first describe sample pitch contours to look at the global prosodic patterns of focus and PFC, separated by focus type (neutral, corrective) and language group (native, advanced, intermediate, low). In the current study, we used ProsodyPro (Xu, 2013) to extract pitch contours at ten equidistant time points of each labeled digit in a digit string. Because the increase of pitch in hertz (Hz) is nonlinear, in contrast to the linear increase in semitones (Nolan, 2003), the pitch in Hz was converted to semitones (st) by applying the following

formula (Lee et al., 2018; Lee et al., 2016), st=$12log_2x$, where x indicates a raw value (Hz), and a reference value is 1 Hz. The semitone scale was then normalized with z-scores for two reasons. Speakers always first produced the target stimuli in neutral focus. And, more importantly, a low group of speakers grew more tired and nervous over time during the experiment. This seemed to affect their performance in producing the stimuli embedded in corrective focus. Therefore, in order to counterbalance the order of presenting the stimuli and to offset speakers' tiredness or nervousness in the production over time, semitones were transformed into z-scores independently by each speaker and each digit string using the following formula, $z=((x_i-M_i)/SD_i)$, in which x_i is a raw value for a certain digit string by each speaker and M_i and SD_i refer to the mean and standard deviation of each digit string for each speaker.

Figure 2 displays time-normalized pitch contours in the two focus conditions, averaged over the digit string (637-686-7664) produced by the five speakers in each group, in which the word "target" indicates a focus position and the dotted line represents a phrase boundary in a digit string. From Figure 2, we observe that the corrective focus condition shows a higher pitch peak in the focus position and a lower pitch valley in the post-focus positions in both the native and the advanced groups, relative to the neutral-focus condition. However, on a closer examination, pitch raising in the focus position and pitch lowering in the post-focus positions seem to be much greater in the native group than in the advanced group in the same phrase to which the target digit belongs. Put it differently, the degree of prosodic modulation by focus is greater in the native group than in the advanced group when the two focus conditions were directly compared. In contrast, both intermediate and low groups of speakers show no such clear indication of prosodic changes in the focus and post-focus positions.

2.5. Acoustic Measurements

After the visual inspection of Figure 2, we obtained acoustic measurements separated by focus positions: focus and post-focus. In
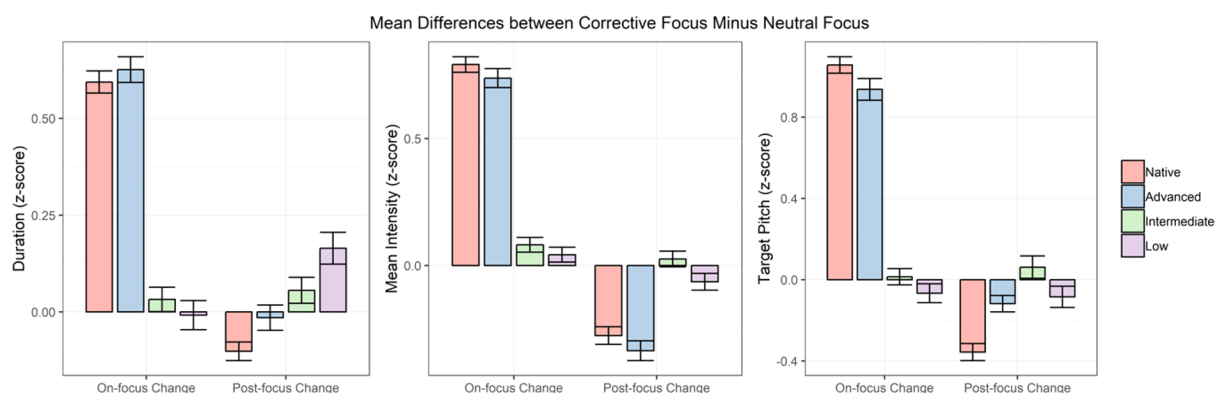
Figure 3. On-focus and post-focus changes by z-scores of duration, intensity, and pitch by four language groups. Bars represent mean differences calculated by corrective focus minus neutral focus and error bars indicate standard errors.

the focus positions, we measured duration (ms), mean intensity (dB) and maximum pitch (Hz). In the post-focus positions, duration (ms), mean intensity (dB) and minimum pitch (Hz) were calculated. Duration and mean intensity were measured directly from each labeled digit using ProsodyPro (Xu, 2013). Z-scores of maximum pitch and minimum pitch were manually obtained from focus positions and post-focus positions, respectively, within pitch contours. Among several pitch-related parameters, maximum pitch was chosen to estimate the pitch peak for focus marking and minimum pitch was selected because it is indicative of PFC. For the same reasons above, duration and intensity were also transferred to z-scores. We also calculated excursion size (z-score) calculated by the difference between maximum pitch in the focus positions and minimum pitch in the post-focus positions because this parameter is expected to signal a pitch trajectory from the peak of on-focus expansion to the valley of post-focus compression.

Before moving onto the next section, it should be noted that we limited post-focus positions to only those within phrases probably due to the possibility that pitch resets at phrase boundaries. As shown in Figure 2, the two focus types seem to converge with each other after the first phrase boundary, regardless of language group. This convergence suggests that PFC can be best expressed within phrases in the case of digit strings. Furthermore, we examined each utterance by spectrogram reading and listening to verify if speakers actually did phrase at the hyphens. Out of 4,000 digit strings, we found only one case (673-**5**29-8998) with no prosodic boundary at the hyphen after the focused digit ("5", the fourth digit). Nevertheless, to achieve consistency across all the digit strings assessed, only "29" were treated as post-focus digits. This treatment is not expected to yield any negative effect on our analysis, because the two digits ("29") are post-focus digits in the digit string. Accordingly, when position 1 was in focus in each digit group of the digit string $(N_1N_2N_3)$-$(N_1N_2N_3)$–$(N_1N_2N_3N)$, positions 2 and 3 were considered post-focus positions. When position 2 was focused, only position 3 was included as a post-focus position. When position 3 was focused, none of the positions were included as post-focus positions. This method was equally applied to the second and third phrases and the tenth position was excluded for further analysis because it is the position showing a falling pitch contour in a declarative sentence.

3. Analyses and Results

Figure 3 exhibits both on-focus and post-focus changes in z-scores of duration, intensity, and pitch, divided by language group. From the figure, we observe that native and advanced groups produced clearly increased on-focus changes in duration, intensity and pitch, and showed an opposite pattern of results in the post-focus positions. Intermediate and low groups, however, seemed to lack such systematic changes by the three acoustic cues in the focus and post-focus positions. In order to examine whether there is a significant difference in the on-focus and post-focus changes across the language group, we conducted a linear mixed-effects model analysis through the *lmerTest* package (Kuznetsova et al., 2013) in R, separately for on-focus and post-focus changes. In the model, there was one fixed effect: language group (native, advanced, intermediate, low) and speaker (5 speakers), string position (1−10), and digit (0−9) were treated as random effects. Dependent variables were aggregated measures of duration, mean intensity, and maximum pitch for the on-focus changes. For the post-focus changes, the aggregated measures of duration, mean intensity, and minimum pitch were used as dependent variables. Unless otherwise stated, we label the maximum pitch and minimum pitch as "target pitch" for the sake of simplicity. An *Anova* function of the *lmerTest* package was implemented to obtain the significance level of the fixed factor. Furthermore, a series of Tukey's tests with Bonferroni correction followed to compare multiple pairs, using the *mcp* function of the *multcomp* package (Hothorn et al., 2008) in R. In what follows, we describe the statistical results of on-focus and post-focus changes in turn.

In regard to the on-focus changes, the effect of language group was significant for all acoustic cues (duration: X^2=261.67, *df*=3, *p*<0.001; mean intensity: X^2=413.92, *df*=3, *p*<0.001, target pitch: X^2=433.32, *df*=3, *p*<0.001). As Table 1 indicates, multiple comparisons of the on-focus changes reveal that native and advanced groups showed similar trends for all acoustic measures, although the advanced group actually produced a significantly longer duration than did the native group. Also, intermediate and low groups can be considered together. More specifically, native and advanced groups produced the longest duration, greatest intensity, and highest pitch across the language group in the focus positions. Intermediate and low groups, on the other hand, demonstrated shorter duration, weaker intensity, and lower pitch in the focus positions than did the native and advanced groups.

Table 1. Pairwise Tukey's tests with Bonferroni correction for on-focus changes by duration, mean intensity, and target pitch

		Native	Advanced	Intermediate
Duration	Advanced	**		
	Intermediate	***	***	
	Low	***	***	0.81
Mean intensity	Advanced	0.99		
	Intermediate	***	***	
	Low	***	***	0.80
Target pitch	Advanced	0.33		
	Intermediate	***	***	
	Low	***	***	0.58

** $p<0.01$, *** $p<0.001$.

Moving onto post-focus changes, language group also had a significant effect on all acoustic measures (duration: $X^2=24.62$, $df=3$, $p<0.001$; mean intensity: $X^2=84.78$, $df=3$, $p<0.001$, target pitch: $X^2=40.51$, $df=3$, $p<0.001$). The results of the multiple comparisons for the post-focus changes were quite different from those found in the on-focus changes, except for mean intensity. As shown in Table 2, native and advanced groups yielded significantly reduced intensity values compared to the other groups. For duration, the native group showed the shortest duration across the language group —although the difference between native and advanced groups did not reach the significance level, and the difference between native and intermediate groups just slightly missed the significance level. No significant difference in duration was evident between advanced and intermediate groups, nor between intermediate and low groups. But the advanced group produced a significantly shorter duration than the low group. With pitch, the native group showed the most compressed pitch in the post-focus positions. There was no significant difference in pitch between advanced and low groups, nor between intermediate and low groups. But a significant difference in the use of pitch was noted between advanced and intermediate groups in the post-focus positions.

Table 2. Pairwise Tukey's tests with Bonferroni correction for post-focus changes by duration, mean intensity, and target pitch

		Native	Advanced	Intermediate
Duration	Advanced	0.73		
	Intermediate	0.06	0.42	
	Low	***	***	0.09
Mean intensity	Advanced	1.00		
	Intermediate	***	***	
	Low	***	***	0.23
Target pitch	Advanced	***		
	Intermediate	***	*	
	Low	***	0.96	0.13

* $p<0.05$, *** $p<0.001$.

Let us now turn to Figure 4, which shows z-scores of excursion size for the four language groups. The values in the figure were calculated and averaged by the difference between maximum pitch in the focus position and minimum pitch in the post-focus positions. We observe that the native group seems to produce the greatest excursion size in the shift from the focus to the post-focus positions, followed by the advanced group. The intermediate and low groups do not appear to differ in pitch excursion size. In order to statistically confirm the visual observation, a linear mixed-effects model analysis using the same package as above was performed using language

group as an independent factor, speaker and digit as random effects, and excursion size as a dependent variable. The result demonstrated a significant effect of language group on excursion size ($X^2=229.13$, $df=3$, $p<0.001$). Post-hoc multiple comparisons with Bonferroni correction were run through the *mcp* function of the package stated above. The results (Table 3) showed the ordering of native> advanced>intermediate, low, from the greatest to the least excursion size, where the symbol '>' indicates a significant difference ($p<0.05$).
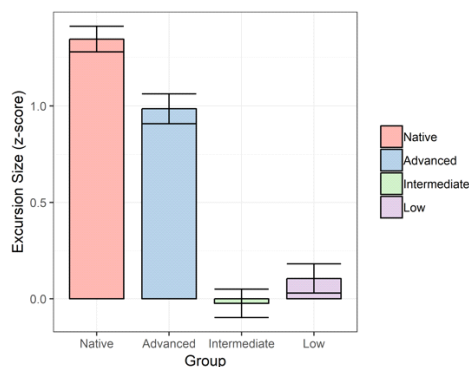

Figure 4. Mean of excursion size by the four language groups. Bars represent mean differences calculated by on-focus maximum pitch minus post-focus minimum pitch and error bars indicate standard errors.

Table 3. Pairwise Tukey's tests with Bonferroni correction for excursion size

		Native	Advanced	Intermediate
Excursion size	Advanced	**		
	Intermediate	***	***	
	Low	***	***	0.59

** $p<0.01$, *** $p<0.001$.

4. Discussion and Conclusion

This study examined whether Korean learners of English showed a native-like performance in a situation where they produced corrective focus in reading a phone number string. Korean learners of English were classified into three groups (advanced, intermediate, low) by their English proficiency and were compared with native groups. Results demonstrated that intermediate and low groups did not clearly express corrective focus relative to neutral focus in both focus and post-focus positions. The advanced group yielded clear prosodic effects of focus comparable to native speakers' focus marking, but their production of PFC was not completely equivalent to that of native speakers, in terms of pitch, and excursion size.

These results therefore provide support for the second hypothesis presented in the Introduction, namely, there is no positive transfer of focus prosody from L1 Korean to L2 English, because neither on-focus expansion nor post-focus compression was produced by intermediate and low groups. This is despite the fact that Korean also shows both on-focus expansion and post-focus compression, though in a weaker form (Lee & Xu, 2010). Instead, Korean learners may have to relearn a focus prosody that is similar to what they already have in their L1, because English-like focus prosody started to emerge only in advanced L2 learners. And even for advanced learners, PFC is still less clear than native English speakers, probably because they have not yet reached full proficiency in their L2.

That is, they did not fully compress all of the acoustic cues after focus, compared to native speakers. More specifically, although they reduced both duration and intensity cues in the post-focus positions, pitch-related parameters including pitch and excursion size were not compressed fully to the extent that native speakers demonstrated. This finding is in line with the view of previous studies (Chen et al., 2012; Wu & Chung, 2011), that PFC does not transfer well from one language to another. This result also enables us to claim that PFC is acquired later than on-focus expansion.

Although the method employed in this study was effective to test our goals, several limitations of the current study need to be considered in the design of future work. First, since digit strings were employed for experimental materials, it is not clear whether our findings would be generalized to regular sentences. Future research needs to be done for comparison with regular sentences. The second limitation concerns the recording procedure, in which neutral-focus recordings always preceded corrective-focus ones for all the speakers. Therefore, we need to switch the order of stimuli for some speakers in order to control for the effect of order. Finally, more speakers will help enhance the reliability and validity of future research.

In conclusion, the current study examined whether Korean learners of English produced a native-like prosodic marking of focus. Our findings revealed that the cut-off point was set quite late between the advanced group and the other learner groups in their acquisition of L2 focus prosody. Both low and intermediate groups did not show on-focus expansion and PFC, indicating that clear interference of L1 was observed in their L2 prosody. Advanced speakers showed clear on-focus expansion but failed to successfully compress all the acoustic cues after focus. This suggests that language proficiency played an important role in the acquisition of L2 prosody to some extent but a certain phenomenon (i.e., PFC in this study) was more difficult to acquire. For future research, it would be interesting to see whether advanced Korean learners of other PFC languages, such as Mandarin, would have similar effects of on-focus expansion and PFC.

References

Alzaidi, M. S., Xu, Y., & Xu, A. (2019). Prosodic encoding of focus in Hijazi Arabic. *Speech Communication*, *106*, 127-149.

Chen, Y. (2015). Post-focus compression in English by Mandarin learners. *Proceedings of the 18th International Congress of Phonetic Sciences* (pp. 0287: 1-5). Glasgow, UK.

Chen, Y., Guion-Anderson, S., & Xu, Y. (2012). Post-focus compression in second language Mandarin. *Proceedings of the Speech Prosody 2012* (pp. 410-413). Shanghai, China.

Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *The Journal of the Acoustical Society of America*, *77*(6), 2142-2156.

de Jong, K. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, *32*(4), 493-516.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, *36*(4), 680-703.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346-363.

Ipek, C. (2011). Phonetic realization of focus with no on-focus pitch range expansion in Turkish. *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 140-143). Hong Kong.

Jun, S. A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, *15*(2), 189-226.

Jun, S. A. (2011). Prosodic markings of complex NP focus, syntax, and the pre-/post-focus string. *Proceedings of the 28th West Coast Conference on Formal Linguistics (WCCFL 28)* (pp. 214-230). Los Angeles, CA.

Kuznetsova, A., Brockhoff, P. B., & Christensen, H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package) (R Package Version) [Computer software]. Retrieved from http://cran.r-project.org/web/packages/lmerTest/index.html

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

Lee, A., & Xu, Y. (2012). Revisiting focus prosody in Japanese. *Proceedings of the Speech Prosody 2012* (pp. 274-277). Shanghai, China.

Lee, Y. (2009). *The phonetic realization of contrastive focus and its neighbors in Korean and English: A cross-language study* (Master's thesis). Hannam University, Korea.

Lee, Y. (2012). Prosodic correlation between the focusing adverb *ozik* 'only' and focus/givenness in Korean. *Journal of Speech Sciences*, *2*(2), 85-111.

Lee, Y. (2015). *Prosodic focus within and across languages* (Doctoral dissertation). University of Pennsylvania, Philadelphia, PA.

Lee, Y. (2017). Prosodic focus in Seoul Korean and South Kyungsang Korean. *Linguistic Research*, *34*(1), 133-161.

Lee, Y., & Xu, Y. (2010). Phonetic realization of contrastive focus in Korean. *Proceedings of the Speech Prosody 2010* (pp. 100033: 1-4). Chicago, IL.

Lee, Y., Nambu, S., & Cho, S. (2018). Focus prosody of telephone numbers in Tokyo Japanese. *Journal of the Acoustical Society of America*, *143*(5), EL340-EL346.

Lee, Y., Wang, B., Chen, S., Adda-Decker, M., Amelot, A., Nambu, S., & Liberman, M. (2015). A crosslinguistic study of prosodic focus. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015)* (pp. 4754-4758). Brisbane, Australia.

Lee, Y., Wang, T., & Liberman, M. (2016). Production and perception of tone 3 focus in Mandarin Chinese. *Frontiers in Psychology*, *7*(1058), 1-13.

Nolan, F. (2003). Intonational equivalence: An experimental evaluation of pitch scales. *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 771-774). Barcelona, Spain.

Swerts, M., & Zerbian, S. (2010). Prosodic transfer in Black South African English. *Proceedings of the Speech Prosody* 2010 (pp. 100198: 1-4). Chicago, IL.

Ueyama, M., & Jun, S.-A. (1998). Focus realization in Japanese English and Korean English intonation. *Japanese and Korean Linguistics 7* (pp. 629-645). Los Angeles, CA.

Wang, B., Wang, L., & Qadir, T. (2011). Prosodic realization of focus in six languages/dialects in China. *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 144-147). Hong Kong, China.

Wang, B., Xu, Y., & Ding, Q. (2018). Interactive prosodic marking of focus, boundary and newness in Mandarin. *Phonetica*, *75*(1), 24-56.

Wu, W. L., & Chung, L. (2011). Post-focus compression in English-

Cantonese bilingual speakers. *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 148-151). Hong Kong, China.

Xu, Y. (2011). Post-focus compression: Cross-linguistic distribution and historical origin. *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 152-155). Hong Kong, China.

Xu, Y. (2013). ProsodyPro － A Tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody* (pp. 7-10). Aix-en-Provence, France.

Xu, Y. (2019). Prosody, tone and intonation. In W. F. Katz, & P. F. Assmann (Eds.), *The Routledge handbook of phonetics* (pp. 314-356). New York, NY: Routledge.

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, *33*(2), 159-197.

Xu, Y., Chen, S., & Wang, B. (2012). Prosodic focus with and without post-focus compression: A typological divide within the same language family? *The Linguistic Review*, *29*(1), 131-147.

• **Jun Liu**
Ph.D. Candidate, Dept. of English Language and Literature,
Cheongju University
298 Daesung-Ro, Cheongwon-Gu
Cheongju, 28503, South Korea
Tel: +82-10-4724-8312
Email: 1748046227@qq.com
Areas of interest: Prosody, L2 acquisition

• **Yi Xu**
Professor, Dept. of Speech, Hearing and Phonetic Sciences,
University College London
Chandler House, 2 Wakefield Street
London WC1N IPF, UK
Tel: 44-20-7679-4082
Email: yi.xu@ucl.ac.uk
Areas of interest: Prosody, Prosody modeling

• **Yong-cheol Lee** Corresponding author
Assistant Professor, Dept. of English Language and Literature,
Cheongju University
298 Daesung-Ro, Cheongwon-Gu
Cheongju, 28503, South Korea
Tel: +82-43-229-8351
Email: soongdora@gmail.com
Areas of interest: Prosody, L2 acquisition