THE ACQUISITION OF KOREAN PROSODIC PROMINENCE BY CANTONESE-ENGLISH BILINGUALS

Xinran Ren and Peggy Mok

The Chinese University of Hong Kong
renxinr@link.cuhk.edu.hk, peggymok@cuhk.edu.hk

ABSTRACT

A prosodically prominent syllable, namely stress, can be contrastive or non-contrastive regarding lexical meanings. Many previous studies predict that learners with fixed-stress L1s will be "stress deaf" to languages with contrastive stress. This study tested the reverse: how speakers of L1 Cantonese (lexical tone), L2 English (lexical stress) produced L3 Korean without lexically contrastive stress but phrasal pitch accent. Both the learners and the natives were examined in producing Korean trisyllabic non-words of varying syllable weight, in isolation and in accentual phrases. Results showed that the learners made the 2nd syllable more prominent like the natives, but lacking boundary effects associated with the accentual phrase. This suggests that learners with word-level prominence can detect the location of Korean stress but to a different domain. Syllable weight effects and the relative use of different acoustic correlates of stress were discussed in relation to language transfer and general prosodic development.

Keywords: stress, Korean Accentual Phrase, L3 acquisition, Cantonese, English

1. INTRODUCTION

In the literature on adult L2 acquisition, many studies examined the challenges when some new features absent in the native languages need to be 'added'. However, our understanding of the situation where certain features from native languages should be 'reduced' is limited to only a few studies. It is not necessarily an easy task, for example, English speakers were found to adopt an overly dynamic pitch movement for L2 German production [1].

In the current study, we aimed to examine such a situation through the acquisition of Korean prosodic prominence by Cantonese-English bilinguals. Cantonese is a tone language, using pitch levels and pitch contours to distinguish lexical meanings. English uses contrastive stress at the word level with various acoustic attributes: duration, intensity, f0 and vowel quality [2, 3, 4]. While both Cantonese and English have word-level prominence which is unpredictable and contrastive in meanings, Korean does not use stress contrastively at the word-level. In Korean, prosodic prominence is associated with the accentual phrase (AP) [5, 6], a prosodic unit smaller than an intonational phrase (IP) but larger than a phonological word, and it does not affect meanings. Korean APs have a default pitch pattern LHLH (HHLH if the AP-initial segment is tense or aspirated) and if the AP contains less than four syllables, one or two of the middle tones will be undershot [6]. As for the phrase-level prominence in Korean, production and perception studies [5, 7] showed that the 2nd syllable in APs was usually perceived as prominent linking to the middle H tone of the LHLH pattern, though controversy still exists for the location of Korean prominence. The prominent syllable is realized with higher f0, increased intensity and longer duration, but these acoustic patterns will be influenced by boundary conditions (IP, AP). Also, prominence assignment was found to be affected by syllable weight, namely, heavy first syllables (CVC) tend to attract stress.

Theoretically, the Stress Parameter Model (SPM) [8] proposes that learners who do not use stress contrastively in their L1 will experience great difficulties in learning a language with lexical stress, due to their "stress deafness". This study tested the reverse: whether the learners of L1 Cantonese (lexical tone) and L2 English (lexical stress) can notice the Korean non-contrastive stress at the AP level, which is realized phonetically but not used phonologically. Importantly, from the perspective of third language acquisition, the language combination adopted in this study is interesting because with L1 Cantonese and L3 Korean typologically closer, it separates typological distance [9] and L2 status [10] as two possible factors influencing language transfer at L3 initial state. We try to figure out whether transfer effects or universal features of non-native prosodic development can account for the learners' production patterns of Korean prosodic prominence.

2. METHOD

2.1. Subjects and materials

5 native speakers of Seoul Korean (4 Females) and 15 native Cantonese university students (14 Females) who have been learning English as their L2 and

Korean as their L3 participated in the production experiment. The learners were highly proficient in English with an age of acquisition around 2-3 years old. They have taken Korean classes at the university for 3 terms (120 hours in total) at the time of testing.

Participants produced 3 repetitions of trisyllabic Korean non-words differing in syllable weight: /pa.ma.ta/, /pa.ma.taŋ/, /pa.man.ta/, /pa.man.taŋ/, /pal.ma.ta/, /pal.ma.taŋ/, /pal.man.ta/, /pal.man.taŋ/, in isolation (IP) and in accentual phrases (AP) (*[AP_] ponɛtɕuseyo*, 'Please send me _.') respectively. The materials were presented in hangul, the Korean written form, and the participants were asked to read as if the non-words were some unknown objects.

2.2. Data analysis

For each token, syllable duration (in ms), average intensity (in dB) and average f0 (in semitone re 100Hz) over the entire syllable were measured using Praat. Boundary of the 3 syllables in a word were defined by the acoustic transition between adjacent segments and for stops, its onset was set at the start of the stop gap silence. F0 measures used 100-500 Hz pitch range for female speakers and 75-300 Hz for males. Using a Praat script ProsodyPro [11], pulse correction was done for abnormal pitch contours.

Some f0 data need to be discarded mostly due to severe creakiness towards the end of IPs. Since the 3 syllables were compared within each token in different conditions, once the acoustic values of any one syllable is missing, the other 2 syllables should be excluded from analysis as well. By this criterion, 2 tokens (0.9%) for the native Koreans and 98 tokens (15%, 90.8% of the excluded tokens are IPs) for the learners were excluded. Consequently, we only examined f0 values at the AP level.

3. RESULTS

Figures 1, 2 and 3 show mean values for each acoustic variable (duration, intensity and f0) respectively. The 4 panels in the figure were divided by different groups (learner vs. native) and levels (IP vs. AP), and each panel shows how the acoustic values change with syllable weight. Standard error bars and mean values were added in the figures as well.

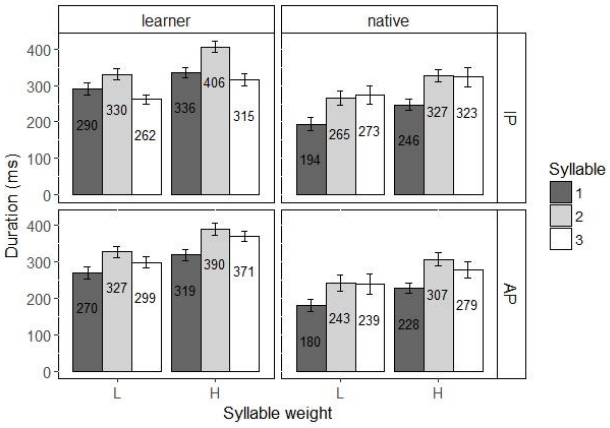
To statistically validate the patterns shown in the figures, linear mixed-effects regression analyses were performed using the lme4 package in R. In the models predicting the acoustic values, fixed effects contain Syllable position (1st, 2nd, 3rd syllable), Weight (light, heavy syllable), Level (IP, AP) and Group (learner, native). In the trisyllabic words, the 2nd syllable was chosen as the intercept, since it was claimed in the literature to be the prominent one, so as to compare with the 1st and 3rd syllable. By-subject random intercepts and all possible slopes were added to the models as random factors [12].

Data were analysed in 2 steps following [13]: first, two separate models were run for the learners and the native group with Syllable position, Weight, Level and their interactions as fixed effects; the next step was to run a full model adding all 4 fixed effects, with the learners mapped on the intercept. This will show whether the L3 Korean learners behave differently from the native group. In Tables 1 (duration), 2 (intensity) and 3 (f0), the first four columns display the outcomes of the analyses for each language group, and the last two columns show whether the difference between the learners and the native speakers is significant in the full model.

3.1. Duration

As is shown in Figure 1, generally, the learners produced the 2nd syllable as the longest while native Korean speakers evenly produced longer 2nd and 3rd syllables. Syllable weight (Heavy) had a lengthening effect on all the syllables for both language groups.

Figure 1: Duration of trisyllabic words produced in IP and AP.



The differences between the learners and the natives were revealed through the results of the mixed-effects analyses in Table 1: (1) the overall duration of syllable 2 was longer in the learners, indicating a slower speaking rate; (2) At the IP level, syllable 3 was shorter than syllable 2 in learners but not different from syllable 2 in the natives; (3) Level and Weight affected the learners' 3rd syllable differently from the natives, i.e., the learners' 3rd syllable was lengthened to be as long as the 2nd syllable in AP ($\beta = 28.33$, SE = 14.99, $t = 1.89$, $p = 0.44$), and the learners' 3rd syllable was lengthened more by heavy syllables in AP compared with IP. For the natives, they showed no effect of level and level × weight, but syllable weight lengthened different syllables equally.

Table 1: Duration (the intercepts estimated the duration values produced in light syllable 2 at IP level).

	L3 learners		Korean native speakers		Learners vs. Korean native	
	B (SE)	t value	B (SE)	t value	B (SE)	t value
(Intercept)	330.12 (12.45)	26.51***	265.3 (19.47)	13.62***	-64.81 (28.08)	-2.31*
Syllable1	-39.97 (9.54)	-4.19***	-70.92 (9.24)	-7.68***	-31.14 (17.53)	-1.78 .
Syllable3	-68.13 (14.99)	-4.54***	7.34 (16.95)	0.43	75.5 (27.99)	2.70*
WeightH	75.8 (6.64)	11.41***	61.85 (7.44)	8.32***	-13.98 (12.32)	-1.14
LevelAP	-2.97 (8.79)	-0.34	-23.08 (14.25)	-1.62	-20.09 (17.29)	-1.16
Syllable1:WeightH	-29.87 (7.51)	-3.98***	-9.75 (10.13)	-0.96	20.3 (14.3)	1.42
Syllable3:WeightH	-23.83 (7.52)	-3.17**	-12.13 (10.13)	-1.2	11.64 (14.3)	0.81
Syllable1:LevelAP	-16.91 (7.51)	-2.25*	8.35 (10.11)	0.83	25.39 (14.28)	1.78 .
Syllable3:LevelAP	39.79 (7.49)	5.31***	-8.91 (10.13)	-0.88	-48.69 (14.29)	-3.41***
WeightH:LevelAP	-12.47 (7.51)	-1.66 .	2.37 (10.11)	0.24	15.01 (14.28)	1.05
Syllable1:WeightH:LevelAP	15.15 (10.62)	1.43	-6.82 (14.3)	-0.48	-22.28 (20.19)	-1.1
Syllable3:WeightH:LevelAP	32.41 (10.62)	3.05**	-13.04 (14.31)	-0.91	-45.51 (20.19)	-2.25*

Significance level: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 2: Intensity (the intercepts estimated the intensity values produced in light syllable 2 at IP level).

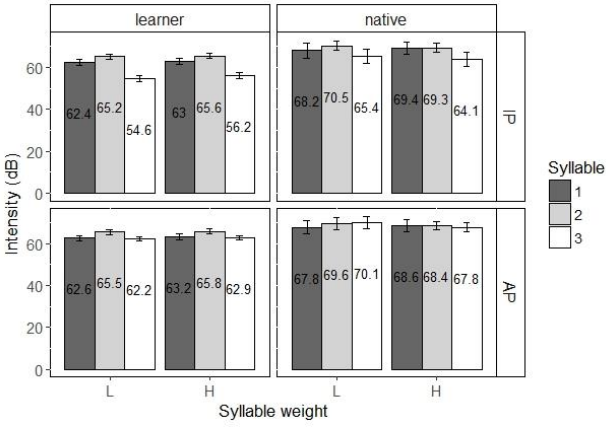
	L3 learners		Korean native speakers		Learners vs. Korean native	
	B (SE)	t value	B (SE)	t value	B (SE)	t value
(Intercept)	65.22 (1.16)	56.01***	70.49 (3.89)	18.15***	5.28 (2.65)	2.00 .
Syllable1	-2.86 (0.53)	-5.40***	-2.22 (1.35)	-1.64	0.64 (1.18)	0.54
Syllable3	-10.6 (0.5)	-21.35***	-5.16 (0.95)	-5.42**	5.45 (1.01)	5.37***
WeightH	0.34 (0.45)	0.76	-1.19 (0.73)	-1.64	-1.53 (0.87)	-1.76 .
LevelAP	0.26 (0.25)	1.03	-0.87 (0.5)	-1.73 .	-1.13 (0.78)	-1.44
Syllable1:WeightH	0.32 (0.35)	0.89	2.31 (0.59)	3.92***	2.01 (0.68)	2.94**
Syllable3:WeightH	1.17 (0.35)	3.31***	-0.24 (0.59)	-0.4	-1.41 (0.68)	-2.06*
Syllable1:LevelAP	-0.06 (0.35)	-0.16	0.45 (0.59)	0.76	0.5 (0.68)	0.73
Syllable3:LevelAP	7.35 (0.35)	20.82***	5.67 (0.59)	9.59***	-1.7 (0.68)	-2.50*
WeightH:LevelAP	0.02 (0.35)	0.07	0.02 (0.59)	0.03	-0.04 (0.68)	-0.05
Syllable1:WeightH:LevelAP	-0.05 (0.5)	-0.1	-0.39 (0.83)	-0.47	-0.32 (0.96)	-0.33
Syllable3:WeightH:LevelAP	-0.87 (0.5)	-1.73 .	-0.86 (0.83)	-1.03	0.04 (0.96)	0.05

Significance level: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

3.2. Intensity

Combining the results from Figure 2 and Table 2, it can be found that the learners consistently produced greater intensity in syllable 2 in both IP and AP (light syllable 2 and 3 in AP: $\beta = 3.25$, SE = 0.50, t = 6.55***), whereas the natives did not employ the intensity cue for prominence production. At the IP level, the natives did produce slightly lower intensity in syllable 3, but this should be caused by the diminishing effect on intensity in IP-final position [14]. For the learners, they intensified syllable 3 with heavy syllables and AP level, while this effect was absent or less strong for the natives.

Figure 2: Intensity produced in IP and AP.



3.3. F0

Contrary to intensity, when it comes to f0, the native speakers produced the 3rd syllable in AP with a markedly higher f0, while the learners used f0 to a much lesser extent. Considering the effect of syllable weight, the difference between the learners and the natives was not large enough to reach statistical significance.

Figure 3: F0 produced in AP.

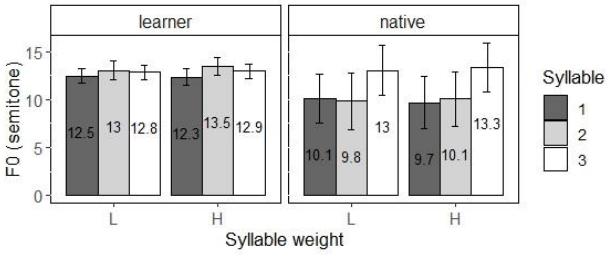


Table 3: F0 in AP (the intercepts estimated the f0 values produced in light syllable 2).

	L3 learners		Native Korean speakers		Learners vs. Korean native	
	B (SE)	t value	B (SE)	t value	B (SE)	t value
(Intercept)	13.01 (0.96)	13.58***	9.78 (2.87)	3.41*	-3.23 (2.3)	-1.4
Syllable1	-0.53 (0.24)	-2.18*	0.32 (0.41)	0.77	0.84 (0.47)	1.77 .
Syllable3	-0.27 (0.5)	-0.54	3.24 (0.37)	8.7***	3.5 (0.88)	3.97***
WeightH	0.47 (0.11)	4.14***	0.27 (0.19)	1.41	-0.19 (0.21)	-0.9
Syllable1:WeightH	-0.67 (0.15)	-4.39***	-0.69 (0.22)	-3.09**	-0.02 (0.29)	-0.07
Syllable3:WeightH	-0.28 (0.15)	-1.82 .	-0.02 (0.22)	-0.07	0.26 (0.29)	0.91

To summarize the group differences in terms of duration, intensity and f0 values, the learners' 2nd syllable was produced as the most prominent one (duration and intensity), whereas the natives' 2nd and/or 3rd syllables were more prominent (duration and f0). Noticeably, the learners can be seen as not using f0 and the natives as not using intensity comparing with the other group. Additionally, the learners showed a trend of strengthening the duration and intensity of the 3rd syllable with AP and with heavy syllable weight.

4. DISCUSSION

Comparing with the natives, the learners successfully detected the location of Korean stress and produced the 2nd syllable as the prominent one consistently across different conditions. However, the learners showed no boundary effects associated with IPs or APs, as can be seen in the natives' production: pre-boundary lengthening reflected by the 3rd syllables produced as long as the 2nd ones, and a rising boundary tone in APs reflected by the highest f0 values for the 3rd syllables. Besides, the learners showed a relatively flat pitch contours in producing the 3 syllables. Since APs in Korean starting with a lax segment are defined by a LHLH tonal pattern and featured by a rising boundary tone and pre-boundary lengthening [15], the L3 learners were not producing a proper AP. Therefore, they did not produce Korean stress in the domain of AP. Instead, they might simply produce a fixed stress at the word level. The failure to associate Korean prosodic prominence to the AP level is probably because the learners have not established the prosodic unit of AP for their L3 Korean. This situation may change after the learners have formed the structure of AP in their Korean phonological system.

Returning to the Stress Parameter Model, our results, as a supplement to the model, illustrated that the task of detecting and producing non-contrastive phonetic prominence is not too difficult, at least for speakers with experience of using prosodic prominence contrastively.

The learners' production patterns will be explained in terms of L3 transfer as well as the universal features of non-native acquisition of prosody.

As for the transfer effects, the learners' L3 may probably be influenced by their previous languages in that the learners consistently used intensity as a cue for prominence, while this cue was not adopted by the native Korean speakers. It was likely that the learners transferred their native correlates for prominence to L3, e.g., intensity, but it is difficult to tease apart L1 and L2 as possible sources of L3 transfer here.

Surprisingly, no obvious transfer of f0 correlate can be observed. The L3 learners in this study, whose native language is a complex tone language, did not produce AP-boundary tones properly and did not adopt f0 cue as much as the other cues for the production of Korean prosodic prominence. This is consistent with several studies such as [16] indicating no advantages for Mandarin speakers to learn f0 as a cue for Korean stop contrast.

The limited use of f0 in L3 Korean may reflect a universal trend of a "reset phrase" where F0 activity is extremely constrained, probably because learners largely attend to grammar and lexicon [17]. In addition, the learners showed a tendency towards learning boundary effects (i.e. 3rd syllables strengthened) in terms of duration and intensity with heavy syllable weight and with AP, indicating that acquisition firstly happens in strong positions.

In conclusion, the current study showed that language speakers with word-level contrastive prominence can detect and produce the non-contrastive phrase-level prominence in terms of duration and intensity in a foreign language. However, it was surprising to find that the experience with a tone language did not help the learning of f0 as a cue for non-contrastive prominence. To possibly confirm the source of L3 initial transfer in the acquisition of prosodic prominence, further research will be conducted to examine the learners' L1 and L2 at the individual level.

6. REFERENCES

[1] Gut, U. 2009. *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German* 9. Bern: Peter Lang.

[2] Fry, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27(4), 765-768.

[3] Fry, D. B. 1958. Experiments in the perception of stress. *Language and speech* 1(2), 126-152.

[4] Fry, D. B. 1965. The dependence of stress judgments on vowel formant structure. In: Bethge, W., Zwirner, E. (eds), *Proceedings of the 5th International Congress of Phonetics Sciences.* Basel: Karger, 306-311.

[5] Jun, S. A. 1995. A phonetic study of stress in Korean. *The Journal of the Acoustical Society of America* 98(5), 2893-2893.

[6] Jun, S. A. 2005. Korean intonational phonology and prosodic transcription. *Prosodic typology: The phonology of intonation and phrasing* 1, 201.

[7] Lim, B. J. 2001. The production and perception of word-level prosody in Korean. *IULC Working Papers* 1(1).

[8] Peperkamp, S., Dupoux, E. 2002. A typological study of stress" deafness. *Laboratory phonology* 7, 203-40.

[9] Rothman, J. 2011. L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27(1), 107-127.

[10] Bardel, C., & Falk, Y. 2007. The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research* 23(4), 459-484.

[11] Xu, Y. 2013. ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. *Proc. Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, Aix-en-Provence, France, 7-10.

[12] Barr, D. J., Levy, R., Scheepers, C., Tily, H. J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language* 68(3), 255-278.

[13] Mitterer, H., Cho, T., Kim, S. 2016. How does prosody influence speech categorization?. *Journal of Phonetics* 54, 68-79.

[14] Cho, T., Lee, Y., & Kim, S. 2011. Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics* 39(3), 344-361.

[15] Jeon, H. S., Nolan, F. 2013. The role of pitch and timing cues in the perception of phrasal grouping in Seoul Korean. *The Journal of the Acoustical Society of America* 133(5), 3039-3049.

[16] Holliday, J. J. 2015. A longitudinal study of the second language acquisition of a three-way stop contrast. *Journal of Phonetics* 50, 1-14.

[17] Nagano-Madsen, Y. 2015. Prosodic Phrasing Unique to the Acquisition of L2 Intonation — an Analysis of L2 Japanese Intonation by L1 Swedish Learners. *Proc. 16th INTERSPEECH*, Dresden, Germany, 820-823.