An investigation of the effect of L1 dialect on the L2 perception of lexical stress

Hyunjung Lee[1], Eun Jong Kong[2], Jeffrey J. Holliday[3]

[1]Incheon National University, [2]Korea Aerospace University, [3]Korea University
[1]hjlee123@inu.ac.kr, [2]ekong@kau.ac.kr, [3]holliday@korea.ac.kr

ABSTRACT

This study explores the effect of Korean learners' native dialect in discriminating English stress patterns. The nature of Kyungsang Korean (KK; lexical pitch-accent dialect) yields a prediction that KK learners of English would outperform Seoul Korean (SK) learners in identifying English stress, as KK speakers are more sensitive to *f0* variation, an acoustic dimension of English stress. We administered two ABX discrimination tasks to three groups of participants (37 KK, 40 SK, 16 L1 English) to examine their accuracy in perceiving English stress location. The tasks consisted of English and Korean nonce words varying in the type of pitch accent. Results showed that the group-averaged accuracies for KK were lower than those for SK, indicating KK's use of *f0* in their native dialect did not positively influence their identification of English stress. Our further acoustic examinations of *f0* support this characterization.

Keywords: pitch accent dialect, stress language, ABX discrimination, non-native perception

1. INTRODUCTION

It is well known that individual learners perform differently in learning L2 sounds depending on various learner-internal and -external factors, with the phonological grammar of the learners' L1 being one of the most important [2, 3, 9, 10]. Given this, it is speculated that phonological variation across dialects of the same L1 might influence L2 learners' speech perception [4]. Many previous studies, however, have assumed some degree of L1 homogeneity, neglecting L1 dialectal variation (although see [4], [6], and [7] for studies examining such effects in L2 vowel perception, and who argue that researchers should pay closer attention to L1 dialectal differences). The purpose of this study is to examine the effect of L1 dialect on the perception of non-native suprasegmental properties.

In this study, we examined whether and how the different tonal systems of Korean dialects affected the non-native perception of English lexical stress by comparing the perception pattern between non-tonal Seoul and tonal Kyungsang dialect speakers of Korean. Kyungsang Korean is different from standard Seoul Korean in that it is a lexical pitch accent variety in which segmental homophones are distinguished by differing the location of an accented syllable or *f0* peak in otherwise similar words (e.g., *kaci* 'type (HL)' vs. *kaci* 'eggplant (LH)') [11, 14]. On the one hand, while Seoul and Kyungsang Korean speakers share the same first language, the dialectal difference regarding the tonal system renders the use of *f0* different for lexical disambiguation. On the other hand, *f0* is used somewhat similarly between Kyungsang Korean and English (i.e., a stress-timed language where *f0* is used to indicate post-lexical prominence) even if they are two separate languages.

The dialectal difference in the L1 (i.e., the tonal differences between Seoul and Kyungsang Korean) and the similarity between one L1 variety and the L2 (i.e., the use of *f0* in Kyungsang Korean and English) raise several hypotheses. If Kyungsang listeners' use of *f0* for lexical disambiguation does not facilitate the perception of English stress, this lack of dialectal effect would result in similar patterns of stress perception between Kyungsang and Seoul Korean learners of English. On the other hand, if the effect of L1 dialect exists, we might expect different patterns of English stress perception between Kyungsang and Seoul Korean listeners. Particularly focusing on the use of *f0* in facilitating the L2 learning of lexical stress, this study compared the perception patterns across three groups native dialect and language listeners: Seoul Korean (SK), Kyungsang Korean (KK), and American English (EN). We adopted an ABX discrimination task [8] with nonce words in Korean and English. Based on cross-linguistic and dialectal comparisons, this study aims to broaden our understanding of the role of individual-level phonology in L2 speech perception.

2. METHOD

2.1. Participants

A total of 93 subjects completed the present perception experiment: 37 KK (16 males), 40 SK (20 males), 16 EN (4 males). Subjects were paid for their participation. The mean age of the KK, SK, and EN groups was 22.7, 23.3, and 21.1 years old, respectively. The English-speaking participants were speakers of midwestern American English. The KK and SK participants were all born and educated in the target dialect regions: the city of Changwon, where the South Kyungsang variety of Korean is spoken, and the Seoul and Kyunggi regions where standard Seoul Korean is used. Subjects reported no language or hearing problems.

2.2. Stimuli

Two sets of stimuli were created: English nonce words and Korean nonce words. For the English nonce word set, we created three disyllabic nonce sequences of segments (/bu.ʧi/, /sər.ʧu/ and /ban.tæk/) produced by two phonetically-trained male native speakers of midwestern American English; the speakers were instructed to produce the three nonce words stressing the first syllable in one version, Strong-Weak (SW) and stressing the second syllable in the other version, Weak-Strong (WS). The two repetitions provided a total of twelve unique nonce word audio stimuli. Similarly, three disyllabic Korean nonce words (/pu.cɔn/, /cʌ.ku/ and /pɔn.mæk/) were presented to two male native KK speakers, who produced the nonce words twice while varying the pitch accent type between HL and LH. A total of 24 English nonce word stimuli (3 nonce words × 2 stress types × 2 speakers × 2 repetitions), and 24 Korean nonce words (3 nonce words × 2 pitch accent types × 2 speakers × 2 repetitions) were used for the ABX discrimination tasks.

2.3. Acoustic measurements

We measured the *f0* peaks of the stimuli used in the perception experiment to examine how the three groups of listeners used *f0* in discriminating stress patterns in their L1 and L2. The vowel portion of each disyllable was first determined from the onset of the first full period to the offset of F2. Peak *f0*s were measured within the vowel for each token. We used a Praat script for the measurement, manually checking each value. *f0* was first measured in Hertz, and then converted into semitones to normalize inter-speaker global pitch differences. The difference between the peak *f0s* (i.e., peak *f0* in S/H minus peak *f0* in W/L) was operationalized as the pitch prominence of the token.

2.4. Tasks and procedure

The three groups of participants completed two sessions of ABX discrimination task (English and Korean) presented on a notebook computer and programmed in E-Prime ver. 3.0 [12]. In both tasks, the stimuli in each trial were triplets of nonce words differing in segments and with either A or B having the same pitch accent/stress pattern as X. For example: A /**bu**.ʧi/, B /sər.**ʧu**/, X /ban.**tæk**/. Participants were instructed to attend to the *rhythmic* pattern of each word in the triplets and indicate whether the rhythm pattern of X was the same as that of A or B by using the mouse to click a button on the screen labelled <A> or <B>.

2.4. Analysis

We examined patterns of accurate responses and the role of *f0* in accounting for the accuracy patterns by constructing two different kinds of logistic mixed-effects regression models. The models of response accuracy were built separately for each pitch accent/stress types (HL vs. LH, SW vs. WS) where correct/incorrect responses (DV) were predicted by a fixed effect of the answer types (X=A, and X=B). The intercept and the slope of answer type were allowed to randomly vary at both the listener level and talker level. The other type of model was built to estimate the coefficient of *f0* in X (*f0*.X: peak *f0* difference between S/H and W/L syllables) in explaining accuracy. We entered the random intercept and slope of *f0*.X at the listener level and the random intercept at the talker level. The models were implemented using the *lmer()* function in the R platform [1, 13].

3. RESULTS & DISCUSSION

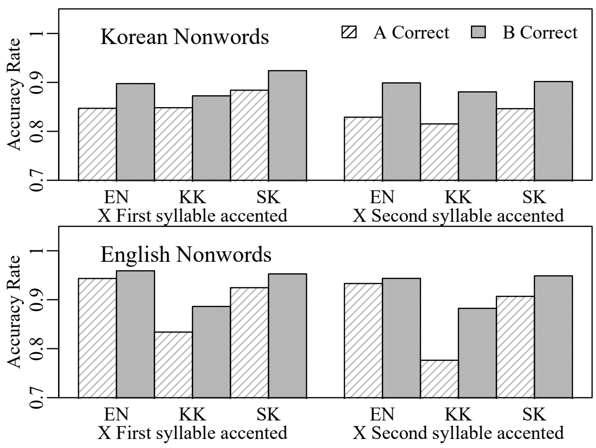
Accuracy patterns were examined by the statistical models of each language session (Korean and English). Figure 1 displays the estimates from the models where the group-averaged accuracies were assessed by the pitch accent/stress type (S/H and W/L) of the word X and answer types (X=A and X=B).

One global tendency across the three listener groups was that accuracies were higher when X was identical to B (solid bar) compared to when it was identical to A (dashed bar), although the magnitudes of the accuracy differences differed across listener groups and the types of pitch accent/stress. Separated by the pitch accent type of X, there were no accuracy differences among the three listener groups in the HL type (left-side bars): $\beta.\text{diff}_{[KK\text{-}EN]} = -.13$, SE = .59, $p = .81$, $\beta.\text{diff}_{[KK\text{-}SK]} = .51$, SE = .47, $p = .28$. That is, neither Kyungsang listeners (KK) nor English native listeners (EN) outperformed Seoul listeners (SK) even though they are lexical pitch-accent dialect and stress language users. Similarly, in the LH type in X, accuracies were estimated to be not significantly different among the listener groups ($\beta.\text{diff}_{[KK\text{-}EN]} = .041$, SE = .54, $p = .93$, $\beta.\text{diff}_{[KK\text{-}SK]} = .19$, SE = .43, p = .64.), although, in all three groups, accuracy significantly differed between the answer types: Accuracy(X=B) > Accuracy(X=A), $\beta.\text{diff}_{[A\text{-}B]}$ = .54, SE = .14, $p < .0001$.

Unlike the patterns in the Korean nonce word task, accuracies differed among the listener groups in the English nonce word task, with KK listeners the least accurate in all analysis conditions. Models yielded significant differences of accuracy estimates between KK listeners and the others in the both stress types: $\beta.\text{SW.B}_{[KK\text{-}EN]}$ = -1.1, SE = .51, $p < .05$; $\beta.\text{SW.B}_{[KK\text{-}SK]}$ = -.95, SE = .38, $p < .05$. Seoul listeners were not shown to be any less accurate than the native English listeners, however, as the accuracy estimates were not significantly different between the two listener groups: $\beta.\text{WS.A}_{[SK\text{-}EN]}$ = -.35, SE = .73, p = .46.

In the current study, the accuracy patterns seemed to reflect a fairly inconsistent role of linguistic experience in prosody discrimination. For one, the fact that the Korean nonce words were spoken by native speakers of their own dialect/language did not help the Kyungsang listeners, but did help the English listeners to perform better in the discrimination tasks. Also, Seoul listeners performed no worse than Kyungsang listeners in the Korean session and English listeners in the English session even though none of the tasks presented audio stimuli spoken by Seoul dialect speakers. To resolve this puzzle of inconsistent role of linguistic experience in performing prosody discrimination, we further investigated the role of *f0* sensitivity in explaining the accuracy scores.

Figure 1: Accuracy rates of ABX tasks averaged across sessions (Korean and English nonwords), accent types of X (first syllable and second syllable accented), and correct answer locations (A and B).



A series of *lmer* models were constructed in order to estimate listeners' sensitivity to *f0* in the word X in accurately discriminating the prosodic properties in the word A and word B. Table 1 summarizes the parameter estimations from the models, and Figure 2 illustrates the distributions of accuracy rates of the stimulus token X against the peak *f0* differences between the two syllables in X: peak *f0* (H or S) minus peak *f0* (L or W).

In the Korean session, the model produced a significant coefficient of *f0*.X for the KK group. The positive coefficient indicates that the responses were likely to be correct as the token X has a greater peak *f0* difference between the H and L syllables. The magnitudes of the *f0*.X coefficients for the groups of SK (β.SK = .096, SE = .047, $p < .05$) and EN (β.EN = .038, SE = .063, p = .60) were smaller than that of KK. This suggests that KK listeners were more sensitive to the acoustic cue of dramatic pitch excursion in performing the prosody discrimination task. It is noted that for the EN group, *f0*.X was not a significant variable in predicting the accuracy patterns.

In contrast, the model for the English session returned no significant coefficient of *f0*.X for the KK (and SK, β.SK = .028) listeners but a significant coefficient of *f0*.X for the EN group (β.EN = .087, SE = .026, $p < .001$). Coefficients of *f0*.X were greater in EN, SK, and KK, in that order. A positive coefficient of the EN group suggests that a bigger pitch expansion in the word helped the EN listeners to correctly discriminate the prominence patterns in the stimuli.

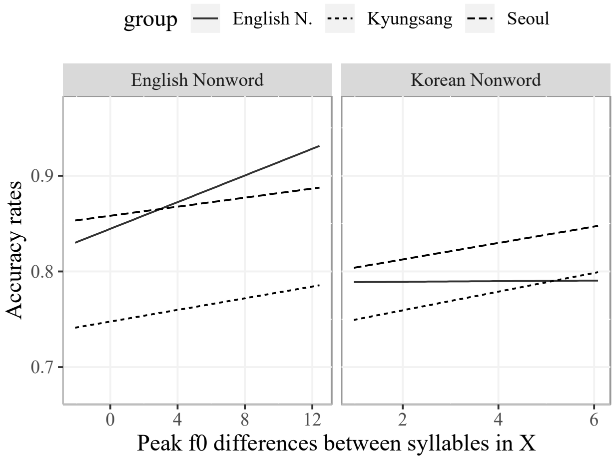
This pattern in the output of the statistical models is well illustrated in Figure 2, where similar degrees of slope steepness were observed in KK and SK. That is, Kyungsang and Seoul Korean listeners patterned similarly in utilizing *f0* information in prosody discrimination of both Korean and English nonce words, which differed from English listeners' use of *f0*. This supports a lack of a dialectal effect in the perception of non-native language prosody.

Although there was no evidence of a dialectal effect in the current study, there seems to be some effect of familiarity with the auditory source in accessing to the *f0* information. While both English and Kyungsang listeners are supposed to be sensitive to *f0* due to the prosodic properties of their native varieties, their full usage of *f0* information was consistently conditioned to the stimulus languages. This might imply that the application of *f0* sensitivity to non-native perception is not necessarily an automatic process.

Table 1: Output of the logistic mixed-effects regression models where *Accuracy* was predicted by the *f0*.X variable with *Listener Group* as an interaction. The reference level of the *Group* variable is Kyungsang Korean. SK = Seoul Korean, and EN = English native listeners.

	Estimate	SE	z value	Pr(>|z|)
<English nonwords>				
(Intercept)	1.568	0.259	6.03	< .0001
f0.X	0.019	0.014	1.39	.162
EN	0.754	0.461	1.63	.102
SK	**0.831**	**0.350**	**2.37**	**.018 ***
f0.X:EN	**0.067**	**0.026**	**2.50**	**.012 ***
f0:SK	0.008	0.019	0.42	.67
<Korean nonwords>				
(Intercept)	1.446	0.253	5.71	< .0001
f0.X	**0.101**	**0.047**	**2.10**	**.035 ***
EN	0.348	0.391	0.89	.37
SK	0.262	0.299	0.87	.38
f0.X:EN	-0.062	0.064	-0.97	.33
f0.X:SK	-0.004	0.050	-0.08	.93

Figure 2: Distributions of averaged accuracy as a function of peak *f0* differences in X.



It should be mentioned that our current analysis is limited to the *f0* aspect of the stress in English because *f0* is a shared parameter between English lexical stress and Kyungsang Korean pitch accent. In addition, the English proficiency of the Seoul and Kyungsang Korean learners was not considered a variable in the current analysis. These limitations might resolve the puzzle of Seoul listeners' high accuracy in the English nonce word task despite their lack of reliance on *f0*. We plan to extend the scope of the acoustic analysis to other acoustic properties of English stress such as intensity, duration and vowel enhancement and reduction, so that the discrimination accuracy of Seoul listeners can be better explained.

4. CONCULSION

This study explored the effect of Korean learners' native dialect in discriminating English lexical stress patterns. The experimental evidence presented in the current study did not support the hypothesis that an L2 learner's first language dialect has a significant effect on the L2 perception lexical stress. The findings imply that sensitivity to phonological cues in a native dialect is not automatically transferred in non-native perception.

These results run somewhat counter to those in [4] and [7], which investigated native dialect effects in L2 vowel perception. Further work is needed to investigate whether segmental and suprasegmental properties of native dialects are accessed differently in L2 perception.

5. ACKNOWLEDGEMENTS

The authors would like to thank Kelly Berkson, Jieun Kang, Hyangwon Lee, and Bokyung Park for assistance in both the recruiting and testing of participants. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2017S1A5A2A03068448).

6. REFERENCES

[1] Bates, D., Maechler, M., Bolker, B., & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

[2] Best, C. T., McRoberts, G. W., & Sithole, N. M. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 345.

[3] Best, C. T., & Tyler, M. D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In J. Munro & O.-S. Bohn (Eds.), *Second Language Speech Learning: The Role of Language Experience in Speech Perception and Production* (pp. 13–34). Amsterdam: John Benjamins.

[4] Chládková, K., & Podlipský, V. J. 2011. Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *Journal of the Acoustical Society of America, 130*(4), EL186-EL192.

[5] Escudero, P., & Boersma, P. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*(4), 551-585.

[6] Escudero, P., Simon, E., & Mitterer, H. 2012. The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *Journal of Phonetics, 40*(2), 280-288.

[7] Escudero, P., & Williams, D. 2011. Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *Journal of the Acoustical Society of America, 131*(5), EL406-EL412.
[8] Gerrits, E., & Schouten, M. E. H. 2004. Categorical perception depends on the discrimination task. *Perception & Psychophysics, 66*(3), 363-376.
[9] Golestani, N., & Zatorre, R. J. 2009. Individual differences in the acquisition of second language phonology. *Brain and Language, 109*(2-3), 55-67.
[10] Kartushina, N., & Frauenfelder, U. H. 2014. On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology, 5*, 1246.
[11] Lee, H., Jongman, A., & Zhang, J. 2016. Variation and change in the nominal pitch-accent system of South Kyungsang Korean. *Phonology, 33*(2), 325-351.
[12] Psychology Software Tools, Incorporated. [E-Prime 2.0]. 2012. http://www.pstnet.com
[13] R Core Team 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
[14] Ramsey, S. R. 1975. *Accent and morphology in Korean dialects: A descriptive and historical study*. Ph.D. dissertation, Yale University.