Perception of Prosodic Prominence by Korean Learners of English

Suyeon Im[a] & José Ignacio Hualde[b]

[a]The Hong Kong Polytechnic University, [b]University of Illinois at Urbana-Champaign
suyeon.im@polyu.edu.hk, jihualde@illinois.edu

ABSTRACT

This study compares the perception of phrase-level prosodic prominence in English by native English speakers and by Korean-speaking learners of English. Native speakers of both languages were asked to indicate which words they perceived as prominent in a recorded excerpt from a lecture. Our results show that, like native English speakers, Korean learners of English rely on duration and pitch to judge prosodic prominence on different words in the phrase. Interestingly, however, unlike native speakers of English, Korean speakers also tend to assign greater prominence to words that start with a voiceless consonant. This can be explained as reflecting the cuing of accentual-phrase-level prominence in Korean, where initial aspirated and fortis consonants attract a high tone and trigger the perception of prominence on the initial syllable.

Keywords: prosody, prominence, transfer, English, Korean

1. INTRODUCTION

Korean and English differ substantially in their prosodic systems [1]. Unlike English, where words may contrast in the position of the stressed syllable, Korean lacks contrastive word-level stress. Within an accentual phrase, Korean speakers tend to perceive prominence on the first syllable bearing a H tone [2, 3, 4, 5]. Prominence at this level is thus cued by pitch. Indirectly, there is also a consonantal effect on prominence. This is because the prominence-lending H usually loosely aligns with the second syllable of the accentual phrase, except that when the first syllable of the accentual phrase starts with a *fortis* or aspirated consonant, it aligns with the first syllable instead, which is thus perceived as prominent [4].

Perhaps because of these differences between the two languages in the functions and realization of prominence, Korean learners of English show a different behaviour from native speakers in their judgments regarding prosodic patterns at the lexical level in English; i.e., their intuitions about which syllable in an English word is lexically stressed [6].

Something that remains to be examined, and we explore here, is whether these differences between Korean learners of English and native English speakers are also found regarding the relative prominence of different words within intonational phrases. It should be noticed that word-level prominence (i.e. the greater prominence of one syllable over the others within the word domain) and phrase-level prominence (i.e. the greater prominence of some words within the domain of the phrase) have very different functions in English. Word-level prominence or word-stress is a matter of lexical identity and may serve to distinguish words (e.g. *trusty*, with initial stress, vs *trustee*, with final stress). Phrase-level prominence (i.e. prenuclear and nuclear accent), on the other hand, may have the function of highlighting certain words for pragmatic purposes and, in part, may also be rhythmic (in the case of prenuclear accents) [7].

A question that arises is whether L1 Korean speakers perceive phrasal prominence in English using the phonetic features that cue accentual-phrase prominence in Korean, namely pitch, and, because there is an association in Korean between pitch and consonant type, also consonantal features. Since Korean lacks lexical stress, we may wonder whether the judgments of prominence on certain syllables in Korean that have been reported in the literature may affect Korean-speakers' perception of prominence in English at the phrasal-level. That is, do Korean-speaking English learners perceive words bearing a high (H*) or rising (L+H*) accent as prominent in the phrase? And, more interestingly, given the correlation in Korean between H tone and word-initial fortis and aspirated consonants, do they perceive English words starting with a voiceless consonant as intrinsically prominent, even if they do not bear a high or rising pitch accent? (in which case, presumably, their judgments concerning which words are prominent may be quite unrelated to the pragmatic highlighting intended by the speaker).

Here we report on an experiment that addresses these questions, using naturally but carefully produced English speech.

2. METHOD

2.1. Participants and experimental task

Thirty Korean speakers and thirty-five speakers of American English participated in the perception experiment reported in this paper. The Korean

speakers were advanced learners of English and were tested at a university in Korea. None of them had lived in English-speaking countries for more than three months. The English-speaking participants were undergraduate students at a university in the U.S. Midwest.

Participants were asked to listen to an excerpt and simultaneously mark on a transcript of the same excerpt those words that they perceived as prominent. The excerpt is a complete and intact public speech from TED Talk titled "Try Something New for Thirty Days" [8], delivered by a male speaker of American English in clear and engaging manner (t = 2'25"). The excerpt contains 361 words and was presented without punctuation or capitalization. We used an acoustically-based, rather than pragmatically-based, definition of prominence, using in this definition those properties that have been found to correlate with phrasal prominence in English. Prominent words were defined for participants as "words that stand out in the speech stream by virtue of being louder, longer, more extreme in pitch, or more crisply articulated than other words in the same utterance." Korean participants completed this task on paper. Native English speakers, tested in the US, completed the same task, but using a computer interface to record their responses [9]. Participants were able to listen to the excerpt twice.

2.2. Analysis

For each word in the sound file, we extracted information on the three acoustic properties mentioned in the instructions to participants: pitch, calculated as maximum f0 in the word, duration, operationalized as mean phone duration (word duration divided by number of phones), and mean word intensity. This information was obtained using ProsodyPro [10]. Raw numbers were converted into z-scores for statistical analysis. In addition, every word was classified as starting with (a) a voiceless obstruent, (b) a voiced obstruent, or (c) a sonorant consonant or a vowel.

An additional analysis of the stimuli involved the prosodic labelling of all words in the sound file in Praat [11] following the ToBI Annotation Conventions [12]. This was done by two expert labellers (including one of the authors), who produced a consensus label for each word. This annotation produced four pitch-accent labels, L* (n = 19), !H* (n = 34), H* (n = 66), and L+H* (n = 65), in addition to a larger class of unaccented words (n = 177). There were other types of pitch-accent labels, H+!H* (n = 1) and L*+H (n=3), and due to small sample size, those labels were assigned to the other labels with the same starred tones (!H* for H+!H*) or with similar contour shapes (L+H* for L*+H).

For the statistical analysis of the results, a mixed-effects binomial logistic regression (g*lmer*) was run using the *lme4* [13] and *afex* [14] packages in R [15]. Participants' prominence markings for each word, as dependent variable, were modelled in relation to L1 group (English vs Korean), maximum f0, mean phone duration, mean word intensity, initial phoneme and the interaction between L1 group and all other fixed factors. Pitch-accent labels were not entered in the model as they may correlate with other acoustic measures, especially, maximum f0. Participant was entered as a random factor (another model with the same structure, but adding word as a second random factor failed to converge).

For further analysis, we added up the number of *prominent* responses for each word in the excerpt. This is reported as a prominence-score or *p-score* ranging from 0 to 1 [16], where a p-score of 1 indicates that 100% of the participants marked the word as prominent and 0 means that none of the participants marked it as prominent. We calculated p-scores for the L1 and the L2 groups. The relation of ToBI labels (which were not entered in the regression) and p-scores will be discussed only informally.

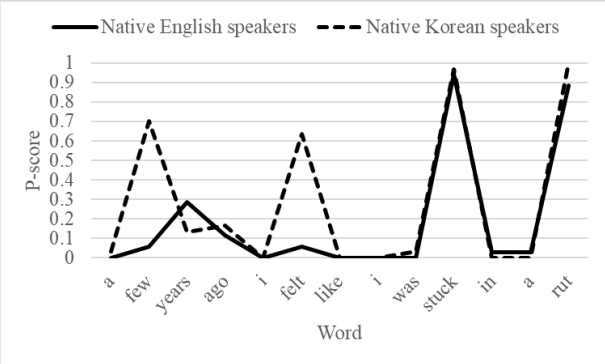
3. PREDICTIONS

For L1-English participants, we expect pitch, duration and intensity to affect their prominence judgments. No effect of word-initial phoneme is expected. Regarding L1-Korean participants, we also expect them to use pitch, and perhaps the other phonetic cues mentioned in the instructions that they received, to assign perceived prominence to words. In addition, we predict that, in their decisions regarding whether a word sounds prominent or not, Korean speakers may take the nature of the word-initial phoneme into account, given the effects of word-initial segments in the intonational phonology of Korean. In particular, words starting with voiceless obstruents (which corresponds to the fortis and aspirated classes of Korean) are predicted to be judged as prominent more frequently than other words, other things being equal, resulting in higher p-scores for these words. Notice that, although participants were asked to judge the prominence of words, not of different syllables, given the very high incidence of monosyllabic words in non-technical English discourse (as can be seen in the example in Figure 1, where almost all words are monosyllabic), participants can be expected to conflate both levels of prominence, especially as Korean lacks word stress.

4. RESULTS

Native English speakers and L1-Korean participants had similar but not identical judgments of prominence. An example is provided in Figure 1, for the phrase "A few years ago, I felt like I was stuck in a rut".

Figure 1: Example of prominence rating of words in a phrase by native English speakers (solid line) and native Korean speakers (dotted line).



As can be seen in this figure, for most words the p-scores obtained from both groups of participants are very similar. For example, both groups of participants overwhelmingly rated the words "stuck" and "rut" as being prominent, and there was also a consensus in both groups that most function words did not carry prominence. Korean speakers (dotted line), however, differed from English speakers (solid line) in judging prosodic prominence for the words "few" and "felt" (which start with a voiceless consonant). Whereas over 60% of Korean-speaking participants marked prominence on these two words, less than 10% of English-speaking participants did.

Table 1 shows the results of the mixed-effects binomial logistic regression model on the prominence judgements provided by native speakers of English and Korean.

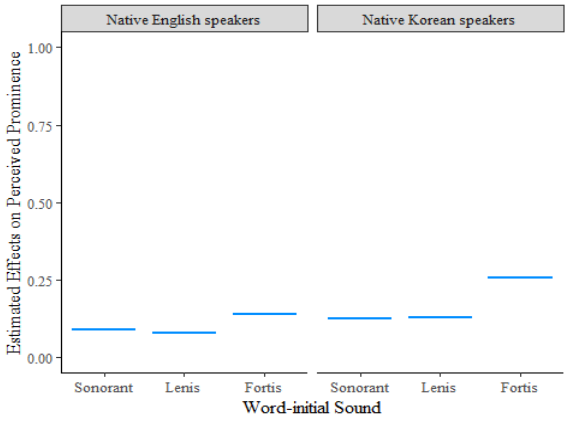
Table 1: *glmer* results for perceived prominence in relation to speaker's L1, and the interaction between speaker's L1 and other factors

	est.	*SE*	*z*	*p*
(Intercept)	-1.96	.09	-21.37	< .01***
L1				
Korean	.39	.13	2.94	< .01***
Word-initial Sound				
Lenis	-.14	.08	-1.79	.07
Fortis	.48	.06	8.15	< .01***
Acoustic Cue				
Max f0	.50	.03	18.31	< .01***
Mean phone duration	.88	.03	32.64	< .01***
Mean word intensity	-.02	.03	-.62	.53
L1:Word-initial Sound				
Korean:Lenis	.17	.11	1.60	.11
Korean:Fortis	.38	.08	4.66	< .01***
L1:Acoustic Cue				
Korean:Max f0	.06	.04	1.65	.10
Korean:Mean phone duration	-.01	.04	-.25	.80
Korean:Mean word intensity	-.07	.04	-1.57	.12

There is a significant effect of L1, initial phoneme, f0 and duration, but not intensity. The interactions show that the two L1 groups differ significantly in the effects of initial phoneme on their prominence ratings, with words starting with a fortis consonant being treated differently from words with other initial sounds for L1-Korean participants. Native Korean speakers crucially differ from native English speakers in that, in addition to using pitch and duration in judging the prosodic prominence of words, they are influenced by their initial phoneme. Korean speakers tend to rate words as prominent if they begin with a voiceless obstruent consonant, while English speakers do not or not to the same extent.

In Figure 2, we further explore the interaction between word-initial phoneme and L1 (with *visreg* [17]). In this figure, on the x-axis, words are classified in three groups according to their word-initial phoneme (Sonorant = sonorant or vowel, Lenis = voiced obstruent, Fortis = voiceless obstruent) and the y-axis shows estimated perceived prominence.

Figure 2: Model estimated effects of word-initial phoneme on L1-English (left) and L1-Korean participants' (right) perceived prominence.

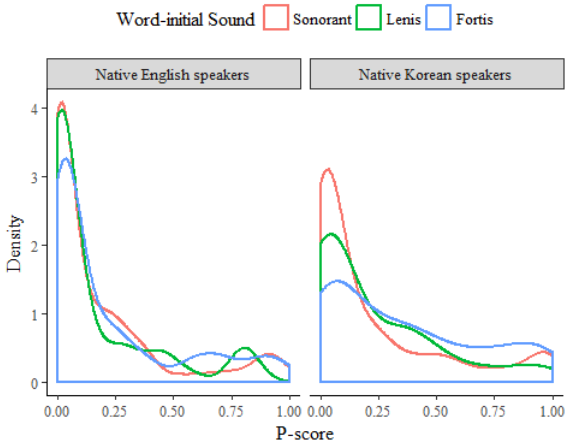


Although this figure shows a somewhat higher estimated prominence for words starting with a voiceless consonant for both groups of speakers, this effect is much greater for Korean speakers. Korean speakers are more likely to perceive as prominent words with an initial voiceless obstruent consonant than words starting with other phonemes.

The effect of word-initial phoneme type on perceived prominence is also shown in the density plots of p-scores (from 0 to 1, on the x-axis) in Figure 3 (made with the r-package *ggplot2* [18]). For L1-Korean participants (right-hand panel), words starting with a fortis (voiceless) consonant (blue line) show a different distribution from other others, with fewer items towards the lower end and more words towards the higher end of the p-score scale. For L1-English participants (left-hand panel), on the other hand, words starting with a fortis consonant show a

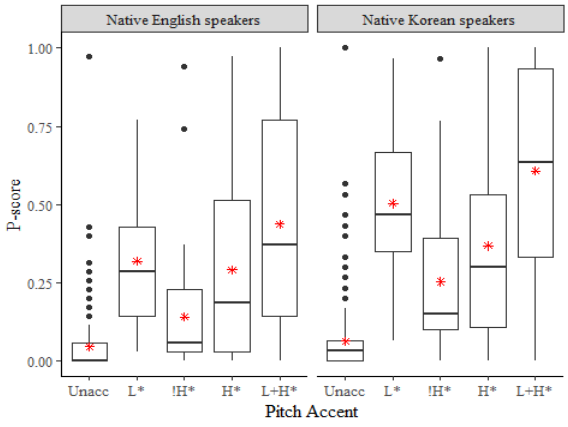
distribution that is much more similar to that of the other two classes of words. (Small differences may be due to interaction with factors not considered here, such as lexical class and information status, which have been shown to affect prominence marking in experiments with the same methodology [16, 20]).

Figure 3: Density plots of p-scores by word-initial phoneme type for L1-English (left) and L1-Korean participants.



To complete the analysis of the data, we now consider the relation between the p-scores obtained for individual words and their ToBI accentual labels. Figure 4 shows words' prosodic prominence as perceived by native English speakers (left panel) and native Korean speakers (right panel), separately for words with each of the accentual labels that were used in our ToBI analysis. The x-axis shows the accentual status of words and the y-axis indicates p-scores.

Figure 4: Prosodic prominence perceived by native English speakers (left panel) and native Korean speakers (right panel) in relation to pitch accent type. *Unacc* = unaccented; accented words are further specified by accent type, L*, !H*, H*, and L+H*.



Both groups of participants judged unaccented words as prominent with much lower frequency than words labelled as accented, resulting in much lower p-scores for unaccented words. With one partial exception, the effect of accent type is also the same for both groups, with the bitonal accent L+H* being judged as the most prominent, and H* words receiving higher p-scores than words labelled as !H*. The interesting difference is that words with an L* accent were given much higher p-scores by Korean speakers than by native English speakers. But, for native English speakers as well, L* resulted in relatively high p-scores, second only to L+H* words, (compare with [20]). The explanation is perhaps that, in the excerpt that our participants judged, L* is very often used at the end of a phrase to indicate continuation in conjunction with a high boundary tone. This interaction between accent and boundary may have triggered the perception of prominence, especially for native Korean speakers.

5. DISCUSSION

This study has investigated how Korean learners of English perceive phrasal prominence in English. We have reported on the results of an experiment where native English speakers and native speakers of Korean were asked to indicate which words they perceived as prominent in a motivational lecture.

Both native speakers of Korean and native English speakers relied on mean phone duration and maximum f0 in judging the prosodic prominence of individual words. In addition, Korean speakers were found to be more likely to associate prosodic prominence with words with an initial voiceless consonant than with words with an initial voiced obstruent or a sonorant or vowel. This makes sense when we consider phrasal tone assignment in Korean, where a high tone is assigned to the initial syllable of words starting with a fortis or aspirated consonant, while a low tone is assigned to the initial syllable of words with an initial lenis consonant or sonorant. This is, of course, not a pattern that we find in English, where the distribution of pitch accents in the phrase is dictated by pragmatic and rhythmic considerations. Interestingly, Korean participants tended to perceive words starting with a voiceless obstruent as prominent even in the absence of strong pitch cues. The association between initial fortis/aspirated consonants and prominence thus appears to be somewhat phonologized in Korean, to the extent that it is used in judgments of prominence in a second language.

6. ACKNOWLEDGEMENTS

This research was supported in part by NSF BCS 12-51343.7. We are thankful to Jennifer S. Cole for input at different stages in the development of this work.

7. REFERENCES

[1] Jun, S.-A. 2014. Prosodic typology: By prominence type, word prosody, and macro-rhythm. In: Jun, S.-A. (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* 2, 520-539.

[2] DeJong, K. 1994. Initial tones and prominence in Seoul Korean. *OSU Working Papers in Linguistics* 43, 1-13.

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

[4] Jun, S.-A. 2005. Korean intonational phonology and prosodic transcription. In: Jun, S.-A. (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing* 1, 201-229.

[5] Lim, B-J. 2001. The role of syllable weight and position on prominence in Korean. *Japanese and Korean Linguistics* 9, 139-50.

[6] Guion, S. G. 2005. Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition* 27(4), 503-533.

[7] Calhoun, S. 2007. *Information structure and the prosodic structure of English: a probabilistic relationship* (Doctoral dissertation). University of Edinburgh.

[8] Cutts, M. 2011. Try something new for 30 days. TED. https://www.ted.com.

[9] Mahrt, T. 2013. Language Markup and Experimental Design software (LMEDS). Retrieved from http://prosody.beckman.illinois.edu/lmeds.html.

[10] Xu, Y. 2013. ProsodyPro – A tool for large-scale systematic prosody analysis. *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)* Aix-en-Provence, 7-10.

[11] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer, version 6.035, www.praat.org [computer program].

[12] Veilleux, N., Shattuck-Hufnagel, S., Brugos, A. 2006. 6.911 Transcribing prosodic structure of spoken utterances with ToBI. https://ocw.mit.edu.

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

[14] Singmann, H., Bolker, B., Westfall, J., Aust, F. 2016. Afex: analysis of factorial experiments. R package version 0.16-1.

[15] R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

[16] Cole, J., Mo, Y., Hasegawa-Johnson, M. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1(2), 425-452.

[17] Breheny, P., Burchett, W. 2017. Visualization of regression models using visreg. *The R Journal* 9: 56-71.

[18] Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

[19] Im, S., Cole, J., Baumann, S. 2018. The mediating effect of information status on acoustic cues to prominence. Presented at Laboratory Phonology 16, Lisbon.

[20] Hualde, J. I., Cole, J., Smith, C., Eager, C., Mahrt, T., Napoleão de Souza, R. 2016. The perception of phrasal prominence in English, Spanish and French conversational speech. *Proceedings of Speech Prosody* 8, 459-463.