



Processing advantages for focused words in Korean

Heather Kember^{1,2}, Jiyoung Choi³, Anne Cutler^{1,2}

¹ The MARCS Institute, Western Sydney University, Australia

² ARC Centre of Excellence for the Dynamics of Language, Australia

³ Hanyang Phonetics and Psycholinguistics Lab, Hanyang University, Korea

h.kember@westernsydney.edu.au, jiychoi@hanyang.ac.kr, a.cutler@westernsydney.edu.au

Abstract

In Korean, focus is expressed in accentual phrasing. To ascertain whether words focused in this manner enjoy a processing advantage analogous to that conferred by focus as expressed in, e.g., English and Dutch, we devised sentences with target words in one of four conditions: prosodic focus, syntactic focus, prosodic + syntactic focus, and no focus as a control. 32 native speakers of Korean listened to blocks of 10 sentences, then were presented visually with words and asked whether or not they had heard them. Overall, words with focus were recognised significantly faster and more accurately than unfocused words. In addition, words with syntactic focus or syntactic + prosodic focus were recognised faster than words with prosodic focus alone. As for other languages, Korean focus confers processing advantage on the words carrying it. While prosodic focus does provide an advantage, however, syntactic focus appears to provide the greater beneficial effect for recognition memory.

Index Terms: prosody; focus; speech processing; Korean

1. Introduction

Marking of salience is argued to be one of only two prosodic universals that occur across languages [1]. Though this may well be the case, there are language specific differences in the way salience is marked. For example, in English and Dutch, a pitch accent is used to mark prosodic focus, whereas in Korean and French, prosodic focus is linked to the position of the word within the phrase. Despite these differences in the method of assigning focus, the assumption is that these varying types of prosodic structure all perform the same function, namely to draw the listener's attention to new or important information within the discourse. The role of focus in sentence processing has however not been widely investigated (other than in English and Dutch). This limits our ability to conclude that effects of prosodic focus are language universal, that is, that listeners across languages are indeed able to exploit the focus information available to them in the speech stream.

In English, words with focus show processing advantages compared to words without focus. For example, words marked by syntactic focus are remembered better compared to words without focus [2, 3]. The same is true for words focused using prosody, with sounds within the salient words recognised more rapidly in comparison to non-salient words [4]. Even when focused words themselves are spliced out of utterances and replaced with neutrally produced words, listeners are still better at detecting phonemes at the position [5, 6] signalled by preceding prosody to be salient. This finding suggests that the

perceptual benefit is not simply due to clearer or more accurate articulations of the words themselves, but that listeners entrain to the utterance prosody; this enables them to predict where focused words will be located, and to direct their attention to that position. Given the differences in method for marking prosodic focus between English and Korean, it is unclear if such processing benefits also occur in Korean.

The Korean accentual phrase (AP) is the smallest prosodic unit demarcated by a pitch contour [7, 8]. It tends to have two to five syllables and can contain more than one lexical item. The initial segment of the first syllable determines whether or not the intonational contour starts at a higher or lower pitch and following that, each AP follows a consistent intonational contour [9]. Therefore the intonational contour of an individual word will depend on its location within an AP, rather than on word-internal metrical structure as seen in Germanic languages [10]. The AP serves an important role in marking prosodic focus, with the focused word placed at the head of a new AP [8]. This is similar to focus marking in French, however the highlighted word in French is placed at the end of an AP [11, 12]. Phonetically, a focused AP tends to have extra strengthening on its left edge, with the subsequent sequence tending to be shorter and produced with a smaller pitch range than in a neutral phrase [13, 14]. This contrasts with prosodic focus in English and Dutch, which uses a pitch accent and is not inherently tied to a position within a phrase. Whether or not the sequence before is altered leading into the focused AP is debated. Some studies have found that the sequence before is shorter [13], whereas others have argued that there is no difference pre-focus, only post-focus [15].

Few studies have investigated the role that focus plays in processing in Korean. Lee and colleagues [16] asked speakers of Korean, English, and Mandarin Chinese to produce number strings (in their native language), and contrastively focus a different number each production. They found that in English and Mandarin Chinese, words with contrastive focus were produced with greater duration, f_0 , and intensity in comparison with words produced without contrastive focus. In Korean however, there did not seem to be a clear acoustic difference between words that were prosodically focused, versus the same words without focus. Lee et al. also asked native listeners of each language to identify the number strings as well as identify which number had received focus. Consistent with their production data, Korean listeners performed far less well (40-50% accuracy) in comparison to listeners of English and Mandarin Chinese (94-97% accuracy). This would suggest that prosody does not play a role in marking salience in Korean, and also that listeners do not use this information during processing. Lee concluded that contrastive prosodic focus is clearly marked in English and

Mandarin Chinese but not marked or recognised accurately in Korean. This is at odds with other data [14, 15, 17] showing acoustic differences between words with and without prosodic focus, and with the general proposal that marking of salience is a prosodic universal [1].

The aim of this study was to investigate the role of focus in sentence processing in Korean by testing recognition memory for words that were focused (either syntactically or prosodically, or both) in comparison to unfocused words.

2. Method

2.1. Participants

32 native speakers of Korean (12M, 20F) were recruited for participation ($M_{age}=25.53$, $SD=7.89$). All participants were born in Korea, excepting one who was born in Australia but returned to Korea shortly after birth. All participants were also proficient in the English language and resided in Australia at the time of the experiment. All participants reported still regularly using Korean (at least weekly). They had no self-reported speech, reading, or hearing issues.

2.2. Stimuli

140 experimental sentences were constructed. Each sentence contained two possible high frequency target words: one target with syntactic focus, and one without syntactic focus. No target words were repeated across sentences. Word frequencies between targets within sentences were not significantly different from each other, $t(139) = .59$, $p = .56$.

A female native speaker of Seoul Korean recorded two versions of each sentence. One version had prosodic focus on the first target and the other version had prosodic focus on the second target. This created four different experimental conditions (two per sentence). For example:

1. 그녀의 선배 (NF) 가 요즘 관심을 두는 것은 **논문** (PS) 이었어.

What her colleague (NF) is currently interested in is a **thesis** (PS).

2. 그녀의 선배 (PR) 가 요즘 관심을 두는 것은 **논문** (ST) 이었어.

What her colleague (PR) is currently interested in is a **thesis** (ST).

(NF = no focus, PR = prosodic focus, ST = syntactic focus, PS = prosodic + syntactic focus. Prosodic focus indicated by bold type) To encourage the speaker to place focus on the correct target, two questions were written for each sentence and provided to the speaker during recording. Participants only heard one version of each sentence to ensure any effect was not simply due to repeated presentation of the same words.

Due to the structure of Korean, words with syntactic focus were constrained to the second target word, so we also created 20 control sentences, similar to the experimental sentences, with target words in early and late positions of the sentence but without syntactic or prosodic focus on either target word. A further 40 filler sentences were created without syntactic or prosodic focus, or target words. For these sentences we selected one foil target word that did not appear in the sentence that would be used in the recognition memory test.

2.3. Procedure

Prior to the experimental session, participants completed an online screening questionnaire that asked questions about their language experience to ensure that they were still regular users of Korean.

Sentences were pseudo-randomised so that all sentence types were evenly dispersed across the experiment. They were presented in blocks of 10, with 20 blocks in total across the experiment. Each block had seven experimental sentences, two filler sentences, and one control sentence. In total, participants listened to 140 experimental sentences, 20 control sentences, and 40 filler sentences.

We created four different versions of the experiment, that varied the order of presentation of stimuli, as well as the sentence versions and target words tested in the experimental sentences. For example, in experiment version one the target word “colleague” was tested in the no focus condition and in version two, it was tested in the prosodic focus condition. In versions three and four, the target word “thesis” was tested instead of the word “colleague”, in syntactic, and prosodic + syntactic conditions respectively. Participants were randomly assigned to one of the four experimental versions, with numbers equal in each of the groups.

Participants were tested individually in either a sound attenuated room or a quiet room, with the study taking around 45 minutes in total. The experiment was presented using E-Prime software on an Acer laptop that was placed in front of the participant at a distance they felt was comfortable for them. Noise-attenuating Sennheiser HD 280 Pro headphones were fitted and adjusted to achieve comfortably audible volume levels for each participant individually.

Participants listened to an entire block of 10 sentences, pressing a space bar to play each sentence. This was to encourage participants to attend to each sentence as it played. During this time, there was a fixation cross on the screen in front of them. After this block was complete, instructions appeared on screen telling participants to prepare for the recognition memory test. Participants were presented with 10 words sequentially on screen (one from each sentence) and asked to respond “yes” or “no” as to whether they remembered hearing that word in the preceding block of sentences. For right-handed participants ($N=31$), the letter “M” was pressed to indicate “YES” and “Z” to indicate “NO”, and for left handed participants ($N=1$) it was the reverse so that the positive response was always indicated by the dominant hand. If participants did not record a response to a particular trial, it timed out after five seconds. Participants were instructed to respond as quickly and as accurately as possible.

After each recognition memory test block, instructions informed the participant that they could take a short break and start the next block of sentences when they were ready. Participants completed one full block of practice items before commencing the experiment to ensure that they understood the experimental procedure; the experimenter was available to answer any questions that arose if needed. Order of sentence and words presented in the recognition memory block from those sentences was kept the same (rather than being scrambled) to ensure that the time delay between presentation of sentences and testing of target words from those sentences was kept as constant as possible across the conditions in which responses to the targets were to be compared.

3. Results

Recognition accuracy scores were aggregated within prosodic condition, to yield a proportion of correct responses for each participant. Figure 1 displays the mean response accuracy across conditions. Because the dependent variable was not a continuous variable, we were unable to use an ANOVA, and instead used a multinomial logistic regression, comparing accuracy across the four conditions. The overall model was significant, Wald $\chi^2(3, N=32) = 33.19, p < .001$.

Responses in each of the conditions with focus were significantly more accurate than responses in the control (no focus) condition. Using follow-up Bonferroni pairwise comparisons, we further found that responses in the prosodic focus and syntactic focus conditions were not significantly different from one another (M difference = .045, $p = .26$), but that the proportion of correct responses in the prosodic + syntactic focus condition was significantly higher than that in the condition with prosodic focus alone (M difference = .09, $p < .001$), while not differing significantly from that in the condition with syntactic focus alone (M difference = 0.048, $p = .17$).

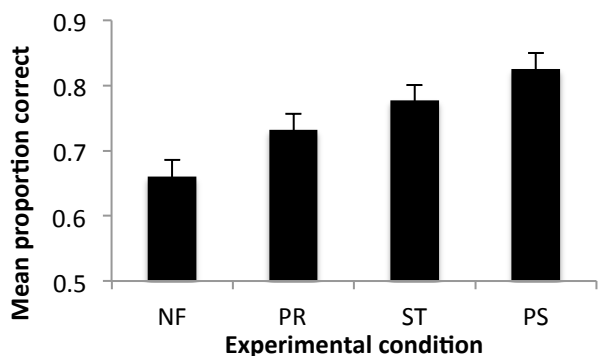


Figure 1. Mean accuracy scores for each of the experimental conditions. Error bars represent standard error of the mean.

We also aggregated reaction time responses for each participant to obtain a mean reaction time dependent variable. The mean reaction times per condition are shown in Figure 2. We analysed these using a within-subjects ANOVA with custom contrasts to compare between conditions. The overall model was significant, $F(3,31) = 15.71, p < .001$. There was a significant effect of focus (no focus compared to all focus conditions), $F(1,31) = 29.53, p < .001$, with words with focus (regardless of focus type) recognised significantly faster than words without focus.

In comparisons between the separate conditions, we found that words with prosodic focus were recognised significantly faster than words in the control condition, $F(1,31) = 10.04, p = .003$. Within the focus conditions, words with some kind of syntactic focus were recognised faster than words with prosodic focus alone, $F(1,31) = 11.58, p = .002$. In line with the accuracy results, however, there was no significant reaction time difference between the two conditions with syntactic focus, $F(1,31) = 1.91, p = .18$.

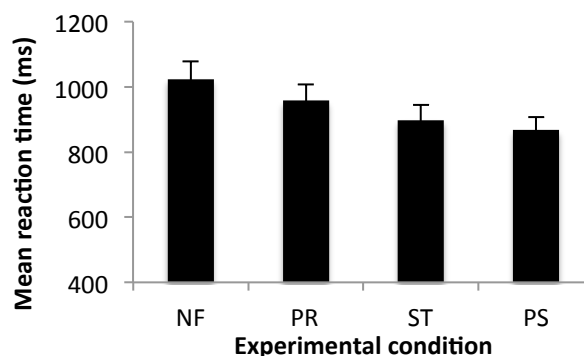


Figure 2. Mean reaction times for each of the experimental conditions. Error bars represent standard error of the mean.

We further compared reaction time for words from the experimental sentences versus the control sentences, to determine whether our results as just described were indeed due to the focus manipulation, or could be ascribed to sentence position. First, we compared reaction time to words that appeared as early targets (no focus condition, prosodic focus condition) with the early targets in control sentences; this comparison used a within-subjects ANOVA with mean reaction time as the dependent variable and condition as the independent variable (3 levels: NF, PR, early control). We followed this with Bonferroni adjusted pairwise comparisons. The overall model narrowly missed significance, $F(2, 31) = 2.99, p = .057$. There was no significant difference between control targets and targets without focus (M difference = 7.99, $p = 1$), but crucially, reaction times for prosodic targets were significantly faster than for control targets (M difference = 65.63, $p = .01$).

We then used a similar within-subjects ANOVA to compare words that appeared as the late target in the experimental conditions (words with syntactic focus, and prosodic plus syntactic focus) and late targets in the control sentences. The overall model was significant, $F(2,31) = 11.89, p < .001$, but crucially, words in both experimental conditions with syntactic focus were recognised significantly faster than control words: M difference = 105.49, $p = .006$, for words with syntactic focus alone; M difference = 135.93, $p = .001$, for words with syntactic plus prosodic focus.

4. Discussion

Our results have shown that Korean words with focus – whether signalled by prosody or by syntax – are more likely to be remembered, and are recognised more rapidly than words without focus. Within those words with focus, if we look at accuracy scores, words in the prosodic focus alone, and the syntactic focus alone conditions, were both more accurately recalled than words without focus, but recall of words with both prosodic focus *and* syntactic focus was significantly more accurate again in comparison to recall for words with prosodic focus alone. Turning to the reaction time scores, again, all focus conditions had significantly faster reaction times compared to words without focus, but both conditions with syntactic focus (ST and PS) had significantly faster reaction times compared to the prosodic focus condition. Thus it seems that while both focus types improve recognition memory for words, syntactic focus may exercise an even stronger effect on recognition memory than prosodic focus.

By comparing reaction times for target words in control sentences with words in the experimental sentences, we ruled out the possibility that our findings here were due solely to an effect of sentence position. In comparison to words in the same position in control sentences (with neither syntactic nor prosodic focus), words with syntactic focus were recognised significantly faster. The same was true for words with prosodic focus – participants responded significantly faster to words with prosodic focus than to early (unfocused) targets in the control sentences. Focused words in the final position in a phrase were not recognised faster because of where they were in the sentence, but their processing advantage seems to be due to the fact that they were focused.

Our results thus point to the processing advantage for focused words being located in the higher levels of utterance recognition (e.g., computation of a semantic representation of the spoken input) rather than the lower levels (representation of acoustic properties or temporal recency of spoken words).

We note that our results here contrast with those reported in the study of Lee et al. [16], in which Korean speakers did not consistently mark focus prosodically on an intended target, and Korean listeners were also far less accurate than English or Mandarin Chinese listeners at identifying the location of focused numbers within a native-language number string. We suggest, however, that the discrepancy may be due to the way focus was elicited in the earlier study. Lee et al. [16] asked their speakers to produce American style telephone numbers (in an XXX-XXX-XXXX format), thus constraining the phrasal structure. If a speaker were asked to place prosodic focus on the fifth number for example, it would be located in mid-phrase rather than at a phrasal edge. Given that, as described earlier, phrase position and focus in Korean are linked [13], this procedure may have led to marking of focus that was hard for speakers to realise and may indeed have been inconsistent. This is not the case for either Mandarin Chinese or English. If participants had been able to determine their own phrase structure in the earlier study, rather than it being defined by the experimenter, then an effect of focus in Korean may have appeared.

The stimuli designed for this experiment were carefully created and recorded. We are in the process of analysing acoustic measurements of the stimuli in order to further verify the relationship between focus and recognition memory.

Despite assessing processing through recognition memory (perhaps less sensitive than online measures), we still found a processing benefit for words with focus compared to words without focus in Korean. Future plans are to use online processing measures to further validate these results.

Across languages, our findings are consistent with work in English showing that words with syntactic focus are better remembered [2, 3], and that words with prosodic focus also receive processing benefits in English and Dutch [5, 6]. No studies to our knowledge have compared the effect of syntactic and prosodic focus on processing within the same experiment in English, so this is also an avenue for future study.

5. Conclusion

These results contribute to the growing body of word assessing prosodic processing across languages. While prosodic focus is marked differently in Korean compared to other languages previously investigated, our findings show that listeners still receive perceptual benefits from increased saliency of words.

This lends weight to the proposition that if some marking of prosodic salience is universal, then it may also be language universal that the speech processing system is optimised to prioritise understanding of words in prosodically salient positions, regardless of how that salience is produced in a specific language.

6. Acknowledgements

This research was conducted with support from the ARC Centre of Excellence for the Dynamics of Language (project ID: CE140100041)[HK, AC].

7. References

1. Bolinger, D., *Intonation across languages*. Universals of human language, 1978. **2**: p. 471-524.
2. Birch, S.L. and S.M. Garnsey, *The Effect of Focus on Memory for Words in Sentences*. Journal of Memory and Language, 1995. **34**(2): p. 232-267.
3. Birch, S.L., J.E. Albrecht, and J.L. Myers, *Syntactic focusing structures influence discourse processing*. Discourse Processes, 2000. **30**(3): p. 285-304.
4. Cutler, A. and D.J. Foss, *On the Role of Sentence Stress in Sentence Processing*. Language and Speech, 1977. **20**(1): p. 1-10.
5. Akker, E. and A. Cutler, *Prosodic cues to semantic structure in native and nonnative listening*. Bilingualism: Language and Cognition, 2003. **6**(02): p. 81-96.
6. Cutler, A., *Phoneme-monitoring reaction time as a function of preceding intonation contour*. Perception & Psychophysics, 1976. **20**(1): p. 55-60.
7. Jun, S.-A., *The accentual phrase in the Korean prosodic hierarchy*. Phonology, 1998. **15**(02): p. 189-226.
8. Jun, S.-A., *K-ToBI (Korean ToBI) labelling conventions*. UCLA Working Papers in Phonetics, 2000. **99**: p. 149-173.
9. Jeon, H.-S. and F. Nolan, *Segmentation of the Accentual Phrase in Seoul Korean*. Proceedings of Speech Prosody 2010, 2010. **10023**: p. 1-4.
10. Jun, S.-A., *Prosody in sentence processing: Korean vs. English*. UCLA Working Papers in Phonetics, 2005. **104**: p. 26-45.
11. Jun, S.-A. and C. Fougeron, *Realizations of accentual phrase in French intonation*. Probus, 2002. **14**(1): p. 147-172.
12. D'Imperio, M., J. German, and A. Michelas, *A multi-level approach to focus, phrasing and intonation in French*. Prosody and Meaning, 2012: p. 11-34.
13. Jun, S.-A. and H.-J. Lee, *Phonetic and phonological markers of contrastive focus in Korean*. in *Icslp*. 1998.
14. Oh, M., *Focus and prosodic structure*. 음성과학, 2001. **8**(1): p. 21-32.
15. Lee, Y.-C. and Y. Xu, *Phonetic realization of contrastive focus in Korean*. Proc. of Speech Prosody 2010, 2010.
16. Lee, Y.-C., et al., *A crosslinguistic study of prosodic focus*. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2015.
17. Jun, S.-A. *Prosodic markings of complex NP focus, syntax, and the pre-/post-focus string*. in *Proceedings of the 28th West Coast Conference on Formal Linguistics*. 2011.