

# PHONETIC FOCUS-MARKING IN KOREAN-SPEAKING 7- to 8-YEAR-OLDS AND ADULTS

Anqi Yang<sup>1</sup>, Taehong Cho<sup>2</sup>, Sahyang Kim<sup>3</sup> and Aoju Chen<sup>1, 4</sup>

<sup>1</sup>Utrecht University, <sup>2</sup>Hanyang University, <sup>3</sup>Hongik University, <sup>4</sup>Max Planck Institute for Psycholinguistics  
a.yang1@uu.nl, taehong.cho@gmail.com, sahyang@gmail.com, aaju.chen@uu.nl

## ABSTRACT

This study investigates how Seoul Korean-speaking children and adults use pitch- and duration-related phonetic cues to mark focus. It was found that to distinguish focus from non-focus, the adults used both the pitch- and duration-related cues, but the children used only the duration-related cues to distinguish focus from post-focus. Further, neither the adults nor the children distinguish narrow focus and broad focus via any of the phonetic cues. However, while the adults did not distinguish contrastive focus from (non-contrastive) narrow focus phonetically, the children distinguish these two using duration in the ‘short’ words.

**Keywords:** focus, phonetic, prosody, L1 acquisition, Korean

## 1. INTRODUCTION

The term ‘focus’ is an information structural category and refers to the new information in a sentence to the receiver [1, 2]. This study involves three types of focus, i.e., narrow focus, broad focus, and contrastive focus. Narrow focus (on a word or words within a sentence) differs from broad focus (over a whole sentence) in the size of the focal constituent, while differs from contrastive focus in that the latter conveys an explicit contrast to an alternative in the context [3].

Focus can be encoded via prosody in many languages either phonetically (i.e. via gradient variation in pitch/duration) or phonologically (e.g. via accentuation/phrasing) [4]. In West Germanic languages like in Dutch, although both phonological and phonetic cues are used, the phonological cues are the primary cues and are acquired earlier than the phonetic ones [5]. To give an example, the 7- to 8-year-old Dutch-speaking children can use accentuation to mark focus and their choice of accent type also become adult-like at this age [5, 6]. However, although they can vary the phonetic realisation of a pitch accent in terms of pitch range for focus-marking purposes, the use of duration for this purpose is still not acquired [7]. In contrast, in

Mandarin Chinese the encoding of focus depends on the phonetic manipulation of pitch and duration [8], and no phonological cues are reported to be used. At the age of seven or eight, Mandarin Chinese-speaking children's use of pitch and duration for focus marking purposes is already in position [9]. These findings together suggest that the use of the major focus-marking cue in each language, like accentuation in Dutch and the use of pitch and duration in Mandarin Chinese, is acquired earlier than the use of the secondary cues [9].

In Seoul Korean, phrasing has been widely considered as a typical focus-marking cue [10], while pitch- and duration-related phonetic cues are also used [11]. In the light of findings on Dutch- and Mandarin-Chinese speaking children, we hypothesise that Seoul Korean-speaking 7- to 8-year-olds may differ from their Mandarin Chinese-speaking peers in their mastery of phonetic focus-marking in terms of rate. To test this hypothesis, we investigated how Seoul Korean-speaking 7- to 8-year-olds used pitch- and duration-related cues (1) to distinguish focus from non-focus (effect of focus), (2) to distinguish narrow focus from broad focus (effect of focal constituent size), (3) and to distinguish contrastive focus from (non-contrastive) narrow focus (effect of contrastivity), in comparison to adults.

## 2. METHOD

### 2.1. Target words and sentences

We aimed to elicit 60 SOV sentences (5 focus conditions × 12 objects) from each participant. The target words were the sentence medial objects (six ‘short’ objects, each with two syllables; six ‘long’ objects, each with four syllables). Five subjects (dog, rat, bear, horse, and cow), and three verbs (draw, touch, and look for) were almost evenly distributed over the sentences. The five focus conditions were: (1) Narrow focus on the sentence initial subject (NF-i); (2) Narrow focus on the sentence-medial object (NF-m); (3) Narrow focus on the sentence final verb (NF-f); (4) Contrastive focus on the sentence medial object (CF-m); (5) Broad focus over a whole

sentence (BF). The target words were on-focus in the NF-m condition, pre-focus in the NF-f condition and post-focus in the NF-i condition and were thus ideal for direct comparisons between focus and non-focus (i.e., pre-/post-focus).

To address the first research question, namely, how focus differs from non-focus, we compared the prosody of the target words in the NF-m condition with that in the NF-i and the NF-f conditions. To address the question about focal constituent size, we compared the prosody of the target words in the NF-m condition with that in the BF condition. To address the question on contrastivity, we compared the prosody of the target words in the NF-m condition with that in the CF-m condition.

## 2.2. Speech elicitation

To elicit the target sentences, question-answer dialogues between the experimenter (Exp) and the speaker (Spe) as illustrated in examples (1) to (5) were embedded in a picture-matching game adapted from [6].

- (1) Exp: Look! A hand, and a loaf of bread. It looks like someone touches the bread. Who touches the bread?  
Spe: [Dog] bread touch. (NF-i)
- (2) Exp: Look! A dog, and it puts out its hand. It looks like the dog touches something. What does the dog touch?  
Spe: Dog [bread] touch. (NF-m)
- (3) Exp: Look! A dog, and a loaf of bread. It looks like the dog will do something to the bread. What does the dog do to the bread?  
Spe: Dog bread [touch]. (NF-f)
- (4) Exp: Look! A dog, and it puts out its hand. It looks like the dog touches something. I will make a guess: Dog egg touch.  
Spe: Dog [bread] touch. (CF-m)
- (5) Exp: Look! This picture is very blurry. I cannot see anything clearly. What happens in the picture?  
Spe: [Dog bread touch]. (BF)

Three piles of pictures were used in the game. The experimenter's pictures always missed some information, e.g. the subject, the object, the action, or all the three pieces of information. The speaker's pictures always contained all the three pieces of information. In every trial, the experimenter showed a picture of hers to the speaker and asked a question

about it, as illustrated in (1) to (5). The speaker took a look at the corresponding picture in his pile and answered the question or made a correction. The experimenter could then look for the right picture in a third pile which was laid around on the table and matched it with her own picture to form a pair.

## 2.3. Speakers

Seoul Korean speakers including eight 7- to 8-year-olds (six girls) and eight adults (four females) were tested individually in Hanyang Phonetics and Psycholinguistics Laboratory, Seoul.

## 2.4. Annotation and acoustic analysis

The audio recording from each speaker was annotated in Praat. Responses deviating from the target sentences in choice of word or sentence structure or produced with self-repairs and hesitations were excluded. In total, 335 sentences from the children and 431 sentences from the adults were analysed. Pitch-related measurements (the maximum and minimum pitch, and pitch range) and duration-related measurements (word duration and first-syllable duration) of the target words were analysed.

## 2.5. Statistical analysis

Mixed-effects modelling was used to assess the effect of fixed factors, i.e., AGE (2 levels: children vs. adults) and FOCUS (2 levels), and the effect of interactions between the fixed factors on the dependent variables, i.e., the maximum pitch, the minimum pitch, pitch range, word duration, and first-syllable duration for the 'short' and the 'long' words separately. The 'short' and 'long' words also differed in other aspects (like in consonantal onsets). To avoid involving more variables than needed to answer the research questions, they were analysed separately. Two random factors SPEAKER and SENTENCE were included. A model with only the random factors was built. AGE and FOCUS was then added to the model one by one to see whether they could significantly improve the model. Last, the interaction between AGE and FOCUS was examined.

### 3. RESULTS

#### 3.1. Focus and non-focus

##### 3.1.1. 'Short' words

Pitch-related measurements were first examined. There was no evidence for the use of the pitch-related cues by either the children or the adults to distinguish focus from either pre- or post-focus.

Duration-related measurements: FOCUS (2 levels: focus vs. post-focus) significantly improved the models in the analyses on word duration ( $p < .01$ ), and first-syllable duration ( $p < .01$ ), and no significant interaction between FOCUS and AGE was found ( $p > .05$ ) in both cases. Both the children and adults thus used a longer word duration and first-syllable duration for the focused words than the post-focal ones. Furthermore, analysing word duration and first-syllable duration, significant interactions between FOCUS (2 levels: focus vs. pre-focus) and AGE were found ( $p < .01$ ). The durational differences between these two conditions were significant in the adults' speech ( $p < .01$ ), but not in the children's speech ( $p > .05$ ) (Figure 1 & Figure 2). Thus, the adults used a longer word duration and first-syllable duration for the focused words than the pre-focal ones, but there was no evidence showing children did so.

Figure 1: Word duration (s): focus vs. pre-focus

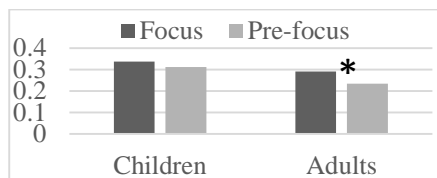
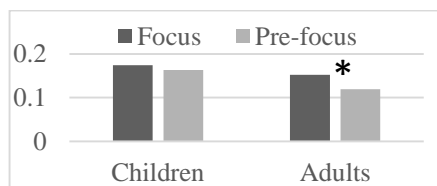


Figure 2: First-syll duration (s): focus vs. pre-focus

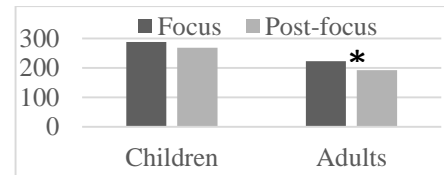


##### 3.1.2. 'Long' words

Pitch-related measurements: Analysing the maximum pitch, a significant interaction was found between FOCUS (2 levels: focus vs. post-focus) and AGE ( $p < .05$ ). The maximum pitch was only significantly higher in the focused words than in the post-focal ones in the adults' speech ( $p < .05$ ), but was not in the children's speech ( $p > .05$ ) (Figure 3). Thus, only the adults used the maximum pitch to

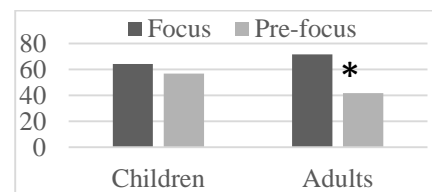
distinguish focus from post-focus. There was no evidence for the use of the minimum pitch or pitch range by either the children or adults to distinguish focus from post-focus.

Figure 3: Max pitch (Hz): focus vs. post-focus



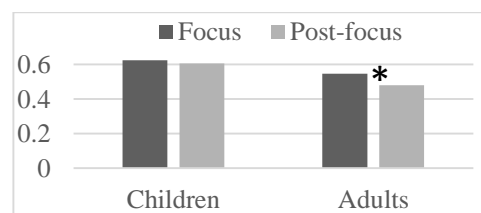
Furthermore, analysing pitch range, a significant interaction was found between FOCUS (2 levels: focus vs. pre-focus) and AGE ( $p < .01$ ). The pitch range was significantly wider in the focused words than in the pre-focal ones in the adults' speech ( $p < .05$ ), but was not in the children's speech ( $p > .05$ ) (Figure 4). Thus only the adults used pitch range to distinguish focus from pre-focus. Besides, there was no evidence showing either the children or adults used either the maximum pitch or the minimum pitch to distinguish focus from pre-focus.

Figure 4: Pitch range (Hz): focus vs. pre-focus



Duration-related measures: Analysing word duration, a significant interaction between FOCUS (2 levels: focus vs. post-focus) and AGE was found ( $p < .05$ ). Word duration was significantly longer in the focused words than in the post-focal ones in only the adults' speech ( $p < .05$ ), but was not in the children's speech ( $p > .05$ ) (Figure 5). Thus, only the adults used word duration to distinguish focus from post-focus. There was no evidence for the use of first-syllable duration by either the children or adults to distinguish focus from post-focus.

Figure 5: Word duration (s): focus vs. post-focus



Besides, there was no evidence showing either the children or adults used either word duration or

first-syllable duration to distinguish focus from pre-focus.

### 3.2. Focal constituent size

There was no evidence showing either the children or the adults used the pitch- or duration-related cues to distinguish narrow focus from broad focus in either the 'short' or the 'long' words.

### 3.3. Contrastivity

#### 3.3.1. 'Short' words

Pitch-related measurements were first examined. There was no evidence for the use of the pitch-related cues by either the children or the adults to distinguish narrow focus from contrastive focus.

Duration-related measures: Analysing word duration, a significant interaction between FOCUS and AGE was found ( $p < .01$ ). The difference in word duration between these two conditions was significant for the children ( $p < .05$ ), but not for the adults ( $p > .05$ ). Analysing first-syllable duration, a significant interaction between FOCUS and AGE was found ( $p < .05$ ). The difference in first-syllable duration between these two conditions approached significance for the children ( $p = .05$ ), but was not significant for the adults ( $p > .05$ ). Thus, the children used a longer word duration and first-syllable duration in contrastive focus condition than in narrow focus condition, but there was no evidence showing that the adults did so.

#### 3.3.2. 'Long' words

There was no evidence showing either the children or the adults used the pitch- or duration-related cues to distinguish narrow focus from contrastive focus in the 'long' words.

## 4. DISCUSSION AND CONCLUSIONS

Regarding the distinction between focus and non-focus (effect of focus), we found that in the 'short' words the adults used only duration to distinguish focus from non-focus, while the children also used duration but only to distinguish focus from post-focus. In the 'long' words, the adults largely relied on pitch to distinguish focus from non-focus. Possibly, it could be difficult to mark focus on the four-syllable 'long' words by making them even longer, so the adults preferred to vary pitch. The children did not even distinguish focus from non-focus using the pitch- and duration-related cues in the 'long' words. These results suggested that first,

the adults may adjust their focus-making strategy in the choice of the phonetic cues although they can vary both pitch and duration. Second, the children have learned to use duration to distinguish focus from post-focus but not from pre-focus. Third, duration seems to be easier to vary than pitch for focus-marking purposes (unless the words to be produced are rather long), and thus is the primary cue to be used by the adults (in the 'short' words) and also the use of duration seems to be acquired earlier than the use of pitch for focus-marking purposes by the children.

Regarding the distinction between narrow focus and broad focus (effect of focal constituent size), we found that focal constituent size was not phonetically distinguished by the children or the adults.

Regarding the distinction between contrastive focus and (non-contrastive) narrow focus (effect of contrastivity), we found that contrastivity was not encoded phonetically by the adults. The children made the distinction between these two types of focus but only using duration-related cues in the 'short' words. The difference between the adults and the children may be caused by that the children were much eager to make corrections in the game. The children's choice of phonetic cues further proves that they have acquired the use of duration but not pitch for focus-marking purposes.

Cross-linguistically, at the age of seven or eight, the Seoul Korean-speaking children could vary duration but not pitch to distinguish focus from post-focus, while the Mandarin Chinese-speaking children could vary both duration and pitch range to distinguish focus from post- and pre-focus [9]. Our hypothesis on differences in rate of acquisition between these two groups of children is thus borne out. Together with findings on Dutch and Mandarin-speaking children, our findings show clearly that whether the phonological cues or the phonetic cues are the primary markers of focus plays a role in how fast children acquire the phonetic marking of focus in their target language.

## 5. REFERENCES

- [1] Lambrecht, K. 1994. *Information structure and sentence form: Topics, focus, and the representations of discourse referents*. Cambridge: Cambridge University Press.
- [2] Gundel, J. K. 1999. On different kinds of focus. In: Bosch, P., van de Sandt, R. (eds), *Focus: linguistic, cognitive, and computational*. Cambridge: Cambridge University Press.

- [3] Gussenhoven, C. 2005. Types of focus in English. In: Buring, D., Gordon, M., Lee, C. (eds), *Topic and focus: intonation and meaning. Theoretical and crosslinguistic perspectives*. Dordrecht: Kluwer.
- [4] Chen, A. 2012. Prosodic investigation on information structure. In: Krifka, M., Musan, R. (eds), *The expression of information structure*. Berlin: De Gruyter Mouton, 251-286.
- [5] Chen, A. 2011a. The developmental path to phonological focus-marking in Dutch. In: Frota, S., Elordieta, G., Prieto, P. (eds), *Prosodic Categories: Production, Perception and Comprehension*. Springer Netherlands, 93-109.
- [6] Chen, A. 2011b. Tuning information packaging: intonational realization of topic and focus in child Dutch. *J. Child Lang*, 38, 1055-1083.
- [7] Chen A. 2009. The phonetics of sentence-initial topic and focus in adult and child Dutch. In Vigário, M., Frota, S., Freitas, M. J. (eds), *Phonetics and Phonology: Interactions and interrelations*. Amsterdam: Benjamins, 91-106.
- [8] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of f0contours. *Journal of Phonetics*, 27(1), 55-105.
- [9] Yang, A., Chen, A. 2014. Prosodic focus marking in Chinese four- and eight-year-olds. *Proc. 7<sup>th</sup> Speech Prosody*, Dublin.
- [10] Jun, S. A. 2006. Intonational Phonology of Seoul Korean Revisited, *Japanese-Korean Linguistics 14*, Stanford: CSLI.
- [11] Lee, Y. C., Xu, Y. 2010. Phonetic Realization of Contrastive Focus in Korean. *Proc. 5<sup>th</sup> Speech Prosody*, Chicago.