

# The acquisition of prosodic focus marking in Central Swedish: Sorting out lexical and post-lexical tones

Anna Sara H. Romøren, Utrecht University  
Aoju Chen, Utrecht University / Max Planck Institute for Psycholinguistics

## ABSTRACT

In many languages prosody is used for marking focus, but the way this is done varies from one language to the next. Previous work on the acquisition of prosodic focus marking has been centred on English, German and Dutch, thus little is known about how prosodic focus is acquired outside West-Germanic varieties. We present a study of prosodic focus marking in Central Swedish, showing how the focus-marking high tone is used in an adult-like way from the age of seven, but that children between four and five are still developing toward adult proficiency.

**Keywords:** Prosody, Swedish, acquisition, information structure, phonology

## 1. INTRODUCTION

Speakers use a range of linguistic means to highlight important information, or focus, and in most Germanic languages, prosody is used for this purpose [18]. North Germanic varieties are particularly interesting for investigating the acquisition of prosodic focus marking, as the presence of a lexical accent contrast makes the means for marking focus work differently from what we observe in English, German and Dutch. In the current study we investigate how Central Swedish - speaking children between 4 and 11 mark focus using prosody.

## 2. BACKGROUND

### 2.1. Information packaging

Information structure relates to the way speakers structure their utterances in accordance with the common ground shared between speaker and listener. The basic idea is that some parts of a sentence anchors it to the previous discourse while others make a new contribution, updating the common ground between speaker and hearer [11]. Following many others, we refer to such new information as the ‘focus’ of a sentence, and three sub-types of focus are relevant for this study. The first is ‘narrow focus’, where one constituent in a

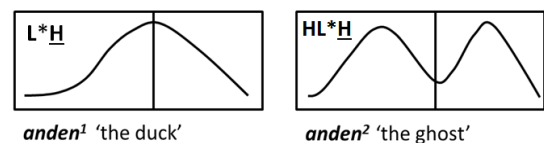
sentence represents new information, while the others refer to given information, as in (1). The second type is ‘contrastive focus’, where alternative candidates to the focal referent are explicitly mentioned, as in (2). The third type is ‘broad focus’, used for sentences in which every constituent refers to new information, as in (3).

- (1) **Speaker 1:** What’s the dog doing with the cake?  
**Speaker 2:** The dog [is throwing]<sub>F</sub> the cake.
- (2) **Speaker 1:** The dog [is eating] the cake.  
**Speaker 2:** The dog [is throwing]<sub>F</sub> the cake.
- (3) **Speaker 1:** What's happening?  
**Speaker 2:** [The dog is throwing the cake.]<sub>F</sub>

### 2.2. Central Swedish (CS)

CS is a lexical pitch accent language, in which all words carry either an accent 1 (early fall, HL\*) or an accent 2 (late fall, H\*L) pitch pattern, aligned with the main stressed syllable. [16]. Focus is marked by adding a separate prominence-marking high tone (H) *after* the lexical part of the contour, resulting in a L\*H contour on accent 1 and a H\*LH contour on accent 2 words (Figure 1)<sup>1</sup>.

**Figure 1:** Lexical accent + focus marking H (underlined) on the minimal pair *anden*<sup>1</sup> (‘the duck’) / *anden*<sup>2</sup> (‘the ghost’), spoken in isolation.



The above situation gives rise to two separate levels of prosodic prominence in CS, one involving the word accents only, and one involving the combination of word accent and prominence marking H. Myrberg & Riad [15] propose the term ‘small accent’ for the former and ‘big accent’ for the latter, and we will adopt these terms in the following.

Prominence-marking H is typically not added to given information, and post-focal lexical accents

are described as downstepped (i.e. [3],[4], [13],[14],[1],[16]).

### 2.3. Acquiring prosodic focus

Studies of prosodic focus marking in English-speaking children have shown that they mark contrastive focus with accentuation by the age of 4 to 5, but that they still develop their proficiency toward the age of 6 [9],[12],[20] or even 11 [19]. Dutch-speaking children at 4 to 5 largely accent focal and de-accent post-focal information in an adult-like way, but differ from adults in their choice of accent type [5],[6]. Non-adult accent type choices have also been reported for German-speaking 5- to 7- year olds, even if their use of accentuation for distinguishing between information states is similar to that of adults [7]. To the best of our knowledge there are no previous studies of prosodic focus in child Swedish, but several studies have shown the lexical accent contrast to be acquired by the age of 2 [8], [10].

## 3. RESEARCH QUESTIONS

Two research questions are investigated in the current study. The first concerns the use of big accent on the focal constituent, and we ask:

1. Do CS-speaking children between 4 and 11 use ‘big accents’ to mark narrow focus in the same way that adults do?

The second research question concerns the post-focal area, and we ask:

2. Do CS-speaking children between 4 and 11 avoid ‘big accents’ post-focally in the same way that adults do?

In order to answer these questions, SVO sentences with varying information structure were elicited as part of a picture game. The use of big accent was then investigated for the sentence-medial and sentence-final position separately. Medial targets were always verbs and final targets were always objects. To answer the first research question we compared conditions rendering the target word narrowly or contrastively focal to the broad focus baseline. To answer the second research question, we compared conditions rendering the target word post-focal to the broad focus baseline.

For the adults we expected a dominance of big accents on narrow focus and avoidance of big accents post focally. In most Germanic languages, broad focus involves aligning maximal

prominence with the right edge of the intonation phrase, thus we expected smaller differences between broad and narrow focus sentence-finally than sentence-medially. Based on previous descriptions of prosodic focus in English-, German- and Dutch-speaking children, we expected our older children to be adult-like in the use of big accent for focus, but that the younger children might be more variable in their big-accent-to-focus mapping.

## 4. METHOD

### 4.1. Participants

Twenty-six CS-speaking children and ten CS-speaking adults participated in the study (Table 1).

**Table 1.** Participant information

Group	N	Age range	Mean age	Gender
Five	10	4;3-5;6	5;0	6 m, 4 f
Eight	8	7;6-8;8	8;3	5 m, 3 f
Eleven	8	10;0-11;0	10;6	4 m, 4 f
Adults	10	20;12-43;10	27;2	5 m, 5 f

The children were divided into three age groups, hereafter referred to as ‘group five’, ‘group eight’ and ‘group eleven’. None of the participants had any history of language disorders, and they were all 1<sup>st</sup> language speakers of CS. The children were recruited from kindergartens and schools in Stockholm, and the adults were recruited at The Royal Institute of Technology (KTH) in Stockholm.

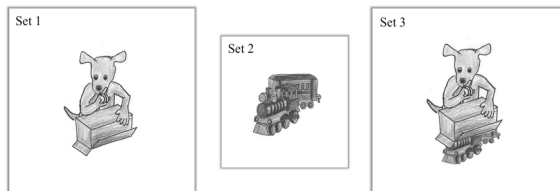
### 4.2. Procedure

The data collected consisted in semi-spontaneous SVO-sentence productions, elicited as part of a game. Sessions were conducted individually in a quiet room using a portable ZOOM H1 handy recorder, with a 44.1 kHz sampling rate and 16-bit accuracy.

#### 4.2.1. The picture-matching game

In the game (based on [6]), the participant’s task was to help the experimenter find correct picture combinations by answering the experimenter’s questions about her pictures. Scripted contexts were combined with wh-questions (see examples 1-3), so as to make focal elements new and non-focal elements given in the responses produced by the participants. Each session started with a picture-naming task, making sure the words used in the game were familiar to the participants.

**Figure 2.** One picture from each set, representing the target sentence *The dog is hiding [THE TRAIN]<sub>F</sub>*.



The materials consisted of three separate sets of pictures (see Figure 2). In the experimenter's set (set 1), the pictures always lacked one piece of information, represented by the subject, the verb or the object. Set 2 depicted these missing information bits individually, and was scrambled face up in a box. The participant's set (set 3) always displayed the complete actions. Sets 1 and 3 were pre-ordered before each session so that corresponding pictures appeared in the same trial.

Each trial was conducted as follows: The experimenter first picked up a picture from set 1, showing it to the participant, uttering scripted context sentences (e.g. *Look, the dog, and it looks like the dog is hiding something*), followed by a critical wh-question (e.g. *What is the dog hiding?*). After the question, the participant could look at his/her complete picture to answer (e.g. *The dog is hiding the train*), after which the experimenter could find the 'missing piece' in the box (set 2).

The game consisted of 30 trials, spread over five sentence conditions; narrow focus on initial constituent (NFI), narrow focus on medial constituent (NFM), narrow focus on final constituent (NFF), contrastive focus on medial constituent (CFM) and broad focus on the whole sentence (BF). The conditions were implemented by presenting non-focal constituents in the trial context, and by asking questions about the focal ones (see Examples 1-3). Six subject nouns, six transitive verbs and six object nouns (half accent 1, half accent 2) were distributed over the five sentence conditions. Each combination of initial, medial and final constituent only occurred once in the experiment; two consecutive trials never realized the same condition and always differed by minimally two constituents. The trials were arranged into two orders, to which the participants were randomly assigned.

### 4.3. Data selection and coding

Each session resulted in a set of SVO sentences, which were analysed using Praat [2]. Responses that did not follow the scripted speech context, or that contained deviant word choices, elided constituents, self-repairs, hesitations, or background noise, were excluded from further analysis. Included responses were segmented into words, and sentence-medial verbs and sentence-final objects were hand coded by the first author according to whether or not they carried a big accent (e.g. resulting in a L\*H contour in the case of accent 1 and a H\*LH contour in the case of accent 2). The coding was extracted to excel sheets using scripts, and the data was checked for extraction errors.

## 5. ANALYSIS AND RESULTS

### 5.1. Analysis

Statistical analyses were conducted building binomial logistic regression models (GLMs) in R [17]. The outcome variable was binary, consisting in the option of big accent or not on the target word. We started with a baseline model (model 0) in which only the intercept was included. We then expanded the model, adding the factor 'focus' in model 1, the factor 'group' in model 2 and the interaction between 'focus' and 'group' in model 3. 'Focus' was always a binary factor, comparing the narrow focus condition with the broad-focus baseline to answer the first research question, and the post-focus condition to the broad focus baseline to answer the second research question. For the medial comparison, NFM and CFM represented the narrow focus condition, and NFI the post-focus condition. Sentence finally, NFF represented the narrow focus condition, CFM and NFM represented the post-focus condition (see Section 4.2.1)<sup>2</sup>.

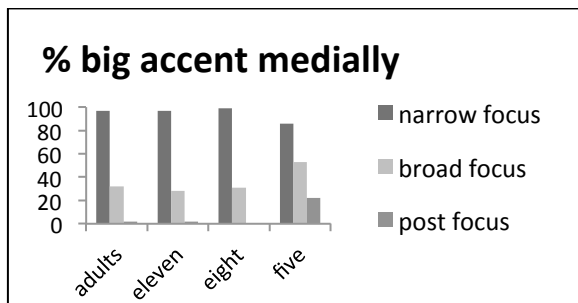
Our factor 'group' involved in four levels, using adults as a baseline to which groups five, eight and eleven were compared. Only factors significantly improving a previous model were included in subsequent models. The model fit was assessed using R's 'anova' function to compare pairs of models.

### 5.2. Big accent and focus sentence-medially

In sentence-medial position (Figure 3), groups eleven and eight were similar to the adults, using big accent at ceiling under narrow focus, at around 30% under broad focus and hardly ever post-

focally. Group five used big accents less consistently on narrow focus, and more for broad and post-focus than the other groups.

**Figure 3.** Percentage big accent sentence-medially across narrow, broad and post focus



Building models on the outcome ‘big accent’ with the factors ‘focus’ (‘broad focus’ vs. ‘narrow focus’), ‘group’ and ‘focus x group’, showed a significant effect of focus ( $p < 0.001$ ) and a significant interaction between focus and group for the comparison between adults and five-year-olds ( $p < 0.002$ ), but no other effects or interactions. In other words, narrow focus generally lead to more big accents, but this effect was significantly weaker in the five-year-olds than in the adults.

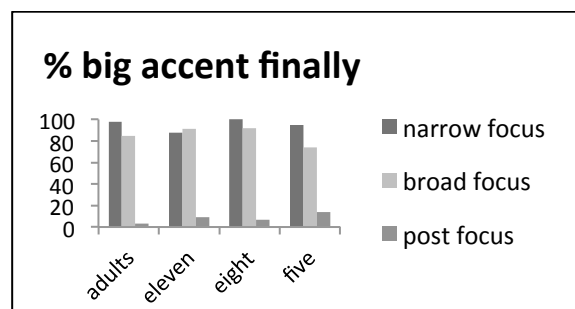
Investigating the effect of post focus on big accents, the factor ‘focus’ was used to compare the NFI condition, rendering targets post-focal, to BF. Because separation in the data from group eight (they never used big accent post focally) made the full models unreliable, this group was excluded from this comparison<sup>3</sup>. Building models for the rest of the data set revealed a main effect of focus ( $p < 0.001$ ) and a main effect of group for the comparison between group five and the adults ( $p < 0.001$ ). In other words, the effect of being post-focal did not differ significantly between the groups, but group five differed significantly from the adults in using more big accent under both broad and post focus.

### 5.3. Big accents and focus sentence-finally

As can be observed in Figure 4, our expectation of BF rendering more big accentuation finally than medially was confirmed. Again, there was separation in the data from group eight (the always used big accent under narrow focus), thus they were excluded from the narrow focus analysis, but not the post-focus analysis. Building models to answer the first research question about big accent for narrow focus revealed a main effect of focus ( $p < 0.009$ ), but no other effects or interactions. In other words, narrow focus generally lead to more

big accents than broad focus, and no between-group differences reached significance.

**Figure 4.** Percentage big accent sentence-finally across narrow, broad and post focus



Turning to the question of post-focal big accents, there was both a main effect of focus ( $p < 0.001$ ) and a significant interaction between focus and group for the comparison between adults and group five ( $p < 0.006$ ). In other words, the effect of being post-focal was significantly weaker in group five than in the adults.

## 6. DISCUSSION

Our most important finding is that CS speaking children between seven and eleven are adult-like in their use of big accents for marking narrow focus sentence medially and finally, using big accents on narrowly focal targets and avoiding these post-focally. We also observed that five-year-olds differed slightly from the adults, showing a weaker effect of narrow focus sentence-medially and a weaker effect of post-focus sentence-finally. However, the difference between five-year-olds and adults was mostly one of degree rather than of category: The children showed the same main effects as the adults, but the effects were weaker. In this sense, our findings mirror descriptions from other languages in showing that the phonological cues to focus are generally in place by four to five, but that they are still developing toward the age of seven. Future studies should involve additional phonetic measures on the accent patterns produced, as the literature suggests that the phonetics of prosodic focus marking might develop later than the ability to use the phonological categories in question [6].

## 7. ACKNOWLEDGEMENTS

This study is supported by NWO through a VIDI grant (276-89-001) awarded to the 2<sup>nd</sup> author. We thank Förskolan Greven, Förskolan Sjöboden, Råsunda Skola and Engelbrektskolan i Stockholm for their cooperation. We are also deeply grateful

for the work put down by research assistants Paula Cox, Frank Bijlsma, Martine Veenendaal and Saskia Versteegen, and to our colleagues Anqi Yang, Tom Lentz and Zenghui Liu, for valuable comments on an earlier draft of this paper. Finally, the biggest ‘thank you’ goes to all the children and adults participating in the study.

## 8. REFERENCES

- [1] Ambrazaitis, G. 2009. Nuclear intonation in Swedish. Lund University.
- [2] Boersma, P., Weenink, D. 2010. PRAAT: Doing phonetics by computer (version 5.1.25) [computer program]. Retrieved from <http://www.praat.org/>.
- [3] Bruce, G. 1977. Swedish word accents in sentence perspective. Lund University.
- [4] Bruce, G. 2007. Components of a prosodic typology of Swedish intonation. In T. Riad, & C. Gussenhoven (Eds.). *Tones and tunes. Typological and comparative studies in word and sentence prosody*. Berlin: De Gruyter, 113-146.
- [5] Chen, A. 2011a. The developmental path to phonological focus marking in Dutch. In S. Frota (Ed.). *Prosodic categories: Production, perception and comprehension. studies in natural language and linguistic theory*, 82. Dordrecht: Springer Science and Business Media B.V, 93-110.
- [6] Chen, A. 2011b. Tuning information packaging: Intonational realization of topic and focus in child Dutch. *Journal of Child Language*, 38, 1055-1083.
- [7] De Ruiter, L. 2010. *Studies on intonation and information structure in child and adult German*. Max Planck Institute for Psycholinguistics, Nijmegen.
- [8] Engstrand, O., Williams, K., Lacerda, F. 2003. Does babbling sound native? Listeners responses to vocalizations produced by Swedish and American 12- and 18-month-olds. *Phonetica*, 60(1), 17-44.
- [9] Hornby, P. A., Hass, W. A. 1970. Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research*, 13, 359-399.
- [10] Kadin, G., Engstrand, O. 2005. Tonal word accents produced by Swedish 18- and 24-month-olds. *Proc. Fonetik 2005*, 67-70.
- [11] Krifka, M., Musan, R. 2012. *The expression of information structure*. Berlin: Mouton De Gruyter.
- [12] MacWhinney, B., Bates, E. 1978. Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behaviour*, 17, 539-555.
- [13] Myrberg, S. 2009. The intonational phonology of Stockholm Swedish. Department of Scandinavian languages, Stockholm University.
- [14] Myrberg, S. 2013. Focus type effects on focal accents and boundary tones. *Proc. Fonetik 2013*, 53-56.
- [15] Myrberg & Riad (submitted). The prosodic hierarchy of Swedish.
- [16] Riad, T. 2013. *The phonology of Swedish* Oxford: Oxford University Press.
- [17] R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>
- [18] Vallduví, E., Engdahl, E. 1996. The linguistic realization of information packaging. *Linguistics*, 34, 459-519.
- [19] Wells, B., Peppé, S., Goulandris, N. 2004. Intonation development from five to thirteen. *Journal of Child Language*, 31, 749-778.
- [20] Wonnacott, E., Watson, D. G. 2008. Acoustic emphasis in four year olds. *Cognition*, 107(3), 1093-1101.

---

<sup>1</sup> One might assume that HL\*H is the natural consequence of adding prominence H to accent 1 (HL\*) words, but the leading H seems not to be consistently realized, thus we follow [13] and [16] in assuming the contour to be L\*H.

The CPM and NFM conditions, both rendering the verbs narrowly focal but differing in contrastivity, were collapsed in the analysis. This was done because no statistically significant differences were found comparing the use of big accent between the two.

<sup>3</sup> The eight-year-olds behave very similarly to both adults and eleven-year-olds, thus we believe it is safe to assume that effects found for these two groups are also present in the data from group eight.