Accentual lengthening in 5-year-old AusE-speaking children: preliminary results

Ivan Yuen, Nan Xu Rattanasone, Elaine Schmidt, Gretel Macdonald, Rebecca Holt, Katherine Demuth

Department of Linguistics, ARC Centre of Excellence in Cognition and its Disorders, Macquarie University

ivan.yuen@mq.edu.au;nan.xu@mq.edu.au;aeis2@cam.ac.uk;gretel.macdonald@gmail.com; rebecca.holt@mq.edu.au;katherine.demuth@mq.edu.au

Abstract

Although contrastive focus is reported in children's productions, their perception remains poor. However, judgement of contrastive focus in production is typically based on perceptual evaluation, and the stimuli in production and perception studies often differ in the number of syllables and sentence positions. Since both factors influence accentual lengthening in adults, this raises questions about children's ability to produce and generalize the durational cue to focus across different items. Eight AusE adults and 8 children participated in an elicited production task. Unlike the adults, children used accentual lengthening only on monosyllabic, not disyllabic words, and showed no additional phrase-final lengthening.

Index Terms: accentual lengthening, phrase-final lengthening, Australian English, contrastive focus

1. Introduction

When prosody is used to emphasize a specific word in an utterance, this function is known as 'focus' or 'accentuation'. A focused word is typically pitch accented, with increased pitch, duration and intensity [1], [2], [3], [4]. However, children show a discrepancy between their production and perception of focus [5]. Children's ability to produce appropriate adult-like acoustic cues to signal focus in previous studies might have been over-estimated. As the target words in previous production and perception studies varied not only in the number of syllables, but also sentence positions, the discrepancy might have resulted from children's inability to generalize cues to focus across number of syllables and sentence positions in production and perception.

In the adult literature, there is evidence that the number of syllables within a word influences the magnitude of accentual lengthening, one of the cues to signal focus. For instance, [6] found that Scottish English-speaking adults exhibited more lengthening on the accented syllable of a monosyllabic word (e.g., bake enforce) than a disyllabic word (e.g., bacon force). In addition, it is also well-documented that words in sentence-final position undergo lengthening and this could interact with number of syllables [7], and accentual lengthening [6].

In examining American English-speaking children's production of contrastive focus, [8] found that the 4-year-olds could use contrastive focus to indicate a new element in a picture description task. Children were shown a pair of pictures which differed only in one element (e.g., a girl petting a cat vs. a girl petting a dog). While the children did produce contrastive focus, it occurred mostly on the subject noun phrase, and less on the verb or the object noun phrase. This usage pattern, however, could be related to sentence positions, since the subject noun phrase occurs sentence-initially and the object noun phrase appears sentence-finally. Given that children as young as 2 years have been shown to use phrase-final lengthening [9], perhaps children at 4 might have confused phrase-final lengthening with accentual lengthening on the object noun phrase in the sentence-final position. As a result, it might have been difficult to detect the sentence-final contrastive focus through perceptual evaluation of the data.

In addition, [10] observed that children aged 5 made more production errors for the non-final focus items in PEPS-C, such as 'I want a GREEN car.' than for the final focus items, such as 'I want a green CAR.' There were also many ambiguous responses in children's productions of monosyllabic focused words in sentence-final position, with children showing a strong trend of not emphasizing utterance-final words even in obligatory contexts. Yet these data were also perceptually scored. It is then not clear if children can produce appropriate acoustic cues to indicate contrastive focus, specifically the use of accentual lengthening as separate from phrase-final lengthening.

In evaluating children's production of contrastive focus on monosyllabic words (e.g. show bob a BOT), American English speaking adults performed worse on items produced by 4-year-olds (50%) than by 7- and 11-year-olds (above 80%) [11]. Adult listeners also appeared to be less successful at identifying contrastive focus in the non-final than the final position. This could also be related to the acoustic cues that these three groups of children used. The 4-year-olds primarily used duration to indicate contrastive focus, the 7-year-olds used f0, and the 11-year-olds used both f0 and duration to do so.

These studies then raise some questions about children's ability to produce an appropriate durational cue to contrastive focus at age 5 when their focus productions are still ambiguous, especially in relation to their use of accentual lengthening on items with different number of syllables.in different sentence positions.

1.1. Predictions

We tested 5-year-old AusE-speaking children and compared their productions to an adult baseline. H_1: we predicted that adults would exhibit accentual lengthening for both mono- and disyllabic words. H_2: we expected children to show accentual lengthening for monosyllabic words, specifically in phrase-final position. If accentual lengthening is a robust cue to focus, we also expected children to generalize accentual lengthening to disyllabic words. H_3: we also predicted that the 5-year-olds

and the adults would implement phrase-final lengthening on focused words in sentence-final position to a greater extent than in sentence-medial position if phrase-final lengthening is separate from accentual lengthening.

2. Method

2.1. Participants

Eight (1M, 7F) monolingual Australian English (AusE) speaking adults were recruited (Age range: 18 - 30 years; Mean = 19;10 years). All were undergraduates at Macquarie University, Sydney, and participated for course credit. Eight monolingual AusE-speaking children (3 M, 5 F) were recruited from the Sydney area, ranging in age from 5;1 to 6;9 years (Mean = 5;8 years).

2.2. Stimuli

The stimuli consisted of an adjective (a colour term) and a noun, forming a noun phrase, for example, '*green ball*'. A set of four adjectives (i.e. *green, grey, orange, yellow*) and a set of eight nouns (i.e. *ball, doll, moon, shoe, bottle, button, pencil, table*) were chosen to generate two types of stimuli: four disyllabic noun phrases (e.g., *green ball*), and four quadri-syllabic noun phrases (e.g., *orange bottle*) to allow us to examine the realization of accentual lengthening across different numbers of syllables. Eight noun phrases were formed to serve as stimuli. Focus location was manipulated to fall either on the adjective or on the noun within the stimuli. These stimuli were embedded in a carrier sentence 'I have a/an X'. Since the adjective must precede the noun of the noun phrase syntactically, the focused adjective ended up being in sentence-medial position and the noun in sentence-final position. To disentangle the potential confound of sentence position from accentual lengthening, we added another condition in which we embedded the focused nouns in the longer carrier sentence 'I have a/an X now'. As a result, there were three experimental conditions, with 8 stimuli in each. This yielded a total of 24 stimuli (Table 1).

Table 1. *Stimuli.*

ADJ-FOC	*I have a GREEN moon/shoe*
	I have a GREY ball/doll
	I have an ORANGE button/table
	I have a YELLOW bottle/pencil
FINAL	
N-FOC	*I have a green MOON/SHOE*
	I have a grey BALL/DOLL
	I have an orange BOTTLE/PENCIL
	I have a yellow BUTTON/TABLE
NON-FINAL	
N-FOC-now	*I have a green MOON/DOLL now*
	I have a grey BALL/SHOE now
	I have an orange BUTTON/PENCIL now
	I have a yellow BOTTLE/TABLE now

These stimuli were presented as coloured pictures on laminated cards. In the first condition, sentence-medial adjectives were focused (hereafter referred to as ADJ-FOC). In the second condition, sentence-final nouns were focused (hereafter referred to as N-FOC). In the third condition sentence medial nouns were focused (hereafter referred to as N-FOC-now).

2.3. Procedure

Participants took part in an elicited production task. The task engaged the participants by inviting them to play a language card game (referred to as 'Snap') with a female AusE-speaking experimenter. Before the test session, the experimenter explained how the game was played and went through three practice trials to familiarize participants with the game procedure before the test session.

During the test session, the experimenter and participant each received a deck of 16 cards with coloured stimuli images on the front of each card and a yellow star on the back of some. All the cards were held so that the images were concealed from the other player. For each trial, the experimenter would reveal the top card of their deck and produce the name of the stimulus item in the carrier sentence with neutral intonation, for example, '*I have a green moon*'. The participant would then reveal the top card of their deck and produce their item in the same carrier sentence. In each case the participant's item would differ from the experimenter's item in either type or colour of object, but not both. The participant was encouraged to emphasize the attribute/property which differed between the two pictures. Thus the participant was expected to produce adjective focus (e.g. '*I have a GREY ball*') on the trials where their item differed from the experimenter's in colour, and noun focus (e.g. 'I have a green SHOE') when their item differed from the experimenter's in object type. After that, both the experimenter and the participant put their cards face-up on the table and counted to three before turning their cards over. When one of the cards had a star marked on the back, whichever player called 'Snap' first received one point. This continued until both players used up all their cards.

The set of 16 stimuli were presented twice in different orders. For the first repetition the carrier phrase 'I have a/an X' was used. The second repetition used the carrier phrase 'I have a/an X now'. The adjective-focused sentences in the second repetition were treated as fillers, as sentence-medial adjective focus had already been elicited in the first repetition, giving a total of 24 sentences for analysis (see Table 1). The presentation order of cards was counterbalanced for all participants such that half the participants were presented the pictures in one order, and the other half in the reverse order. The responses were audio-recorded onto a PC using Audacity (audio recording software) at a sampling rate of 44.1 kHz, with a Behringer C2 condenser microphone.

2.4. Acoustic coding

Productions of the adjective + noun phrases were annotated and segmented in Praat [12], coding for the onset and offset of both the adjectives and the nouns. The onset consonants consisted of five types: (a) a stop/plosive, (b) a palatal glide, (c) a nasal (d) a fricative, and (e) no onset consonant. The coda consonants consisted of the following four types: (a) an affricate, (b) a lateral, (c) a nasal, and (d) no coda consonant. The coding criteria were therefore based on the ease of identifying the beginning and the end of the adjective and noun stimuli. When the onset consonant was a plosive/stop consonant, the beginning of the onset was indicated by the onset of the burst release. In items containing an onset glide, we used a pause (if present) and voicing to identify the onset. An additional cue was to use F2 transition to the palatal glide from a schwa in the preceding word. When the item contained a nasal consonant, onset of the nasal resonance was used as the cue. The beginning of high energy noise was used to identify items containing a fricative onset consonant. When there was no onset consonant,

we used the onset of clear F2 and voicing to mark the beginning of the word.

To identify the end of words, we used the offset of the fricative portion for affricates. The end of the lateral coda consonant was based on the offset of voicing and F2 with minimal energy. The offset of nasal resonance was used to identify the end of words containing a nasal coda. Voicing and F2 offset were used as cues to the end of words containing no coda consonants.

3. Results

3.1. Adult duration

Word durations of the target adjectives and nouns in focus vs. non-focus positions were extracted. A repeated measures ANOVA was conducted, with 'Focus', 'Number of syllables' and 'Word category' as factors. Alpha was set at .05.

There was a significant main effect of 'Focus' ($F = 52.148$, $df = 1, 7$, $p < .0001$), and a significant main effect of 'Number of syllables' ($F = 53.913$, $df = 1, 7$, $p < .0001$). There was also a significant 3-way interaction among 'Focus', 'Number of syllables' and 'Word category' ($F = 6.818$, $df = 1, 7$, $p = .035$).

Not surprisingly, disyllabic target words were generally longer than monosyllabic words (344ms vs. 286ms). As predicted, focused word duration was longer than its non-focused counterpart (343ms vs. 287ms) (see Figure 1). The 3-way interaction resulted from larger accentual lengthening for monosyllabic nouns than monosyllabic adjectives (74ms vs. 44ms). However, the pattern was reversed for disyllabic focused words, with less accentual lengthening on disyllabic nouns than disyllabic adjectives (31ms vs. 73ms). Since monosyllabic nouns also occurred utterance-finally, this suggests that phrase-final lengthening might interact with focus-related lengthening on monosyllabic nouns.

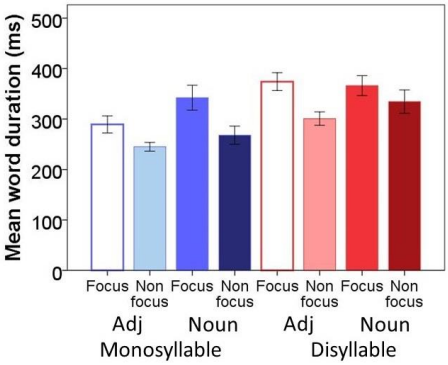

Figure 1. *Mean duration (ms) of focused vs. non-focused adjectives and nouns containing either monosyllables or disyllables in the adult group, with +/- 1 SE.*

To tease apart these two sources of lengthening, we performed another repeated measures ANOVA on focused nouns in utterance-final vs. utterance-medial positions. Two independent variables were included in the analysis: 'Position' and 'Number of syllables'. Alpha was set at 0.05. If focus-related lengthening is further modulated by utterance-final position, we expected the focused nouns to be longer in final positions than non-final positions.

There was a significant main effect of 'Number of syllables' ($F = 15.015$, $df = 1, 7$, $p < .006$). However, counter to our prediction, no significant main effect of 'Position' was found. Yet there was a significant 2-way interaction between 'Number of syllables' and 'Position' ($F = 7.592$, $df = 1, 7$, $p = .028$). The interaction arose because the focused noun was longer in final than non-final position only when the nouns were monosyllabic (342ms vs. 284ms). However, when the nouns were disyllabic, there was no durational difference between final and non-final position (366 ms vs. 348 ms) (see Figure 2).

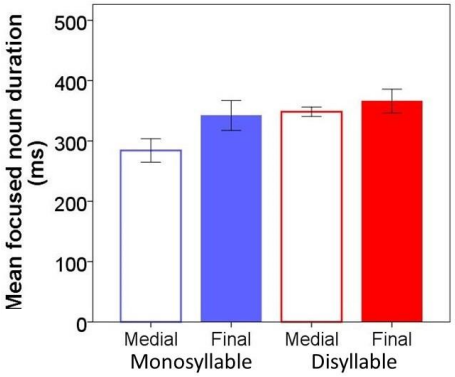

Figure 2. *Mean duration (ms) of focused noun in sentence-medial vs. sentence-final positions in the adult group, with +/- 1 SE.*

3.2. Child duration

Using word durations of focused adjectives and nouns as the dependent variable, we again conducted a repeated measures ANOVA to examine accentual lengthening on target adjectives and nouns in the children's data, with the same three factors: 'Focus', 'Number of syllables' and 'Word category'. Alpha was set at 0.05.

Unlike in adults, there was no significant main effect of 'Focus' ($F = .007$, $df = 1, 7$, $p = .937$). However, similar to the adults, there was a significant main effect of 'Number of syllables' ($F = 59.402$, $df = 1, 7$, $p < .0001$). Like the adults, there was also a significant 2-way interaction between 'Focus' and 'Number of syllables' ($F = 11.984$, $df = 1, 7$, $p = .011$), and a significant 3-way interaction among 'Focus', 'Number of syllables' and 'Word category' ($F = 6.623$, $df = 1, 7$, $p = .037$).

Unlike the adults, children did not show robust accentual lengthening. As expected, they did show longer duration for disyllabic words than monosyllabic words (410 ms vs. 315 ms). In contrast to adults, children showed an interaction between 'Focus' and 'Number of syllables'. This interaction arose because accentual lengthening only took place for monosyllabic, not disyllabic target words.

The significant 3-way interaction arose because all target words underwent accentual lengthening except the focused disyllabic adjectives. Instead of lengthening, the focused disyllabic adjectives were shortened relative to their non-focused counterparts. (see Figure 3)

Since the focused nouns occurred utterance-finally, the observed accentual lengthening could be utterance-final lengthening in disguise. Therefore, a separate repeated measures ANOVA was performed to investigate children's ability to use the durational cue to signal focus. Two independent variables were included in the analysis: 'Position' and 'Number of syllables'. Alpha was set at 0.05.

There was a significant main effect of 'Number of syllables' ($F = 79.724$, $df = 1, 7$, $p < .0001$). Counter to the adult patterns and our prediction, there were neither a main effect of 'Position' nor a 'Position-Number of syllables' interaction (see Figure 4).

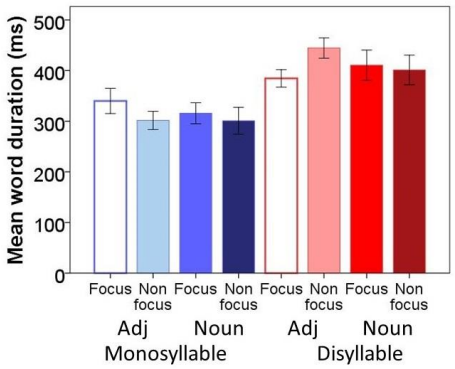

Figure 3. *Mean duration (ms) of focused vs. non-focused adjectives and nouns containing either monosyllables or disyllables in the child group, with +/- 1 SE.*

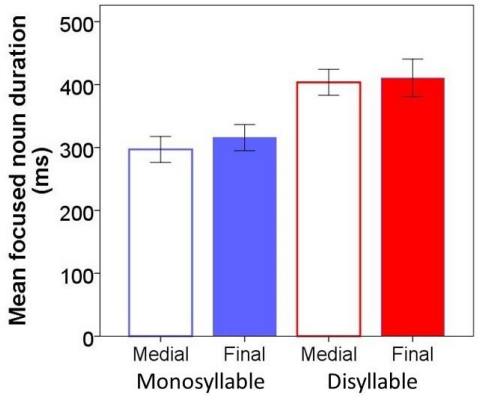

Figure 4. *Mean duration (ms) of focused noun in sentence-medial vs. sentence-final positions in the child group, with +/- 1 SE.*

4. Discussion

The findings showed that children can employ accentual lengthening to signal focus; however, this ability is not the same as or as robust as that observed in the adults. As predicted in H_1, adults exhibited accentual lengthening in both monosyllabic and disyllabic words. In partial support of H_2, children, however, implemented accentual lengthening only on monosyllabic, not disyllabic words. In addition, the adults showed accentual lengthening on focused disyllabic adjectives; whereas the children shortened them. This indicates that children are still inconsistent in using the durational cue to signal focus. In other words, children are still learning how to generalize accentual lengthening across the number of syllables, and its subtle acoustic realization.

It is interesting that the interaction of accentual lengthening and phrase-final lengthening is contingent on the number of syllables within a word. The adults showed additional phrase-final lengthening effects on focused monosyllabic words, but not focused disyllabic words. This partially supports H_3. It is possible that phrase-final lengthening might be present on the second syllable of the disyllabic word, which requires further analysis. Unlike the adults, *no* additional phrase-final lengthening was observed in the children. When children implement accentual lengthening, phrase-final lengthening seems to disappear. This suggests that children are still learning how to weigh duration to signal focus and sentence position separately in an adult-like manner.

5. Conclusion

This study showed that children aged between 5 and 6 years can use accentual lengthening, however, this ability is not as robust as what previous studies have suggested, and far from being adult-like. Perhaps the disparity between the production and perception of focus in children is not as anomalous as previously thought, given that children are still learning how to use the appropriate acoustic cues to signal focus in production.

6. Acknowledgements

This work was supported by the ARC Laureate Fellowship [FL130100014] awarded to the last author. We would like to thank our participants for taking part in the research, Amy German and Cecily Mccowage for their assistance in data collection and annotation, members of the Child Language Lab and the two anonymous reviewers for their helpful comments

7. References

[1] Cooper, W. E., Eady, S. J. and Mueller, P. R. “Acoustical aspects of contrastive stress in question-answer contexts”, Journal of Acoustical Society of America, 77:2142-2156, 1985

[2] Fry, D. B. “Duration and intensity as physical correlates of linguistic stress”, Journal of the Acoustical Society of America, 27: 765-768, 1955.

[3] Kochanski, G., Grabe, E., Coleman, J., and Rosner, B. “Loudness predicts prominence: fundamental frequency lends little”, Journal of the Acoustical Society of America, 118:1038-1054, 2005.

[4] Eady, S. J. and Cooper, W. E. “Speech intonation and focus location in matched statements and questions”, Journal of the Acoustical Society of America, 80: 402-416, 1986.

[5] Speer, S. R. and Ito, K. “Prosody in first language acquisition – acquiring intonation as a tool to organize information in conversation”, Language and Linguistics Compass, 3:90-110, 2009.

[6] Turk, A. E. and White, L. “Structural influences on accentual lengthening in English”, Journal of Phonetics, 27:171-206, 1999.

[7] Turk, A. E. and Shattuck-Hufnagel, S. “Multiple targets of phrase-final lengthening in American English words,” Journal of Phonetics, 35:445-472, 2007.

[8] Hornby, P. A. and Hass, W. A., “Use of contrastive stress by preschool children”, Journal of Speech and Hearing Research, 13: 395-399, 1970.

[9] Snow, D. “Phrase-final syllable lengthening and intonation in early child speech”, Journal of Speech and Hearing Research, 37:831-840, 1994.

[10] Wells, B., Peppé, S. and Goulandris, N. “Intonation development from five to thirteen”, Journal of Child Language, 31: 749-778, 2004.

[11] Patel, R. and Brayton, J. T. “Identifying prosodic contrasts in utterances produced by 4-, 7- and 11-year old children”, Journal of Speech, Language and Hearing Research, 52:790-801, 2009.

[12] Boersma, P. and Weenink, D. Praat: doing phonetics by computer. http://www.praat.org/