

Interaction of vowel length, utterance-position and focus accent in 3-year olds' Australian English

Ivan Yuen, Felicity Cox, Katherine Demuth

Linguistics Department, CCD, Macquarie University

ivan.yuen@mq.edu.au, felicity.cox@mq.edu.au, katherine.demuth@mq.edu.au

Abstract

Durational contrasts can be used to signal prosodic boundaries (utterance-final lengthening) and discourse focus (accentual lengthening). Durational differences can also signal phonemic vowel length contrast such as in non-rhotic Australian English where vowel length distinguishes lexical items (such as, *cart* /kɑ:t/ and *cut* /kʌt/). This raises the question of how and when children acquire phonemic vowel length contrasts along with various prosodic length requirements. An elicited imitation task was conducted with nine 3-year-old Australian English-speaking children. Our findings show that the children distinguished long and short vowels in spite of utterance-final and accentual lengthening. The scope of the two types of prosodic lengthening differed, suggesting they are in the process of learning about these prosodic contexts.

Index Terms: prosody, utterance-final lengthening, accentual lengthening, vowel acquisition, phonemic vowel length

1. Introduction

Spoken language is organized prosodically into phrasal units to aid communication. Such prosodic organization can manifest acoustically in pitch, intensity and/or duration. Duration can be used as a cue to signal a phrasal boundary (utterance-final lengthening), or to highlight/focus a constituent in an utterance (accentual lengthening). At the same time, duration can be used on a segmental level to distinguish one word from another (phonemic contrast). For instance, Australian English (AusE) relies on vowel duration contrasts to phonemically distinguish the word *cart* /kɑ:t/ from the word *cut* /kʌt/ [5,6]. This raises the question of how and when children acquire phonemic vowel length distinctions given interactions with durational lengthening that occurs as a function of prosodic context. The goal of this paper is therefore to examine how utterance-final lengthening and accentual lengthening interact with the production of phonemic vowel length contrasts in AusE-speaking 3-year-olds.

The boundary-related lengthening phenomenon we examine in this study is utterance-final lengthening [1, 2, 14, 15, 17]. In adult speech production, words occurring in utterance-final position exhibit longer duration than those in utterance-medial position. Similar boundary-related lengthening has been observed in children as young as 2 years [9]. Recent studies have shown that utterance-final lengthening is progressive [14, 15]. That is, segments closer to the boundary will be lengthened more than those further away. In other words, in a C₁VC₂ # syllable, C₂ will be lengthened more than V. In addition, placement of focus accent, which is used to highlight a focused word, increases

word duration. An accented word/syllable is longer than its unaccented counterpart [13].

On a segmental level, American English-speaking children are accurate in producing vowel quality and respective tense/lax vowel durations around three years of age [7, 12]. The ability to adjust vowel duration manifests very early (by 1;6 years) to signal contrastive phonemic differences as well as coda voicing contrast [10, 16]. These findings reflect children's intrinsic and extrinsic uses of duration.

On the basis of these findings, we would predict that, by the age of three, children should be able to distinguish phonemically long and short vowels in Australian English (AusE). They should also exhibit utterance-final lengthening while maintaining phonemic vowel length contrasts. As adults exhibit progressive utterance-final lengthening, indicating that the effect resides primarily in the coda [1, 14], we also predict that children will exhibit a greater degree of utterance-final lengthening in the coda than in the vowel. At the same time, if children are attuned to the use of focus accent, we predict accentual lengthening in children's productions for both long and short vowels.

2. Methodology

2.1. Participants

Nine monolingual children (2M; 7F) from Sydney were analyzed in the study. Their mean age was 2;7 years, ranging from 2 to 3 years of age. According to parental report, they were healthy and typically developing, with no speech or hearing difficulties. The parent completed the MacArthur Communicative Development Inventories (CDI) 100-word checklist to estimate each participant's vocabulary size [8]. This was used only as a screening tool to ensure typical language development. The raw mean CDI score was 94.3 ranging from 85 to 100 with a standard deviation of 4.44. An additional three participants were excluded from the analysis because they failed to produce 80 percent of the experimental items (fewer than 13 out of the 16 total items).

2.2. Stimuli

There were sixteen target stimuli, constructed on the basis of an AusE vowel-pair contrastive in vowel length but not quality: /e:/-/ɛ/. These vowels were then used to generate four monosyllabic nonce words: /kɛ:s/- /kɛs/; /gɛ:s/ - /gɛs/. All nonce words had C₁VC₂ syllable structure in which C₁ was either a voiceless or voiced velar stop. The choice of a stop onset was based on the relative ease of segmentation for subsequent acoustic analysis. C₂ was a voiceless alveolar fricative. The fricative coda was chosen to ensure a continuous acoustic signal for durational measurement purposes. AusE is

a non-rhotic dialect and therefore the items did not contain any postvocalic /ɹ/. The four nonce words were associated with four pictures of novel objects. Each word was then embedded in a carrier sentence, appearing in different prosodic contexts.

2.2.1. Prosodic Contexts

We manipulated the prosodic contexts in which the words appeared in terms of 1) Utterance Position: *Medial* versus *Final*, 2) Focus Accent: *Accented* versus *Unaccented*. This created four test contexts for the nonce words as illustrated below, resulting in a total of 16 tokens for each participant.

Utterance-Medial		Utterance-Final	
Accented	Unaccented	Accented	Unaccented
/kɛ:s-/kɛs/	/kɛ:s-/ kɛs/	/kɛ:s-/kɛs/	/kɛ:s-/kɛs/
/gɛ:s-/gɛs/	/gɛ:s-/gɛs/	/gɛ:s-/gɛs/	/gɛ:s-/gɛs/

To manipulate utterance position, we embedded each target word in a carrier sentence. In the utterance-medial context, the target word was embedded in ‘*There X goes*’; in the utterance-final context, the target word was embedded in ‘*Here comes X*’.

We then recorded an adult AusE-speaking female producing these nonce words in the carrier sentences in child-directed speech at a sampling rate of 16kHz. The recorded sentences then served as the auditory prompts to elicit productions from each child.

2.3. Procedure

We tested children using an elicited imitation task. In this task children were invited to play a ‘language game’ in a sound-attenuated recording studio. There were two computers in the studio: one for displaying stimuli and the other for recording. The child participant was seated in front of a computer used for presenting the stimuli, with a Behringer C2 microphone placed on the table in front of the child to capture his/her speech. Two Sony SRS 55 speakers were placed on both sides of the computer to play the pre-recorded auditory prompts. The child was invited to look at the pictures on the computer with accompanying pre-recorded auditory prompts and repeat what was said. Prior to the main experiment, there was a warm-up and familiarization period in which the experimenter showed the child what to do. Children were encouraged through praise and sticker rewards to repeat what was said. The productions were recorded directly onto a computer through a pre-amplifier at a sampling rate of 44kHz. The recordings were then downloaded and analyzed off-line in Praat [3].

2.4. Acoustic Coding

We first checked the produced tokens for 1) acoustic quality, and 2) target accuracy. We excluded 3 items containing incorrect segments, 5 items with omitted coda, and 5 items of poor acoustic quality with overlapping noise. A total of 131 tokens were segmented and annotated in Praat [3].

As we are interested in the vowel and coda durations, the target words were segmented into ‘vowel’ and ‘/s/ coda’ regions and labeled as illustrated in Figure 1. With the help of spectrograms and waveforms, the following criteria were used to determine the beginning and the end of the respective regions. The onset and offset of F2 energy and voicing were employed to determine the beginning and end of the target vowel. As the onset plosive preceded the vowel, we also used

the end of aspiration as an additional cue to identify the beginning of the vowel. In cases when vowels were breathy or heavily fricated, F2 energy took precedence over voicing as the criterion. As regards the ‘/s/ coda’ region, the criterion was the beginning and end of high frequency noise around 3-5K Hz [11]. The duration of the target vowels and /s/ codas were then extracted for further analysis using Praat.

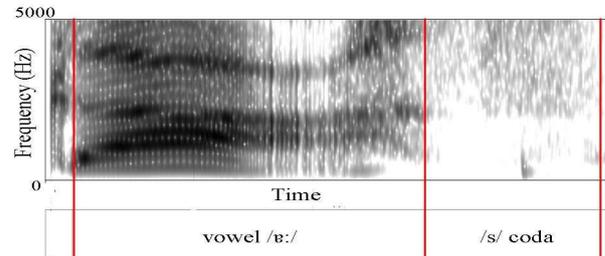


Figure 1. Spectrographic illustration of the annotation points for vowels and /s/ codas in /gɛ:s/ from a child

3. Results

We conducted repeated measures ANOVAs on vowel duration and fricative /s/ duration separately, with **position**, **focus accent** and **vowel type** as factors.

3.1.1. Vowel Duration

As regards vowel duration, there were statistically significant main effects for **position** ($F = 43.14$, $df = 1,8$; $p < 0.0001$), **focus accent** ($F = 5.507$, $df = 1,8$; $p = 0.047$), and **vowel** ($F = 27.106$, $df = 1,8$; $p = 0.001$). As predicted, the vowels in utterance-final position were longer than their counterparts in utterance-medial position. Accented vowels were also longer than unaccented vowels. More importantly, the long and short vowels were distinct from one another. These main effects can be seen in Figure 2.

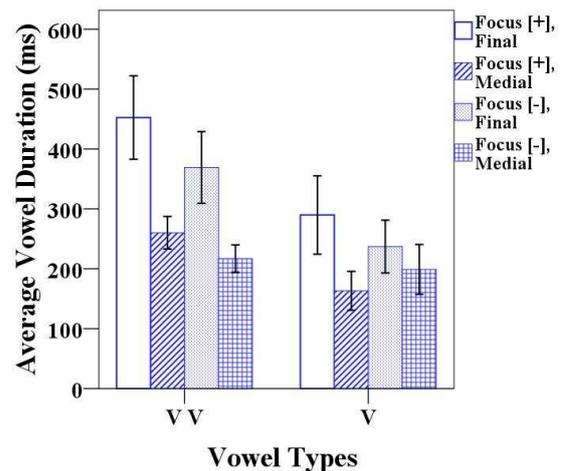


Figure 2: Average vowel duration for 2 utterance positions and 2 focus conditions with +/- 2 standard errors

All three interactions reached statistical significance: **position by focus accent** interaction ($F = 6.317$, $df = 1,8$; $p = 0.036$), **focus accent by vowel** interaction ($F = 6.531$, $df = 1,8$; $p =$

0.034) **vowel by position** interaction ($F = 7.038$, $df = 1,8$; $p = 0.029$).

In the **position by focus accent** interaction, there was a mean difference of 95 ms between medial and final utterance positions for unaccented vowels, but a mean difference of 159 ms for accented vowels. Thus, accented vowels were lengthened more in final position than in medial position. In the **focus accent by vowel** interaction, the magnitude of accentual lengthening was less for the short vowels than the long vowels. There was a mean difference of 9 ms in accentual lengthening for the short vowels, but 63 ms for the long vowels. That is, the long vowels were lengthened more when they were accented. In the **vowel by position** interaction, there was a mean difference of 82 ms between the medial and the final utterance positions for the short vowels, but a mean difference of 172 ms for the long vowels. Thus, once again, the children lengthened the long vowels more than the short vowels in the final position.

3.1.2. Coda Duration

As regards fricative coda /s/ duration, we found only a significant main effect of **position** ($F = 50.428$, $df = 1,8$; $p < 0.0001$). The /s/ fricative was on average 147 ms longer in the utterance-final position compared to the utterance-medial position (Figure 3).

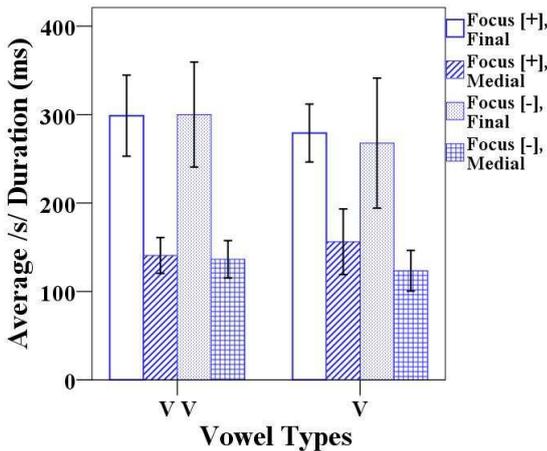


Figure 3: Average /s/ coda duration for 2 utterance positions and 2 focus conditions, with +/- 2 standard errors

3.1.3. Progressive utterance-final lengthening

To test whether utterance-final lengthening is progressive in children's production, we calculated the durational difference score (DD_score) between two utterance positions for each segment in the target word according to the formula in (1), where *segment* is either the target vowel or the fricative /s/ coda.

$$DD_score = \frac{segment(final) - segment(medial)}{segment(medial)} \quad (1)$$

If utterance-final lengthening progressively decreases with the distance from the phrase boundary, we predict a difference in the DD_score between the fricative /s/ coda and the vowel because the former lies immediately adjacent to the boundary, and the latter not. The larger the DD_score, the larger the

effect of utterance-final lengthening on the segment of interest.

We then conducted a repeated measures ANOVA using DD_score as the dependent variable with **focus accent (+/-)**, **vowel (long/short)** and **syllable unit (vowel/coda)** as factors. There was a statistically significant main effect of **syllable unit** ($F = 11.815$, $df = 1, 8$, $p = 0.009$). As predicted, the DD_score was larger in codas than in vowels. An interaction between **focus accent** and **syllable unit** also reached statistical significance ($F = 5.537$, $df = 1,8$, $p = 0.046$). This means that the magnitude of progressive final lengthening interacts with accentual lengthening. As illustrated in Figure 4, progressive utterance-final lengthening was observed in both accented and unaccented long vowels with the coda lengthened more than the vowel. However, short vowels exhibited progressive utterance-final lengthening only in the unaccented condition. In the accented condition, utterance-final lengthening affected both the short vowel and coda duration to a comparable extent as reflected in the similar mean DD_score. This suggests that utterance-final lengthening contributed equally to the accented short vowel and the coda. Note also the relatively large interspeaker variability for each condition. This indicates that individual children are at different stages of realizing boundary-related and focus-related lengthening on one hand, while maintaining vowel length contrasts in Australian English on the other. This suggests that phonemic vowel length contrasts may be acquired by the age of 3, whereas interactions with utterance-final lengthening and accentual lengthening may take longer to fully master.

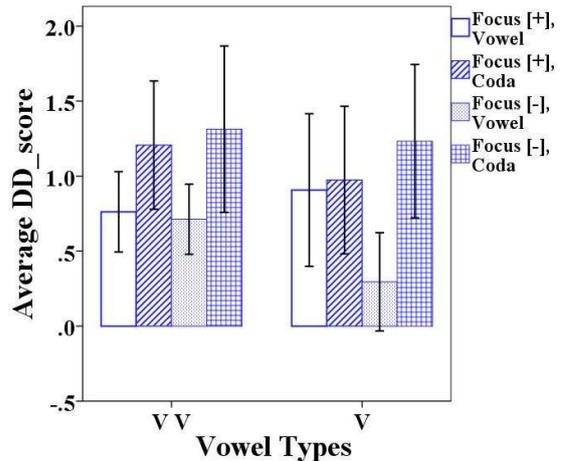


Figure 4: Average DD_score for 2 focus conditions in vowels and codas with +/- 2 standard errors

4. Discussion

This study investigated the multiple factors that can affect vowel duration in English (intrinsic vowel length, focus accent, and utterance position) and how these are acquired. It was found that phonemic vowel duration and accentual lengthening are both well controlled by the age of three. Furthermore, utterance position impacted on both vowel and coda durations. However, the effect of utterance position on vowels and codas differed. The coda appears to be the primary locus of utterance-final lengthening. This may be related to the nature of the *source* of lengthening. In utterance-final lengthening, the source comes from the boundary which spreads leftwards. That is why the lengthening effect manifests predominantly in the coda of the target word in the

current study. On the other hand, the *scope* of accentual lengthening encompasses the vowel. Thus, 3-year-old AusE-speaking children appear to be sensitive to the difference in the scope of boundary-related and focus-related lengthening.

Interestingly, focus accent also modulates progressive boundary-related lengthening, particularly in short vowels. Though these children generally exhibited progressive utterance-final lengthening as in adults, they did not do so for the accented short vowels where utterance position affected their short vowels and codas equally. When the short vowels are unaccented, boundary-related lengthening progressively diminished from coda to vowel. However, when short vowels are accented, boundary-related lengthening did not diminish progressively. This finding indicates that the scope of focus-related lengthening is sensitive to whether the vowel is long or short.

Children also robustly distinguished long and short vowels when utterance position and focus accent were kept constant. This is consistent with Buder and Stoel-Gammon's [4] report that Swedish children produced long and short vowels accurately around 3 years of age. Note, however, that phonemic vowel length contrasts are much more robustly attested in the vowel inventory of Swedish than AusE. This makes our finding all the more interesting, showing that learners of AusE are acutely aware of the different types of factors that can influence vowel duration.

Our findings also show that children did not treat long and short vowels in the same way across contexts. Utterance-final lengthening and accentual lengthening are stronger for long vowels than short vowels, indicating a constraint on increasing the duration of short vowels. In addition, focus accent exaggerated utterance-final lengthening, indicating the cumulative vowel lengthening effect.

5. Conclusion

To conclude, AusE-speaking children distinguish phonemic vowel length contrasts in their speech productions by age 3. They are also sensitive to both boundary-related and focus-related lengthening phenomena even though they are still learning how to realize intrinsic vowel length contrasts in these varying prosodic contexts. Interestingly, they already exhibit progressive utterance-final lengthening, as do adults. This suggests that children of this age have begun to learn the source of lengthening variation and its scope. This results in their beginning to sound more adult-like in their speech planning and production. Further work will extend this to real words for generalisability.

Little is known, however, about how and when aspects of these phonemic distinctions and interactions with prosodic context are learned across populations. One might expect that learning in these domains would be more challenging in children growing up bilingually and/or in adults learning a second language, especially in the case of two languages with different prosodic structures. One might also wonder about how and when these aspects of phonology are learned in children with specific language impairment (SLI), or in those with hearing loss, where temporal organization of sounds and words may be impaired. The results presented here provide a first pass at beginning to provide a baseline of development for investigating these issues across populations.

6. Acknowledgements

We thank Child Language Lab for comments and assistance. This research was supported in part by funding from Macquarie University, ARC DP110102479, ARC CE110001021, and NIH R01 HD057606.

7. References

- [1] Berkovits, R. "Progressive utterance-final lengthening in syllables with final fricatives". *Language and Speech*, 36(1), 89-98, 1993.
- [2] Berkovits, R. "Durational effects in final lengthening, gapping and contrastive stress". *Language and Speech*, 37, 237-250, 1994.
- [3] Boersma, P., Weenink, D. Praat: doing phonetics by computer (version 5.1.29) Available from <http://www.praat.org/>, 2010.
- [4] Buder, E., Stoel-Gammon, C. "American and Swedish children's acquisition of vowel duration: effects of vowel identity and final stop voicing". *Journal of Acoustical Society of America*, 111(4), 1854-1864, 2002.
- [5] Cox, F. "The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers". *Australian Journal of Linguistics*, 26, 147-179, 2006.
- [6] Cox, F., Palethorpe, S. "Timing differences in the VC rhyme of Standard Australian English and Lebanese Australian English". *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, 2011.
- [7] Davis, B., MacNeilage, P.F. "Acquisition of Correct Vowel Production". *Journal of Speech, Language and Hearing Research*, 33, 16-27, 1990.
- [8] Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. "Short-form versions of the MacArthur Communicative Development Inventories". *Applied Psycholinguistics*, 21, 95-116, 2000.
- [9] Snow, D. "Phrase-final syllable lengthening and intonation in early child speech". *Journal of Speech and Hearing Research*, Volume 37, 831-840, 1994.
- [10] Song, J.Y., Demuth, K. "Compensatory vowel lengthening for omitted coda consonants: a phonetic investigation of children's early representation of prosodic words". *Language and Speech*, 51(4), 382-399, 2008.
- [11] Stevens, K. N. *Acoustic Phonetics*. MIT Press. Cambridge, MA, 1998.
- [12] Stoel-Gammon, C., Herrington, P.B. "Vowel systems of normally developing and phonologically disordered children". *Clinical Linguistics and Phonetics*, 4(2), 145-160, 1990.
- [13] Turk, A., Sawusch, J. "The domain of accentual lengthening in American English". *Journal of Phonetics*, 25, 25-41, 1997.
- [14] Turk, A., Shattuck-Hufnagel, S. "Word-boundary-related duration patterns in English". *Journal of Phonetics*, 28, 397-440, 2000.
- [15] Turk, A., Shattuck-Hufnagel, S. "Multiple targets of phrase-final lengthening in American English". *Journal of Phonetics*, 35 (4), 445-461, 2007.
- [16] Weismer, G., Dinnsen, D., Elbert, M. "A study of the voicing distinction associated with omitted word initial stops". *Journal of Speech and Hearing Disorders*. 46, 320-328, 1981.
- [17] Wightman, C., Shattuck-Hufnagel, S., Ostendorf, M. "Segmental durations in the vicinity of prosodic phrase boundaries". *Journal of Acoustical Society of America*, 91(3), 1707-1717, 1992.