FIXED TEMPORAL PATTERNS IN CHILDREN'S SPEECH DESPITE VARIABLE VOWEL DURATIONS

Melissa A. Redford[1] and Grace E. Oh[2]

[1]University of Oregon and [2]Konkuk University
[1]redford@uoregon.edu, [2]grace1111@konkuk.ac.kr

ABSTRACT

The current study compared children's and adults' ability to produce inherent and context-specific vowel duration differences with their ability to repeatedly produce the same vowel in the same context. Children (5- and 8-year-olds) and adults produced real English words in a frame sentence multiple times. Mean vowel duration and variability in vowel duration were analysed as a function of the manipulated factors. Results were that children produced exactly the same contrasts as adults despite also exhibiting more variability in their production of individual vowels. The results are consistent with a model where the 'plan' is remembered relative timing information and execution is the achievement of motor goals at specified temporal intervals.

Keywords: articulatory timing; temporal variability; vowel duration; speech acquisition.

1. INTRODUCTION

It is well established that children have poorer articulatory timing control than adults, as evidenced by their more variable productions of the same segment in the same context across multiple repetitions (see, e.g., [3], [4], [9], [10]). In fact, work in this area indicates that adult-like articulatory timing control is not acquired until either age 12 or 14 years, depending on whether the measurements are taken in the acoustic or kinematic domain (*cf.* [4] and [9]).

It is equally well established that children acquire linguistically relevant temporal patterns very early. For example, the vowel duration pattern that is a correlate of lexical stress in English is acquired by age 2 years (see, e.g., [2], [5], [8]). Note, though, that it is only the pattern that is acquired; absolute durational values are substantially different from those produced by adults. To wit, young English-speaking children produce shorter unstressed vowels than stressed vowels within a word, but the difference between their unstressed and stressed vowels may be smaller than in adults' speech [8]. This is likely because young children's unstressed vowel durations are so much longer than adults' [5], [8], consistent with the aforementioned age-dependent differences in articulatory timing control.

A reasonable explanation for the slow acquisition of articulatory timing control compared to the early acquisition of linguistically relevant temporal patterns is that they reference separate processes. The acquisition of timing control follows from motor learning; the acquisition of relative timing patterns from language learning. But where is the dividing line between these two types of learning? More specifically, given that relative timing patterns require the coordination of articulators through time (i.e., articulatory timing) how do we make sense of variability at one level and stability at another?

One answer assumes that the relative timing of speech motor goals (acoustic and/or sensorimotor) is abstracted and remembered during word learning, and thus is part of the lexical representation [6]. This answer is consistent with the view from schema theory [7] and Articulatory Phonology [1], both of which suggest that the representation of sequential action (non-speech or speech) is fundamentally temporal. If this answer is correct then it would mean that the production of temporal patterns follows from a plan. Execution is thus the instantiation of speech motor goals at the temporal intervals specified in the plan, with achievement of individual goals subject to the effects of on-line motor control. And it is on-line motor control that is immature in children, not the lexical representations that provide the plan for sequential action.

Although the scenario developed here accounts neatly for the separate observations in the literature on children's speech-language, its adequacy depends on whether the observations hold true within speakers. The current study was designed to investigate this. We compared children's and adults' ability to produce inherent and context-specific duration differences in vowels with their ability to repeatedly produce the same vowel. In line with our hypothesis that relative timing patterns are learned and remembered as part of lexical acquisition, we expected no difference between children's and adults' ability to realize the sub-phonemic durational contrasts. Age-related differences in temporal variability were expected, consistent with hypothesized differences in on-line motor control.

2. METHOD

2.1. Participants

Three groups of American-English speaking participants were recruited for the study. As of this writing we have measured and analysed productions from 21 speakers: 7 typically developing 5-year-old children (3 males); 7 typically developing 8-year-old children (5 males); 7 college-aged adults (4 males). 5-year-old children had a mean age of 5;4 years. 8-year-old children had a mean age of 8;1 years. Prior work on children's speech production indicates major changes in articulatory timing abilities between age 5 and 8 years, see, e.g., [4]. At the same time, typically developing 5-year-old children will have acquired a reasonably large vocabulary and an adult phonology; that is, one that no longer includes the many simplifying processes observed in younger children's speech-language [11].

2.2. Stimuli

The stimuli were designed to investigate the production of 3 different durational contrasts: contrasts due to vowel quantity; those due to final consonant voicing; and those due to polysyllabic shortening. The real word stimuli used to encode each contrast are shown in the Tables below.

Table 1: The quantity contrast was elicited in a /bVt/ frame.

Diphthongs	Tense	Lax
bite	beet	bit
bait	bought	bet
boat	boot	but

Table 2: The contrast due to final voicing was elicited for the vowel /æ/ using a /b/ or /k/ onset.

Onset	Voiced	Voiceless
/b/	bad	bat
	bag	back
/k/	cab	cap
	cad	cat

Table 3: The contrast due to polysyllabic shortening was elicited for the vowel /æ/ in the initial stressed syllable of different words. All words had a /b/ or /k/ onset and a post-vocalic /t/.

Syllable #	/b/	/k/
1	bat	cat
2	batty	catty
3	battery	catalogue
4		caterpillar

The stimuli were recorded by a female speaker of West Coast American English, aggregated and presented auditorily in the frame sentence "I said ___ again." The participants' task was to respond to this presentation with the phrase "She said ___ again." Auditory presentations were used to control for age-dependent differences in reading level. The change in the frame sentence for elicitation was meant to make the speech task slightly more meaningful. The items were presented in random order in 6 blocks to elicit 6 repetitions of each item.

2.3. Measurement and Analyses

Because all vowels of interest were bounded by stop consonants, acoustic measures of duration were straightforward. Vowel duration was measured from the onset of voicing in words beginning with /k/ and at closure release in words beginning with /b/. The dependent variables in the analyses were the per word mean vowel durations (averaged across the 6 repetitions) and the per word coefficient of variability (standard deviation in duration / mean duration). Linear mixed effects modelling was used to assess the fixed effects of age group, a between-subjects factor, and contrast, a within-subjects factor. The analyses of quantity included vowel quality as an additional within-subjects factor; and those of final voicing, a place of articulation factor. The analyses of contrasts due to final consonant voicing and polysyllabic shortening were split by onset consonant. Participant was always treated as a random factor and the intercepts included.

3. RESULTS

The results were largely consistent with the predictions. In every case, children produced the expected contrast. In almost every case, children produced individual vowels in the same context with greater temporal variability than adults. The detailed results on effects of age, quantity, coda voicing, and polysyllabic shortening on vowel production are presented below.

There was no effect of age on vowel duration in the production of diphthongs, tense and lax vowels. There was the expected significant effect of both quantity, $F(1, 144) = 100.74$, $p < .001$, and quality, $F(1, 144) = 46.78$, $p < .001$, on vowel duration, but no interaction between age and the other fixed factors. Figure 1 shows that speakers of all ages produced diphthongs with longer mean vowel durations than tense vowels, which in turn were produced with longer mean vowel durations than lax vowels.

Figure 1: Mean vowel duration as a function of age group and vowel quantity.

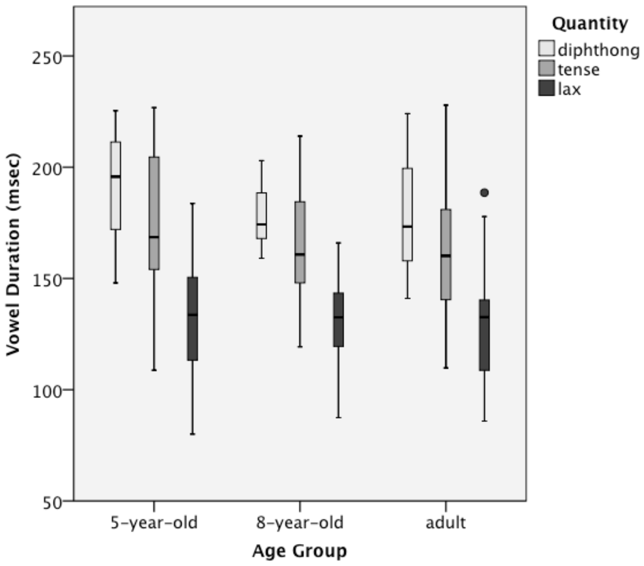
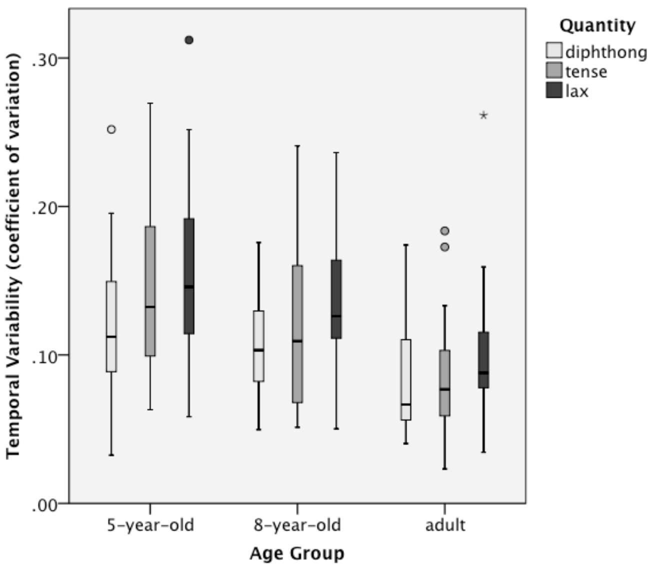


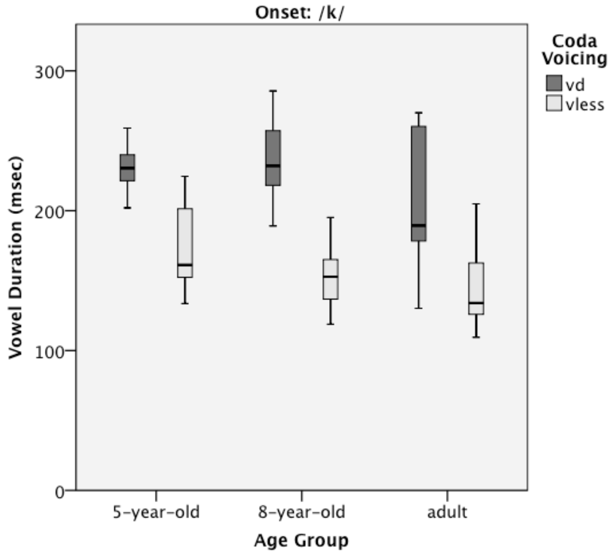
Figure 2: Temporal variability as a function of age and vowel quantity.



In contrast to the results on vowel duration, there was a significant effect of age on temporal variability, $F(2, 18) = 4.23$, $p = .031$, but no effect of quantity or of quality or any interaction of age with these factors. That is, speakers produced every vowel with roughly the same amount of temporal variability regardless of vowel quantity and quality. Pairwise comparisons between age groups revealed that it was only the 5-year-olds that differed significantly from adults on the measure, mean difference = .049, SE = .017, $p = .027$.

In contrast to the results on vowel quantity, there were effects of age on vowel duration in words that began with both /b/, $F(2, 18) = 4.09$, $p = .034$, and with /k/, $F(2, 18) = 7.99$, $p = .003$; there was also a main effect of coda voicing for both the /b/-onset words, $F(1, 60) = 25.63$, $p < .001$, and the /k/-onset words, $F(1, 56) = 20.83$, $p < .001$; and an effect of coda place, for the /k/-onset words, $F(1, 56) = 8.88$, $p = .004$. The effect of age interacted with the effect of coda voicing in words that began with both /b/, $F(2, 60) = 4.29$, $p = .019$, and with /k/, $F(2, 56) = 6.79$, $p = .004$. The interaction is shown for words with /k/ onsets in Figure 3.

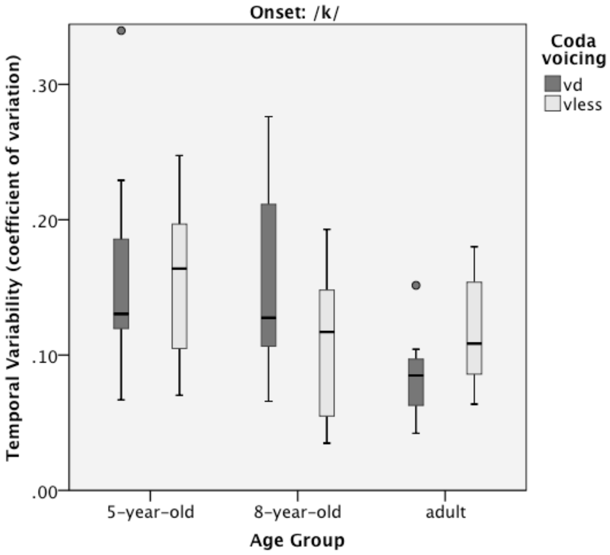
Figure 3: Mean vowel duration as a function of age group and coda voicing in /k/-onset words.



The data shown in Figure 3 suggest that the interaction between age and coda voicing on vowel duration might have been due to a somewhat greater voicing-dependent duration contrast in children's speech compared to adults' speech. To test whether this was true, we calculated the difference between vowels that preceded voiced and voiceless codas, then tested whether this difference varied systematically with age. It did not. Thus, in spite of the interaction between age and voicing, it would seem that children and adults both produce roughly the same voicing-dependent durational contrast.

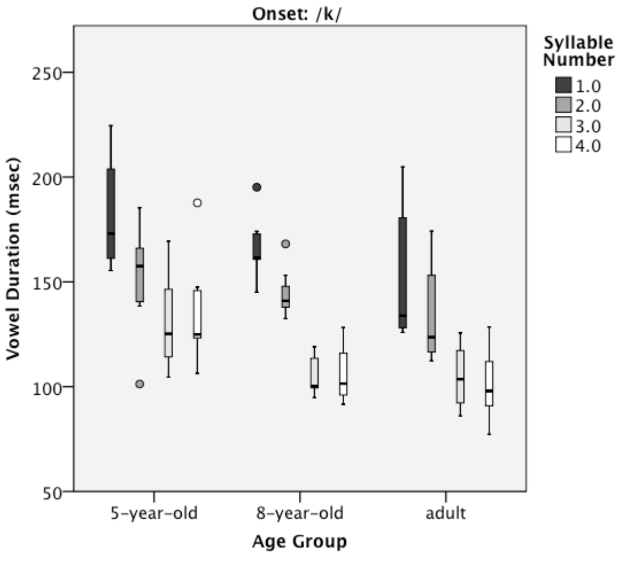
The analysis on temporal variability showed only a significant effect of coda voicing for words with /b/ onsets, $F(1, 59) = 7.79$, $p = .007$; mean variability was somewhat higher for words with voiced codas than for those with voiceless codas. The effect of age was, however, significant in words with /k/ onsets, $F(2, 18) = 6.65$, $p = .007$. As shown in Figure 4, younger children produced vowels with more temporal variability than older children and adults, mean difference = .049, SE = .017, $p = .027$. There was also an interaction between age and coda voicing, $F(2, 54) = 4.05$, $p = .023$, which appears to have been driven by the adult data (see Figure 4).

Figure 4: Temporal variability as a function of age and coda voicing in /k/-onset words.



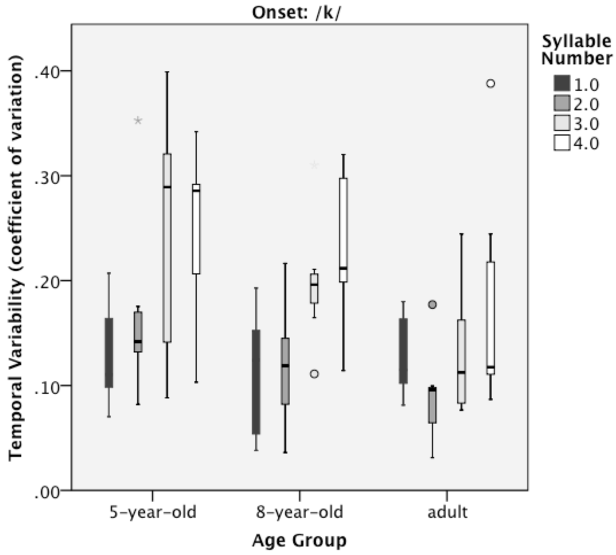
The final analyses investigated lexically stressed /æ/ production in words of different lengths. The expected effect of syllable number was observed for both /b/-onset words, $F(2, 36) = 34.53$, $p < .001$, and /k/-onset words, $F(3, 54) = 71.88$, $p < .001$. There was also an effect of age for /k/-onset words, $F(2, 18) = 4.00$, $p = .037$, but no interaction between age and syllable number in either set of words. The results for /k/-onset words are shown in Figure 5.

Figure 5: Mean vowel duration as a function of age and syllable number in /k/-onset words.



As for variability, the effect of age did not reach significance for either set of words, though the trends were clear and in the expected direction (see, e.g., Figure 6). There was also an effect of syllable number for /k/-onset words, $F(2, 54) = 11.07$, $p = .037$, but no interaction between age and syllable number. Across speakers, /æ/ duration was more variable in longer words compared to shorter words.

Figure 6: Temporal variability as a function of age and syllable number in /k/-onset words.



4. GENERAL DISCUSSION

The current study examined children's ability to produce sub-phonemic inherent and context-dependent duration contrasts in real English words, and especially to investigate whether they produced these contrasts in an adult-like manner in spite of poorer articulatory timing control. Overall the results confirm that children do indeed acquire fine-grained temporal patterns prior to acquiring adult-like articulatory timing control. The findings are therefore consistent with existing literature, where the acquisition of linguistic patterns, including temporal ones, has been largely investigated and discussed independently of articulatory timing control and its development (but see the discussion in Kent and Forner [3] for an exception to this generalization). We suggest that the present results, along with those that have been previously reported in the literature, are best understood in the context of a model of production where the 'plan' is remembered relative timing information and execution is the achievement of motor goals at plan-specified temporal intervals.

5. ACKNOWLEDGMENTS

The work was supported by a grant from the National Institute of Child Health and Human Development (NICHD), #R01HD061458. The content is solely the responsibility of the authors and does not necessarily reflect the views of NICHD.

6. REFERENCES

[1] Browman, C. P., & Goldstein, L. M. 1986. Towards an articulatory phonology. *Phonology Yearbook* 3, 219–252.

[2] Kehoe, M., Stoel-Gammon, C., & Buder, E.H. 1995. Acoustic correlates of stress in young children's speech. *J. Speech & Hear. Res.* 38. 338–350.

[3] Kent, R. D., & Forner, L. L. 1980. Speech segment duration in sentence recitations by children and adults. *Journal of Phonetics*, 8, 157–168.

[4] Lee, S., Potamianos, A. & Narayanan, S. 1999. Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am.* 105, 1455–1468.

[5] Pollock, K. E., Brammer, D. M., & Hageman, C. F. 1993. An acoustic analysis of young children's productions of word stress. *Journal of Phonetics* 21, 183–203.

[6] Redford, M.A. 2015. Unifying speech and language in a developmentally sensitive model of production. Resubmitted.

[7] Schmidt, R. A. (2003). Motor schema theory after 27 years: Reflections and implications for a new theory. *Res. Q. Exerc. Sport* 74, 366-375.

[8] Schwartz, R. G., Petinou, K., Goffman, L., Lazowski G., & Cartusciello, C. 1996. Young children's production of syllable stress: An acoustic analysis. *J. Acoust. Soc. Am.* 99, 3192–3200.

[9] Smith, A., & Zelaznik, H. N. 2004. Development of functional synergies for speech motor coordination in childhood and adolescence. *Dev. Psychobiol.* 45, 22–33.

[10] Smith, B. L. 1992. Relationships between duration and temporal variability in children's speech. *J. Acoust. Soc. Am.* 91, 2165–2174.

[11] Stoel-Gammon, C., & Dunn, C. 1985. *Normal and disordered phonology in children*. Austin, TX: Pro Ed.