

# TOWARDS UNDERSTANDING THE PROTRACTED ACQUISITION OF ENGLISH RHYTHM

*Hema Sirsa & Melissa A. Redford*

University of Oregon, USA

hsirsa@uoregon.edu; redford@uoregon.edu

## ABSTRACT

Several global and specific rhythm metrics and speech rate were used to characterize differences in the rhythms of 5- and 8-year-olds' spoken English. The results were that speech rate and the rate-normalized Pairwise Variability Index (nPVI) differentiated between 5- and 8-year-olds' speech. A further result was that the variance in nPVI values was better explained by a specific measure devised to capture patterns of supralexical accenting than by the factor of age expressed in months. These results are taken to suggest that the protracted acquisition of English rhythm may be due in part to the slow rate at which children acquire prosodically conditioned vowel reduction.

**Keywords:** rate, rhythm, stress timing, speech acquisition

## 1. INTRODUCTION

A number of recent studies have used global metrics to assess the acquisition of language rhythm and found that children take significantly longer to acquire the stress-timed pattern of English compared to the syllable-timed pattern of French or Spanish [2, 7]. For example, Post and Payne [9] have reported that English-speaking 6-year-olds show less overall variability of vocalic interval durations than English-speaking adults. The current study was motivated by two questions that arise from these findings: (1) Are global rhythm metrics sensitive enough to also distinguish between younger and older school-aged children's speech? (2) What accounts for the protracted acquisition of stress timing that is suggested by previous studies?

Possible explanations for why stress timing takes longer to acquire than syllable timing are derived from explanations for the perception of distinct language rhythms. For example, one explanation follows from the idea that stress and syllable timing emerge from language phonotactics and the presence/absence of vowel reduction [3, 10]. Specifically, the idea is that stress timing takes

longer to acquire than syllable timing because the complex syllabic structures and vowel reduction of stress-timed languages take longer to acquire than the simpler syllable structures and full vowels of syllable-timed languages (see e.g., [7, 9]).

While phonotactics and vowel reduction likely provide part of the explanation for the protracted acquisition of stress timing [9], White and Mattys [15] point out that:

Language-specific prosodic timing processes, such as accentual lengthening, word-initial lengthening and phrase-final lengthening, should clearly also be considered in a full model of influences on vocalic and consonantal interval durations (p. 518).

Whereas English lexical stress patterns may be acquired quite early (see e.g., [1]), at least some of the phrase-level prosodic patterns that White and Mattys refer to are not [1, 14]. Since phrase-level prosody disproportionately affects vowel/rhyme durations [13], the slow acquisition of these higher-level prosodic patterns might explain why English-speaking 6-year-olds show less overall variability of vocalic interval durations than adults: children may be less able to create maximal contrasts between accented and unaccented syllables.

To this list of explanations for the protracted development of stress timing, we might add a third—speech rate.

Children's average speaking rate differs substantially from that of adults', and speech rate changes slowly across childhood. For example, Sabin and colleagues [11] reported only two significant rate increases between kindergarten (age 5 or 6 years) and adulthood, with the first increase occurring between kindergarten and 2<sup>nd</sup> grade (age 7 or 8 years).

Speech rate may affect the acquisition of language rhythm because it is another factor that distinguishes between stress- and syllable-timed languages [5]. And, like other global metrics, it interacts with word- and phrase-level prosody. Specifically, vowel durations in function words

and other unstressed syllables are compressed more at faster rates of speech than stressed vowels in content words [6]. Thus stress-timing may be characterized by greater variability in the sequential duration of vowels in adult speech than child speech because adults have acquired the ability to disproportionately, but appropriately, reduce unstressed and unaccented vowels in order to gain articulatory speed.

## 2. METHODS

### 2.1. Participants

Ten typically developing 5-year-olds ( $M = 5;5$ ,  $SD = 4$  months) and 10 typically developing 8-year-olds ( $M = 8;4$ ,  $SD = 6$  months) participated in the study. All children were monolingual native American-English speakers. The children had no speech/language or hearing problems as determined by parental report and a pure-tone hearing screening.

### 2.2. Procedure and materials

Children were audio recorded while completing the Recalling Sentences subtest from the Clinical Evaluation of Language Fundamentals (CELF-4) [12]. The subtest uses delayed imitation of sentences to evaluate a child's linguistic development. The initial sentences in the test are shorter and syntactically simpler than later sentences, and the test always proceeds from the simpler sentences to the more complex ones. We administered the test in the normal way, except that we always began with the first sentence in the test regardless of a child's age and we used prerecorded stimuli to control for sentence prosody and other phonological and phonetic characteristics of the model sentences.

The first 6 sentences (out of 32) from the Recalling Sentences subtest were almost always reproduced without error and were always produced readily and fluently by the children in the study. These sentences, which had a mean length of 9 syllables ( $SD = 1.9$ ), provided the basis for rhythm measurement.

### 2.3. Acoustic segmentation

Each of the 120 sentences was displayed as an oscillogram and spectrogram and hand segmented into consonantal and vocalic intervals. Only intervals that showed clear evidence of oral closure were considered consonantal. Those portions of sonorant consonants that could not be

differentiated from the vowel by a drop in overall amplitude and/or the diminishment of higher formants were considered vocalic. Thus, glides and liquids very frequently contributed to the duration of vocalic intervals.

### 2.4. Rhythm metrics

The consonantal and vocalic interval durations were extracted and used to calculate several global rhythm metrics, speech rate, and several specific rhythm metrics. The global rhythm metrics we chose were the rate-normalized measures of overall variability in vocalic and consonantal durations, nPVI [8] and varcoC [4] and the overall percentage of vocalic duration in an utterance, %V [10]. These were calculated for each sentence, and then averaged to provide a single nPVI, varcoC, or %V value for each child.

The specific rhythm metrics we devised were as follows: the average duration ratio of strong to weak vowels (S2W) from two non-final trochaically-stressed words (*tractor* and *rabbit*); the average duration ratio of a content word vowel to a function word vowel (C2F) in 6 non-final determiner noun phrases with monosyllabic nouns; and the average duration ratio of the ultimate and penultimate vowels (U2P) in 2 sentences with phrase-final determiner noun phrases (DP) with monosyllabic nouns. S2W was meant to capture the aspect of rhythm due to trochaic stress within a word, C2F was meant to capture the aspect due to accenting within supra-lexical prosodic units, and U2P that due to phrase-final lengthening.

Speech rate for each child was calculated as the average number of vocalic intervals per second of speech across the 6 sentences.

## 3. RESULTS

Global rhythm measures and speech rate were uncorrelated with one another, but the rate-normalized global measures correlated with a subset of the specific rhythm measures. Of the various metrics, only nPVI and speech rate distinguished between 5-year-old and 8-year-old speech. A regression analysis indicated that a significant proportion of nPVI variance was explained by DP accenting. These results are given in greater detail below.

### 3.1. Correlations between metrics

Table 1 shows the intercorrelations between the global rhythm measures (nPVI, varcoC, %V), the

specific rhythm measures (S2W, C2F, U2P), and speech rate.

**Table 1:** The intercorrelations between the global rhythm metrics, specific rhythm metrics, and speech rate.

	varco-C	%V	S2W	C2F	U2P	Rate
nPVI	.076	-.046	.343	.677	.517	.387
varco-C		-.357	.727	.353	.527	.077
%V			-.218	-.285	-.146	-.281
S2W				.430	.377	.219
C2F					.527	.340
U2P						.448

None of the global metrics correlated significantly with each other or with speech rate, but the metrics based on vocalic and consonantal interval durations correlated significantly with several of the specific rhythm metrics. nPVI was significantly correlated with DP accenting (C2F) and with final lengthening (U2P). VarcoC was significantly correlated with trochaic stress (S2W) and with final lengthening. Final lengthening was also significantly correlated with DP accenting.

**3.2. Rhythmic differences by age group**

Due to the large number of intercorrelations between measures, we assessed the effect of age group on rhythm in a MANOVA with Speaker added as a covariate.

**Table 2:** Means and standard deviations of the global rhythm measures (nPVI, VarcoC, %V) and speech rate (number of vocalic intervals per second) for each age group.

Age		Metrics			
		nPVI	varcoC	%V	Rate
5	M	48.33	61.24	50.69	3.46
	SD	5.4	5.67	2.91	0.21
8	M	55.00	62.41	49.01	3.99
	SD	5.81	7.32	2.72	0.19

**Table 3:** Means and standard deviations of the specific rhythm metrics for each age group. The metrics capture trochaic stress within a word (S2W), the accenting of a determiner noun phrase (C2F), and final lengthening (U2P).

Age		Metrics		
		S2W	C2F	U2P
5 yo	M	1.62	3.13	3.69
	SD	0.38	0.86	1.14
8 yo	M	1.84	3.55	4.72
	SD	0.44	0.60	1.39

The omnibus analysis showed that 5-year-olds' rhythm patterns differed significantly from 8-year-olds' rhythm patterns [ $F(7,11) = 5.52, p = .006$ ].

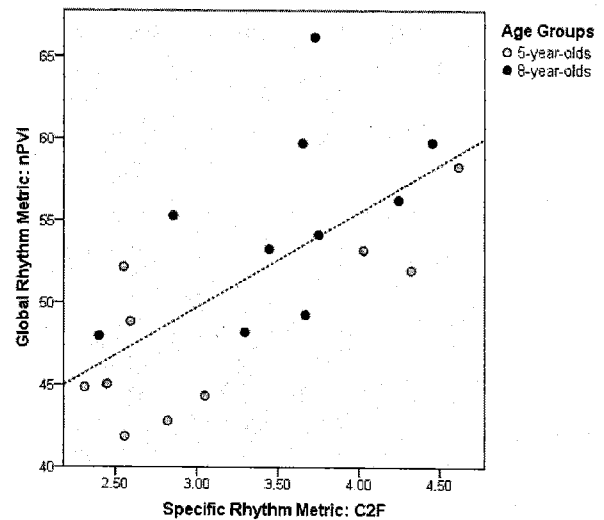
Speaker had no effect on the dependent measures. The direction of difference between the groups is evident in Table 2 and 3 below, which display the group means and standard deviations for each of the measures by age group.

It should also be evident from Tables 2 and 3 that 5-year-olds and 8-year-olds were best distinguished by measures of overall variance in vocalic interval durations (nPVI) and by speech rate: nPVIs were smaller and speech rates were slower in 5-year-olds' speech compared to 8-year-olds'. These differences were found to be significant in Type III tests of the group effect [nPVI,  $F(1,17) = 8.47, p = .010$ ; rate,  $F(1,17) = 30.82, p < .001$ ].

**3.3. Sequential vocalic interval durations**

The nPVI primarily reflects the summed differences of sequential vowel durations across an utterance. As noted in the introduction, multiple factors must contribute to this index of rhythm, which differentiates younger and older children's speech. The high correlations between nPVI, C2F, and U2P suggest 2 possible sources for the vowel-to-vowel duration variability measured by the nPVI; namely, DP accenting and final lengthening. We used a regression analysis to determine whether one or both of these factors independently predicted nPVI when age in months was also entered as a predictor variable.

**Figure 1:** The relationship between nPVI and DP accenting (C2F).



The model with 3 predictors explained 58% of the variance in nPVI values. Final lengthening (U2P) contributed very little towards this result, and without it the model explained 52% of the

variance in nPVI values. DP accenting (C2F) was the best predictor of nPVI values [ $b = 4.48$ ,  $t(16) = 2.74$ ,  $p = .014$ ], surpassing age-in-months, which was only a significant predictor of nPVI values when U2P was removed [ $b = .13$ ,  $t(17) = 2.16$ ,  $p = .046$ ]. By itself, DP accenting explained 46% of the variance in nPVI values. This strong relationship between vowel-to-vowel duration differences and patterns of supralelexical accenting is evident in Figure 1, which shows the relationship between nPVI and C2F values.

#### 4. GENERAL DISCUSSION

The current study sought to understand the relationship between global and specific rhythm measures, how these measurements might combine to distinguish younger and older children's speech, and whether such combinations might help to answer the question of why rhythm acquisition is so protracted in English. A principal finding was that only speech rate and the nPVI differentiated between 5- and 8-year-olds' speech. It is possible that additional differences between 5- and 8-year-old speech might have emerged had we used a spontaneous speech task rather than a delayed imitation task to assess language rhythm.

The other principal finding of the study was that variance in nPVI values was better explained by a specific measure devised to capture patterns of supralelexical accenting than by age. We find this latter result to be consistent with the argument that rhythm emerges not only from the lexical stress patterns and phonotactics of a language, but also from patterns of accenting across a phrase. And from this explanation for language rhythm, we also find an explanation for the protracted acquisition of stress timing: it requires the acquisition of most aspects of English prosody.

Stress timing also requires the mastery of vowel reduction and vowel lengthening. The finding that younger and older English speaking children did not differ significantly on measures of phrase-final lengthening may suggest that lengthening is mastered fairly early. Young children's slower speech rates may indicate that they have yet to master vowel reduction.

#### 5. ACKNOWLEDGEMENTS

We are grateful to the 2009-2011 Speech & Language lab testing teams for their help with data collection. This work was supported by NIH grant 5R01HD061458.

#### 6. REFERENCES

- [1] Allen, G., Hawkins, S. 1978. The development of phonological rhythm. In Bell, A., Hooper, J. (eds.), *Syllables and Segments*. Amsterdam: North-Holland. 173-185.
- [2] Bunta, F., Ingram, D. 2007. The acquisition of speech rhythm by bilingual Spanish- and English-speaking four- and five-year-old children. *Journal of Speech, Language and Hearing Research* 50, 999-1014.
- [3] Dauer, R. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11, 51-62.
- [4] Dellwo, V. 2006. Rhythm and speech rate: A variation coefficient for deltaC. In Karnowski, P., Szigeti, I. (eds.), *Language and Language Processing: Proceedings of the 38th Linguistic Colloquium Piliscsaba 2003*. Frankfurt: Peter Lang, 231-241.
- [5] Dellwo, V. 2008. The role of speech rate in perceiving speech rhythm. *Speech Prosody* Campinas, Brazil. [http://www.isca-speech.org/archive/sp2008/papers/sp08\\_375.pdf](http://www.isca-speech.org/archive/sp2008/papers/sp08_375.pdf)
- [6] Gay, T. 1978. Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America* 63, 223-230.
- [7] Grabe, E., Gut, U., Post, B., Watson, I. 1999. The acquisition of rhythm in English, French, and German. In Barriere, I., Morgan, G., Chiat, S., Woll, B. (eds.), *Current Research in Language and Communication: Proceedings of the Child Language Seminar*. London: City University, 157-163.
- [8] Grabe, E., Low, E. 2002. Durational variability in speech and the rhythm class hypothesis. *Laboratory Phonology 7*. Cambridge: Cambridge University Press.
- [9] Post, B., Payne, E. 2010. Phonological factors in rhythmic development: A cross-linguistic study. *Workshop on Prosodic Development* Barcelona, Spain.
- [10] Ramus, F., Newport, M., Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265-292.
- [11] Sabin, E., Clemmer, E., O'Connell, D., Kowal, S. 1979. A pausological approach to speech development. In Siegmán, A., Feldstein, S. (eds.), *Of Speech and Time: Temporal Speech Patterns in Interpersonal Contexts*. Hillsdale, NJ: Lawrence Erlbaum, 35-55.
- [12] Semel, E., Wiig, E.H., Secord, W.A. 2003. *Clinical Evaluation of Language Fundamentals, Fourth Edition (CELF-4)*. San Antonio, TX: The Psychological Corporation/A Harcourt Assessment Company.
- [13] Turk, A.E., White, L. 1999. Structural influences on accentual lengthening in English. *Journal of Phonetics* 27, 171-206.
- [14] Wells, B., Peppe, S., Goulandris, N. 2004. Intonation development from five to thirteen. *Journal of Child Language* 31, 749-778.
- [15] White, L., Mattys, S.L. 2007. Calibrating rhythm: First language and second language studies. *Journal of Phonetics* 35, 501-522.