Segmetal timing of young children and adults

Minjung Kim and Carol Stoel-Gammon
University of Washington, Dept. of Speech and Hearing Sciences
E-mail : minjungk@u.washington.edu

Abstract

Young children's speech is compared to adult-to-adult speech and adult-to-child speech by measuring durations and variability of each segment in CVC words. The results demonstrate that child speech exhibits an inconsistent timing relationship between consonants and vowels within a word. In contrast, consonant and vowel durations in adult-to-adult speech and adult-to-child speech exhibit significant relationships across segments, despite variability of segments when speaking rate is decreased. The results suggest that temporal patterns of young children are quite different from those of adults, and provide some evidence for lack of motor control capability and great variance in articulatory coordination.

1. Introduction

Segmental duration and variability of children's speech have been of interest in view of development of motor control for speech. Various studies have found that young children display longer segmental durations and greater temporal variability than older children and adults, and these two variables decrease over age and are assumed to change due to neuromuscular maturation and experience [6], [9],[10].

One issue associated with duration and temporal variability involves the relationship between these two parameters; it has been suggested that duration is highly correlated with variability (e.g., [3], [7]). With regard to the development in motor control for speech, some literature has suggested that temporal variability is simply as a consequence of longer duration by slow speech rate and might not necessarily function as a developmental criterion associated with neuromuscular maturation [4], [6]. However, Smith reported that duration tends to become adult-like earlier than temporal variability in the process of development, and concluded that variability should not be considered as a function of duration
but as an equivalent measure for speech motor control [9].

The present study compares the segmental timing patterns of young children, aged 21 to 33 months, with adult timing patterns. The main question is whether young children's speech resembles adult "slow" speech, or whether it displays distinct timing patterns, possibly related to lack of motor control for speech. This study examines intraword segmental timing through measuring duration and variability of each segment and determining the temporal relationships between segments in CVC monosyllable words.

2. Method

2.1 Participants

Child speech samples were collected from six typically developing children, three boys and three girls, who were involved in a series of previous studies of early phonological acquisition [2].All children were monolingual English speakers and met the inclusion criterion of proficiency greater than the 15^{th} percentile on the MacArthur Inventory of Communicative Development [5].

Adult-to-child (A-C) speech samples were gathered from four female "experimenters"who interacted with the children during data collection (one experimenter per session). The participants for adult-to-adult (A-A) speech samples were four female native speakers of American English. All adult speakers had no history of speech, language, or hearing impairment

2.2 Materials and procedures

For all samples, the target words were specific CVC words: *fish, sheep, Pete, Kit.* These words were elicited from children through naming activities with toys or objects in picture books and produced multiple times at each age. Most child speech productions were drawn from spontaneous speech, but imitations of the experimenters' productions were also included in the dataset.

A-C speech was recorded in the same way as child productions; all the target tokens were digitized at a 20-kHz sample rate using Kay Elemetrics Computerized Speech Laboratory (hardware Model 4300B). For A-A speech, short phrases or sentences with or without target words were elicited. Participants were instructed to produce each target word in response to simple questions at two speaking rates; normal and slow speaking rates. All A-A speech samples were recorded in a soundproof room at the Department of Speech and Hearing Sciences, University of Washington using a high-quality microphone and a digital CD recorder.

2.3 Acoustic analysis

The duration of each segment in CVC target words was measured on the basis of waveforms and 260Hz broad bandwidth spectrograms with a 10-kHz range in the Praat 4.1.6. [1]. Only CVC tokens in which the onset and offset of each segment were well identified spectrographically were included in the analysis. Tokens with a segment obscured by ambient noise were excluded. In defining the portion of each segment in Pete and Kit, only Voice Onset Time (VOT) was measured for the duration of initial consonants. For the duration of the final consonants in sheep, *Pete*, and *Kit*, a stop-gap which corresponds to the temporal distance between the offset of the preceding vowel and the release of the stop consonant was measured.

Reliability was assessed for 10 % of the tokens (randomly selected) by two other analyzers: one for A-A speech, and the other for child speech and A-C speech.Duration measures made by the primary researcher and the other analyzers differed in absolute value by 6.1 ms for A-A speech, 7.8 ms for A-C speech, and 6.9 ms for child speech.

A total of 320 word tokens for A-A speech, 228 for A-C speech, and 291 for child speech were analyzed.

2.4 Statistical analysis

In order to compare segmental timing patterns from each group, mean durations and standard deviation of all segments were calculated. Coefficient of variance (i.e., the ratio of standard deviation to mean) was calculated to compare variability because the range of segmental durations was large. Regression analysis was conducted using SPSS to investigate whether there is any linear relationship between durations of initial consonant and vowel in *fish* and *sheep*.

3. Results

3.1 Segmental durations and variability

Table I summarizes the results of the paired t-tests comparing segmental durations in adult normal and slow rate speech by pooled subjects. Since subjects A-A1 and A-A 2 did not release the final stop in many cases, data from A-A 3 and A-A 4 were pooled for the stop-gap in *Pete* the duration of the stop-gap in *Kit* is not included. The systematic lengthening effect of segments in slow speech is evident, and the durational differences between normal rate speech and slow rate speech are significant for all segments.

Table II presents numerical values of segmental durations in milliseconds with standard deviations (ms) and coefficients of variation (%) in parentheses. The overall word duration both in A-C speech and child speech (CS) is longer than that in adult "normal rate" speech. The durational patterns of A-C and CS productions are different from those of A-A slow speech. Notably, A-C speech exhibits longer vowels and final consonants relative to initial consonants. Children produced substantially longer vowels than initial consonants; however, in contrast with the A-C productions, they also produced relatively long initial consonants. The vowel and stop-gap was also substantially longer in children's productions in comparison with A-C speech.

TABLE I. Results of paired t-tests comparing segmental durations of adult speech at normal and slow speaking rates by pooled subjects: IC = initial consonant, Vl = vowel, FC = final consonant, VOT = voice onset time, SG = stop gap. Standard errors are shown in parentheses.

			Fish				*sheep*	
		IC	*Vl*	*FC*		*IC*	*Vl*	*SG*
Mean (SE)	Normal	122.5 (3.95)	127.2 (3.08)	236.3 (5.76)		158.7 (4.13)	112.0 (3.54)	103.6 (4.56)
	Slow	210.0 (7.70)	186.6 (6.32)	306.7 (9.16)		241.7 (8.38)	171.6 (7.14)	144.1 (6.37)
t	Paired	-11.328*	-9.043*	-9.000*		-10.573*	-9.949*	-5.705*
			Pete				*Kit*	
SG		*VOT*	*Vl*	*SG*		*VOT*	*Vl*	
Mean (SE)	Normal	61.0 (3.29)	112.7 (5.63)	95.3 (4.22)		70.7 (3.02)	99.7 (4.45)	
	Slow	94.8 (4.20)	166.5 (7.93)	146.7 (9.56)		99.5 (3.49)	126.6 (6.07)	
t	Paired	-7.063*	-9.281*	-6.132*		-6.331*	-4.591*	

* p < 0.001, df = 39 for all pairs with exception of SG in *Pete* (df= 19; only A-A 3 and A-A 4 pooled for this word)

Table II also illustrates variability measure across subjects by speech group. Average durations, pooled across speakers, for each segment were more variable in A-A slow speech than in A-A normal rate speech. Segmental durations in A-C speech were also highly variable. Variability was not necessarily correlated with segmental durations: some short segments were more variable compared to other long segments. All the segments in child speech were variable as expected by long segmental durations; VOT was also very variable.

3.2. Durational relationship between initial consonants and the following vowels

The results showed that durations of both initial consonants and following vowels increase as the whole word becomes longer. The correlations for the two variables in A-A productions are significant ($p<.001$), as shown in Table III. A-C speech also exhibits correlations between consonant and vowel durations in most instances.

In contrast to the linear relationships observed in A-A and A-C speech, children's productions exhibit no temporal relationships between durations of the initial consonant and following vowel except for FC 3 who shows a moderate correlation in fish productions. This pattern is maintained even without the outliers in sheep by FC 3.

4. Conclusion

The results of the present investigation reveal a general trend of increase in segmental duration and variability as a speaking rate decreases in adult speech. This tendency was even apparent in stop consonants, which are reported to be relatively free from tempo effects [3].

Both adult-to-child (A-C) speech and child speech exhibited long word durations in general; timing patterns for these groups differed from the adult-to-adult (A-A) slow speech group in that the proportion of vowel duration was greater. It appears that child speech is more similar to A-C speech than A-A slow speech with regard to vowel elongation, but consonants in child speech were noticeably elongated relative to A-C speech. This elongation of consonants might be associated with the strategy in learning speech sounds through hyper-articulation (e.g., the VOT study [8]).

There was a significant correlation between durations of the initial consonant and vowel in both A-A and A-C speech whereas child speech exhibited an inconsistent relationship between the two variables although both A-C speech and child speech shared some features of A-A slow speech. This finding suggests that variability across segments in young children's productions is not merely a result of slower speaking rate, contrary to Kent and Forner's assumption [6]. We hypothesize that the inconsistent timing patterns across segments in children's productions are associated with a lack of speech motor control resulting in variance in articulatory coordination.

TABLE II. Segmental durations (ms) by speech group: A-A N = normal speech addressed to adults by pooled subjects, A-A S = slow speech addressed to adults by pooled subjects, A-C = adult speech to children by each subject, FC/MC = female/male child speech by each subject. IC = initial consonant, Vl = vowel, FC = final consonant, VOT = voice onset time, SG = stop gap. Standard deviation (ms) and coefficient of variation (%) are in parentheses

	fish			*sheep*		
	IC	*Vl*	*FC*	*IC*	*Vl*	*SG*
A-A N	*123 (25,20.3)*	*127 (19,15.3)*	*236 (36,15.4)*	*159 (26,16.5)*	*112 (22,20.0)*	*104 (29,27.9)*
A-A S	*210 (49,23.2)*	*187 (40,21.4)*	*307 (58,18.9)*	*242 (53,21.9)*	*172 (45,26.3)*	*144 (40,27.9)*
A-C 1	*159 (36,22.7)*	*256 (93,36.5)*	*307 (46,15.1)*	*192 (41,21.2)*	*194 (49,25.2)*	*160 (24,15.3)*
A-C 2	*136 (54,39.9)*	*204 (98,48.1)*	*272 (70,25.8)*	*197 (45,22.9)*	*175 (36,20.6)*	*126 (22,17.2)*
A-C 3	*144 (41,28.6)*	*196 (50,25.6)*	*299 (68,22.7)*	*206 (77,37.3)*	*182 (68,37.2)*	*104 (22,21.0)*
A-C 4	*169 (34,20.1)*	*199 (59,29.9)*	*301 (50,16.7)*	*182 (61,33.4)*	*173 (50,28.8)*	*101 (19,19.0)*
CS-F1	*190 (55,28.7)*	*269 (87,32.4)*	*318 (75,23.5)*	*227 (68,30.2)*	*200 (66,32.8)*	*231 (57,24.8)*
CS-F2				*232 (56,24.0)*	*186 (42,22.6)*	*191 (59,31.1)*
CS-F3	*185 (58,31.6)*	*258 (96,37.2)*	*261 (76,29.1)*	*241 (50,20.9)*	*189 (81,42.8)*	*168 (68,40.7)*
CS-M1	*194 (64,32.7)*	*326(101,31.0)*	*242 (98,40.4)*			
CS-M2	*163 (57,35.0)*	*236 (96,40.8)*	*259 (56,21.6)*	*185 (41,22.6)*	*157 (46,29.4)*	*137 (34,24.9)*

	Pete			*Kit*		
	VOT	*Vl*	*SG*	*VOT*	*Vl*	*SG*
A-AN	*61 (21,34.0)*	*113 (36,31.6)*		*71 (19,27.0)*	*100 (28,28.2)*	
A-AS	*95 (27,28.0)*	*167 (50,30.1)*	*140 (36,26.0)*	*100 (22,22.2)*	*127 (38,30.3)*	*146 (45,30.6)*
A-C1	*63 (18,28.4)*	*206 (42,0.2)*	*134 (22,16.2)*	*83 (29,34.3)*	*147 (32,21.8)*	*141 (36,25.6)*
A-C2	*66 (25,37.9)*	*200 (55,27.7)*	*104 (30,28.9)*	*114 (35,30.6)*	*166 (32,19.0)*	*131 (32,24.2)*
A-C3	*66 (22,33.2)*	*164 (34,20.7)*	*84 (12,14.1)*	*112 (48,42.7)*	*148 (49,32.9)*	*107 (18,16.9)*
A-C4	*93 (42,45.2)*	*216 (65,30.2)*	*103 (26,24.9)*	*89 (26,28.9)*	*141 (76,53.9)*	*124 (25,0.4)*
CS-F1				*76 (31,41.2)*	*154 (62,40.3)*	*249 (109,43.7)*
CS-F2	*105 (51,48.7)*	*226 (58,25.5)*	*259 (86,33.3)*	*86 (33,38.3)*	*188 (89,47.3)*	*251 (80,32.0)*
CS-F3	*110 (53,48.6)*	*263 (131,49.9)*	*167 (72,43.3)*	*116 (46,40.1)*	*222(160,72.2)*	*268 (139,51.7)*
CS-M1	*98 (45,46.6)*	*304 (106,34.9)*	*173 (72,41.7)*			
CS-M3	*95 (27,28.9)*	*267 (150,56.2)*	*185 (50,26.9)*	*91 (33,36.5)*	*207 (106,51.2)*	*218 (94,42.9)*

TABLE III. *Correlation coefficient (r)* between the durations of the initial consonant and vowel in fish and sheep, by subject.

	fish	sheep
A-A 1	0.729**	0.851**
A-A 2	0.772**	0.788 **
A-A 3	0.578*	0.646**
A-A 4	0.903**	0.789**
A-C 1	0.245	0.526*
A-C 2	0.834**	0.724*
A-C 3	0.818**	0.374(0.722**)[a]
A-C 4	0.743*	0.408
CS-F 1	0.215	0.136
CS-F 2		0.039
CS-F 3	0.592*	0.077 (0.313)[a]
CS-M 1	0.284	
CS-M 2	0.395	0.022

* $p<0.05$, ** $p<0.01$, *** $p<0.001$
[a] r value without an outlier

Reference

[1] Boersma, P., and Weenink, D. (2003). Praat Version 4.1.6., Institute of Phonetic Sciences (Amsterdam, The Netherlands).

[2] Buder, E. H., and Stoel-Gammon, C. (2002). "American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing,"J. Acoustic. Soc. Am. 111, 1854-1864.

[3] Crystal, T. H., and House, A. S. (1988a). "Segmental durations in connected speech signals: Current results," J. Acoustic. Soc. Am. 83, 1553-1573.

[4] Crystal, T. H., and House, A. S. (1988b). "A note on the variability of timing control," J. Speech Hear. Res. 31,497-501.

[5] Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S., and Reilly, J. (1991). MacArthur Communicative Development Inventories (San Diego State Univ., San Diego)

[6] Kent, R. D., and Forner. L. L. (1980). "Speech segment durations in sentence recitations by children and adults," J.Phonetics 8, 157-168.

[7] Klatt, D. H. (1974). "The duration of [s] in English words," J. Speech Hear. Res. 17, 51-63.

[8] Macken, M. A., and Barton, D. (1980). ''The acquisition of the voicing contrast in English: A study of voice onset in word-initial stop consonants,'' J. Child Lang. 7, 4174.

[9] Smith, B. L. (1992). "Relationships between duration and temporal variability in children's speech," J. Acoust. Soc. Am. 91, 2165-2174.

[10] Tingley, B. M., and Allen, G. D. (1975). "Development of speech timing control in children," Child Dev. 46, 186-194.