DEVELOPMENTAL ASPECTS OF SECOND FORMANT TRAJECTORIES

Megan M. Hodge
Glenrose Rehabilitation Hospital, Edmonton, Alberta

1. Introduction

This investigation was based on the premise that the development of spatial-temporal control for speech production is governed by principles similar to those that characterize the development of spatial-temporal control for coordinated action in other skilled movements. Accordingly, it is predicted that as speech skill increases, an increase in the accuracy, speed, consistency and economy of speech movements should be evident. Correspondingly, measures of perceptual and acoustic correlates of these same phenomena should be sensitive to developmental changes in speech skill.

This paper presents the results of an acoustic investigation of the relationship between increasing talker age and the "speed" of speech movements, i.e., the speed at which the size and shape of the vocal tract changes to produce the sequences of consonant and vowel sounds in the speech signal. The speaking rates of young children are typically slower than those of adults [1]. One factor contributing to the slower speaking rates of children is their greater syllable durations, compared to adults. The purpose of this investigation was to determine if the greater syllable durations of children are attributable to slower rates of articulatory movement as indexed by second formant (F2) transition rate. The specific hypotheses to be tested were that as talker age increased to young adulthood (a) an increase in F2 transition rate (corresponding to an expected increase in speed of articulator movement), and (b) a decrease in the intra-subject variability, in F2 transition rates, would be observed.

Changes or "transitions" in the locations of the first (F1) and second (F2) formants in the speech signal over time have been used to make qualified inferences about the relative extents, durations and rates of change in vocal tract position as talkers produce a sequence of speech sounds [2, 3]. Changes in the frequency of the energy region corresponding to F1 are associated with changes in the degree of major vocal tract constriction and those corresponding to F2 are associated with changes in the anterior-posterior site of major vocal tract constriction. Over the course of an utterance, a trajectory of the changes in formant frequency can be obtained.

A prerequisite to testing predictions about relationships between measures of speech spectra such as formant frequencies, and increasing talker age is to identify satisfactory spectral normalization procedures. This is necessary to eliminate the absolute resonant frequency differences that result from vocal tract size differences so that interage differences in magnitudes of spectral change that are due to vocal tract size alone are not misinterpreted as developmental differences in speech production behaviors. The normalization procedure used involved a transformation of frequency measures in Hertz to the Bark scale using the method described by Syrdal [4]. The spectral measures compared across speakers were formant transition extents, defined as the amount of change or difference from the onset to the offset of the consonant-vowel transition for the opening gesture of the first syllable in the word "baby".

2. Outline of the Experiment

Twenty males in each of four age groups (2.5-3.5 years; 5-6 years; 9-10 years and 20-30 years) and 10 males between 14-22 months served as subjects (N=90). All subjects included passed a hearing, speech and language screening appropriate to their chronological age. Multiple productions of the word "baby" were recorded from each subject. Recordings were obtained in a sound-treated booth using a Shure Model 10A head-mounted microphone or a Sony Electret 150 microphone connected to a Sony Model 399 reel-to-reel tape recorder running at 19 cm/s.

Spectrograms were prepared for five repetitions of the word for each subject in the four older groups and for 3 to 5 repetitions for each subject in the youngest age group. The Kay 7800 and 5500 Digital Sonographs were used to generate the spectrograms, using filter bandwidth and dynamic range settings that provided the clearest resolution of the formant pattern. Each spectrogram was prepared for further analysis by marking the vowel segment in the first syllable from the first to the last clearly defined glottal pulse. The interval between these two points provided the duration of the vocalic portion of the syllable. The mid-point of F2 was then hand-traced from the begining to the end of this interval. These marked spectrograms were then analyzed by means of a computer program that processed and stored data on a graphics tablet connected to a mainframe Harris computer. The data for each utterance were stored in a file that included an identifier for each measure and a listing of time (x axis) and frequency (y axis) points. Following entry of the formant trajectories, the frequency data in the files were converted to Bark values. Formant transition rate was determined by calculating the extent of spectral change from the onset to the offset of the formant transition of interest and then dividing this extent by the time interval between these two points, i.e., the formant transition duration. The onset value was defined as the first spectral value in the trajectory and the offset value was defined as the first occurence of maximal spectral change, in the direction of interest, rounded to the nearest 0.01 Bark.

3. Summary of Results

a) As shown in Table 1, durations of the vocalic portion of the syllable were greatest for the one year-olds (M=208 s) and least for the adults (M=164).

b) Mean F2 transition extents were greatest for the three youngest age groups and least for the nine year-olds. Similarly, F2 transition durations were greatest for the three year-olds and least for the nine year-olds (see Table 1).

c) Mean F2 transition rates were similar across the five age groups, ranging from 0.113 Bark/ms for the one year-olds to 0.0096 Bark/ms for the three year-olds (see Table 1). The adults and the one year-olds had the greatest F2 transition rates. A perceptual analysis of the tokens revealed that many of the one year-olds' productions were characterized by an /i/ rather than an /e/ vowel quality in the first syllable, i.e., /bibi/ rather than /bebi/.

d) The intra-subject standard deviations for F2 transition rates across the multiple repetitions of the word exhibited the expected age trend, i.e., as age increased, the mean intra-subject standard deviation decreased. However, when the intra-subject standard deviations were compared for a subset of the time-matched subjects from each age group (vowel durations ranging from 155-200 ms), the two youngest groups still demonstrated the greatest mean intra-subject standard deviations, but the four older age groups did not show a well-defined age trend.

Age (yrs)	Extent (Bark)	Duration (ms)	Rate (Bark/ms)	Vowel Duration (ms)
1	1.21	103	.0013	208
	(.41)	(29)	(.0031)	(64)
3	1.22	128	.0096	203
	(.51)	(33)	(.0035)	(34)
5	1.20	124	.0101	205
	(.33)	(24)	(.0027)	(32)
9	0.91	95	.0100	172
	(.26)	(32)	(.0031)	(42)
Adults	1.14	106	.0111	164
	(.29)	(27)	(.0023)	(32)

Table 1. Group means and standard deviations of the F2 transition rates for the first CV in "baby".

Age (yrs)	All Subjects		Time-Matched Subjects	
	M	SD	M	SD
1	.0053	.0022	.0061	.0035
	(n=10)		(n=4)	
3	.0033	.0019	.0058	.0038
	(n=20)		(n=6)	
5	.0026	.0017	.0017	.0008
	(n=20)		(n=6)	
9	.0025	.0013	.0023	.0011
	(n=20)		(n=6)	
Adults	.0023	.0014	.0027	.0013
	(n=20)		(n=6)	

Table 2. Group means and standard deviations of the intra-subject standard deviations of F2 transition rates in the first CV of "baby". Data for a subset of time-matched subjects are also reported.

4. Conclusions

a) First it is obvious that the results did not support the initial hypothesis, i.e., that an increase in the rate of spectral change, and inferred increase in rate of articulatoy speed, would be observed as talker age increased. Rather, F2 transition rates appeared relatively stable across the four older age groups. As F2 slope has been associated with the intelligibility of an utterance [5], F2 transition rate may be determined by phonetic identity and as such, not influenced by increasing control of the articulators. However, F2 transition extents and durations appeared to be postively related, i.e., the greater the extent, the greater the duration, and these values generally decreased with increasing age. Thus, the longer durations of the vocalic portions of the younger children's syllables appeared to be associated with greater extents of F2 frequency change rather than slower F2 transition rates, compared to the adults. Stated a different way, the adults exhibited spectral and temporal "vowel reduction" compared to the children.

b) The high F2 transition rates of the one year-olds appear to reflect the greater F2 values associated with the high front vowel /i/ and suggest that these children were not producing the more finely graded movement for the mid-front vowel /e/ in the first syllable of the word "baby".

c) As expected, the intra-subject trial to trial variability in F2 transition rates decreased monotonically as age increased. However, when vowel duration was controlled, this relationship was not evident for the four older groups. This finding supports the point [6] that caution needs to be exercised when using variability measures to index speech sensorimotor development when (a) there is a dependent relationship between the size of the mean and standard deviation of the distribution of the measure of interest, and (b) there are potential differences in the shapes of the distributions to be compared.

5. References

1. Haselager, G, Slis, I Rietveld, A. An alternative method of studying the development of speech rate. Clinical Linguistics and Phonetics. 1991; 5:53-64
2. Weismer, G, Kent, RD, Hodge, M, Martin, R. The acoustic signature for intelligibility test words. Journal of the Acoustical Society of America. 1988;84:1281-1291
3. Kent, RD, Weismer, G, Kent, J, Rosenbek, J. Toward phonetic intelligibility testing in dysarthria. Journal of Speech and hearing Disorders: 54:482-500
4. Syrdal, A. Aspects of a model of the audiotory representation of American English vowels. Speech Communication. 1985;4:121-135
5. Kent, RD, Kent, J, Suffit, G, Brooks, B, Rosenbek, J. Relationships between speech intelligibility and the slope of second formant transitions in dysarthric subjects. Clinical Linguistics and Phonetics. 1989;3:347-358
6. Munhall, K. Articulatory variability. In P. Square-Storer (Ed) Acquired apraxia of speech in aphasic adults: Theoretical and clinical issues. 1989:64-83