667

Journal of Speech and Hearing Research, Volume 32, 120–132, March 1989

THE EMERGENCE OF PHONETIC SEGMENTS: EVIDENCE FROM THE SPECTRAL STRUCTURE OF FRICATIVE-VOWEL SYLLABLES SPOKEN BY CHILDREN AND ADULTS

SUSAN NITTROUER* MICHAEL STUDDERT-KENNEDY** RICHARD S. McGOWAN
Haskins Laboratories, New Haven, CT

A variety of evidence, including the speech errors of normal and aphasic speakers, and the metalinguistic skills of literate individuals, demonstrates that speech has an underlying phonemic organization. However, we know little about how this organization develops in the child. The purpose of the present study was to test the hypothesis that phoneme-sized phonetic segments emerge as functional units of perceptuomotor control from the child's gradual reorganization of the gestures forming its early words or syllables. We investigated the acoustic structure of syllables produced by young children and adults. Fricative-vowel syllables spoken by 40 subjects (eight adults and eight children at each of the ages 3, 4, 5, and 7 years) were analyzed acoustically to determine how well different syllables-initial fricatives were contrasted and how strongly they were affected by vocalic context. Results indicated two independent developmental trends: The extent to which speakers differentiated between /ʃ/ and /s/ increased with age, while the extent to which they coarticulated each fricative with its following vowel decreased. The results support the hypothesis that children initially organize their speech gestures over a domain at least the size of the syllable and only gradually differentiate the syllable into patterns of gestures more closely aligned with its perceived segmental components.

The status of the phoneme-sized phonetic segment in speech perception and production is notoriously problematic.[1] On the one hand, we have ample evidence for segmental function in speech errors (e.g., Shattuck-Hufnagel, 1983), backward talking (Cowan, Leavitt, Massaro, & Kent, 1982), aphasic deficit (e.g., Blumstein, 1981), and, not least, the alphabet itself. On the other hand, coarticulation, or the interleaving of gestural patterns both within and across syllable boundaries, usually precludes isolation of phonetic segments in either the articulatory or the acoustic records of a spoken utterance. This apparent lack of isomorphism between surface and underlying form raises the question of how the developing child comes to perceive and produce the phonetic segments of language.

Several studies of early phonological development suggest an approach to this question (e.g., Ferguson, 1986; Ferguson & Farwell, 1975; Menn, 1983a, b; Menyuk & Menn, 1979; Waterson, 1971). One theme of these studies is that the child's entry into language is mediated by meaning, and the earliest unit of segmental contrast in child speech is neither the phoneme nor the feature, but the word or formulaic phrase, consisting of one or a few syllables. As outlined by Studdert-Kennedy (1987), at least three lines of evidence support this view. First is the observation that phonetic forms mastered in one word are not necessarily mastered in another. For example, a 15-month-old child may execute [n] correctly in *no* but substitute [m] for [n] in *night* and [b] for [m] in *moo* (Ferguson & Farwell, 1975). Thus, the child does not contrast [b], [m], and [n], as in the adult language, but the three words with their insecurely grasped onsets.

A second point is that particular words may vary widely in their execution from one occasion to another. For example, Ferguson & Farwell (1975) report ten radically different attempts by a 15-month-old girl to say *pen* within one half-hour session. The attempts included such diverse structures as: [dɛᵈⁿ̩], [hɪn], and [ᵐbõ]. A striking property of the list is that it includes all the gestures of the adult model (lip closure, tongue raising and fronting, alveolar closure, velum lowering/raising, glottal opening/closing), but in each attempt gestures are omitted and/or incorrectly phased with respect to one another. For example, lip closure for the initial [p] in [pɛn] is properly executed with an open glottis and raised velum, but if glottal closure for [ɛ] and velum lowering for [n] are initiated at the same time as lip closure, tens of milliseconds earlier than in the correct utterance, the result will be [ᵐb], as in [ᵐbõ]. Here, then, it would seem that the child analyzed the adult model not into a sequence of integral segments, but into a collection of gestures which, if correctly timed, would have yielded the appropriate segments (cf. Browman & Goldstein, 1986).

A third line of evidence supporting the word as the unit of contrast is the widely attested difficulty that a child may have in switching place or manner of consonantal articulation within a syllable. Children often have idiosyncratic strategies for such words: Some children omit, others harmonize one or another of the consonants. For example, one child may attempt *fish* with [fɪ'], another with [ɪʃ]; faced with *duck*, one child may try [gʌk], another [dʌt]. Whether the child deletes or harmonizes a discrepant consonant, the evident difficulty in switching place of articulation suggests that a word is "assembled before it is spoken" as a single prosodic unit (Menn,

---

*Currently affiliated with Boys Town National Institute, Omaha, NE.

**Also affiliated with University of Connecticut and Yale University.

[1] Throughout this paper, the term *phonetic segment* refers to a segment having the articulatory, acoustic, and perceptual domain of the phoneme.

0022-4685/89/3201-0120$01.00/0

1983a, p. 16). It would seem, in fact, that a child with such idiosyncratic habits has "learned an articulatory program of opening and closing her mouth that allows her to specify two things: the vowel and one point of oral closure" (Menn, 1983a, p. 5). Reframing this in terms of gestures, we may say that the child has learned to coordinate glottal closing/opening and tongue positioning with raising/lowering the jaw, in order to approximate an adult word.

This description of a child's early words as articulatory routines leads neatly into an account of phonetic segments in production as "crystalized," or emergent, spatiotemporal patterns of gesture (cf. Studdert-Kennedy, 1986, 1987). The account we propose is consistent with the approach of several earlier studies in its emphasis on the development of temporal coordination (e.g., Gilbert & Purves, 1977; Menyuk & Klatt, 1975). However, the account differs from those approaches because it proposes to derive segments rather than assume them. We regard gestures as the primitive units to be coordinated, phonemes and their featural descriptors as the abstract, systemic products of establishing routines of coordination.

We may view the developmental process as driven by the conflicting demands of articulatory "ease" and lexical accumulation (Lindblom, MacNeilage, & Studdert-Kennedy, 1983). As long as the child has only a few words, only one or two articulatory routines are needed. Initially, the child exploits these routines by adding to his or her repertoire only words composed of gestural patterns similar to those already "solved," and by avoiding words with markedly different patterns. Once the initial routines have been consolidated, new routines begin to emerge under pressure from the child's accumulating vocabulary. New routines emerge either to handle a new class of adult words not previously attempted, or to break up and redistribute the increasing cohort of words covered by an old articulatory routine.

Early phonological development would seem then to be a process of: (a) diversifying articulatory routines to encompass more and different classes of adult models, and (b) gradually narrowing the domain within a word to which an articulatory routine applies. The logical end of the process is a coherent articulatory routine for each phonetic segment in a word (Menn, 1983a). The present study was an attempt to test the foregoing account of phonological development by examining patterns of articulatory organization as evidenced in the acoustic records of young children's utterances. We hypothesized that, if the account has any merit, young children whose phonological development was not yet complete would contrast phonemic minimal pairs less effectively and would show more evidence of intrasyllabic coarticulation than older children or adults.

In an earlier investigation of fricative-vowel syllable perception, we found that young children, as compared to adults, were relatively less sensitive to the segmental composition of a syllable and more sensitive to the transitional acoustic patterns that assure syllabic coherence (Nittrouer & Studdert-Kennedy, 1987). We drew this conclusion from the perceptual weights assigned by children and adults to fricative spectra and vocalic formant transitions in identifying stimuli distributed along an /ʃ/ to /s/ continuum. Young children (3 to 5 years) tended to place relatively more weight on the vocalic formant transitions and less on the fricative noise spectrum than did adults and older children (7 years).

The acoustic structure of a consonant-vowel syllable reflects the pattern of articulatory gestures in moving from a constricted or closed vocal tract to a more open tract. Thus, the acoustic onset of the syllable largely indicates the manner and place of closure, as determined by the consonant, while the acoustic offset largely indicates the shape of the open configuration, as determined by the vowel. By contrast, formant transitions reflect a pattern of gestures that execute or are concurrent with the act of opening itself, the pattern being determined by both consonant and vowel. Evidently, then, the young children of our study tended to focus relatively more perceptual attention on the pattern of consonant release gestures distributed over the syllable than on the initial consonantal constriction.

Given this difference between children and adults in perceptual organization, we might expect a corresponding difference in production. In a preliminary investigation of anticipatory coarticulation in children's speech, Repp (1986) analyzed /ə/-consonant-vowel sequences produced by one 4-year-old and one 9-year-old. He found that the younger child demonstrated greater intrasyllabic coarticulation than the older: Characteristics of the C segment were influenced more strongly by the V segment. This result is consistent with the hypothesis that phonological development involves learning to reorganize syllabic patterns of gesture into sequences of phonemically based patterns. The results also suggest that this development may still be going on in early childhood.

An alternative account of early speech development holds that a child's utterances are more, rather than less, segmentally organized than an adult's. A corollary of this notion is that coarticulation increases with age, perhaps as a way of making production more efficient. Kent (1983), for example, supports this claim with spectral displays of 3 children's and 3 adults' productions of the word *box*: The children's second formant trajectories during the vocalic portion display less spectral movement in anticipation of the following /ks/ than do adults'. The reduced spectral movements clearly did not result from the children's distributing equivalent closing gestures for /k/ over longer trajectories (although their utterances were nearly one and a half times as long as the adults'). Rather, the children's second formant frequencies were disproportionately higher than those of adults throughout the vocalic portion of their utterances. We will return to this point in the discussion.

The purpose of the present study, then, was to test these alternative accounts of development by analyzing and comparing the acoustic structures of fricative-vowel syllables spoken by children and adults. The goal of our analysis was to determine how well different initial consonants were contrasted and how strongly they were affected by vocalic context. The fricatives were /ʃ/ and /s/, the vowels /i/ and /u/.

As an index of how well speakers distinguished articulatorily between the fricatives, we used the centroid, or center of gravity. The centroid is the first moment of the spectral distribution, that is, its mean frequency, weighted by amplitude.[2] Because the resonances of the back cavity are poorly coupled with the atmosphere (Heinz & Stevens, 1961), fricative centroids primarily reflect details concerning the front cavity size and constriction shape. These components of vocal tract configuration, which primarily determine the high frequency parts of the spectra, are themselves determined by constriction shape, constriction placement, and lip posture. Thus, the centroid serves as a measure of anticipatory lip rounding, as well as an index of the details of fricative production, involving tongue tip configuration.

However, as a measure of fricative-vowel coarticulation, other than anticipatory lip rounding, the centroid tends to be a poor measure. The lingual aspects of coarticulation are better measured shortly before the onset of voicing, by a local detail in the fricative spectrum, namely, the beginnings of a peak that continues into the vocalic second formant (F2) (Soli, 1981). This fricative F2 is a resonance of the back cavity, and its measured amplitude rises as the constriction is released, due to excitation of the back cavity by glottal turbulence and to increasing coupling with the atmosphere (McGowan & Nittrouer, 1988). F2 frequency then varies with both fricative and following vowel, largely due to variations in constriction placement and lingual posture. Fricative F2 therefore serves as an index of both the extent to which speakers differentiate constriction placement of /ʃ/ and /s/ and the extent to which they anticipate lingual vowel gestures. Thus, from centroid differences between /ʃ/ and /s/ and from variations in fricative F2 produced before different vowels, one may infer how precisely a speaker has specified, and how distinctly he or she has segregated, the consonant in a fricative-vowel syllable.

Two contrasting outcomes of the study seemed possible, each consistent with one of the alternative views of development outlined above. First, if young children organize their speech production as a series of phonetic segments to a greater extent than adults do, we would expect individual segments to be produced more distinctly by children than by adults, with less influence of the phonetic environment. Alternatively, if young children are still in the process of mastering the segmental structure of speech and therefore tend to organize their speech in terms of syllables (or words), we would expect less distinction between phonetic segments and a greater influence of the surrounding phonetic context.

METHOD

Subjects

One group of eight adults (four females, four males) and four groups of eight children each at the ages 3, 4, 5, and 7 years participated in this study. No greater than a 3/5 split existed for the number of females/males in any age group. All adults were 20 or 21 years old, and all children were within −1 and +5 months of their birthdays. All subjects had hearing thresholds of better than 20 dB HL (American National Standards Institute, 1969) at each of the frequencies 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz.

Tympanometric results for all subjects displayed normal pressure peaks between +100 daPa and −150 daPa. All subjects were monolingual English speakers from the Middle Atlantic region, and all the children were judged to have normal articulation skills for their ages by two speech pathologists listening to recordings of spontaneous speech. The subjects also participated in a perception experiment (Nittrouer & Studdert-Kennedy, 1987).

Equipment

A Uher model 4200 portable tape recorder with an Electrovoice model 635A microphone was used to collect speech samples. This system has a flat frequency response to 10 kHz. A Digital Equipment Corporation VAX 780 computer and a Kay Elemetrics Digital Spectrograph were used for acoustic analyses.

Procedures

Speech samples were collected after the subjects had completed the perception experiment. Ten tokens each of the reduplicated syllables /ʃiʃi/, /sisi/, /ʃuʃu/, and /susu/ were collected in randomized groups of four. We used reduplicated syllables to maximize opportunities for consonant-vowel coarticulation: The first fricative would be open to anticipatory coarticulation, while the second would be open to both anticipatory and perseveratory coarticulation. Utterances were spoken in response to pictures of a girl, referred to by the pronoun *she*, of a boy pointing and saying *see*, of a *shoe*, and of a girl named *Sue*. The pictures had been used in the perception experiment, so that subjects were familiar with them and knew the appropriate labels. The experimenter provided a model for speaking the disyllables, using approximately the same rate, loudness, and intonation pattern with each subject. Subjects then had a brief practice period in which they were encouraged to approximate the experimenter's style of speaking. In other words, the young children were encouraged not to scream into the microphone or to speak too softly, quickly, or slowly.

From the ten tokens of each utterance, eight with minimal extraneous noise were chosen. These 32 tokens

---

[2] Centroids are computed over a frequency range equal to one half the sampling frequency. A centroid is the sum over the bins of the Discrete Fourier Transform of the frequencies, weighted by the absolute value of the amplitude.

per subject were digitized with a 20-kHz sampling rate, using a low-pass filter with a cut-off of 9.6 kHz. By means of the Haskins Laboratories Waveform Editing and Display program, the onsets of both vowels in all tokens were identified as points at which periodicity appeared in the waveform. Two measurements were then made on each token: centroids and second formant frequencies.

Centroids. Discrete Fourier Transform (DFT) spectra were computed using a 25.6-ms Hamming window at two points in the waveforms: beginning at 100 ms before the onset of the first vowel (VO1-100 ms), and beginning at 30 ms before the onset of the second vowel (VO2-30 ms). These points were chosen because they seemed to be, respectively, the least and the most likely to show effects of vowel context. The first section ended at 74.4 ms before the first vowel began, well before the region in which Soli (1981) reported seeing changes in the fricative spectrum due to the upcoming vowel. The second section ended close to the onset of the second vowel, well within the region in which Soli (1981) reported spectral changes due to the upcoming vowel. Centroids were computed on these spectral sections.

Second formants. A 210-ms section beginning at 60 ms before the onset of the second vowel and extending 150 ms into the second vowel was extracted from each token. Linear Predictive Coding (LPC) analysis was then performed on these sections, using a 24-coefficient model and a 20-ms Hamming window with 10-ms updates. We originally hoped to display formant trajectories similar to those of Soli (1981). However, LPC analysis failed to isolate the second formant at all in the fricative spectra of three adult male speakers, and only did so reliably from about 30 ms before vowel onset in the fricatives of the remaining five adult speakers. Also, for many subjects, even over the first 100 ms of the vowel, F2 was difficult to trace precisely.

We therefore decided to derive F2 values centered at two discrete locations in the section chosen for analysis: (a) 30 ms before the onset (VO2-30ms), and (b) exactly at the onset of laryngeal vibration for the second vowel (VO2). The first location was chosen both because it was the location farthest from vowel onset at which F2 peaks were present for most adults, and because it was the earliest point in Soli's analysis at which a difference in F2 values appeared among vowel contexts. The second location was chosen because it was the earliest point in the syllable at which F2 peaks were present for all adults.

Formant values for young children are difficult to determine. Descriptions of the formant structure of adult speech often do not match the apparent structure of children's speech (Bickley, 1986). We therefore applied strict criteria in assessing each second formant value. Wide-band spectrograms of at least one token of each syllable type for each subject provided a rough estimate of a speaker's F2 values and served as a guide to interpretation of the LPC analysis. In the LPC display, we required that a formant be traceable for at least 50 ms into the vowel, with a bandwidth of less than 400 Hz and an amplitude of at least 10,000 ILS (Interactive Laboratory System, Signal Technology, Inc.) units[3] in both fricative and vowel, and that the residual be less than 1000 ILS units, once voicing began. With these criteria, we obtained F2 values from at least 4 of the 8 tokens of each type (at least 16 of the 32 tokens) for 31 of the 32 children.[4]

Durations. Although we tried to have all subjects speak at approximately the same rate, we were concerned that differences in spectral structure for children and adults might simply reflect differences in speaking rate. As a check on this possibility, we measured segment and syllable durations for half of the subjects (4 in each age group), chosen randomly. Duration measures were generally more variable in individual children than in individual adults. However, analyses of variance on these segment and syllable durations revealed no significant mean differences as a function of age, so that the spectral differences between children and adults to be reported below are not likely to have arisen from differences in speaking rate.

RESULTS

Overall Spectra

To illustrate the overall spectra on which centroids were computed, Figures 1 and 2 display DFT spectra for a typical female adult and a typical 3-year-old, averaged over eight tokens of each syllable type at VO1-100 ms, and smoothed with a 400-Hz rectangular window. The /ʃ/ spectra are displayed in the left panels, the /s/ spectra in the right panels; dotted lines represent fricatives spoken in the /u/ context; solid lines, those spoken in the /i/ context. Centroid values are indicated below the figures.

For the adult (Figure 1), the /ʃ/ spectra rise sharply and are then flat or falling, with centroid values close to 6 kHz. The /s/ spectra, by contrast, rise relatively slowly across the whole frequency range and have centroid values close to 8 kHz. Thus, for this adult speaker the two fricative spectra clearly contrast in both form and centroid values, with /s/ spectra more heavily weighted in the high frequencies than the /ʃ/ spectra. While using the higher moments in addition to the first may generally provide a more detailed description of obstruent spectra (Forrest, Weismer, & Milenkovic, 1987), the first moment, or cen-

---

[3] This is strictly a relative amplitude value, derived from the ILS analysis package. Amplitude is measured as dB* 100 + 10,000, so that an ILS amplitude value of 12574 = 25.74 dB, although these units have no physically meaningful reference. However, we consistently found that when amplitude dipped below 10,000 ILS amplitude units (0 dB), formant frequency values fluctuated unpredictably across frames and bandwidths became quite broad, suggesting that the values given did not represent a strong spectral peak.

[4] We were unable to obtain F2 values for the youngest child in the study. This failure is consistent with the proposal that very young children's productions may not be accurately characterized in terms of adult formant patterns (Bickley, 1986).

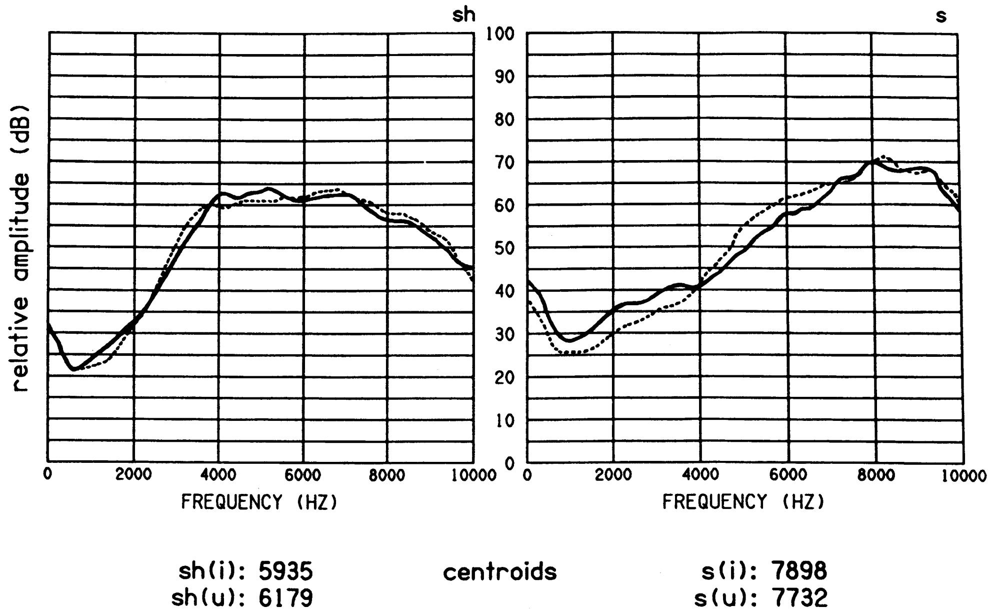

FIGURE 1. DFT spectra for one female adult, averaged over eight tokens of each utterance taken at VO1-100 ms, and smoothed using a 400 Hz rectangular window. The /ʃ/ spectra are on the left, the /s/ spectra on the right. The solid lines represent the /i/ context, the dotted lines represent the /u/ context. Mean centroids (in Hz) for the eight tokens are listed below the spectra.

troid, reliably distinguishes between spectral shape for these two fricatives.

For the child (Figure 2), the /ʃ/ spectra rise more slowly to a greater amplitude in the high frequencies and yield higher centroid values than the adult's do. For /s/, on the other hand, the child's spectra display greater amplitude over the low frequencies and less amplitude over the high frequencies than the adult's, so that (despite a presumably smaller vocal tract) they have lower centroid values. Overall, differences in both form and centroid value between /ʃ/ and /s/ spectra are less for the child than for the adult.

Centroids

VO1-100 ms. Mean centroids at VO1-100 ms for each age group are listed in the top portion of Table 1. Here (as

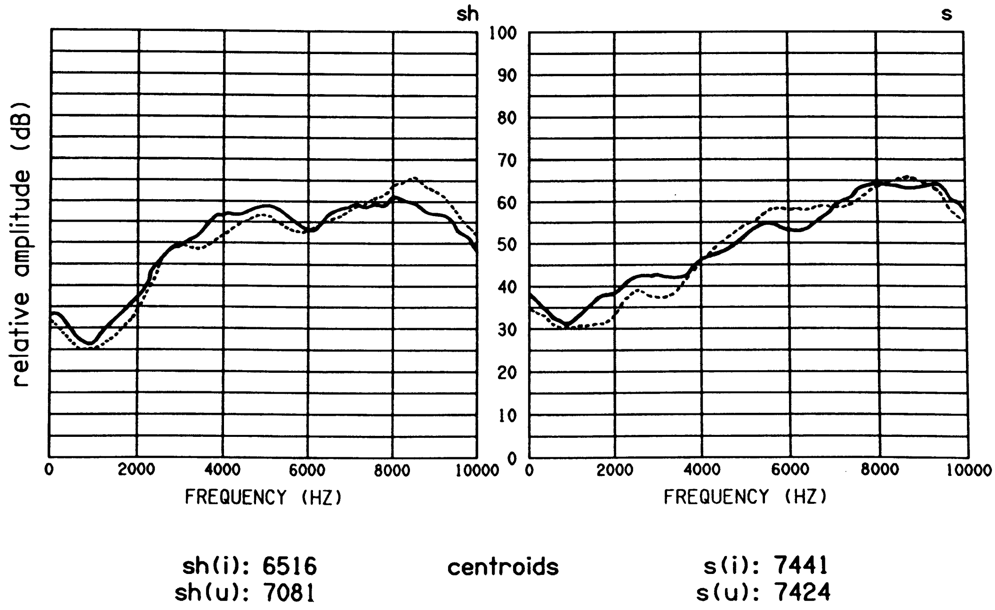

FIGURE 2. DFT spectra for one 3-year-old. See Figure 1 legend for details.

TABLE 1. Mean centroids in Hz at VO1-100 ms, mean centroid fricative ratios (s/ʃ), and mean centroid vowel-context ratios (i/u) for each age group.

	Age (years)				*Adult*		
	3	*4*	*5*	*7*	*Female*	*Male*	*Mean*
Centroids							
si	7234	7099	7263	6867	7820	6950	7385
ʃi	6189	6492	6011	5857	5819	5487	5653
su	6967	6819	6850	6689	7484	6863	7173
ʃu	6280	6382	6072	5750	5864	5656	5760
Fricative ratios							
si/ʃi	1.17	1.10	1.21	1.17	1.34	1.27	1.31
su/ʃu	1.11	1.07	1.13	1.16	1.28	1.21	1.25
Vowel-context ratios							
si/ʃu	1.04	1.04	1.06	1.03	1.04	1.01	1.03
ʃi/ʃu	0.99	1.02	0.99	1.02	0.99	0.97	0.98

in Tables 2, 3, and 4) mean values for adult males and females are listed both separately and combined. For statistical analysis and for graphical display, the adult data were combined.

Figure 3 graphs mean centroids at VO1-100 ms as a function of age. The relatively low centroids for /s/, noted in the 3-year-old spectra above, also appear in the group results for children of all ages. Nonetheless, centroids take higher values for /s/ than for /ʃ/ for both children and adults. The differences between fricatives increase with age, due to a steady decrease in the /ʃ/ centroids from the 3- and 4-year-olds to adults, combined with the low /s/ centroids for the children and an abrupt increase between 7-year-olds and adults. An effect of vowel context (lower centroids before /u/ than before /i/) is clear for /s/ but absent for /ʃ/, at all ages. Analysis of variance (for which the significant results are tabulated in Table 5) confirms these impressions: The analysis shows significant effects of fricative and of vowel context, and significant interactions of fricative-by-age and of fricative-by-vowel. A set of four orthogonal *t* tests was carried out to assess the age effects in this and all subsequent analyses: 3- versus 4-year-olds; 3- and 4- versus 5-year-olds; 3-, 4-, and 5- versus 7-year-olds; all children combined versus adults. Only those results in which the probability of obtaining the observed *t* value by change was less than .01 are reported. These tests traced the fricative-by-age interaction to stronger differentiation of the fricatives by the 7-year-olds than by the younger children, and by the adults than by all the children combined.

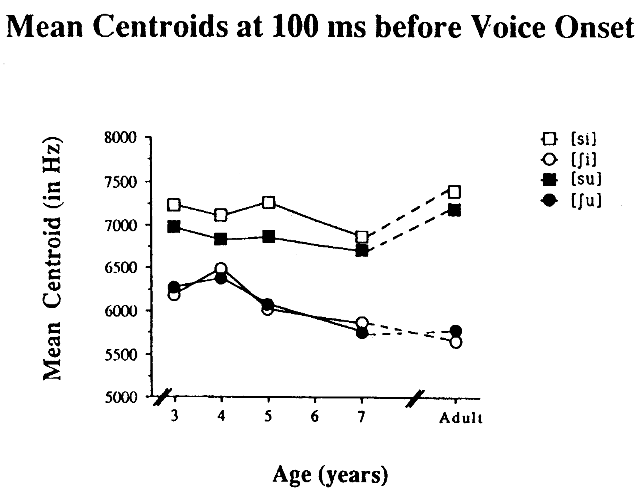

FIGURE 3. Mean centroids in Hz at VO1-100 ms as a function of age.

We can illustrate these effects perspicuously and, at the same time, remove any effects of age due to changes in vocal tract size, if we normalize the data by computing ratios of the mean centroids for each speaker on the relevant comparisons. Mean fricative ratios (si/ʃi and su/ʃu) and mean vowel-context ratios (si/su and ʃi/ʃu) are listed for each age group in the lower portions of Table 1. These ratios capture the contrasting effects of age on fricative differentiation and on anticipatory lip rounding as measured by the centroids.

Fricative ratios estimate the extent to which speakers differentiated the fricatives in each vowel context: The more a ratio deviates from a value of 1, the stronger is the differentiation between fricatives. All fricative ratios are greater than 1, and these deviations are significant, as indicated by the significant fricative effect in the analysis of variance. Mean fricative ratios also increase with age (the fricative-by-age interaction) and are larger before /i/ than before /u/ (the fricative-by-vowel interaction). Thus, by the centroid measure, all speakers differentiated the fricatives more strongly before /i/ than before /u/, and adult speakers differentiated the fricatives more strongly than child speakers did in both vowel contexts.

The vowel-context ratios (si/su and ʃi/ʃu) estimate the differential anticipatory effect of vowel context on each fricative. Ratios greater than 1 indicate that centroids are higher before /i/ than before /u/, while ratios less than 1 indicate the reverse. Table 1 shows that, across all age groups, ratios are greater than 1 for /s/, but fluctuate around 1 for /ʃ/ (the fricative-by-vowel interaction). Thus, for all speakers, by the centroid measure, the anticipatory effect of the vowel on the fricative is stronger for /s/ than for /ʃ/.

VO2-30 ms. Mean centroids at VO2-30 ms for each age group are listed in the top portion of Table 2 and are

TABLE 2. Mean centroids in Hz at VO2-30 ms, mean centroid fricative ratios (s/ʃ), and mean centroid vowel-context ratios (i/u) for each age group.

	Age (years) 3	4	5	7	*Adult* *Female*	*Male*	*Mean*
Centroids							
si	6734	6612	6719	6469	7078	6115	6596
ʃi	6080	6022	5685	5704	5335	5317	5326
su	6555	6255	6226	6351	6688	6147	6417
ʃu	5845	5898	5641	5545	5385	5460	5423
Fricative ratios							
si/ʃi	1.11	1.10	1.19	1.14	1.33	1.15	1.24
su/ʃu	1.12	1.06	1.12	1.15	1.24	1.13	1.18
Vowel-context ratios							
si/su	1.03	1.06	1.08	1.02	1.06	0.99	1.03
ʃi/ʃu	1.04	1.02	1.01	1.03	0.99	0.97	0.98

graphed in Figure 4. At this point in the second fricative of the disyllable, just before vocalic onset, the overall spectral weight is lower for both fricatives, for both vowel contexts, and for all ages than at VO1-100 ms. Presumably, the downward shift reflects jaw lowering and incipient release of the fricative constriction close to vocalic onset rather than a general difference between the first and second fricatives. The shift is greater for /s/ than for /ʃ/ and for adults than for children. Nonetheless, the overall pattern of higher centroids for /s/ than for /ʃ/, and before /i/ than before /u/, is preserved. Analysis of variance (Table 5) shows significant effects of fricative and of vowel context, and significant interactions of fricative-by-vowel and of fricative-by-vowel-by-age. Orthogonal *t* tests traced the three-way interaction to greater fricative differentiation by the 5-year-olds than by the 3- and 4-year-olds before /i/ but not before /u/.

Mean fricative and vowel-context ratios are listed in the bottom portions of Table 2. All fricative ratios are greater than 1, and these deviations are again significant by the analysis of variance, but the effects of age and vowel context are smaller than at VO1-100 ms. Comparison of Figures 3 and 4 (or of the fricative ratios in Tables 1 and 2) suggests that the failure to find an age-by-fricative interaction at VO2-30 ms, corresponding to that at VO1-100 ms, was due to a larger decrease in adult than in child /s/ centroids from the first measurement point to the second. Mean vowel ratios at VO2-30 ms are, with the exception of some adult values, somewhat greater than 1 for both fricatives. The difference in vowel effect between fricatives is less consistent than at VO1-100 ms, as indicated by the reduced significance level of the fricative-by-vowel interaction, but the overall anticipatory effect of vowel context is again significantly greater for /s/ than for /ʃ/.

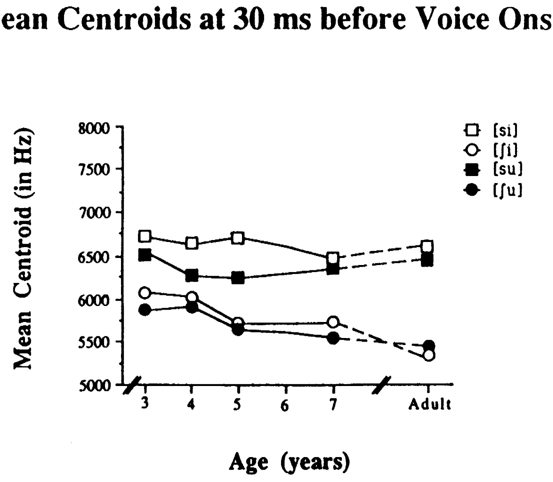

FIGURE 4. Mean centroids in Hz at VO2-30 ms as a function of age.

Second Formant Frequencies

VO2-30 ms. Mean F2 frequencies at VO2-30 ms for each age group are listed in the top portion of Table 3, and are graphed in Figure 5. As would be expected from the course of vocal tract growth, F2 frequencies decrease with age and show a particularly abrupt drop between 7-year-old children and adults. F2 frequencies are also higher for /ʃ/ than for /s/ before both vowels, and higher before /i/ than before /u/ for both fricatives. The higher frequencies for /ʃ/ than for /s/ reflect excitation of the cavity behind the constriction, which is shorter for /ʃ/ than for /s/ (McGowan & Nittrouer, 1988). Analysis of variance (Table 5) shows significant effects of age, fricative and vowel context, and a significant vowel-by-age interaction. Orthogonal *t* tests traced (a) the main effect of age to the sharp drop in F2 frequencies from 7-year-olds to adults and (b) the vowel-by-age interaction to significantly greater effects of vowel context at 3 than at 4 years; at 3, 4, and 5 than at 7 years; and for all children combined than for adults.

Mean fricative and vowel-context ratios are listed in the bottom portions of Table 3. Fricative ratios, as indicated by the analysis of variance, are significantly less than 1 at all ages, due to higher frequencies for /ʃ/ than for /s/. Mean vowel-context ratios are significantly greater than 1 for both fricatives, but decline with age (the vowel-by-age interaction). Evidently, younger speakers tended to anticipate each vowel more than older speakers did in both fricatives.

TABLE 3. Mean F2 frequencies in Hz at VO2-30 ms, mean F2 fricative ratios (s/ʃ), and mean F2 vowel-context ratios (i/u) for each age group. (Note: F2 frequencies at this point in the fricative could not be reliably determined for three of the adult males. The adult male values are therefore for a single subject).

	Age (years)				*Adult*		
	3	*4*	*5*	*7*	*Female*	*Male*	*Mean*
F2 frequencies							
si	2789	2710	2691	2495	2180	1676	2079
ʃi	2960	2948	2863	2745	2281	1812	2187
su	2332	2432	2326	2227	2056	1523	1949
ʃu	2430	2633	2487	2539	2159	1698	2067
Fricative ratios							
si/ʃi	0.94	0.92	0.94	0.91	0.96	0.92	0.95
su/ʃu	0.97	0.93	0.94	0.88	0.95	0.90	0.94
Vowel-context ratios							
si/su	1.20	1.11	1.16	1.12	1.06	1.07	1.07
ʃi/ʃu	1.23	1.12	1.15	1.08	1.06	1.07	1.06

VO2. Mean F2 frequencies at VO2 for each age group are listed in the top portion of Table 4 and are graphed in Figure 6. The general pattern across ages and fricatives and vowel contexts is similar to that for F2 frequencies measured 30 ms earlier. Analysis of variance (Table 5) shows significant effects of age and fricative and vowel context, and significant interactions for vowel-by-age and for fricative-by-age. Orthogonal *t* tests revealed (a) significantly higher overall F2 frequencies for the younger (3, 4, 5 years) than for the older (7 years) children and for all children combined than for adults; (b) significantly greater effects of vowel context for 3- than for 4-year-olds, for 3-, 4-, and 5- than for 7-year-olds, and for all children combined than for adults; (c) significantly less differentiation between fricatives for 3- than for 4-year-olds, and for 3-, 4-, and 5- than for 7-year-olds.

Mean fricative and vowel-context ratios are listed in the bottom portions of Table 4. Fricative ratios are again significantly less than 1 at all ages but also tend to decline with age (the fricative-by-age interaction). This age effect appears to be due primarily to a higher mean fricative ratio for the 3-year-olds than for other age groups, a trend also present, though to a lesser extent, for the /u/ fricative ratio obtained at VO2-30 ms. Mean vowel-context ratios display essentially the same pattern as those measured 30 ms earlier: All ratios are greater than 1 for both fricatives, but decline with age (the vowel-by-age interaction). Again, younger speakers evidently anticipated each vowel to a greater extent than did older speakers with both fricatives.

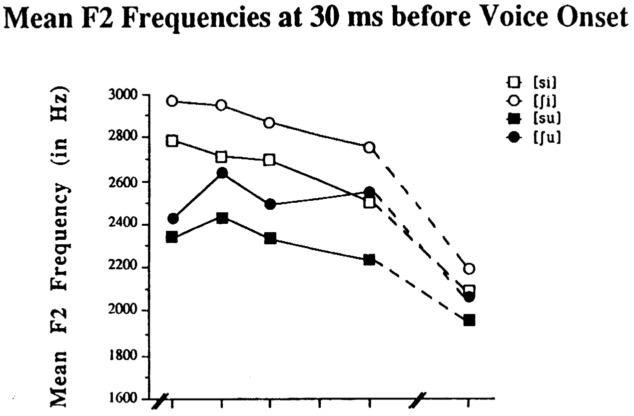

FIGURE 5. Mean F2 frequencies in Hz at VO2-30 ms as a function of age.

Correlations Between Fricative and Vowel Context Ratios

In the preceding analyses, centroids and F2 frequencies, when entered into measures of fricative differentiation (fricative ratios) or fricative-vowel coarticulation (vowel-context ratios), displayed somewhat complementary patterns with respect to the effects of age. For centroids, fricative ratios tended to increase with age, while vowel-context ratios remained stable; for F2 frequencies, fricative ratios tended to remain stable (apart from a small effect at VO2), while vowel-context ratios decreased with age. Thus, centroids showed age-related changes in fricative differentiation; F2 frequencies showed age-related changes in fricative-vowel coarticulation. This dissociation between F2 frequencies and centroids evidently arises, as noted earlier, because the two measures assess different aspects of articulation: The centroid is primarily an index of front cavity size, as determined by tongue tip position and lip rounding, while F2 frequency is primarily an index of the length of the cavity behind the fricative constriction. At the same time, since fricative differentiation increases with age, while fricative-vowel coarticulation decreases, one may wonder whether the former effect is not simply a consequence of the latter: Perhaps speakers contrast the two fricatives, spoken before the same vowel, only to the extent that they reduce anticipatory vowel coarticulation.

TABLE 4. Mean F2 frequencies in Hz at VO2, mean F2 fricative ratios (s/ʃ), and mean F2 vowel-context ratios (i/u) for each age group.

	Age (years)				Adult		
	3	4	5	7	Female	Male	Mean
F2 frequencies							
si	2931	2756	2707	2487	2145	1754	1950
ʃi	2952	2993	2901	2742	2274	1898	2086
su	2276	2409	2262	2168	1890	1530	1710
ʃu	2372	2576	2406	2458	2096	1806	1951
Fricative ratios							
si/ʃi	0.99	0.92	0.93	0.91	0.94	0.92	0.93
su/ʃu	0.96	0.94	0.94	0.88	0.90	0.85	0.88
Vowel-context ratios							
si/su	1.29	1.15	1.20	1.15	1.13	1.15	1.14
ʃi/ʃu	1.25	1.17	1.21	1.12	1.08	1.05	1.07

To test this possibility, we computed Pearson product-moment correlation coefficients across all age groups combined, first between all possible combinations of centroid fricative ratios with F2 vowel-context ratios, and second between all centroid vowel-context ratios and all F2 fricative ratios. Table 6 lists the results, with the second set of correlations in parentheses. Both pairs of centroid ratios are significantly correlated, despite the fact that one member of each pair derives from measures taken on the first fricative in the reduplicated syllables, the other from measures taken on the second. Both pairs of F2 ratios, derived from measures taken at points 30 ms apart within the same syllable, are also significantly correlated. Since all the measures entering into these correlations were independently taken, their statistical significance attests principally to the reliability of the measurements. All but one of the eight correlations between fricative and vowel-context ratios are negative, but none is significant. We conclude that the degree to which speakers differentiated between syllable-initial /s/ and /ʃ/ was largely independent of the degree to which they anticipated the following vowel.

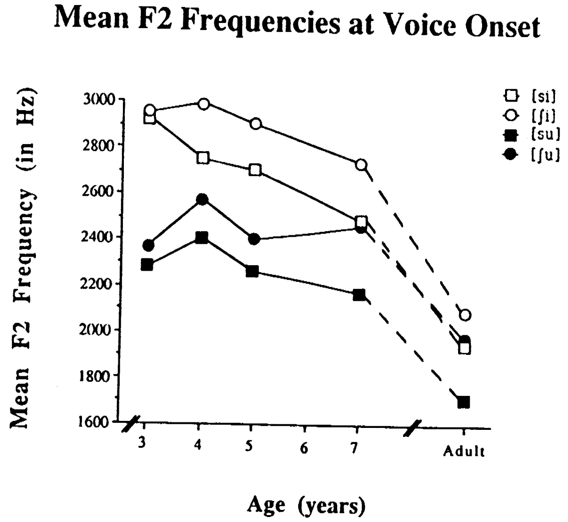

FIGURE 6. Mean F2 frequencies in Hz at VO2 as a function of age.

DISCUSSION

We hypothesized that young children whose phonologies are still developing have a broader minimal domain of articulatory organization than adults have, because they have not yet fully mastered the segmental structure of speech. We therefore predicted that such children would contrast phonetic segments less clearly than adults do and would display a greater influence of surrounding phonetic context. The results confirm this prediction and afford some insight into the details of child-adult differences in fricative contrast and in fricative-vowel coarticulation.

Fricative Contrast

By the centroid measure at VO1-100 ms, adults differentiate between the fricatives more strongly than 7-year olds, and the 7-year olds more strongly than younger children. The following sections argue that the age-related increase in fricative contrast is primarily due to improved control over constriction shape, and that the younger children already execute constriction placement quite largely, and lip rounding entirely, in an adult fashion. We will consider each factor in turn.

Lip rounding. To separate the effects of labial and lingual action on centroid values, we take advantage of the fact that lip rounding contributes to the articulation of both /ʃ/ and /u/. We may then ask whether lip rounding as a component gesture in the articulation of /ʃ/ and anticipatory lip rounding before /u/ are additive or redundant. For centroids at VO1-100 ms the answer comes from a comparison of the frequency-lowering effects due to /ʃ/ and to /u/ (Figure 3). For /s/, the centroid value before /u/ is some 200–400 Hz lower than before /i/ across all age groups, yielding /si/-/su/ vowel-context ratios significantly

TABLE 5. Summary of significant effects in analyses of variance of centroids at VO1-100 ms and VO2-30 ms, and of second formant frequencies at VO2-30 ms and VO2.

Measurement point	*Dependent variable*	*Independent variable*	*Degrees of freedom*	*F*	p
VO1-100 ms	Centroid	Fricative	1,35	238.39	<.0001
		Vowel	1,35	16.47	<.0003
		Age × Fricative	4,35	6.99	<.0003
		Fricative × Vowel	1,35	21.72	<.0001
VO2-30 ms	Centroid	Fricative	1,35	112.09	<.0001
		Vowel	1,35	22.15	<.0001
		Fricative × Vowel	1,35	5.84	<.01
		Age × Fricative × Vowel	4,35	2.78	<.05
VO2-03 ms	F2	Age	4,30	12.05	<.0001
		Fricative	1,30	49.69	<.0001
		Vowel	1,30	191.49	<.001
		Age × Vowel	4,30	7.20	<.0003
VO2	F2	Age	4,34	25.38	<.0001
		Fricative	1,34	87.07	<.0001
		Vowel	1,34	409.47	<.0001
		Age × Fricative	4,34	3.12	<.03
		Age × Vowel	4,34	13.83	<.0001

greater than 1. For /ʃ/, on the other hand, the effect of /u/ ranges unsystematically from a drop of roughly 100 Hz to a rise of about 90 Hz, yielding /ʃi/-/ʃu/ vowel-context ratios close to 1, again with no effect of age. Evidently, then, anticipatory lip rounding before /u/ does not increase the lip rounding already present as a gestural component of /ʃ/. The stronger differentiation of the fricatives before /i/ than before /u/ across all ages (the fricative-by-vowel interaction) illustrates the same point: The lower centroid fricative ratios before /u/ than before /i/ reflect the consistently lower /s/ centroids before /u/ and the absence of a systematic vowel-context effect on /ʃ/ centroids. The pattern of results is similar for the centroids at VO2-30 ms, though complicated by age-related effects due to tongue action at a point close to vocalic onset (discussed below). We conclude that lip rounding for /ʃ/ and for /u/ are redundant rather than additive and that they are already in the articulatory repertoire of a 3-year-old.

Constriction placement. To separate the effects of constriction shape and constriction placement on fricative contrast, we first take advantage of the fact that the fricative F2 is largely a resonance of the cavity behind the constriction, of which the frequency is determined by a combination of constriction position in the front-back dimension, the shape of the tongue body behind the constriction, and variations in overall vocal tract length (largely a function of lip posture). Accordingly, Tables 3 and 4 show higher F2 values for /ʃ/ than for /s/, an effect of constriction position, and for /i/ than for /u/, an effect of the shape of the tongue body behind the constriction and of tract length. (Note that if the vowel difference were due to tongue tip placement, we would expect F2, a back cavity resonance, to be higher before /u/ than before /i/.) As a result, F2 fricative ratios are significantly less than 1 at both measurement points before both vowels, while F2 vowel-context ratios are significantly greater than 1 for both fricatives. Soli (1981) reports similar results for a single adult speaker. He gives F2 frequencies averaged over five utterances at a point 30 ms before vocalic onset, as 1750 Hz and 1550 Hz for /s/ and 1940 Hz and 1800 Hz for /ʃ/, before /i/ and /u/ respectively [values estimated from Figures 3 and 7 in Soli (1981)]. These frequencies yield fricative ratios of 0.88 for si/ʃi and 0.89 for su/ʃu, vowel-context ratios of 1.09 for si/su and 1.10 for ʃi/ʃu. These values are close to the mean values for adults at VO2-30 ms in the present study (Table 3).

TABLE 6. Pearson product-moment correlation matrix (all age groups combined) for mean centroid fricative ratios and F2 vowel-context ratios and, in parentheses, for mean centroid vowel-context ratios and F2 fricative ratios, averaged across contexts.

	Centroid VO1-100	*Centroid VO2-30*	*F2 VO2-30*
Centroid VO2-30	.77* (.54*)	—	—
F2 VO2-30	−.21 (.01)	−.23 (−.13)	—
F2 VO2	−.29 (−.04)	−.22 (−.28)	.91* (.87*)

*$p<.001$

As we have already remarked, the higher F2 values for /ʃ/ than for /s/ reflect excitation of the cavity behind the constriction, this cavity being smaller for /ʃ/ than for /s/ (McGowan & Nittrouer, 1988) so that all F2 fricative ratios are less than 1. The younger children have somewhat higher ratios than the 7-year-olds and adults, but this effect is significant only at VO2, not at VO2-30 ms. Evidently, the younger children are somewhat unstable in their tongue placement as they approach vowel onset. Whether they tend toward a relatively fronted /ʃ/ or toward a relatively retracted /s/, or both, we cannot immediately determine.

Constriction shape as indexed by centroids. Because the resonant frequencies of the cavity behind the fricative constriction contribute little to the spectrum until just before vowel onset, centroid values at VO1-100 ms and at VO2-30 ms primarily reflect the shape of the constriction and the length of the cavity in front of it. The relative lengths of the back cavities for /s/ and /ʃ/ and, therefore, the relative placements of the tongue tip are indexed by F2 fricative ratios. We have just seen that these ratios do not differ significantly between children and adults until vowel onset (VO2) (compare Tables 3 and 4). We therefore infer that children and adults do not differ significantly in the two main determinants of front cavity length during fricative production, namely, relative place of constriction and (as already argued above) relative degree of lip rounding. Accordingly, any differences in centroid fricative ratios as a function of age must be due to differences in constriction shape.

The increase in centroid fricative ratios with age (Tables 1 and 2) indicates then that children do not differentiate constriction shapes for /s/ and /ʃ/ as sharply as do adults. However, we cannot tell from this pattern of ratios whether children tend to produce more /s/-like constrictions for /ʃ/ than do adults, more /ʃ/-like constrictions for /s/, or some roughly intermediate constriction for both.

Fricative-Vowel Coarticulation

Lip rounding and constriction shape. We have argued from the patterns of /s/ and /ʃ/ centroids and of F2 frequencies before /i/ and /u/ that 3-year-old children already execute lip rounding and coordinate it with tongue and jaw action in an essentially adult fashion. The pattern of age-related decline in F2 frequencies at both measurement points lends further support to the argument. For if a normal decline with age due to vocal tract growth were supplemented before /u/ by an added decline due to increased anticipatory lip rounding, we would expect the overall decline in F2 frequencies to be greater before /u/ than before /i/, so that F2 vowel-context ratios would increase with age. In fact, the pattern is exactly the reverse. Vowel-context ratios decrease with age, due to a much larger drop in F2 frequencies before /i/ than before /u/ (see Tables 3 and 4). Moreover, the pattern is present for both fricatives (there is no fricative-by-vowel interaction) and is therefore independent of the child-adult differences in fricative constriction shape, described above, an interpretation borne out by the pattern of correlations in Table 6. Thus, the children's stronger fricative-vowel coarticulation (as indicated in the age-by-vowel interactions for F2 frequencies) is evidently not due to child-adult differences in either constriction shape or lip rounding.

The conclusion that adults and children do not differ in their degree of anticipatory lip rounding before /u/ was also reached by Sereno, Baum, Marean, and Lieberman (1987). These authors analyzed the syllables /si, su,ti,tu,di,du/ spoken by four adults and eight children. For /si,su/, their analysis included an estimate of the "major spectral peak . . . between 1.5 and 3.0 kHz for the adult stimuli and between 1.5 and 3.5 kHz for the child stimuli in anticipation of the second formant of the following vowel" (p. 515). Although their estimates of F2 frequencies correspond moderately well with our own, their analysis did not permit them to separate the effects of lip rounding from those due to lingual coarticulation, nor did they observe an age-related decline in the differences between /si/ and /su/. Computation of vowel ratios (/si/-/su/) for their data yields a mean of 1.11 for the adults and of 1.12 for the children, with the highest ratios (1.28 and 1.32) coming from the two oldest children.

These values are not far from those of the present study (see Table 3) but the Sereno et al. (1987) samples were evidently too small and too unsystematic for an age-related pattern to emerge. The child sample included one 3-year-old, one 4-year-old, two 5-year-olds, three 6-year-olds, and one 7-year-old; from these subjects, between 1 and 6 tokens of /si/ and /su/ were gathered (the mean number of tokens was 2.9). The analysis then treated this small heterogeneous group as a homogeneous sample of "children."

Several other possible reasons for the discrepancy in age effects between the two studies suggest themselves. First, Sereno et al. (1987) seem not to have established spectral continuity between fricative and vowel second formants, so their measurements may be somewhat unreliable. Second, "each utterance was read from a 3 × 5 card by the adult subjects and repeated by the children" (p. 513). Since children's immediate repetitions (or imitations) are often more phonetically precise than their spontaneous utterances, this procedure is likely to have reduced adult-child differences. Finally, Sereno et al. (1987) analyzed "only those child tokens judged by the experimenters to be highly intelligible" (p. 513), so perhaps only the most adult-like tokens were included. Nonetheless the data of Sereno et al. (1987) are consistent with the claim that adults and children do not differ in their degree of anticipatory lip rounding before a rounded vowel.

Overlap of consonant and vowel gestures. If children and adults do not differ in anticipatory lip rounding, the children's stronger fricative-vowel coarticulation must be due to greater overlap between their consonant and vowel lingual gestures, that is, to greater fronting of the tongue body before /i/ and greater backing of the tongue body before /u/. This interpretation receives particularly strong support from the age-related decline in F2 vowel-context ratios for /ʃ/, since we have already seen that lip rounding for /ʃ/ is no greater before /u/ than before /i/. The declining vowel-context ratios for /ʃ/ must therefore be almost entirely due to a decline in anticipatory lingual action.

This outcome is consistent with the hypothesis that young children organize their speech over a wider temporal domain than adults do, and this hypothesis can nicely explain certain results of Kent (1983), noted earlier. The F2 values for the three 4-year-old children of his study were markedly higher throughout the vocalic portions of the syllable than the three adults' F2 values, even at the end of the vocalic portions just before the /k/ segments, when F2 rises with the onset of consonantal closure. The difference was probably not simply due to the expected higher formant frequencies of children over those of adults. For, although the author reports no F2 values for isolated productions of /ɑ/ in the 3 children of his study, F2 values in the utterance /bɑks/ were very much higher than the F2 values listed for children's productions of /ɑ/ by Peterson and Barney (1982). Perhaps these children already had their tongue bodies raised close to their positions for /k/ from the onset of the vocalic segment. In fact, Kent (1983) himself describes such a pattern of whole syllable organization, and labels it the "everything moves at once" principle (p. 70). He suggests further that this principle of

> synchronous movements of the structures involved in a motor objective . . . might in fact be more common in developing or impaired neuromotor systems than in a highly practiced and mature system. At least in the case of speech, fluent motor execution and a high event rate may depend on an overlapping of movements rather than synchronization of movements. (pp. 70, 71)

Implicit here is the notion that development consists in reducing the synchronous activity of the articulators. Thus, for the 15-month-old child mentioned earlier, a synchrony principle neatly captures the substitution of [mb] for [p] in *pen*: At this age the child still has difficulty in the relative phasing of the actions of discrete articulators (tongue, lips, velum, glottis) and so executes lip closure, velum lowering, and glottal narrowing at the same time. Perhaps, indeed, it is precisely where the adult language permits such synchrony (as in the lip-jaw-tongue actions of the fricative-vowel syllables in the present study) that the child demonstrates adult-like performance. However, spatiotemporal coordination of discrete articulators is presumably no longer a problem, even in the 3-year-old. For by this age, the child normally has escaped from consonant harmony and can comfortably execute a CVC syllable with two different places of lingual articulation for the consonants. What the young child evidently cannot always do is to phase, in a fully adult fashion, diverging consonantal and vocalic gestures of the tongue in a CV dyad.

We may be tempted to attribute the children's relatively poor fricative differentiation and relatively strong fricative-vowel coarticulation to their immature systems of motor control. Perhaps they were trying to produce intrasyllabic segmental contrasts identical to those of adults, but lacked the motor skill to do so. However, the results of our earlier perception experiment on these same children (Nittrouer & Studdert-Kennedy, 1987) argue that their perceptual organizations were also less segmental than those of adults. The children's fricative judgments, presumably unconstrained by their motor control abilities, were relatively more strongly influenced by fricative-vowel transitions and less strongly by the steady-state fricative spectrum than were adults' judgments.

Taken together, the two studies suggest that a child's phonology is grounded in both perceptual and motoric constraints. Certainly, perceptual capacity is logically prior to and must lead productive capacity, but perhaps the two are never far apart. At each point in development, an organism is complete and sufficient, adapted to its present needs. And at each point in language development, we may suppose, the child has the phonology that its perceptuomotor skills permit and assure.

We are not, of course, proposing that the children in these studies were insensitive to differences in the initial portions of the syllables (i.e., differences between /ʃ/ and /s/). Perception experiments have demonstrated that even infants can discriminate between syllables differing only in the initial fricative constriction (Kuhl, 1980). Rather, we are proposing that the initial domain of perceptuomotor organization is a meaningful unit of one or a few syllables, a coherent acoustic structure formed by the interleaved actions of partially independent articulators. As the number and diversity of the words in a child's lexicon increase, words with similar acoustic and articulatory patterns begin to cluster. From these clusters there ultimately precipitate the coherent units of sound and gesture that we know as phonetic segments. Precipitation is probably a gradual process perhaps beginning as early as the second to third year of life when the child's lexicon has no more than 50–100 words. But the process is evidently still going on in at least some regions of the child's lexicon and phonological system as late as 7 years of age.

ACKNOWLEDGMENTS

Some of the work reported here was part of a thesis completed by the first author and submitted to the City University of New York in partial fulfillment of the requirements for the doctoral degree in Speech and Hearing Sciences. We are grateful to Arthur Boothroyd and Lawrence Raphael, who served as thesis advisors. We thank Kevin Munhall, Bruno Repp, and D. H. Whalen for suggestions regarding analysis procedures and for comments on earlier drafts of the paper. The instructive comments of Margaret Cheesman and Winifred Strange are also gratefully acknowledged. This work was supported by NIH Grants HD-01994 and NS-07237 to Haskins Laboratories.

REFERENCES

AMERICAN NATIONAL STANDARDS INSTITUTE. (1969). (ANSI S3.6-1969). *Specifications for audiometers*. New York: ANSI.

BICKLEY, C. (1986). Formant estimation of high fundamental frequency speech. *Journal of the Acoustical Society of America*, 79(Suppl. 1), S38(A).

BLUMSTEIN, S. E. (1981). Phonological aspects of aphasia. In M. T. Sarno (Ed.), *Acquired aphasia* (pp. 129–155). New York: Academic Press.

BROWMAN, C. P., & GOLDSTEIN, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*, 219–252.

COWAN, N., LEAVITT, L. A., MASSARO, D. W., & KENT, R. D. (1982). A fluent backward talker. *Journal of Speech and Hearing Research, 25*, 48–53.

FERGUSON, C. A. (1986). Discovering sound units and constructing sound systems: It's child play. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 36–51). Hillsdale, NJ: Lawrence Erlbaum.

FERGUSON, C. A., & FARWELL, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*, 419–439.

FORREST, K., WEISMER, G., & MILENKOVIC, P. (1987). Statistical representation of word-initial obstruents: Adult data. *Journal of the Acoustical Society of America, 81*(Suppl. 1), 578(A).

GILBERT, J. H. V., & PURVES, B. A. (1977). Temporal constraints on consonant clusters in child speech production. *Journal of Child Language, 4,* 417–432.

HEINZ, J. M., & STEVENS, K. N. (1961). On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America, 33,* 589–596.

KENT, R.D. (1983). The segmental organization of speech. In P. MacNeilage (Ed.), *The production of speech* (pp. 57–89). New York: Springer-Verlag.

KUHL, P.K. (1980). Perceptual constancy for speech-sound categories in early infancy. In G. H. Yeni-Konshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Vol. 2. Perception* (pp. 41–66). New York: Academic Press.

LINDBLOM, B., MACNEILAGE, P., & STUDDERT-KENNEDY, M. (1983). Self-organizing processes and the explanation of phonological universals. In B. Butterworth, B. Comrie, & O. Dahl (Eds.), *Language universals* (pp. 181–203). The Hague: Mouton.

MCGOWAN, R., & NITTROUER, S. (1988). Differences in fricative production between children and adults: Evidence from an acoustic analysis of /ʃ/ and /s/. *Journal of the Acoustical Society of America, 83,* 229–236.

MENN, L. (1983a, March). *Language acquisition, aphasia, and phonotactic universals.* Paper presented at the 12th Annual University of Wisconsin-Milwaukee Linguistics Symposium, Milwaukee, WI.

MENN, L. (1983b). Development of articulatory, phonetic, and phonological capabilities. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 3–50). London: Academic Press.

MENYUK, P., & KLATT, D. (1975). Voice onset time in cluster production by children and adults. *Journal of Child Language, 2,* 223–231.

MENYUK, P., & MENN, L. (1979). Early strategies for the perception and production of words and sounds. In P. Fletcher & M. Garman (Eds.), *Language acquisition* (pp. 49–70). Cambridge: Cambridge University Press.

NITTROUER, S., & STUDDERT-KENNEDY, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research, 30,* 319–329.

PETERSON, G., & BARNEY, H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24,* 175–184.

REPP, B. H. (1986). Some observations on the development of anticipatory coarticulation. *Journal of the Acoustical Society of America, 79,* 1616–1619.

SERENO, J. A., BAUM, S. R., MAREAN, G. C., & LIEBERMAN, P. (1987). Acoustic analyses and perceptual data on anticipatory labial coarticulation in adults and children. *Journal of the Acoustical Society of America, 81,* 512–519.

SHATTUCK-HUFNAGEL, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 109–136). New York: Springer-Verlag.

SOLI, S. D. (1981). Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *Journal of the Acoustical Society of America, 70,* 976–984.

STUDDERT-KENNEDY, M. (1986). Sources of variability in early speech development. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 58–76). Hillsdale, NJ: Erlbaum.

STUDDERT-KENNEDY, M. (1987). The phoneme as a perceptuomotor structure. In A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds), *Language perception and production* (pp. 67–84). London: Academic Press.

WATERSON, N. (1971). Child phonology: A prosodic view. *Journal of Linguistics, 7,* 179–211.

Received February 8, 1988
Accepted April 22, 1988

Requests for reprints should be sent to Susan Nittrouer, Boys Town National Institute for Communication Disorders in Children, 555 North 30th Street, Omaha, NE 68131.