689

The perceptual effects of child–adult differences in fricative-vowel coarticulation

Susan Nittrouer[a] and D. H. Whalen
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

(Received 9 September 1988; accepted for publication 7 June 1989)

Earlier work [Nittrouer *et al.*, J. Speech Hear. Res. **32**, 120–132 (1989)] demonstrated greater evidence of coarticulation in the fricative-vowel syllables of children than in those of adults when measured by anticipatory vowel effects on the resonant frequency of the fricative back cavity. In the present study, three experiments showed that this increased coarticulation led to improved vowel recognition from the fricative noise alone: Vowel identification by adult listeners was better overall for children's productions and was successful earlier in the fricative noise. This enhanced vowel recognition for children's samples was obtained in spite of the fact that children's and adults' samples were randomized together, therefore indicating that listeners were able to normalize the vowel information within a fricative noise where there often was acoustic evidence of only one formant associated primarily with the vowel. Correct vowel judgments were found to be largely independent of fricative identification. However, when another coarticulatory effect, the lowering of the main spectral prominence of the fricative noise for /u/ versus /i/, was taken into account, vowel judgments were found to interact with fricative identification. The results show that listeners are sensitive to the greater coarticulation in children's fricative-vowel syllables, and that, in some circumstances, they do not need to make a correct identification of the most prominently specified phone in order to make a correct identification of a coarticulated one.

PACS numbers: 43.71.Es, 43.70.Ep, 43.70.Fq

INTRODUCTION

Coarticulation, the temporal overlap of the production and acoustic realization of phonemes, helps explain the difficulty in finding acoustic invariance in the speech signal and the inappropriateness of taking acoustic segments to correspond to linguistic segments (Liberman *et al.*, 1967; Studdert-Kennedy, 1976, 1985; Liberman and Mattingly, 1985). Not only are these effects obvious and far-reaching in the acoustic record, but there is also ample evidence that listeners are sensitive to the information coarticulation provides. Of particular interest for the present work is the vowel information present in neighboring fricative noises (e.g., LaRiviere *et al.*, 1975; Soli, 1981; Yeni-Komshian and Soli, 1981; Whalen, 1983). For adult listeners, there is often enough vowel information in these fricative noises for reasonably accurate identification of the vowel from the fricative noise alone.

While coarticulation, and its perceptual effect, has been well studied for adults, only recently has it been examined developmentally (Nittrouer, 1985; Nittrouer and Studdert-Kennedy, 1987; Sereno *et al.*, 1987; Flege, 1988; Nittrouer *et al.*, 1989). The examination by Nittrouer *et al.* (1989) of fricative-vowel syllable production has revealed several interesting similarities and differences between children and adults. One coarticulatory effect, the co-occurrence of liprounding for the vowel with a preceding fricative gesture, was found to be the same for both groups of speakers. Positioning of the tongue during fricative production, however, was found to be more vowel dependent for the children than for the adults. As for the fricatives themselves, there was less differentiation of the fricatives /ʃ/ and /s/ for the children's productions than for the adults'. The first of these findings (i.e., that the magnitude of liprounding was similar for child and adult speakers) was reported also by Sereno *et al.* (1987). However, their results failed to reveal any difference between children and adults in the magnitude of lingual coarticulation. The reason for this discrepancy in findings may be that Sereno *et al.* (1987) limited the frequency region in which they looked for and measured the spectral consequence of lingual coarticulation, whereas Nittrouer *et al.* did not. Sereno *et al.* did not investigate the precision of children's fricative production relative to that of adults.

Two acoustic measurements led to the three major conclusions of Nittrouer *et al.* (1989). First, centroid measurements demonstrated greater spectral similarity for children's samples of /ʃ/ and /s/ noises than for adults'. The centroid, or first spectral moment, is a frequency mean of the whole spectrum, weighted by amplitude, and seems to be a good index of the distinctiveness of fricative production (Forrest *et al.*, 1988; Nittrouer *et al.*, 1989). Of particular concern to the present study, centroids appear to reflect reliably the frequency of the fricative spectral prominence, which is the formant (or formants) affiliated with the constriction and the cavity anterior to it. Nittrouer *et al.* found children's centroids for /s/ and /ʃ/ to be intermediate between adults' /s/ and /ʃ/ centroids, with both relatively closer to adults' /s/ values than to their /ʃ/ values.

Centroid measurements also showed that children and adults liprounded in anticipation of upcoming rounded vow-

[a] Present address: Boys Town National Institute for Communication Disorders in Children, 555 North 30th St., Omaha, NE 68131.

els to the same extent during fricative production. Liprounding lengthens the cavity anterior to the constriction, lowering centroid values. Similar shifts in these values were found for children and adults for syllables containing the vowel /u/ as compared to those containing the vowel /i/, and these shifts were greater for /s/ than for /ʃ/ noises across all age groups.

The finding of greater coarticulation of lingual vowel gestures for children's productions than for adults' was supported by measurements of fricative second formants ($F2$'s). Fricative $F2$ frequencies varied as a function of the upcoming vowel to a greater extent for children's samples than for adults'. Across speaker age, however, there was overlap in fricative F_2 frequencies between the two vowel categories: Children's samples for both /i/ and /u/ fell within or above the adult /i/ range. Because fricative $F2$ is affiliated with the cavity posterior to the fricative constriction, it primarily indexes the amount of anticipatory lingual posturing. Relatedly, McGowan and Nittrouer (1988) showed that the amplitude of fricative $F2$, relative to the higher-frequency components, was greater for children's samples. In fact, it was difficult to identify an $F2$ resonance at all in adults' samples until 30 ms before vowel onset (cf. Soli, 1981). Therefore, it was not possible to determine how much earlier, if at all, this difference in lingual coarticulation extended back into the fricative.

It is conceivable that the two differences between children's and adults' FV-syllable production described by Nittrouer *et al.* (1989) were related, especially because both involve tongue gestures. That is, either the greater anticipatory coarticulation of the vowel tongue–body gesture may have prohibited proper tongue–tip placement for the fricative, or the less-precise fricative gesture may have permitted greater assimilation of the vowel gesture. In either case, high negative correlations would be expected between measures of the magnitude of vowel coarticulation (i.e., fricative $F2$'s) and those of the precision of fricative production (i.e., centroids). However, Nittrouer *et al.* failed to find significant correlations between these measures. Thus the precision of fricative production and the magnitude of coarticulation were not related in individual speakers.

In the present study, we extended the work of Nittrouer *et al.* (1989) to answer three questions regarding the perception of coarticulatory effects in the FV syllables studied by them: (1) Do the enhanced coarticulatory effects of children's FV syllables provide usable perceptual information, or are they just noise? (2) If children's enhanced coarticulation is perceptible, do the perceptual tasks indicate that the age-related differences are greater in temporal extent or only in magnitude? (3) Are there interactions between fricative and vowel judgments when those judgments are based on the fricative noise alone?

The first question is of interest because the acoustic measurements of Nittrouer *et al.* (1989) not only indicated that coarticulation was greater for children's FV syllables, but also that the acoustic difference between the two fricatives studied (/s/ and /ʃ/) was less. Sereno *et al.* (1987) presented adults' and children's fricative noises in isolation and found that vowel judgments by adult listeners were more accurate for the adult talkers. Therefore, it may be that any vowel information provided by the fricative $F2$ is uninterpretable without correct identification of the fricative. If this is the case, we would expect identification of both the fricative and the vowel for children's fricative noises to be impaired. Conversely, if the greater acoustic evidence of vowel information in the fricative noise is, in fact, perceivable as vowel information, then vowel identification for children's samples should be better than for adults'. As noted above, decisions based on the $F2$ resonance will require a correct identification of the speaker as a child or as an adult because children's /u/ values are as high or higher than adults' /i/ values.

The second question has already received a partial answer in Nittrouer *et al.* (1989). They had difficulty even finding an $F2$ in some of the adult's productions, but found $F2$'s, and vowel-related frequency effects on these values, quite early in the fricative noises of children's samples. McGowan and Nittrouer (1988) also found, for a subset of the stimuli used in the present experiments, that the intensity differences between fricative $F2$'s in children's and adults' productions decreases from the beginning of the noise to vowel onset. Taken together, these results suggest that, at least for lingual gestures, children coarticulate more from the very onset of the syllable, and that age-related differences in production are less just before release of the fricative constriction. The effects of liprounding, however, have not been studied throughout the fricative noise, so it may be that there are differences at the beginning of the noise for this effect as well. Specifically, children may have coarticulated liprounding to the same degree as adults, but starting ealier.

The final question concerns sources of information used for making fricative and vowel judgments, and how these judgments interact. Two acoustic sources of information were of interest here: fricative $F2$ and the higher spectral prominence. Because an earlier study (Harris, 1958) demonstrated that judgments for /ʃ/ and /s/ are not strongly dependent on $F2$ frequency just after vowel onset, it is reasonable to think that these fricative judgments would not be strongly dependent on $F2$ frequency within the fricative noise itself. However, the vowel also affects the main spectral prominence of the fricative noise (Mann and Repp, 1980; Whalen 1981), and fricative judgments of the noise can affect how the vowel is perceived (Whalen, 1989). Two of the phones (/u/ and /ʃ/) are associated with lower frequency noise, while the other two (/i/ and /s/) are associated with higher frequency noise. When all four combinations of fricative and vowel are considered, we find that two combinations (/ʃi/ and /su/) allow for a trade between fricative and vowel judgments: Noises lacking extreme frequency values can be heard either as an /i/-colored /ʃ/ or as a /u/-colored /s/. Such a trade has been found for synthetic speech (Whalen, 1989), and so we will be looking for it in the responses to our truncated natural speech.

Each of these questions was studied by having adult listeners make phonetic judgments of syllable components based, often, on just a portion of the fricative noise. In general, we expected to confirm that children differ in their organization of speech gestures (Nittrouer *et al.*, 1989), but do so in a way that nonetheless leads to correct recovery of phonetic information by adult listeners. Our results will thus substantiate or refute the age-related acoustic and articulatory differences reported by Nittrouer *et al.* and by McGowan and Nittrouer (1988). They will begin also to tease apart the processes of recognizing coarticulated vowels and consonants.

I. GENERAL METHOD

A. Stimuli

All stimuli for the present set of experiments came from speech samples collected for Nittrouer *et al.* (1989), and additional details can be found there. Samples consisted of the disyllables /ʃiʃi/, /ʃuʃu/, /sisi/, and /susu/, spoken by eight adults (four female, four male) 20 or 21 years old, and eight children at each of the ages 3, 4, 5, and 7 years. No greater than a 3/5 ratio existed between the numbers of boys and girls in any of the children's age groups. All speakers had hearing thresholds better than 20 dB HL at each of the frequencies 500 Hz, and 1, 2, and 4 kHz, and had normal articulation for their ages, as judged independently by two speech pathologists. The samples 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." All pictures had been used with these subjects in an earlier perception experiment and so were familiar to them. Although a short practice session was provided in which speakers imitated the experimenter's model of producing the syllables at a constant rate with equal stress on each syllable, recorded samples were not imitated, but were instead spoken in response to the pictures. This method was chosen to ensure samples as close as possible to each speaker's normal production. Samples were collected in randomized groups of four, and ten samples of each syllable were obtained. Of these, the first eight samples with no extraneous noise were digitized. Thus, 32 samples from each speaker were available for use in the present study.

A Uher model 4200 portable tape recorder with an Electrovoice model 635A microphone was used to collect the samples. This system had a flat frequency response out to 10 kHz. The eight samples of each utterance per speaker were digitized onto a VAX 11/780 computer at a sampling rate of 20 kHz with low-pass filtering below 9.6 kHz. Using the Haskins Laboratories Waveform Editing and Display program (WENDY), the onsets of both vowels in all tokens were identified as the points at which periodicity first appeared in the waveform.

WENDY was also used to extract the short speech slices to be used in the present experiments. Which tokens were used in each experiment depended on the purpose of the individual experiment. However, slices for all three experiments were extracted from the first syllable only and were recorded on tape. No carrier phrase was used. The interstimulus interval (ISI) was always 2.5 s, with a 5-s pause after every ten tokens.

B. Listeners

All listeners in the present experiments were native speakers of American English between the ages of 20 and 40 years with no history of speech or hearing disorders and were naive to the purpose of the experiment in which they participated.

C. Procedure

Listeners sat in a quiet room and listened to stimuli at a comfortable level over TDH-39 headphones. Instructions were given verbally prior to the start of the experiment, with the response alternatives indicated to the subjects. All responses were written.

II. EXPERIMENT 1

This experiment was designed primarily to answer the first question posed in the Introduction: Did the enhanced coarticulatory effects found by Nittrouer *et al.* (1989) for children's FV syllables provide usable information? It also provided preliminary data regarding the other two questions.

A. Method

The speakers (one male and one female) in each age group from the Nittouer *et al.* (1989) study were chosen at random. Two tokens of each utterance from each of these speakers were selected randomly from the set of eight, and five slices were extracted from the first syllable of each: (1) the first half of the fricative (1/2F), (2) the first three quarters of the fricative (3/4F), (3) the entire fricative (allF), (4) the fricative, plus the first quarter of the vowel (F + 1/4V), and (5) the entire syllable (allFV).[1]

These 40 slices per speaker (4 utterances × 2 tokens × 5 slices) were recorded 6 times each in randomized blocks of 400 (40 slices × 10 speakers), giving a total of 2400 items. In order to determine whether or not fricative and vowel judgments interact to any extent, the five listeners participating in this experiment were asked to identify the syllable from which they thought the slice was extracted (i.e., "see," "she," "Sue," or "shoe"), not just the upcoming vowel. If listeners had not been required to make fricative judgments, or if fricative identity had been given to them, it would not have been possible to determine whether or not correct vowel judgments were obtained in the absence of correct fricative judgments. Also, it would not have been possible to substantiate or refute the claim of Nittrouer *et al.* (1989) that children were less precise than adults in fricative production.

B. Results

Analyses of variance were performed on the data from experiment 1 using speaker age, consonant, vowel, and slice as main effects, and adults versus children as a planned comparison.[2] For the analyses reported here, data were collapsed over repetitions of the same token, tokens of the same utterance, and speakers within an age group so that the degrees of freedom would not be artificially inflated. Scores were submitted to an arcsine transform because fricative judgments were nearly 100% correct.

1. Whole-syllable identification

Figure 1 shows the percentage of correct whole-syllable responses to each slice for each age group. Responses were

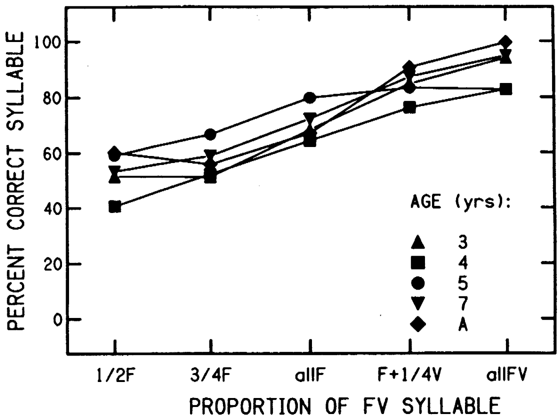

FIG. 1. Mean percent correct whole-syllable identification scores from experiment 1 for slices extracted from FV syllables spoken by children of four ages and by adults. Closed symbols indicate scores that were significantly above chance ($p < 0.05$).

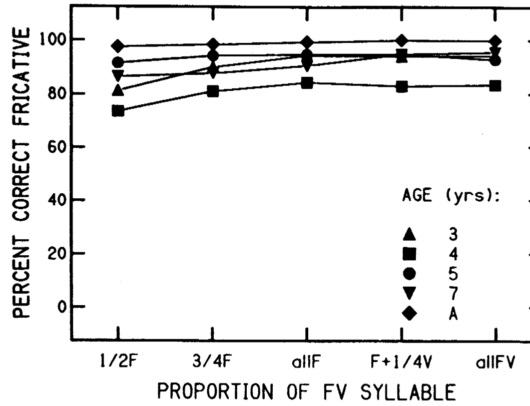

FIG. 2. Mean percent correct fricative identification scores from experiment 1 (see Fig. 1 legend for details).

significantly above chance, as indicated by closed symbols, for all slices for all age groups, with scores improving as slices increased in length. Scores were not statistically different for adult speakers than for child speakers for slices consisting only of fricative noise: The mean number of correct responses to these three slices was 61% for adult samples and ranged between 52% and 69% for samples from the children's groups. This nonsignificant finding suggests that the enhanced fricative information of the adults' samples over the children's samples and the enhanced anticipatory vowel information of the children's samples over the adults' contributed about equally to whole-syllable judgments made from the fricative noise alone. Statistically significant results were obtained for the main effects of fricative [$F(1,240) = 58.21$, $p < 0.01$], vowel [$F(1,240) = 7.31$, $p < 0.01$], and slice [$F(2,240) = 72.00, p < 0.01$], indicating that syllables were more accurately recognized when the fricative was an /s/ rather than an /ʃ/ (66% vs 55%), when the vowel was an /i/ rather than an /u/ (62% vs 58%), and when more of the fricative noise was presented (70% for allF vs 53% for 1/2F).

For the two slices containing some portion of the vocalic segment (F + 1/4V and allFV), responses to adults' samples were more accurate than responses to children's samples: The mean number of correct syllable judgments for these two slices was 95% for adults' samples and ranged between 79% and 91% for samples from children's groups. An analysis of variance of whole-syllable scores for these two longest slices showed a significant main effect of age [$F(4,160) = 24.34$, $p < 0.01$] and a significant result for the planned comparison of adults versus children [$F(1,160) = 58.22, p < 0.01$]. In addition, significant main effects were found for fricative [$F(1,160) = 44.66$, $p < 0.01$], vowel [$F(1,160) = 7.50$, $p < 0.01$], and slice [$F(1,160) = 52.29, p < 0.01$]. Again, more accurate syllable judgments were obtained for tokens with the fricative /s/ (91% vs 84% for /ʃ/) and with the vowel /i/ (88% vs 87% for /u/), and for longer slices (91% for allFV vs 84% for F + 1/4V).

2. Fricative identification

Figure 2 displays the percentages of correct fricative responses (i.e., whole syllable correct plus fricative only correct). Again, all scores were clearly well above chance, but here it can be seen that responses to adults' slices were always superior to those for children's slices: Averaging across all slices, the mean number of correct fricative judgments was 99% for adults' samples and ranged between 81% and 94% for samples from children's groups. An analysis of variance performed on fricative scores for all five syllable slices demonstrated that the main effect of age [$F(4,400) = 64.32$, $p < 0.01$] and the planned comparison of adults versus children [$F(1,400) = 182.01$, $p < 0.01$] were both significant. In addition, the main effects of fricative [$F(1,400) = 68.57$, $p < 0.01$], vowel [$F(1,400) = 32.25$, $p < 0.01$], and slice [$F(4,400) = 18.18$, $p < 0.01$] were all found to be significant, reflecting more accurate fricative judgments for /s/ (95% vs 87% for /ʃ/), for the /u/ context (94% vs 89% for the /i/ context), and for longer slices (93% for allFV vs 86% for 1/2F).

3. Vowel identification

In Fig. 3, the percentages of correct vowel responses (i.e., whole syllable correct plus vowel only correct) are shown, all of which are significantly above chance. It can be seen that, for slices consisting only of fricative noise, responses to children's tokens were generally more accurate than responses to adult's tokens. For slices consisting of 3/4 of the fricative noise, only responses to 3-year-olds' samples were as poor as those to adults' samples. For slices consisting of the whole fricative, responses were better to samples from all children's groups than to those from adults: Percentage of vowels correctly recognized was 67% for adults and 78% for children, averaged over all four age groups. An analysis of variance performed on vowel identification scores for the three slices consisting only of fricative noise demonstrated significant results for the main effect of age [$F(4,240) = 9.25, p < 0.01$] and the planned comparison of adults versus children [$F(1,240) = 13.43$, $p < 0.01$]. Also found to be significant were the main effects of fricative [$F(1,240) = 8.04$, $p < 0.01$], vowel [$F(1,240) = 75.12$,

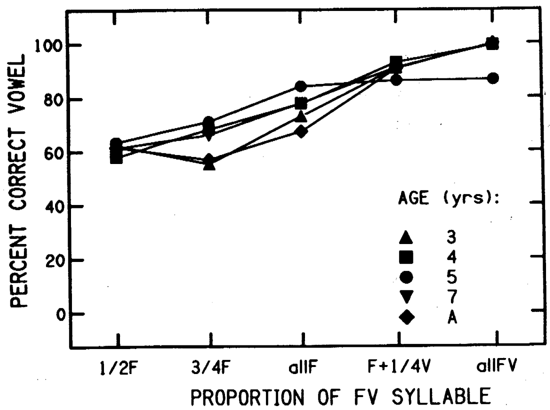

FIG. 3. Mean percent correct vowel identification scores from experiment 1 (see Fig. 1 legend for details).

$p<0.01$], and slice [$F(2,240) = 79.53, p<0.01$]. These significant effects indicate that listeners more often identified vowel context correctly for /s/ than for /ʃ/ noises (69% vs 65%), for the /i/ than for the /u/ context (72% vs 61%), and for longer than for shorter slices (76% correct for allF vs 61% for 1/2F).

For the two slices that contained some part of the vocalic segment, the greater accuracy of vowel judgments for children's samples disappears. Obviously, these slices provide much more information about vowel identity than the three slices consisting only of fricative noise. Puzzlingly, identification of the 5-year olds' vowels did not improve with the addition of the vocalic segment.

C. Discussion

This first experiment showed that adult listeners were indeed able to interpret correctly the vowel information in the fricative noises of children's productions. This was true despite the fact that identification of the children's fricatives was consistently worse than that of the adults'. Identification of the vowel was generally better for children's productions when only the fricative noise was given, but was of essentially the same accuracy for adults' and children's samples when some or all of the vocalic segment was present. This confirms the acoustic results (Nittrouer *et al.*, 1989) that indicated greater anticipatory coarticulation in children's FV syllables than in adults'. In addition, the lower identification scores for children's fricatives support the claim of less distinct fricative production in children's FV syllables. The somewhat greater accuracy for recognition of /s/ than of /ʃ/ may reflect the fact that, although children's centroids for both /s/ and /ʃ/ were intermediate between adults' /s/ and /ʃ/ values, children's /ʃ/ centroids showed disproportionately higher values. However, it would be inappropriate to place too much importance on these significant fricative effects because, as will be seen, they were not found in experiment 2.

The basis for the correct vowel judgments must be inferred, as we do not have any direct measure of it. For the adult productions, especially for judgments on the first half of the fricative noise, the effect of the vowel on the main spectral prominence of the fricative noise is the likeliest source. Adults show very little, if any, trace of $F2$ that early in the noise, yet vowel judgments were above chance even for that slice. The finding of slightly more accurate vowel recognition for /s/ noises than for /ʃ/ noises also suggests that listeners could have used the vowel-related shifts in the fricative spectral prominence: These shifts were greater for /s/ than for /ʃ/ (Nittrouer *et al.*, 1989). Of course, use of these shifts in the fricative spectral prominence to correctly identify vowel context required correct fricative identification, which was not always the case for children's samples. For the children's productions, it seems more likely that $F2$ was used throughout, given that the fricative was often misidentified. Even if the fricative is misidentified, $F2$ provides fairly accurate information about vowel context. Thus the finding that listeners could correctly identify the vowel at all for adults' samples coupled with the difference in identification accuracy between adults' and children's productions indicates that both sources of vowel information were used.

In interpreting the $F2$ resonance in the fricative noise, our listeners must have been doing some type of normalization. This is because the absolute values for the $F2$ in children's fricative noises are consistently higher than those of adults', regardless of vowel category. Unless vocal tract size was taken into account, an $F2$ from a child's production would always be identified as an /i/ resonance. Thus, even when only a fricative noise is present, listeners correctly account for vocal tract size, despite the absence of a clear formant structure.

It could be argued that the finding of a significant main effect of vowel identity on vowel judgments suggests this very bias of perceiving all children's samples as coming from an /i/ context. Since there were four times as many children's as adults' samples in this experiment, it may actually account for some portion of this effect. However, if it were completely responsible for the enhanced recognition of /i/ over /u/, then vowel recognition for /u/ would have been significantly below chance, which it was not. More likely, because fricative $F2$ is higher in frequency preceding /i/ than preceding /u/, it is closer to the main spectral prominence of the fricative and so is also greater in relative amplitude. Therefore, fricative $F2$ preceding /i/ is more easily accessible to listeners.

III. EXPERIMENT 2

Having answered the first question we posed in the Introduction with experiment 1, we designed experiment 2 to investigate more carefully the next two questions concerning the time course of children's enhanced coarticulation and the interaction of fricative and vowel judgments. Although some indication of temporal effects had been found in the first experiment, overall identification scores were too high to see specifically what those effects were. The fact that overall identification scores were so high also prohibited close inspection of the interaction between fricative and vowel judgments. Moreover, the fact that the duration of slices varied as a function of the length of the fricative noise could have been a confounding factor. We overcame these difficulties in the second experiment by presenting very brief slices of the same absolute duration from different locations within the fricative noise.

In addition, we wanted to examine more closely the relation between the perceptual results and earlier acoustic findings in this second experiment. Although we had found a relation in the first experiment between speaker age and identification scores, both for fricative and for vowel, we were not able to discern exactly how these perceptual differences were related to the age-dependent acoustic (and therefore, articulatory) differences reported by Nittrouer *et al.* (1989) and by McGowan and Nittrouer (1988). One reason for this shortcoming was the selection of speakers was random, and, as happens with random selection, some of the age groups were represented by less-adept speakers. In the present experiment, then, we selected children on the basis of the acoustic measure of fricative differentiation, which experiment 1 showed to be a possible factor in the perception of the fricative, since children's fricatives were less accurately perceived.

A. Method

Three groups of speakers were chosen from those participating in Nittrouer *et al.* (1989). One group consisted of four adults (two male and two female), selected randomly. A second group consisted of four children, one child of each of the ages sampled (3, 4, 5, and 7 years) who demonstrated less than a 1-kHz difference in mean centroids between /ʃ/ and /s/. Because this lack of spectral differentiation between fricatives was typical of children's samples, this was labeled the "childlike" group. The third speaker group consisted of two children (4 and 5 years old) who demonstrated greater than a 1-kHz difference between /ʃ/ and /s/ centroids.[3] Because this pattern was more similar to adult production, this was labeled the "adultlike" group. The two groups of children did not differ from each other in either the extent to which they coproduced the following vowel with the fricative or the amplitude of fricative $F2$ relative to the higher formants. Thus samples from the adultlike children were similar to those from the adults on the extent to which the fricative spectral prominence provided information about fricative identity, but were similar to samples from the childlike children on the extent to which $F2$ provided information about vowel identity.

All eight tokens of each utterance from each speaker were used. Slices 60 ms in duration were extracted beginning just before vowel onset, and then working backwards into the fricative in 40-ms steps until the beginning was reached. Between two and seven slices were extracted from each fricative (i.e., fricatives were between 100 and 300 ms in duration). Children demonstrated more variability in fricative duration than adults, who consistently produced fricatives between 150 and 220 ms long. Extracted slices were randomized, and a test tape consisting of 1273 slices of fricative noise was made. Eight new listeners participated in this experiment, and their task was to write the syllable, both fricative and vowel, from which they thought the slice was extracted.

B. Results

Scores were collapsed over tokens of the same slice (i.e., taken from the same location) within the same utterance and over speakers within a group as a way of restricting the number of degrees of freedom. These values were entered into analyses of variance with the main effects of speaker group, fricative, vowel, and slice. In addition, orthogonal-planned comparisons were performed for adults versus children and adultlike versus childlike children. Arcsine transforms were not used in this experiment because identification scores were not close to either 100% or 0%.

While all slices extracted from each token were presented in the perception experiment, only responses to the five slices closest to vowel onset (− 220 to − 60 ms from vowel onset) were included in the analyses. Informal inspection for the few slices earlier than those revealed little difference between them and the − 220-ms slice. Because their inclusion would have greatly increased the number of empty cells in the analysis, they were, instead, excluded altogether.

1. Whole-syllable identification

Figure 4 shows scores for whole-syllable identification. All were significantly above chance, as indicated by the closed symbols, but there were differences among speaker groups. Although listeners were able to identify the syllable from which slices were extracted with equal accuracy for samples from the adultlike speakers and the adults (45% correct across all slices for each group), they were less accurate in syllable identification for slices from the childlike speakers (36% correct). This difference shows up in the statistical analysis as a significant main effect for speaker group [$F(2,420) = 14.04, p < 0.01$] and as a significant result for the comparison between childlike and adultlike children [$F(1,420) = 20.93, p < 0.01$]. It is not clear why a difference in syllable identification was found in this experiment for fricative noises from adults and from one group of children when no difference was found for adult and child speakers in experiment 1. It could be that enough children in the previous experiment differentiated between /ʃ/ and /s/ in an adultlike manner that they could be consid-

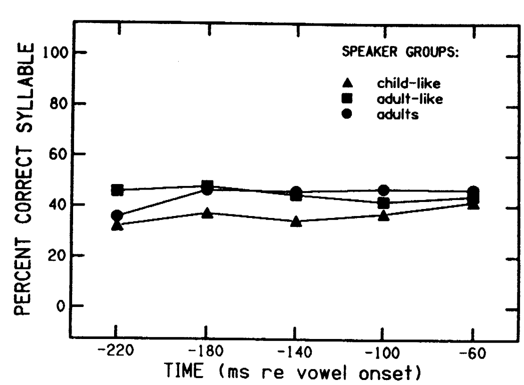

FIG. 4. Mean percent correct syllable identification scores from experiment 2 for slices of fricative noise extracted from *FV* syllables spoken by children and adults. Closed symbols indicate scores that were significantly above chance ($p < 0.05$).

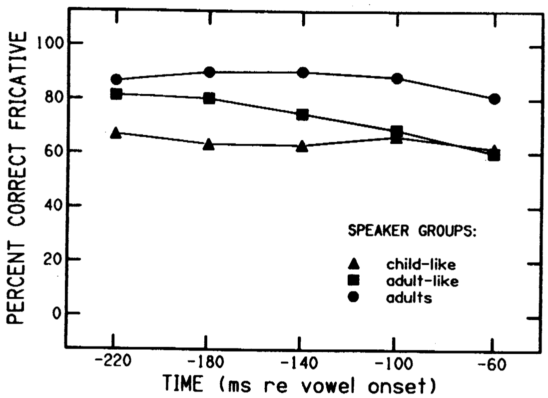

FIG. 5. Mean percent correct fricative identification scores from experiment 2 (see Fig. 4 legend for details).

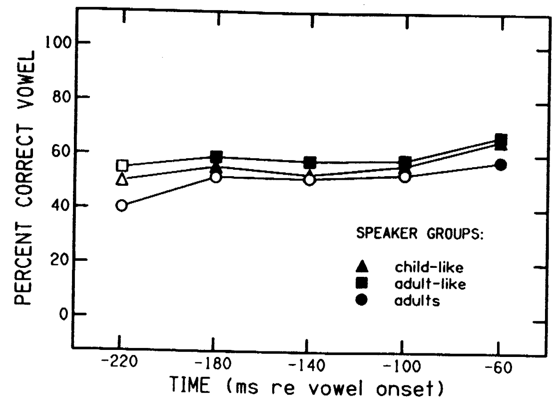

FIG. 6. Mean percent correct vowel identification scores from experiment 2 (see Fig. 4 legend for details).

ered adultlike speakers. However, the more likely explanation for the discrepancy in results between the two experiments is that the use of very brief slices in the present experiment made the task more difficult and so more sensitive to certain listener effects. This explanation is supported by the fact that, for the briefest slices in experiment 1, syllable identification scores appear to be more accurate for slices from adults' samples than for those from any of the children's groups (see the 1/2F data of Fig. 1). The only other effect to demonstrate statistical significance for syllable identification scores in experiment 2 was the main effect of vowel [$F(1,420) = 13.76$, $p < 0.01$], reflecting the trend also found in experiment 1 of more accurate syllable recognition in the /i/ context than in the /u/ context (45% vs 39%).

2. Fricative identification

Figure 5 shows scores for fricative identification. All were significantly above chance, as indicated by the closed symbols. Listeners had the easiest time recognizing the fricative for adults, followed by adultlike children, and lastly by childlike children, and this effect of speaker group was statistically significant [$F(2,420) = 81.59, p < 0.01$], as were the comparisons of adults versus children [$F(1,420) = 137.75$, $p < 0.01$], and adultlike versus childlike children [$F(1,420) = 25.44$, $p < 0.01$]. Furthermore, overall fricative judgments decreased in accuracy as vowel onset was approached, as indicated by a significant main effect of slice [$F(4,420) = 6.48$, $p < 0.01$]. Neither the main effect of fricative nor that of vowel was found to be significant.

3. Vowel identification

Figure 6 shows scores for vowel identification. Only scores for those slices closest to vowel onset were significantly above chance for all speaker groups. Identification scores at this location were 57% for adults, 67% for adultlike children, and 65% for childlike children. So, the 10% difference in vowel identification scores found in experiment 1 is replicated here. In addition, scores for the childlike group for slices at − 180 and − 100 ms before vowel onset were significantly above chance, as were scores for slices between − 180 and − 100 ms of vowel onset for the adultlike group. This difference in vowel identification scores for speaker groups was statistically significant [$F(2,420) = 7.91$, $p < 0.01$], as was the comparison of adults versus children [$F(1,420) = 9.69, p < 0.01$]. The main effect of how close to vowel onset the slice was located was also significant [$F(4,420) = 7.12$, $p < 0.01$]. Averaging across speaker groups, 50% of all vowel judgments were accurate for slices starting at − 220 ms from vowel onset, while 63% of these judgments were accurate for slices starting at − 60 ms. As found in experiment 1, vowel judgments for /i/ were more accurate than those for /u/ (60% vs 51%), and this difference was statistically significant [$F(1,420) = 28.20$, $p < 0.01$]. No difference was found between /s/ and /ʃ/ for accuracy of vowel judgments.

C. Discussion

In general, the results of experiment 2 support the interpretations of experiment 1. Fricative identity was easier to recover from adults' productions than from children's, but vowel context was recovered both more reliably and earlier from children's. The difference in fricative identification between the adultlike and the childlike speakers confirms that the acoustic difference found previously (Nittrouer *et al.* 1989) has a straightforward effect on the perception of the fricatives. For vowel identification, samples from the adultlike children were not adultlike, but rather demonstrated more accurate judgments beginning earlier in the fricative noise, as did samples from the childlike group. This finding makes it even more likely that the greater accuracy for vowel judgments was due to the $F2$ resonance, and not to the fricative spectral prominence. On the acoustic measures of fricative $F2$, the two groups of children were similar.

From the results of experiment 2, it seems that fricative information decreases over time: There was a general decline in fricative perceptibility as vowel onset was approached. The vowel information shows almost the reverse pattern. While the earliest time slice showed no significant vowel identification for any age group, the next three time slices showed a slight trend for increasing accuracy, and on the last time slice, vowel identification scores improved for all three groups (to being the only slice from the adult speakers to

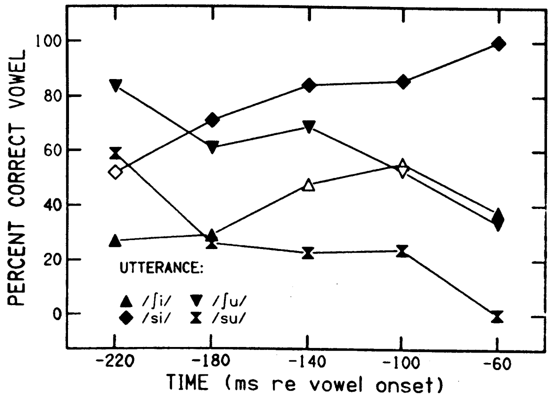

FIG. 7. Mean percent correct vowel identification when the fricative was misidentified for samples from adult speakers. Data are from experiment 2, and scores significantly different from chance are indicated by closed symbols.

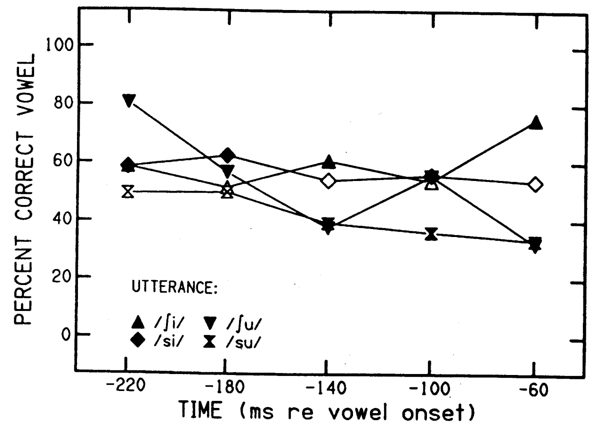

FIG. 9. Mean percent correct vowel identification when the fricative was misidentified for samples from childlike speakers (see Fig. 7 legend for details).

elicit above-chance performance). These temporal effects of decreasing accuracy for fricative judgments with increasing accuracy for vowel judgments seem to be due to changes in posturing of the tongue, both in preparation for vowel production and in termination of fricative production. Specifically, the tongue body takes on more of the position for the vowel, and the tongue-tip constriction is widened. This lowers the relative amplitude of the fricative spectral prominence while increasing the relative amplitude of the $F2$ resonance.

As mentioned previously, correct identification of the vowel for children's productions was found in the face of misidentifications of the fricative. For the most part, then, the two decisions were independent of each other. Yet the effect of the vowel on the main spectral prominence of the noise is such that an interaction between the decisions might be expected. The main prominence of the noise is lower in frequency before /u/ than before/i/, and /ʃ/ noises are lower in frequency than are /s/ noises. If, then, fricative and anticipatory vowel judgments made from the acoustic slices presented here were to any extent interdependent, the following pattern of predictions could be made: When the fricatives of /ʃu/ or /si/ were misidentified, the probability of correct vowel judgments would have increased. When the fricatives of /ʃi/ or /su/ were misidentified, the probability of correct vowel judgments would have decreased. That is, a low noise from /ʃ/ heard as /s/ should have led to /u/ vowel judgments, while a high noise from /s/ heard as /ʃ/ should have led to /i/ vowel judgments. Whether these judgments were correct or not depended, of course, on what the original context actually was

Figures 7–9 display the percent correct vowel identification for all instances of incorrect fricative identification. Responses to each syllable type are shown separately. Closed symbols indicate that performance was significantly greater or lesser than chance. Adult speakers (Fig. 7) and adultlike speakers (Fig. 8) show good agreement with the predictions. Of the 40 possible comparisons, 31 are significantly different from chance, and 28 of the 31 are in the predicted direction: Scores for /ʃi/ and /su/ are significantly poorer than chance, while scores for /ʃu/ and /si/ are significantly better. Thus there is substantial evidence for an interaction between the vowel and the fricative judgments. Even with the childlike speakers (Fig. 9), 9 of 14 significant comparisons are in the predicted direction, although not as strongly.

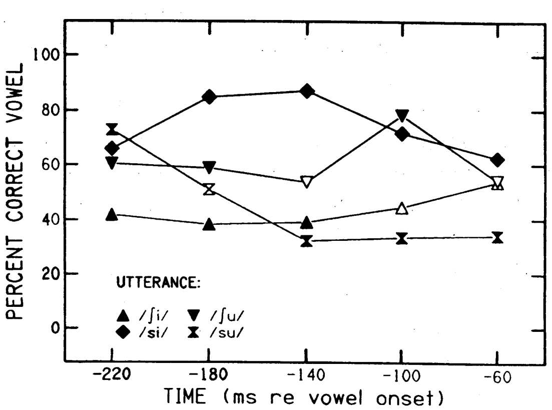

FIG. 8. Mean percent correct vowel identification when the fricative was misidentified for samples from adultlike speakers (see Fig. 7 legend for details).

IV. EXPERIMENT 3

Having established in experiment 2 that there was an interaction between fricative and vowel judgments, experiment 3 was designed to investigate the magnitude of this interaction. To do this, we identified the fricative to listeners. For those vowel judgments based on the $F2$ resonance, this should have little effect because these judgments were largely independent of fricative identification to begin with. However, the interactions found in experiment 2 lead us to predict that those syllables which demonstrated a relation between fricative and vowel judgments should show substantial change in accuracy when fricative identity is known.

A. Method

Two child speakers of each age group (one male and one female) and four adult speakers (two male and two female) were randomly selected. Children were not distinguished based on centroid differences for /ʃ/ and /s/ noises because listeners would be told the identity of the fricative. Slices 90 ms in length were extracted from each of the 32 tokens of each speaker, starting 120 ms before vowel onset. Therefore, slices ended 30 ms before vowel onset, the point at which vowel-related changes in fricative $F2$ are first observed in adult fricative spectra. Separate audiotapes were made for /ʃ/ and for /s/ slices. Stimuli were blocked by speaker, with order of blocks randomized and four practice items starting each block (two slices from the /i/ context, and two from the /u/ context), resulting in tapes consisting of 240 items each (12 speakers$\times 20$ items). Although the two categories of fricatives were presented separately and stimuli were blocked by speaker, listeners were not told whether speakers were children or adults. The order of presentation of the blocks (and therefore of speakers) was randomized, so that adult and child utterances occurred throughout the experiment.

Eight listeners participated in this experiment, half hearing the /ʃ/ tape first, and half hearing the /s/ tape first. The task was simply to write what they thought the vowel context was.

B. Results and discussion

Figure 10 shows the percent correct vowel identifications for each of the four syllables, divided by speaker group. Superimposed on each column are the most comparable values from experiment 2, which are the correct vowel identification scores for the − 100-ms stimuli. For the child speakers, these scores were averaged across the two types of child speakers of experiment 2 (since we did not distinguish between them in this experiment). Collapsing across syllable type, vowel identification scores in this third experiment were once again more accurate for child (69%) than for adult (60%) speakers [$F(1,7) = 38.30, p < 0.01$]. The size of the difference is almost exactly that found in the first two experiments, about 10%, adding further support to the interpretation of this age effect as due to enhanced $F2$ information for children's samples.

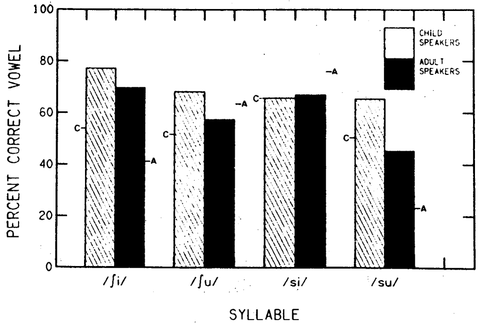

FIG. 10. Mean percent correct vowel identification from experiment 3, in which the identity of the fricative was known. Tic marks next to each column represent the corresponding value from experiment 2, in which identity of the fricative was not known.

For those syllables which showed an interaction between fricative and vowel judgments in experiment 2 (that is, /ʃi/ and /su/, which are on the extreme left and right, respectively, in Fig. 10), knowing the fricative identity resulted in an average gain in vowel identification scores of 22 percentage points for adults' samples and 17 percentage points for children's samples. For the syllables that did not demonstrate an interaction between fricative and vowel judgments (that is, /ʃu/ and /si/), knowing the identity of the fricative contributed little: Scores for adults' samples actually decreased by 7.5 percentage points, while scores for children's samples improved by 5 percentage points. Thus, in just those cases where removing the source of interaction should have improved performance, performance was improved. Moreover, vowel judgments for all children's samples improved to some extent when fricative identity was given. Apparently in the earlier two experiments, the decreased ability of listeners to accurately judge fricative identity for children's samples actually degraded vowel identification scores somewhat.

V. GENERAL DISCUSSION

On the whole, we have found the following answers to the questions posed in the Introduction: (1) The greater coarticulation between vowel and fricative gestures reported by Nittrouer *et al.* (1989) for children's speech is indeed perceivable by adult listeners as phonetic information, and not just as noise. (2) The enhanced coarticulation of children's FV syllables is greater in temporal extent, as well as in magnitude. (3) Perception of a coarticulated phone can include both processes that are dependent on recognition of the prominent phone, as well as processes that are independent of recognition of that phone.

These conclusions serve to support and extend the findings of Nittrouer *et al.* (1989), and to provide new insights into the processes of speech perception. With children's speech, listeners were able to separate the enhanced vowel information from the degraded consonant information. It cannot be the case that listeners heard an /ʃ/ or an /s/ and then determined vowel context from the allophonic variation of this phonetic segment because, for children's tokens, listeners were often unable to determine whether the fricative was an /ʃ/ or an /s/.

At the same time though, evidence was also found for the idea that phonetic judgments of adjacent segments interact under certain conditions: specifically, in those cases where both segmental judgments are based on the same acoustic information. In fact, other work by one of us (Whalen, 1989) reached the same conclusion. In that study, listeners were asked to identify both the fricative and the vowel of synthetic fricative-vowel syllables. The frequencies of the fricative spectral prominences were deliberately chosen to be ambiguous between /ʃ/ and /s/, and the vocalic formants were chosen to be ambiguous between /i/ and /u/. The fricative noises had no $F2$, and the vocalic portions had no formant transitions. Consequently, it was impossible to use the fricative $F2$ to identify the upcoming vowel, or to use the vocalic $F2$ to identify the preceding fricative. When a fricative noise was identified as /ʃ/, vowel judgments were more often /i/ than when the same noise was identified as /s/. Listeners must have used the fricative spectral prominence to make judgments about both fricative and vowel identity. Thus, the use of the same acoustic information to make phonetic judgments about adjacent segments apparently dictates that these judgments be dependent (i.e., that they interact), while the use of different acoustic information allows those judgments to be independent.

The present results are at variance with those of Sereno *et al.* (1987), who found no evidence of greater coarticulation in the productions of children. Several reasons for the difference on the acoustic side have been given in Nittrouer *et al.* (1989), as well as in the Introduction here. On the perceptual side, though, we now have an alternative explanation for at least some of their results. In Sereno *et al.*, the listeners were required to give only a vowel judgment. Thus, for the fricative noises, we do not know from their data whether the same type of interaction that was found in the present study also occurred, namely, that their listeners in fact misheard the fricative in some cases and therefore misheard the vowel as well. Whether or not this happened, there is a further difference between the two studies, namely, that Sereno *et al.* grouped the stimuli of adult and child speakers separately, while ours were always randomized together. Furthermore, the adult stimuli were always presented first in Sereno *et al.*, perhaps leading listeners to adopt a successful strategy for adult sounds that subsequently failed with children's productions. In any case, the issue is of enough importance that we can expect other studies to further illuminate these differences.

One additional result of interest was observed in the present data: The $F2$ frequency *per se* did not determine perceptual judgments. Nittrouer *et al.* (1989) showed that the absolute values of children's fricative $F2$ frequencies were always higher than those of adults. If the identification of vowel context was dependent on this acoustic parameter, all slices from children's samples would have been identified as coming from the /i/ context. The fact that this result was not observed indicates that listeners do not perform direct acoustic-to-phonetic mappings. Neither could listeners have been doing some sort of acoustic normalization of the kind proposed by various authors (e.g., Assman *et al.*, 1982; Gerstman, 1968; Lieberman *et al.*, 1972; Nearey, 1978; Strange *et al.*, 1976; Syrdal and Gopal, 1986). The parameters considered necessary for vocal tract normalization by these proposed algorithms were not usually available in these experiments: Only one of the first two formants was present; $F2$ frequencies for only two vowels per speaker were present; stimuli were not presented in a carrier phrase; many stimuli were extremely brief in duration and often did not include dynamic information; and usually the stimuli were not blocked by speaker. Instead, the acoustics must provide information about something else. Direct accounts of speech perception (Fowler, 1984, 1986) hold that the acoustic signal directly specifies articulatory events, and all results from the present set of experiments are in keeping with this idea: When gestures were poorly produced, as Nittrouer *et al.* (1989) found for constriction shape of children's fricatives, perceptual results were poor, when gestures were produced well, as was found to be the case for the enhanced vowel tongue gestures within the children's fricatives, perceptual results were good. Thus, within these extemely brief acoustic slices, some information was apparently available about the size of the vocal tract generally and about the vowel-related posture of the tongue specifically.

Finally, it seems that the articulatory events perceived are not simply articulatory movements, such as liprounding, but are more global characteristics of the vocal tract, such as volume of the cavity anterior to a fricative constriction. Moreover, when different articulatory gestures affect the same global vocal tract parameter, listeners apparently cannot separate these gestures; however, when different parameters are affected, listeners can. In the present study, both anticipatory liprounding and characteristics of the fricative constriction affected the volume of the cavity anterior to the fricative constriction, and so the frequency of the fricative spectral prominence. Listeners seemed unable to judge how cavity volume, as indicated by the frequency of this prominence, was being affected separately by constriction and by liprounding characteristics. Consequently, if an erroneous decision was made concerning one, an erroneous conclusion was often reached concerning the other. On the other hand, when different articulatory events primarily affected different aspects of the acoustic signal, then listeners apparently could separate these events. So, in the present study, anticipatory lingual gestures affected fricative $F2$, which is not as important a cue, if one at all, to fricative identity. Listeners could, to a large extent, separate this information concerning vowel identity from information associated with fricative identity. It seems that when simultaneous articulatory events shape different portions of the acoustic signal, listeners divide the signal into temporally overlapping but qualitatively separate units, about which they make independent phonetic judgments. However, when simultaneous articulatory events shape the same part of the acoustic signal, listeners do not divide the signal into qualitatively separate units, but instead make dependent phonetic judgments of each.

In conclusion, we have demonstrated that the enhanced anticipatory vowel coarticulation previously found in children's fricative-vowel syllables is perceptible. Furthermore, children's coarticulation is not only greater in magnitude, but in extent as well. Listeners are able to separate information about this enhanced vowel coarticulation from that of fricative production for children's speech because, to some extent, the two articulatory events affect different acoustic sources of information. However, it was shown also that judgments of the two phones interact when they are based on the same acoustic information.

ACKNOWLEDGMENTS

This work was completed while the first author was a postdoctoral fellow at Haskins Laboratories, supported by NIH Grant NS-07237 to Haskins Laboratories. It was also supported by NIH Grant HD-01994 to Haskins Laboratories. Manuscript preparation was partially supported by NIH Grant NS27433 to the first author. We thank Carol Fowler, John Mott, Donna Neff, Michael Studdert-Kennedy, and two anonymous reviewers for their helpful comments.

[1]Mean syllable and fricative-noise durations were not significantly different across age groups (mean syllable duration = 298 ms; mean fricative-noise duration = 166 ms). However, children demonstrated more variability in these durations, both across children and among tokens from individual children.

[2]This planned comparison was done because the demonstration of a significant age effect would show only that there was a difference among scores for some of the speaker groups; it would not isolate this difference to one between children and adults. Conversely, the lack of a significant age effect would not necessarily ensure that there was no difference between the four children's groups and the one adult group.

[3]Ideally, one child from each of the four age groups demonstrating greater than a 1-kHz difference in mean centroids between /ʃ/ and /s/ would have been selected, but software constraints limited us to ten speakers in experiment 2. There was an absolute upper limit on the number of slices that could be included, and the large number of slices extracted from each individual token quickly brought us near this limit. We wanted to include four adult speakers to avoid the great inequality between the numbers of adult and child speakers that existed in the first experiment; so we were limited to six child speakers. We decided that it would be most reasonable to select the larger number of child speakers who demonstrated a small difference in centroids between /ʃ/ and /s/ because all adult speakers demonstrated greater than a 1-kHz difference, and the possibility existed that children demonstrating a large centroid difference would perform identically to adults.

Assman, P. F., Nearey, T. M., and Hogan, J. T. (1982). "Vowel identification: Orthographic, perceptual, and acoustic aspects," J. Acoust. Soc. Am. 71, 975–989.

Flege, J. E. (1988). "Anticipatory and carry-over nasal coarticulation in the speech of children and adults," J. Speech Hear. Res. 31, 525–536.

Forrest, K., Weismer, G., Milenkovic, P. M., and Dougall, R. N. (1988). "Statistical analysis of word-initial voiceless obstruents: Preliminary data," J. Acoust. Soc. Am. 84, 115–123.

Fowler, C. A. (1984). "Segmentation of coarticulated speech in perception," Percept. Psychophys. 36, 359–368.

Fowler, C. A. (1986). "An event approach to the study of speech perception from a direct-realist perspective," J. Phon. 14, 3–28.

Gerstman, L. J. (1968). "Classification of self-normalized vowels," IEEE Trans. Audio Electroacoust. AV-16, 78–80.

Harris, K. S. (1958). "Cues for the discrimination of American English fricatives spoken in syllables," Language Speech 1, 1–7.

LaRiviere, C. J., Winitz, H., and Herriman, E. (1975). "The distribution of perceptual cues in English prevocalic fricatives," J. Speech Hear. Res. 18, 613–622.

Liberman, A. M., Cooper, F. S., Shankweiler, D. S., and Studdert-Kennedy, M. (1967). "Perception of the speech code," Psychol. Rev. 74, 431–461.

Liberman, A. M., and Mattingly, I. G. (1985). "The motor theory of speech perception revised," Cognition 21, 1–36.

Lieberman, P., Crelin, E., and Klatt, D. (1972). "Phonetic ability and the related anatomy of the newborn, adult human, Neanderthal man and the chimpanzee," Am. Anthropol. 74, 287–307.

Mann, V. A., and Repp, B. H. (1980). "Influence of vocalic context on perception of the [ʃ]-[s] distinction," Percept. Psychophys. 28, 213–228.

McGowan, R. S., and Nittrouer, S. (1988). "Differences in fricative production between children and adults: Evidence from an acoustic analysis of /ʃ/ and /s/," J. Acoust. Soc. Am. 83, 229–236.

Nearey, T. (1978), "Phonetic feature systems for vowels," Ph.D. thesis, Indiana University Linguistics Club, Bloomington, IN.

Nittrouer, S. (1985). "The role of coarticulation in the perception and production of speech by young children (3 to 7 years)," unpublished Ph.D. dissertation, City University of New York.

Nittrouer, S., and Studdert-Kennedy, M. (1987). "The role of coarticulatory effects in the perception of fricatives by children and adults," J. Speech Hear. Res. 30, 319–329.

Nittrouer, S., Studdert-Kennedy, M., and McGowan, R. S. (1989). "The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults," J. Speech Hear. Res. 32, 120–132.

Sereno, J. A., Baum, S. R., Marean, G. C., and Lieberman, P. (1987). "Acoustic analyses and perceptual data on anticipatory labial coarticulation in adults and children, "J. Acoust. Soc. Am. 81, 512–519.

Soli, S. D. (1981). "Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation," J. Acoust. Soc. Am. 70, 976–984.

Strange, W., Verbrugge, R., Shankweiler, D., and Edman, T. (1976). "Consonant context specifies vowel identity," J. Acoust. Soc. Am. 60, 213–224.

Studdert-Kennedy, M. (1976). "Speech perception," in *Contemporary Issues in Experimental Phonetics*, edited by N. J. Lass (Academic, New York), pp. 243–293.

Studdert-Kennedy, M. (1985). "Perceiving phonetic events," in *Persistence and Change*: *Proceedings of the First International Conference on Event Perception*, edited by R. Warren and R. E. Shaw, (Erlbaum, Hillsdale, NJ), pp. 139–156.

Syrdal, A. K., and Gopal, H. S. (1986). "A perceptual model of vowel recognition based on the auditory representation of American English vowels," J. Acoust. Soc. Am. 79, 1086–1100.

Whalen, D. H. (1981). "Effects of vocalic formant transitions and vowel quality on the English [s]–[š] boundary," J. Acoust. Soc. Am. 69, 275–282.

Whalen, D. H. (1983). "Vowel information in postvocalic fricative noises," Language Speech 26, 91–100.

Whalen, D. H. (1989). "Vowel and consonant judgments are not independent when cued by the same information," Percept. Psychophys. 46, 284–292.

Yeni-Komshian, G. H., and Soli, S. D. (1981). "Recognition of vowels from information in fricatives: Perceptual evidence of fricative-vowel coarticulation," J. Acoust. Soc. Am. 70, 966–975.