9-1993

# The perceptual weighting of speech-related acoustic cues for 3 & 1/2-year-old children differs from that of adults: Results using natural and synthetic stimuli

Carol J. Manning
*University of Nebraska at Omaha*

# THE PERCEPTUAL WEIGHTING OF SPEECH-RELATED ACOUSTIC CUES FOR 3 & 1/2-YEAR-OLD CHILDREN DIFFERS FROM THAT OF ADULTS: RESULTS USING NATURAL AND SYNTHETIC STIMULI

A Thesis

Presented to the

Department of Special Education

and the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

University of Nebraska at Omaha

by

Carol J. Manning

September 1993

## THESIS ACCEPTANCE

Acceptance for the faculty of the Graduate College, University of Nebraska, in partial fulfillment of the requirements for the degree Master of Arts, University of Nebraska at Omaha.

Committee

| Name | Department |
|---|---|
| [illegible] | Special Education and Communication Disorders |
| [illegible] | Boys Town National Research Hospital |
| | |

[illegible]
Chairperson

Sept 20, 1993
Date

## Abstract

Previous studies have found that children's judgments of syllable-initial /s/ and /ʃ/ are more related to the vocalic F2 transition and less related to the fricative-noise spectrum than are adults' judgments [Nittrouer & Studdert-Kennedy, JSHR, 30 (1987); Nittrouer, J. Phon., 20 1992]. These results have been taken as evidence that young children organize linguistic input in units more closely approximating syllable size than phoneme size. Furthermore, such results have led to a model of speech development proposing that children's weighting of the acoustic cues for phonemic categories changes as they gain linguistic experience, with a general shift in weighting away from dynamic acoustic parameters (those associated with overall syllable production) towards more static acoustic parameters (those associated with the individual phonemic segments of which the syllable is composed). The present study investigated identification by adults and by 3 & 1/2-year-olds of syllable-initial fricatives for stimuli with either natural or synthetic vocalic portions. The two goals of this work were (1) to see if previous findings indicating children's enhanced weighting of formant transitions and diminished weighting of fricative-noise spectra could be replicated for stimuli with natural vocalic portions, and (2) to see if these same patterns would be demonstrated for stimuli with synthetic vocalic portions. Results for children showed that the previously observed patterns of weighting of speech-relevant acoustic information held for stimuli with natural vocalic portions only. For stimuli with synthetic vocalic portions, children's results resembled those of adults. That is, their judgments of fricative identity were more strongly related to the fricative-noise spectrum and less strongly related to the vocalic F2 transition than those judgments had been for stimuli with natural vocalic portions. It was concluded that children's weighting schemes are not simply the consequence of immature psychoacoustic capacities, and that certain schemes are specific to speech stimuli.

## Acknowledgments

I wish to express appreciation to Dr. Susan Nittrouer, thesis chairperson, for her expert assistance throughout every phase of this study. Her experience in the areas of speech development, speech perception and production, and scientific research proved to be an invaluable resource to me. I am grateful that she has become a friend as well as an advisor.

I would also like to thank the members of the thesis committee, Dr. Walt Jesteadt and Dr. John Christensen, for sharing their knowledge and expertise. In addition, I wish to thank Dr. Donal Sinex for his suggestions and comments, and Gina Meyer for her continual support and assistance. Finally, special appreciation is given to my husband, Robert Manning, for his unfailing patience and understanding.

# Table of Contents

# List of Tables

# List of Figures

*Figure*

# INTRODUCTION

The perception of speech requires that listeners have certain abilities: the ability to recognize the organizational units of the speech signal (e.g., words, syllables, phonemes), and the ability to categorize those units as belonging to linguistic classes (Jusczyk, 1992). While speech has traditionally been transcribed as sequences of symbols representing individual phonetic segments, it does not necessarily follow that listeners extract these individual segments in normal, everyday speech perception (Jusczyk & Derrah, 1987; Repp, 1981). The acoustic speech signal cannot be divided into strings of acoustic segments corresponding to discrete phonetic segments. As Pisoni (1985) states, "Research has demonstrated that it is extremely difficult, if not impossible, to segment speech into acoustically defined units that are independent of adjacent segments and free from the contextual effects that occur in sentence environments" (p. 382). Examination of the acoustic correlates of speech reveals that the acoustic information associated with an individual phonetic segment is overlapped with the information associated with others (e.g., Daniloff & Hammarberg, 1973; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Furthermore, the precise qualities of the acoustic properties associated with each segment are influenced by the phonetic context in which that segment occurs (e.g., Cooper, Delattre, Liberman, Borst, & Gerstman, 1952; House & Fairbanks, 1953).

In order to identify the perceptual unit, many investigators have attempted to discover which aspects of the speech signal are used in perception (e.g., Harris, 1958; Kewley-Port, 1983; Liberman, Delattre, & Cooper, 1958; Lisker, 1975; Mack & Blumstein, 1983; Stevens & Blumstein, 1978). As a result of that work, some have suggested that it is static information that is primarily used; that is, information that remains stable across several tens or hundreds of milliseconds. For example, Stevens and Blumstein (1978) proposed that the primary cue used in the perception of stop-initial CV syllables was the onset spectrum of those consonants. Others have proposed that the primary acoustic information for making phonemic decisions is contained in the dynamic portions of speech (Kewley-Port, 1983); that is, those portions of the signal with time-varying acoustic properties. Although these theoretical positions differ with respect to the nature of the acoustic information used to make phonemic decisions, they are similar in that both suggest that listeners extract abstract phonetic features, and subsequently use these features to derive phonetic segments. However, as suggested earlier, listeners do not necessarily extract these segmental units from the signal. In fact, there is evidence that at least one group of listeners clearly may not: young children.

Research exploring young children's perception of speech has revealed differences in children's and adults' perceptual strategies. These age-related differences suggest that young children may not recognize phonetic segments in the acoustic speech signal as clearly as adults do, but instead may pay more attention to some larger unit, such as the syllable. For example, using an identification task, Nittrouer (1992) examined the fricative judgments of children (ages 3 to 7 years) and of adults for fricative-vowel syllables, as a function of the fricative-noise spectrum and the vocalic second-formant (F2) transition. Stimuli consisted of synthetic fricative noises, concatenated with natural vocalic portions. The fricative noises were single-pole noises, with center frequencies ranging from 2200 Hz (most /ʃ/-like) to 3800 Hz (most /s/-like) in 200-Hz steps. These noises were concatenated with natural /ɑ/ and /u/ portions, taken from samples of an adult male saying /ʃɑ/, /sɑ/, /ʃu/, and /su/. The vocalic portions thus had F2 transitions appropriate either for /ʃ/ or for /s/, and so served as an additional cue to fricative identity. (Because the spectrum of the fricative noise remains stable for roughly 100 to 200 msec, it is a static cue to fricative identity. Because the vocalic F2 changes over some portion of the initial vocalic segment, it is a dynamic cue to fricative identity.) Five tokens of each vocalic portion, from each fricative context, were used so that irrelevant acoustic differences among the tokens (such as duration and fundamental-frequency differences) would be randomly distributed across stimulus presentations. Thus, there were 180 stimuli (9 fricative noises x 2 vowels x 2 fricative contexts x 5 tokens). Each stimulus was presented twice, and results were collapsed across the five tokens each of /(ʃ)ɑ/, /(s)ɑ/, /(ʃ)u/, and /(s)u/),[1] providing a total of ten responses for each fricative noise with each of the four types of vocalic portion. Mean identification functions across the fricative noises for the /u/ vocalic portion are shown in Figure 1 for adults and for 3 & 1/2-year-olds. Fricative noise is represented on the abscissa, and percentage of 's' responses is represented on the ordinate. Results showed that children demonstrated greater separation in phoneme boundaries (defined as the 50% point on the identification functions) for the two formant conditions and shallower slopes than adults. The larger separation for boundaries indicates that F2 transition (the 'dynamic' cue) was weighted more heavily in children's than in adults' phonemic decisions. At the same time, their shallower slopes indicate that the fricative noise (the 'static' cue) was weighted less heavily. This second conclusion follows from the work of Berg (1989), who showed that the shallower the slope, the less attention the listener was paying to the dimension

[1] *Throughout this manuscript, the fricative label in parentheses will indicate that the vocalic portion had an F2 transition appropriate for a vowel produced after that fricative.*

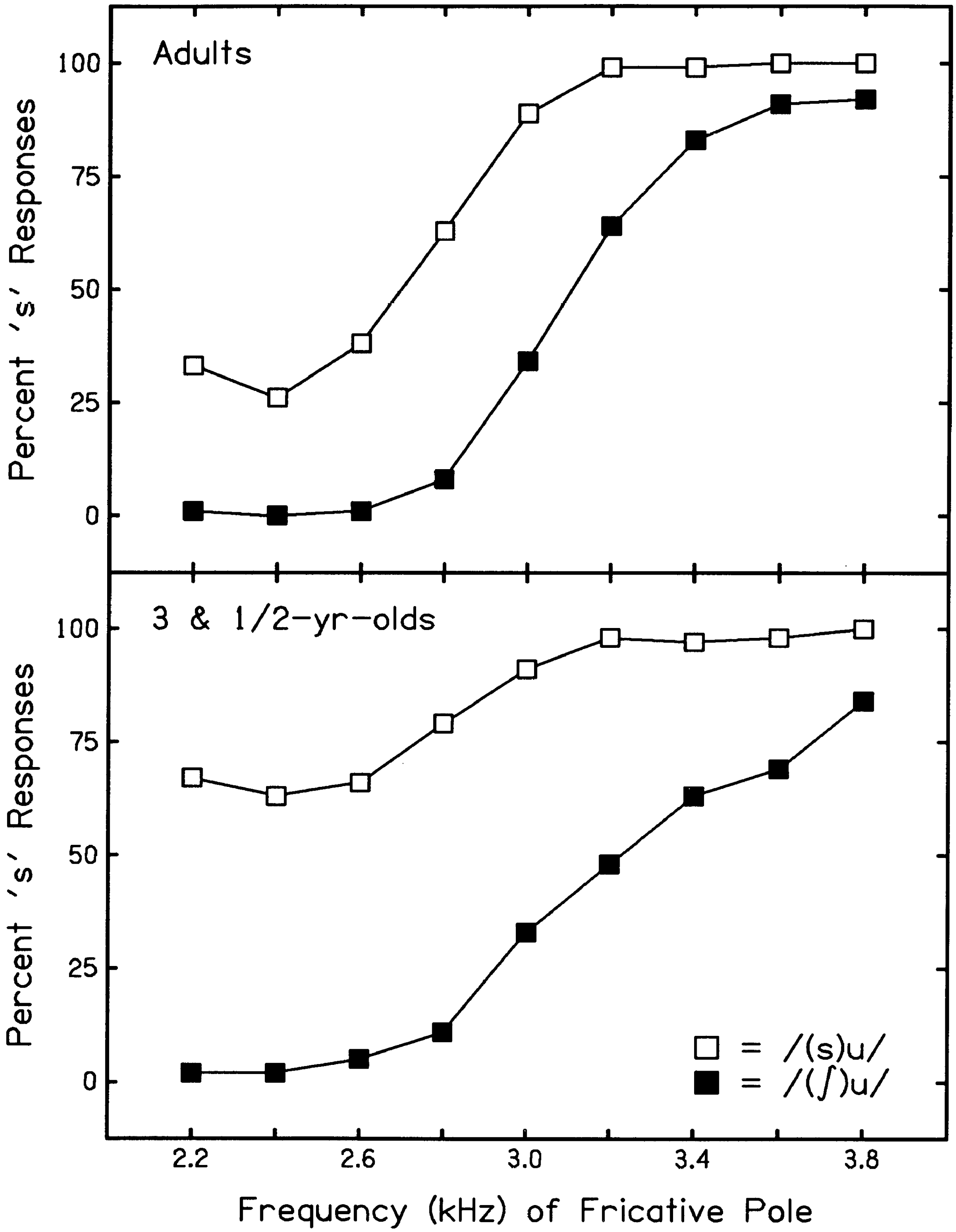


Figure 1: Identification functions for adults and 3 & 1/2-year-olds from Nittrouer (1992), for stimuli with /u/ vocalic portions.

represented on the abscissa. These results replicated those obtained by Nittrouer and Studdert-Kennedy (1987). Others have reported similar findings: Morrongiello, Robson, Best, and Clifton (1984) found that 5-year-olds weighted the vocalic first-formant (F1) transition more than adults in making judgments of 'say' versus 'stay.' Parnell and Amerman (1978) found that when adults and children were asked to identify voiceless stops in stop-vowel syllables, the younger children (mean age 4 years, 6 months) weighted the information provided by the aperiodic (static) noise less than 11-year-olds and adults.

These studies reveal that the weighting of acoustic information in phonemic decisions differs for children and adults. Specifically, children seem to weight information that spans the syllable to a greater extent than adults, and weight information temporally constrained to individual acoustic segments, and associated with specific phonetic segments, to a lesser extent. Nittrouer and Studdert-Kennedy (1987) hypothesized that "Perhaps young children are not as adept as adults at recovering the individual phonemes from the syllable, but instead tend to perceive syllables as relatively undifferentiated wholes" (p. 321). In other words, enhanced sensitivity to formant transitions may reflect a tendency on the part of children to organize incoming speech signals more as syllabic units, rather than as individual phonetic segments. This suggestion is supported by the findings of Nittrouer and Studdert-Kennedy showing that younger children seemed less able than older children and adults to use (for phonemic decisions) the acoustic portion of the syllable associated with an individual segment (i.e., steady-state fricative noise). Instead, they seemed to be more perceptually attentive to the vocalic formant transitions, which are the dynamic components of speech that tie the static portions of the syllable together. Thus, one aspect of speech development may involve a shift in the relative weighting of the static and dynamic components of speech signals, reflecting a concommitant shift in children's sensitivity to, or 'awareness of,' the segmental structure of speech.

Others have similarly proposed that speech perception is initially based on the syllabic unit, but that this organizational tendency changes over time. It has been found that the relative weighting of the acoustic information in speech changes with development (Greenlee, 1980; Krause, 1982; Morrongiello et al, 1984; Nittrouer, 1992; Nittrouer & Studdert-Kennedy, 1987; Wadrip-Fruin & Peach, 1984), as does sensitivity to phonetic units. Supporting this latter point, Fox and Routh (1975) showed that children at the age of 4 years could segment sentences into words and words into syllables, but the ability to segment words or syllables into phonemes did not appear until roughly 7 years of age.

The studies mentioned above indicate that children's sensitivity to the segmental structure of speech undergoes a developmental 'refining' of sorts. However, other studies using discrimination procedures have shown that infants' perception can be as sensitive to acoustic detail as that of adults, or very nearly so. Fodor, Garrett, and Brill (1975) found that 14- to 18-week-old infants could group syllables based upon whether they shared a similar consonant, leading these authors to conclude that the infants were able to analyze syllables into phonetic segments. Similarly, Eilers, Wilson, and Moore (1977) found that infants as young as 1 to 3 months could discriminate between syllables that differed by only one phoneme. Miller and Eimas (1983) also concluded that infants' and adults' categorization of speech sounds is very similar. However, none of the studies described above manipulated the acoustic parameters in such a way that the relative weighting of those parameters for infants and adults could be evaluated. It may be that these two groups of listeners arrive at similar decisions, but do so based on different perceptual strategies, or weighting schemes, of the available acoustic information.

While it seems that there exist opposing views regarding whether or not children are sensitive to the segmental structure of speech in the same way adults are, it may be that the results of the various studies reveal differences in the utilization of perceived information. It is possible that the tasks involved (discrimination and identification) are measuring different skills. Categorical discrimination of speech sounds may be accounted for in terms of sensitivity to acoustic changes, whereas different weighting schemes may be used when the decisions to be made are not merely whether the sounds are the same or different, but instead require that phonemic labels be attached to the segments (Jusczyk, 1986; Jusczyk, Rossner, Cutting, Foard, & Smith, 1977). Jusczyk, Pisoni, Reed, Fernald, and Myers (1983) used a high-amplitude sucking procedure with two-month-old infants to determine whether they processed nonspeech sounds differently from speech sounds. This was an extension of a study by Eimas and Miller (1980), using synthetic speech stimuli. If results indicated that the infants processed nonspeech sounds differently from speech sounds, this would provide evidence that a specialized mode of speech processing had been used in the analogous speech experiment. Stimuli were sine-wave analogs of the synthetic speech syllables /bɑ/ and /wɑ/, used by Eimas and Miller. Two sets of these stimuli were used, differing in overall duration. Within each set, stimuli differed in the durations of their initial frequency transitions. Data revealed that the infants " ... not only discriminated differences in duration of frequency transitions in nonspeech sounds, but they displayed a pattern of discrimination that was both relational and categorical and therefore directly comparable to the findings obtained by Eimas and Miller with synthetic speech stimuli" (p. 176). Thus, the infants in this particular study seemed to be using the same processing skills in the discrimination of nonspeech sounds that the infants had used during the speech discrimination task of Eimas and Miller, a finding that led to the conclusion that speech perception requires nothing more than general auditory capabilities. However, another possibility exists. It may be that discrimination of speech stimuli primarily measures the sensitivity of the listener to the acoustic information on which phonemic categories may be based; identification tasks may be required to determine how the listener uses that information in making decisions about phonemic identity. For each of these tasks, and for different acoustic signals, different weighting schemes may be employed.

Similarities in discrimination results between infants and adults (and between nonspeech and speech stimuli) may exist because the task reveals sensitivity to acoustic changes, and those sensitivities in infants are sufficient for discriminating the phonemic changes being presented. In contrast, certain perceptual strategies used by children during identification tasks are not necessarily adult-like. This contrast can possibly explain the findings that infants' discrimination of some speech sounds is similar to that of adults, while, at the same time, not discounting the differences in young children's and adults' weighting of the static and dynamic aspects of speech during identification tasks. Although discrimination may be achieved using strategies involving general auditory skills or sensitivities, identification tasks may require the use of specific speech-processing skills. These processes may involve the weighting of the various characteristics of the speech signal (which were derived using general auditory skills), and then assignation of a linguistic label based on the weighted sum of these characteristics. It is during the weighting process that children's and adults' strategies may differ, as they weight the various acoustic properties according to their own perceptual schemes. This suggested account of differences in children's and adults' speech perception receives support from Jusczyk's (1992) model of word recognition, illustrated in Figure 2. Support is also received for this account from Simon and Fourcin (1978). These investigators have stated that discrimination tasks merely explore the discriminatory capabilities of the auditory mechanism, but "A labeling paradigm, on the other hand, requires that subjects abstract from the stimuli the relevant distinctive acoustic patterns, evaluate them in terms of a functional system, and make a linguistic decision about the category to which they belong" (p. 926). In other words, another kind of processing occurs during identification tasks.

One objective of the present study was to investigate the conclusions reached by Nittrouer (1992), namely, that young children weight the information in the speech acoustic signal differently than adults do in making phonemic decisions. The following null hypothesis was tested:

> In identification tasks, adults and 3 & 1/2-year-olds would demonstrate functions with similar phoneme boundaries and slopes. In other words, children and adults would weight similarly the acoustic parameters of speech.

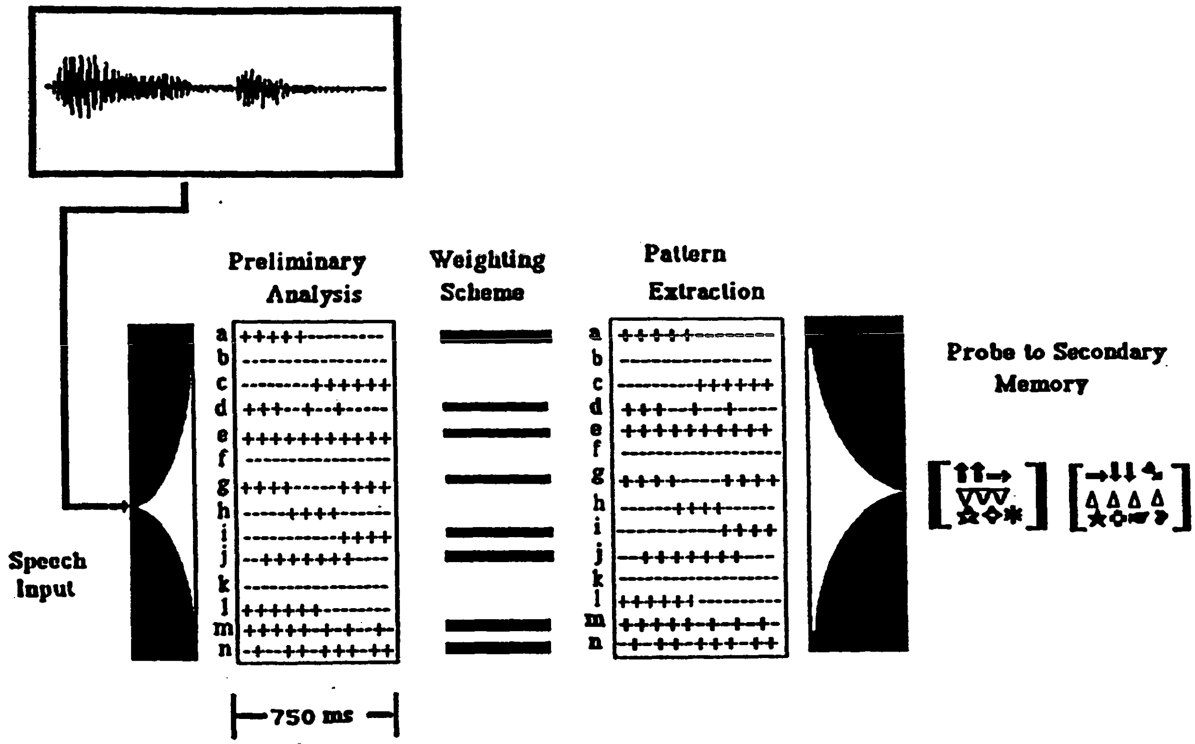


Figure 2: Diagram of Jusczyk's (1992) Word Recognition Model. Input to the system is the wave form of the English utterance "baby". First stage of processing involves a preliminary analysis of the signal by the auditory system. Because the auditory system is constantly monitoring the input, a continuous reading of the presence of activity is available for a given timeslice, here 750 msec. Certain outputs of this stage of analysis are more closely monitored than others due to the Weighting Scheme associated with the native language (designated by the bold bars). The sound pattern is extracted from the emphasized processes (in bold face) and then serves as a probe to secondary (lexical) memory. The bold brackets indicate the stressed syllable, the symbols inside are meant to indicate that some featural description of the sound structure is present, but not one that is explicitly segmented into phonemes.

A second objective of this experiment was to see if the adults and children in this study would demonstrate similar phoneme boundaries and slopes for stimuli with natural vocalic portions and for those with synthetic vocalic portions. This objective seemed worthwhile because speech perception experiments with adults, children, and infants usually assume (even if just implicitly) that the same speech-processing strategies are used with natural and synthetic stimuli. However, it is possible that listeners (either all or just the young) require the acoustic characteristics unique to sound produced by a human vocal tract for these speech-processing strategies to be invoked. The second objective tested the following hypothesis:

> In identification tasks, adults and 3 & 1/2-year-olds would demonstrate functions with similar phoneme boundaries and slopes for stimuli with both natural and synthetic vocalic portions. In other words, children and adults would weight similarly the acoustic characteristics of natural and synthetic speech.

Regarding the first objective, it was predicted that a pattern of results would be obtained that was similar to those of Nittrouer and Studdert-Kennedy (1987) and to those of Nittrouer (1992); that is, that children, compared to adults, would display shallower slopes (when frequency of the fricative pole is represented on the abscissa) and greater separation in phoneme boundaries (as a function of whether the formant transition is appropriate for /s/ or for /ʃ/). Regarding the second objective of this study, it was difficult to predict what the effect would be of using synthetic, instead of natural, vocalic portions. It might be that all listeners, or just the children, would not use the weighting schemes normally used with speech signals to make phonemic decisions about these completely synthetic stimuli. If this were the case, a different pattern of results for phoneme boundaries and slopes should be observed for the stimuli with natural and synthetic vocalic portions. The onset frequencies of F2 in the synthetic vocalic portions were set to the most extreme values found in the natural stimuli. In both Nittrouer and Studdert-Kennedy and in Nittrouer, two vowel contexts had been used: /i/ and /u/ in Nittrouer and Studdert-Kennedy, and /ɑ/ and /u/ in Nittrouer. In both studies, the vowel context that showed the greater difference in F2 onset as a function of the preceding fricative displayed the greater separation in phoneme boundaries. In Nittrouer and Studdert-Kennedy, for example, the F2 onset for /(ʃ)i/ was 200 Hz higher than the F2 onset for /(s)i/, and the F2 onset for /(ʃ)u/ was 320 Hz higher than that of /(s)u/. Adults in that study showed a 64-Hz separation in /i/ phoneme boundaries, and a 420-Hz separation for /u/ phoneme boundaries.

## METHOD

### Subjects

Two groups of listeners participated in this experiment: adults between 20 and 40 years of age, and preschool children between 3 years;2 months (3;2 years) and 4;3 years of age. All subjects were right-handed, native speakers of American English who had no history of speech or hearing problems. In addition, all adults had at least an eleventh grade competency for sight reading a word list. For children, it was required that no child have a medical or family history that would put them at risk for a speech or language problem. Specifically, no member of any child's immediate family had ever been seen for a speech or language problem, pregnancies and deliveries were normal for all children, and no child had a history of frequent middle-ear infections. In theory, the criterion for being considered free of a history of frequent middle-ear problems was that the child could not have had three or more ear infections within the first year of life or within the twelve months just prior to testing. In practice, absolute number of episodes was not found to be a good criterion for judging if a child might be at risk for transient hearing loss due to middle-ear infections. The risk associated with middle-ear problems seemed related to how quickly parents sought medical attention when an infection occurred. Therefore, parents were questioned carefully if they reported any history of middle-ear problems, and decisions were made concerning who to accept based on those answers. In one case, a child was eliminated who had less than the criterion number of infections because it was determined that it was probably a long-standing, serious problem by the time medical attention was sought. In four cases, children were accepted who surpassed the criterion number of infections because it was clear that medical attention was sought before any serious effects on hearing would have been expected. Finally, no member of any child's immediate family was "strongly left-handed," defined as using the left hand to perform all everyday activities. Handedness of adults was assessed by asking if they considered themselves right- or left-handed, and then inquiring further if they replied that they were left-handed. Often it turned out that individuals who categorized themselves as left-handed actually performed some activities (e.g., rolling a ball or holding a racket) with their right hand. Sixteen adults and 26 children meeting the specified criteria participated.

## Equipment and Materials

A Madsen audiometer was used to screen subjects' hearing. Stimuli were presented free-field using an IBM compatible computer, a 12-bit digital-to-analog converter (Data Translation 2801A), a Frequency Devices 901F filter, a Tascam amplifier (model PA 30-B), and a JBL Control-1 speaker. Responses were registered directly to the computer by a box with three buttons (although only two buttons were actually used for recording responses). Trials were initiated by pressing a fourth button, attached to the box by a long cable. Children indicated their responses by pressing one of two large, colored buttons mounted on a board. These buttons were not attached to anything. Children held two handles on the outermost edges of the board between trials. A picture of a girl served as the prompt for 'Sue' and a picture of a shoe served as the prompt for 'shoe.' Reinforcement was presented following each response, with one of four devices: three Plexiglass boxes, each containing a mechanical animal and two lightbulbs, and a CGA monitor that could produce one of four displays of brightly colored shapes.

## Stimuli

Two sets of stimuli were used. Both sets were digitized at a 10-kHz sampling rate, and consisted of nine synthetic fricative noises similar to those used by Nittrouer (1992). These noises were synthesized using a KLATT software serial synthesizer. They were single-pole noises whose center frequencies varied along a continuum from 2200 Hz to 3400 Hz in 150-Hz steps, and were 230 ms in length.

For one set of stimuli, natural vocalic portions were concatenated with the synthetic fricative noises. (This set of stimuli will hereafter be known as the 'natural stimuli'.) These were the same vocalic portions used by Nittrouer (1992). Five vocalic portions were taken from samples of a male speaker saying /ʃu/ and five vocalic portions were taken from samples of the same speaker saying /su/. Consequently, there were five /u/ tokens with formant transitions appropriate for each fricative. For the five /(ʃ)u/ portions, mean duration was 348 ms and mean fundamental frequency (f0) was 97 Hz. For the five /(s)u/ portions, mean duration was 347 ms and mean f0 was 99 Hz. All ten vocalic portions had F2 transitions that fell in frequency through the entire portion. For /(ʃ)u/ portions, mean starting frequency for F2 was 1706 Hz and mean ending frequency was 903 Hz. For /(s)u/ portions, mean starting frequency for F2 was 1520 Hz and mean ending frequency was 962 Hz. Each of these vocalic portions was concatenated with each of the nine fricative noises, making a total of 90 stimuli. During testing, each stimulus in this set was presented to each listener once, and responses to each fricative noise were collapsed across the five tokens of the /(ʃ)u/ and across the five tokens of the /(s)u/ vocalic portions.

For the second set of stimuli, two synthetic vocalic portions were concatenated with each of the fricative noises. (This set of stimuli will hereafter be known as the 'synthetic stimuli'.) Each synthetic vocalic portion was 270 ms in duration, with an f0 of 100 Hz. The first formant (F1) remained constant throughout the vocalic portion at 250 Hz, and the third formant (F3) remained constant at 2100 Hz. For both, F2 fell through the entire vocalic portion to a final value of 850 Hz. One of these portions had an F2 onset of 1800 Hz (the most /ʃ/-like). This matched the highest F2 onsest found for the five natural /(ʃ)u/ tokens. The other had an F2 onset of 1480 Hz (the most /s/-like), the lowest onset found for the five natural /(s)u/ portions. Each of these vocalic portions was concatenated with each of the nine fricative noises, making a total of 18 stimuli. During testing, each stimulus was presented to each listener five times.

Both the amount of change in F2 frequency over the course of the vocalic portion and the slope of that change varied between the two sets of stimuli. For the natural stimuli, F2 in /(ʃ)u/ changed, on average, by 803 Hz over the course of the vocalic portion, at a rate of 2.31 Hz per ms. F2 in natural /(s)u/ changed, on average, by 558 Hz, at a rate of 1.61 Hz per ms. The synthetic /(ʃ)u/ portion changed by 950 Hz, at a rate of 3.52 Hz per ms, while the synthetic /(s)u/ changed by 630 Hz, at a rate of 2.33 Hz per ms. Thus, the synthetic stimuli demonstrated more total change in F2 over the course of the vocalic portion than their natural counterparts, and this frequency change occurred at a faster rate in the synthetic than in the natural stimuli.

## Procedures

Pretest. Parents with children of the appropriate age for this study were contacted either by mail, by announcements made on the local University radio station, or by flyers distributed by their child's daycare facility. Regardless of how they first heard of the study, all parents received flyers that described the goals and procedures of the study, a questionnaire designed to determine if the child met the criteria for the study, and a prepaid envelope addressed to the Speech Development Laboratory at the University of Nebraska at Omaha prior to coming to the laboratory. If parents were interested in participating, they completed the questionnaire and returned it in the prepaid envelope. This questionnaire asked information about the birth and medical history of the child, the child's personal and family history of speech or language problems, the child's history of ear infections, the child's handedness, and the handedness of immediate family members. As part of the questionnaire, parents signed a statement indicating that they understood that the study did not serve as a speech/language screening. Upon receiving the completed questionnaire, an experimenter called the parent to check on any unclear responses, and to schedule the child, if appropriate. On the first day of participation, one experimenter engaged the child in a brief conversation, while a second experimenter listened. This procedure served as a confirmation that there was no reason to suspect a speech or language problem for that child. If a problem was suspected, the child was dismissed. The parent was told of any suspicions, and was encouraged to have a formal evaluation done. This situation occurred with only one child.

Adults participating in the study completed a questionnaire asking about potential speech, language, or hearing problems, and about handedness on the first day of participation. In addition, adults also took the reading portion of the Wide Range Achievement Test-Revised (WRAT-R [Jastak & Wilkinson, 1984]). This test consists of orally reading a word list, and provides grade-level norms.

On the first day of participation, all listeners had their hearing screened. This screening consisted of the presentation of pure tones of the frequencies of 0.5 kHz, 1.0 kHz, 2.0 kHz, 4.0 kHz, and 6.0 kHz at 25 dB HL (ANSI, 1969). The pure tones were presented free-field. Subjects who did not pass the screening were informed of this fact and dismissed. This situation occurred for two adults, but, all children passed the hearing screening.

General testing. All stimuli were presented at 68 dB SPL peak intensity. Each stimulus was presented repeatedly until a response was recorded, at an onset-to-onset rate of 2 sec. However, listeners almost always responded after one stimulus. During testing, the listener sat at a table in a sound-attenuated room. Children sat facing the Plexiglass boxes and graphics monitor that would present reinforcement. The computer controlling presentation of stimuli and recording of responses was in an adjacent room. An experimenter in the adjacent room controlled the software, and, in the case of children, recorded responses. That experimenter was able to see the listener through a one-way mirror.

All listeners participating in the identification experiment also participated in a second, discrimination experiment. Adults participated in a total of three sessions, each approximately 50 minutes in length: one and a half sessions of identification and one and a half sessions of discrimination. All three sessions were completed within the same week. Children participated in a total of six sessions, each of approximately 25 minutes in length: three sessions of identification and three sessions of discrimination. Identification and discrimination tasks were conducted during different (but temporally adjacent) weeks. For adults, the identification experiment was always conducted during the first complete session, and the first half of the second session. This was because the goal of the identification experiment was to determine how listeners weight the two cues relevant to fricative judgments (i.e., fricative-noise spectrum and F2 transition) in normal, everyday speech perception. Concern existed that participating in the discrimination task just prior to participating in the identification task (as would happen on the second day of testing with adults, if the discrimination experiment were conducted prior to identification) might alter those response patterns. For this same reason, adults were always tested with the natural stimuli before the synthetic stimuli in identification. This concern did not exist for testing with children because a minimum of two days always intervened between one kind of task and the other. Instead, some concern existed that children's interest in the tasks might diminish the second week. Therefore, the order of presentation of the identification and the discrimination tasks varied across children, and the order of presentation of natural and synthetic stimuli varied across children in the identification task.

Testing with children. An experimenter was in the room with the child at all times, sitting to one side of the child. Usually, one of the child's parents was also in the room during testing. The only exceptions were when siblings came to the laboratory and required the parents' attention. The parent and the experimenter in the room with the child listened to taped monologues of a male radio personality over earphones during testing.

The hearing screening was presented on the first day of testing. Next, the child was trained to push one button (corresponding to one response category) at a time, using a procedure similar to that of Tallal (1980). The board with the two large buttons was placed horizontally in front of the child, and one picture placed above the corresponding button. (Each picture was in a frame with a border that matched the corresponding button in color.) The best exemplar of that category was presented ten times. (The best exemplar of 'shoe' consisted of the 2200-Hz fricative noise, concatenated with vocalic portions with F2 transitions appropriate for /ʃ/. The best exemplar of 'Sue' consisted of the 3400-Hz noise, concatenated with F2 transitions appropriate for /s/.) At first, the child received help and verbal encouragement to press the button when the stimulus was heard. After this procedure was completed for one response category, that picture was removed and the procedure repeated for the other response category. During this simple training procedure, children also learned to hold the handles on the board between responses. This procedure helped to keep children sitting quietly, and prevented them from resting their hands on the buttons, which makes it difficult to see when a button-press actually occurs. This training procedure was included only for the first set of stimuli (natural or synthetic) presented to a child.

Next, both pictures were positioned above the corresponding buttons at the same time, and the child was trained to choose the appropriate button when that stimulus was heard. To do this, the best exemplars of each response category were presented ten times each, in randomized blocks of five each. The experimenter in the room with the child would provide help during the first block, if necessary, but not during the second block. In order to continue in the identification experiment, each child was required to respond correctly to nine out of the last ten stimuli. (This training procedure was used with both stimulus sets.) One boy (3;3 years) did not train on the synthetic stimuli, and training was not subsequently tried with natural stimuli. The subject was dismissed at this point. Figure 3 provides a summary regarding children who were dismissed before completing testing with both sets of stimuli.

Next, testing with all stimuli in the set was conducted. Stimuli were presented in five randomized blocks of 18 (9 fricative noises x 2 F2 transitions).

For the natural stimuli, tokens of each vocalic portion (i.e., /(ʃ)u/ and /(s)u/) were randomized within each block. Each vocalic portion was presented only once with each fricative noise. Children were required to demonstrate 80% correct responses to the endpoint fricative noises, when presented with vocalic portions having F2 transitions appropriate for the response category corresponding to that noise. This requirement insured that only data from listeners who maintained general attention throughout the task were included in the final analysis. Six children failed to meet this criterion for one set, but met it for the other: three children met this criterion for the natural stimuli, but not for the synthetic; and three children met it for the synthetic stimuli, but not the natural. Five children failed to meet the 80% criterion with either stimulus set. Two children failed to meet this criterion for the natural stimuli, and were not subsequently tested with the synthetic stimuli. One child did not meet the criterion for the synthetic stimuli, and was not subsequently tested with the natural stimuli. Eleven children achieved the 80% criterion for both sets of stimuli. For the final analysis then, data from fourteen children were used for each of the natural and the synthetic stimulus sets.

Testing with adults. Testing procedures used with adults differed from those used with children in four ways. First, adults did not have an experimenter in the room with them. They initiated trials themselves. Second, adults responded by pointing to one of the pictures. Third, adults were not reinforced after responding. Finally, adults had less training than children. Adults received one block of training (10 stimuli) consisting of the best exemplars of both response categories at the start of testing with both stimulus sets. Data from all sixteen participating adults were included in the final analysis.

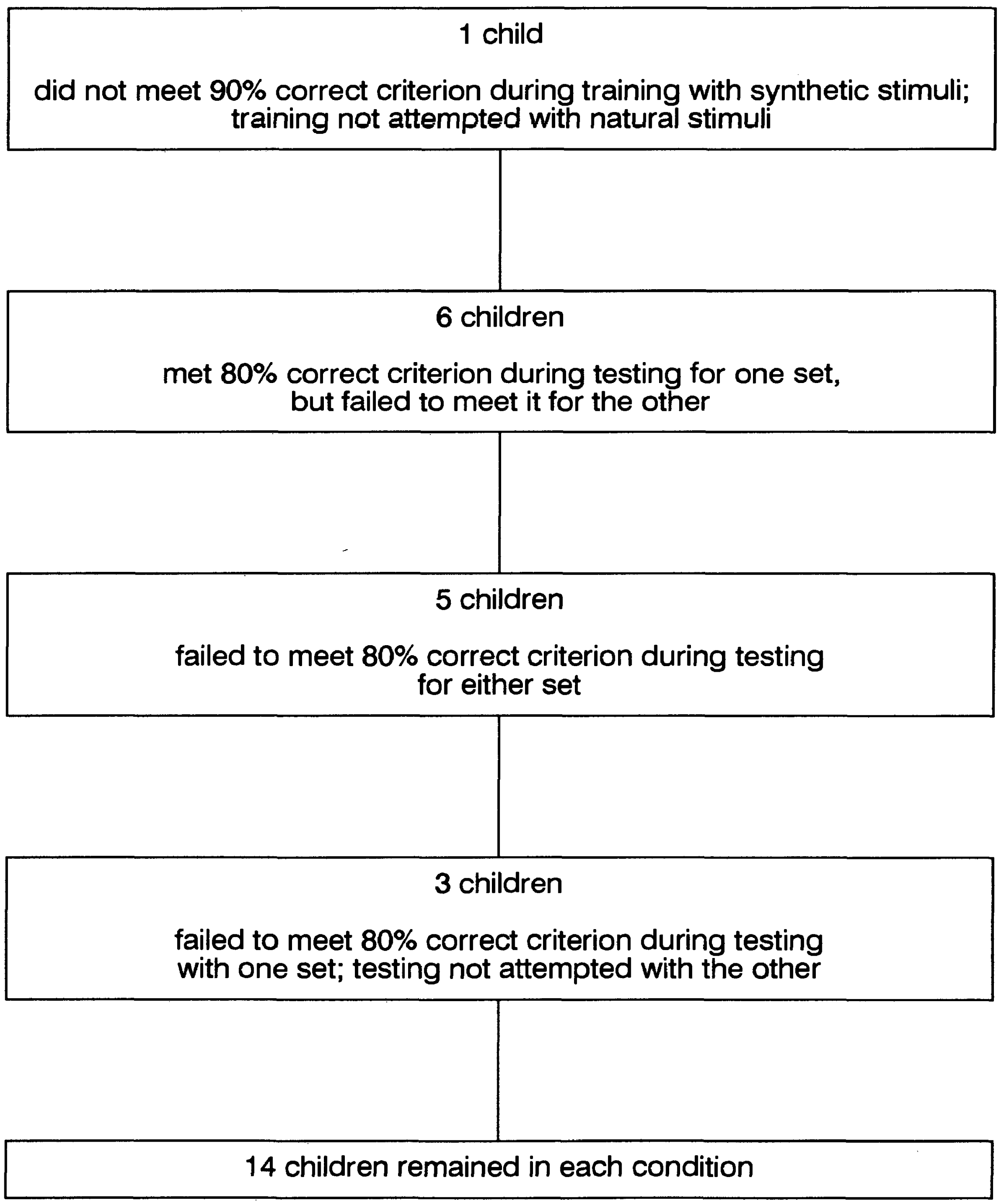


Figure 3: Chart of children who were dismissed.

Differences in procedures for this study, compared to Nittrouer (1992). The major difference between these two studies was the extent of the fricative-noise continuum. Noises in Nittrouer's study varied along a continuum from 2200 Hz to 3800 Hz in nine 200-Hz steps. Thus, the continuum in the present study was a truncated version of that used in the earlier study, in that the noises used in this study had 3400 Hz as the highest value. It was predicted that this difference might result in generally lower phoneme boundaries for subjects in the present study than for those in the 1992 study. Another difference between Nittrouer's study and the present study concerned the number of presentations of each stimulus: Subjects in the present study were presented with five repetitions of each stimulus, while Nittrouer's subjects were presented with ten repetitions. It was predicted that this difference would result in generally steeper functions in the present than in the 1992 study. Also, the subjects in the current study were presented with the stimuli free-field, rather than over headphones, as in Nittrouer's study. Finally, subjects in the 1992 study heard each stimulus only once, whereas subjects in the present study heard stimuli repeatedly. Possible effects of these last two manipulations could not be predicted.

## RESULTS

Comparison of previous and current results with natural stimuli

Results with natural stimuli were compared to those obtained by Nittrouer (1992). Sixteen adults between 20 and 40 years of age and nine 3 & 1/2-year-olds participated in Nittrouer's study. Figure 4 displays identification functions for the natural stimuli for both studies for adults, and Figure 5 displays the identification functions for both studies for children. For both experiments, the

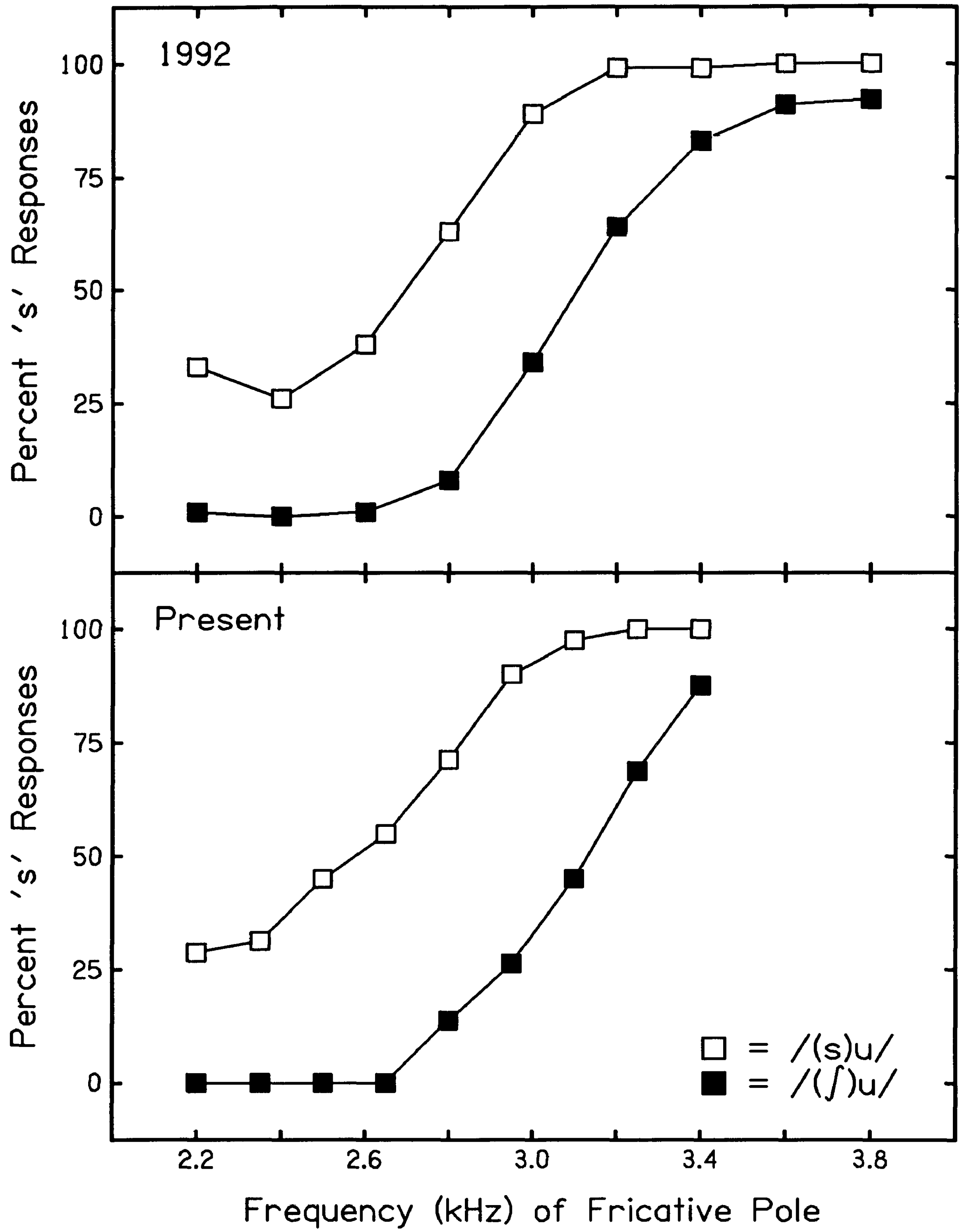


Figure 4: Identification functions for adults from Nittrouer (1992) and from the present study for natural stimuli.

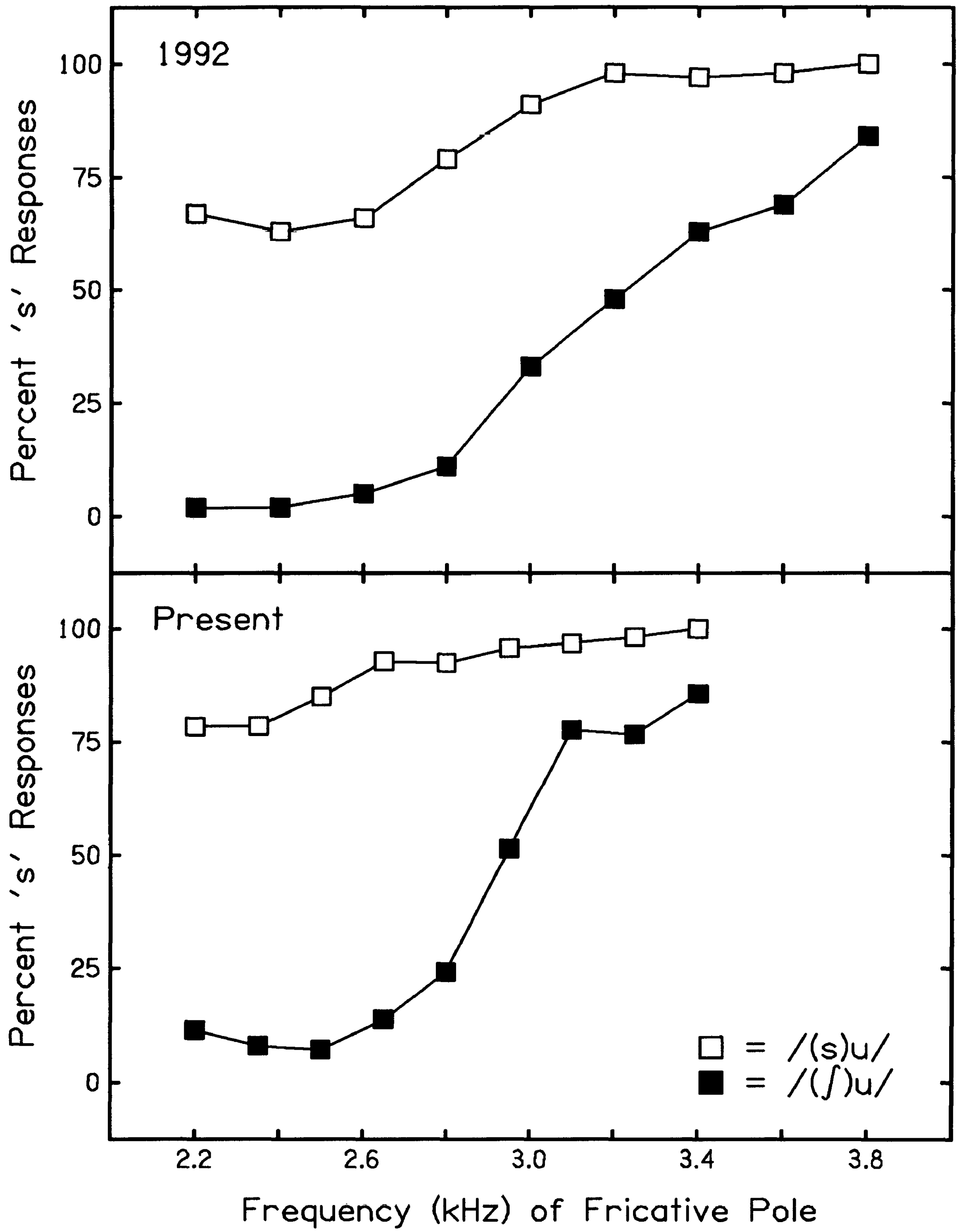


Figure 5: Identification functions for 3 & 1/2-year-olds from Nittrouer (1992) and from the present study for natural stimuli.

percentage of /s/ responses were tallied for each subject for both the /s/ and the /ʃ/ transition conditions. Probit analysis (Finney, 1964) was performed on these data. This analysis fits a straight line to the cumulative probability function (probit score = z + 5). From these scores, a distribution mean was calculated. This distribution mean represents the subject's phoneme boundary between /ʃ/ and /s/. Slope was also calculated, and is defined as the change in probit units per kiloHertz of change in the fricative noise.

*Table I:* Mean phoneme boundaries for Nittrouer (1992) and for the natural stimulus set in the present study

| | Phoneme Boundary | | | |
|---|---|---|---|---|
| **transition:** | /s/ | | /ʃ/ | |
| | **3-yr-olds** | **adults** | **3-yr-olds** | **adults** |
| 1992 study: | 1998 | 2517 | 3332 | 3163 |
| present study | 1890 | 2494 | 3019 | 3111 |

Table I provides mean phoneme boundaries for both age groups for both studies. The 3 & 1/2-year-olds' lower phoneme boundaries for vocalic portions with /s/ transitions (found in both the 1992 and 1993 data) reflect their greater weighting, relative to adults, of the F2 transition in making these phonemic decisions. A two-way Analysis of Variance (ANOVA) was performed separately on each data set (1992 and present), using age as the between-subjects' factor and transition (appropriate for either /s/ or /ʃ/) as the within-subjects' factor. For the 1992 data, the main effect of transition was statistically significant [$F(1,23)=97.95$, $p<0.001$]. The Age x Transition interaction was also found to be statistically significant [$F(1,23)=13.35$, $p=0.001$]. For the present data, the main effects of age and transition were found to be statistically significant, [$F(1,28)=12.42$, $p=0.002$] and [$F(1,28)=120.51$, $p<0.001$], respectively. The Age x Transition interaction was also statistically significant [$F(1,28)=10.74$, $p=0.003$]. Thus, similar trends were observed for the two studies, with one exception: a significant main effect of age was found for the present study, but not for the 1992 study.

A significant main effect of age was undoubtedly found in the present study (whereas it was not found in the 1992 study) because 3 & 1/2-year-olds did not display higher-frequency phoneme boundaries for the /ʃ/ transition condition, as they had in the 1992 study. Those higher phoneme boundaries balanced the lower-frequency phoneme boundaries observed for children for the /s/ transition condition (in the 1992 study), resulting in similar mean phoneme boundaries (across the two transition conditions) for adults and 3 & 1/2-year-olds. In the present study, this failure to find raised phoneme boundaries for the /ʃ/ transition condition for 3 & 1/2-year-olds, compared to adults, meant that their mean phoneme boundaries were lower in frequency than those of adults.

A three-way ANOVA was also performed on these data, using study (1992 or present) and age as the between-subjects' factors, and transition as the within-subjects' factor. The main effect of age was found to be statistically significant [$F(1,51)=15.71$, $p=0.002$]. The main effect of transition was also found to be statistically significant [$F(1,51)=218.29$, $p<0.001$]. Finally, the Age x Transition interaction was found to be statistically significant [$F(1,51)=23.25$, $p<0.001$]. The fact that there was no statistically significant main effect for study, as well as no significant interactions involving study, would indicate that the present study demonstrated similar phoneme boundaries to those obtained by Nittrouer (1992).

*Table II:* Mean slopes for Nittrouer (1992) and for the natural stimulus set in the present study

| | Slopes | | | |
|---|---|---|---|---|
| **transition:** | /s/ | | /ʃ/ | |
| | **3-yr-olds** | **adults** | **3-yr-olds** | **adults** |
| 1992 study: | 1.45 | 3.45 | 2.68 | 4.06 |
| present study: | 2.17 | 4.26 | 3.64 | 6.29 |

Table II provides mean slopes from both studies. It can be observed that results from both studies display a similar pattern: children's slopes are shallower than those of adults. This reflects their lesser weighting of the fricative noise during the task, relative to the adults. A two-way ANOVA was performed on the 1993 data, with age as the between-subjects' factor and transition as the within-subjects' factor. The main effect of age was significant [$F(1,28)=14.95$, $p=0.001$]. The main effect of transition was also found to be significant [$F(1,28)=11.42$, $p=0.002$], reflecting the fact that both groups demonstrated shallower slopes for the /s/ than for the /ʃ/ transition condition.

A significant age effect for slope had also been found by Nittrouer (1992). However, slopes were not compared between that study and this one because of the expectation of steeper slopes in the present study. Despite this difference between the two studies, it is important to note that both revealed shallower slopes for children's functions than for those of adults.

Comparison of results from natural and synthetic stimuli

Results from the sets of stimuli with natural vocalic portions and with synthetic vocalic portions were compared. Figure 6 displays identification functions for the natural and synthetic stimuli for adults, and Figure 7 displays these functions for children.

*Table III:* Mean phoneme boundaries for the natural and synthetic stimulus sets

| | Phoneme Boundary | | | |
|---|---|---|---|---|
| **transition:** | /s/ | | /ʃ/ | |
| | **3-yr-olds** | **adults** | **3-yr-olds** | **adults** |
| natural stimuli: | 1890 | 2494 | 3019 | 3111 |
| synthetic stimuli: | 2621 | 2627 | 3154 | 3006 |

Table III provides mean phoneme boundaries for the natural and synthetic stimuli. Examination of the phoneme boundaries for the synthetic stimuli reveals similar results for children and adults, indicating that the weighting of the F2 transitions by the two groups was similar during this task. A two-way ANOVA was performed on the synthetic data, with age as the between-subjects' factor and transition as the within-subjects' factor. Only the main effect of transition was found to be significant [$F(1,28)=78.34$, $p<0.001$]. Failure to find a significant Age x Transition interaction suggests that children and adults weighted the F2 transition similarly during this task.

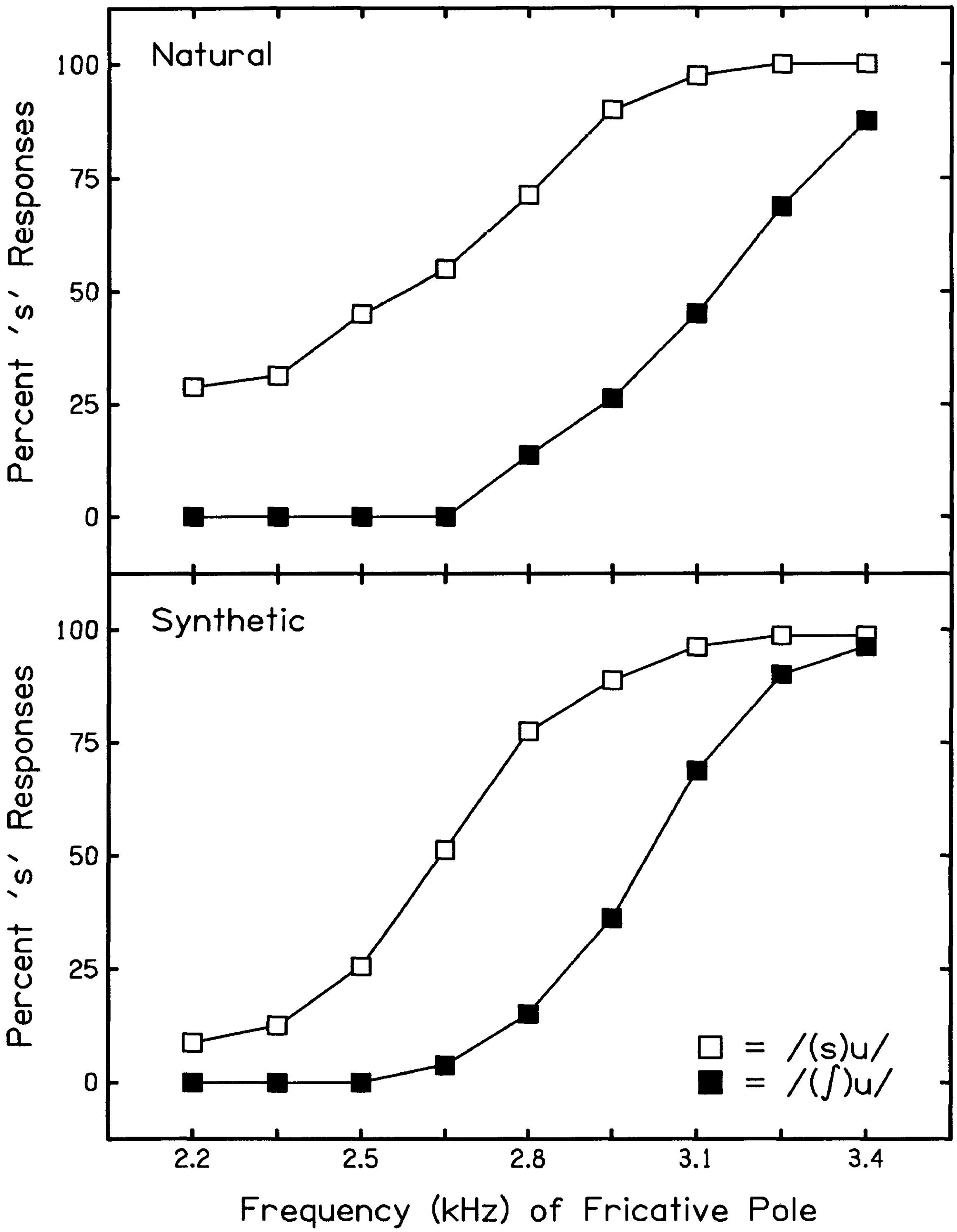


Figure 6: Identification functions for adults from the present study for natural and synthetic stimuli.

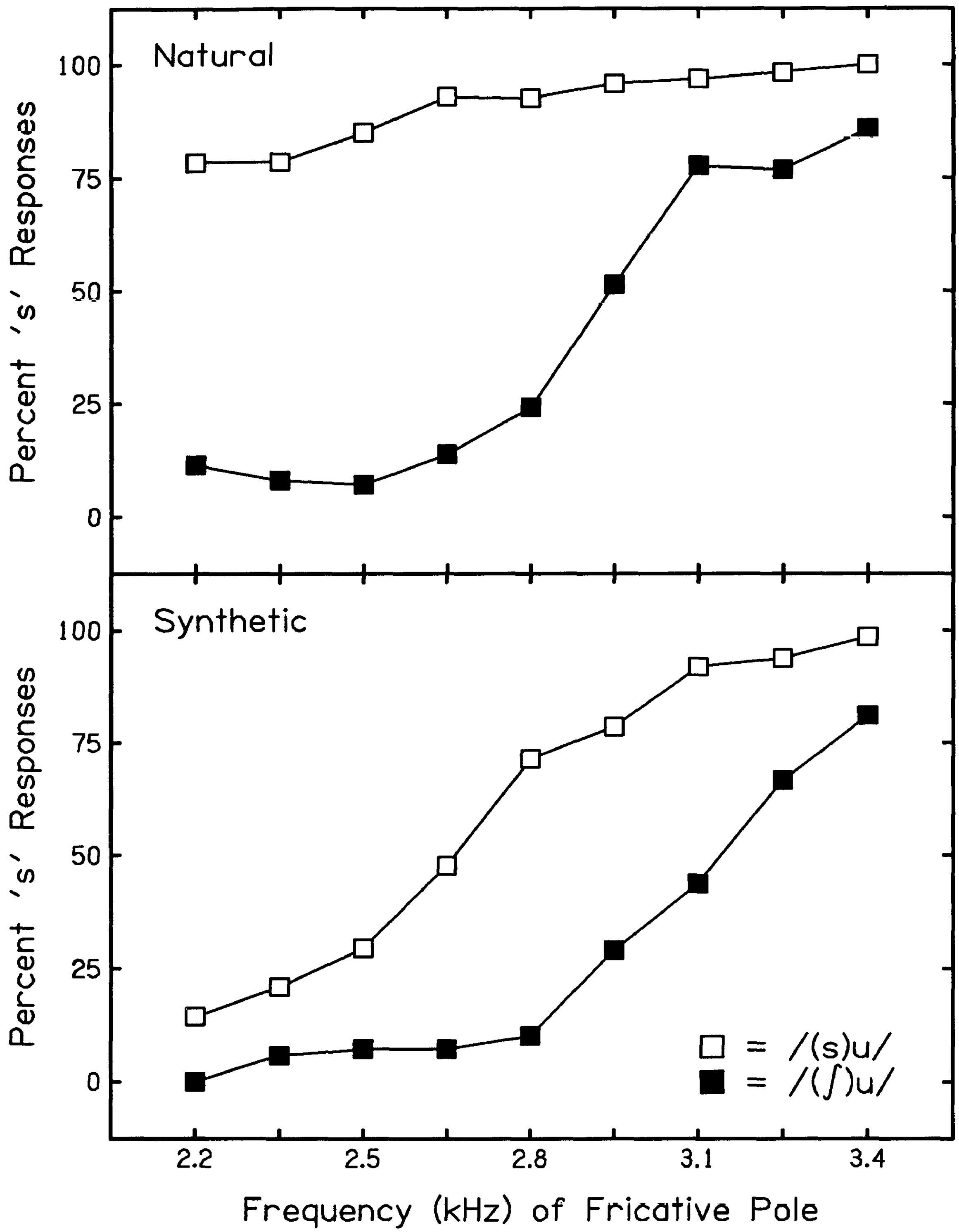


Figure 7: Identification functions for 3 & 1/2-year-olds from the present study for natural and synthetic stimuli.

A three-way ANOVA was performed on the combined results from the natural and synthetic stimuli, with age as the between-subjects' factor and stimulus set (natural or synthetic stimuli) and transition as the within-subjects' factors. Because only 11 children participated in both the natural and synthetic identification tasks in the present study, only the data from those 11 were included in this analysis. The main effects of stimulus set and transition were found to statistically significant: [$F(1,25)=7.47$, $p=0.011$] and [$F(1,25)=137.25$, $p<0.001$], respectively. Three two-way interactions were also found to be significant: Age x Stimulus Set [$F(1,25)=9.00$, $p=0.006$], Age x Transition [$F(1,25)=8.83$, $p=0.007$], and Stimulus Set x Transition: [$F(1,25)=13.16$, $p=0.001$]. Of most interest to the present study was the finding of a significant Age x Stimulus Set interaction because it indicates that the pattern of results across the two stimulus sets differed for the two listener groups.

Simple Effects Analysis was done on the phoneme boundaries for natural and synthetic stimuli as a way of investigating the pattern of results for each listener group across the two sets of stimuli. This analysis looked at the effects of stimulus set and transition for each listener group separately. Results revealed that only 3 & 1/2-year-olds displayed a significant effect of stimulus set [$F(1,25)=16.43$, $p<0.001$], indicating that these children demonstrated higher mean phoneme boundaries (across the two transition conditions) for the synthetic than for the natural stimuli. This result was most likely due to the fact that the /s/ transition in the synthetic stimuli did not result in the extremely low-frequency phoneme boundaries for children observed for this transition condition in the natural stimulus set. In support of this suggestion is the finding of a significant Stimulus Set x Transition interaction for children [$F(1,25)=11.52$, p=0.002]. Thus, the F2 transition had different effects on children's phoneme boundaries for the two stimulus sets.

*Table IV:* Mean slopes for the natural and synthetic stimulus sets

| | **Slopes** | | | |
|---|---|---|---|---|
| **transition:** | /s/ | | /ʃ/ | |
| | **3-yr-olds** | **adults** | **3-yr-olds** | **adults** |
| natural stimuli: | 2.17 | 4.26 | 3.64 | 6.29 |
| synthetic stimuli: | 4.67 | 5.42 | 3.38 | 6.66 |

Table IV shows the slopes for children and adults for both the natural and synthetic stimuli. A two-way ANOVA was performed on the slope data for the synthetic stimuli, with age as the between-subjects' factor and transition as the within-subjects' factor. The main effect of age was found to be statistically significant [$F(1,28)=14.94$, $p<0.001$], suggesting that children did not weight the fricative noise to the same extent that adults did. In addition, the Age x Transition interaction was statistically significant [$F(1,28)=10.49$, $p=0.003$], reflecting the fact that adults demonstrated shallower slopes for the /s/ than for the /ʃ/ transition condition, (as both listener groups had for the natural stimuli), but that 3 & 1/2-year-olds demonstrated the opposite pattern for synthetic stimuli (i.e., shallower slopes for /ʃ/ than for /s/). That is, adults demonstrated shallower slopes for /s/ than for /ʃ/ for both stimulus sets. Children, on the other hand, demonstrated shallower slopes for /s/ than for /ʃ/ for the natural stimuli, but the opposite transition-related pattern of slopes for the synthetic stimuli.

A three-way ANOVA was performed on the combined slope data for the natural and synthetic stimuli, with age as the between-subjects' factor and stimulus set and transition as the within-subjects' factors. The main effects of age and transition were found to be statistically significant: [$F(1,25)=18.21$, $p<0.001$] and [$F(1,25)=10.70$, $p=0.003$], respectively. Finally, the Age x Transition interaction was significant [$F(1,25)=5.37$, $p=0.03$]. This interaction reflects the fact that, if mean slopes are computed for each transition condition across the two stimulus sets, children's slopes look similar for /s/ and /ʃ/ transition conditions (due to the reversal in patterns across the natural and synthetic stimulus sets), but adults clearly demonstrate shallower slopes for /s/ than for /ʃ/ for both stimulus sets.

## DISCUSSION

Nittrouer (1992) found that children demonstrated greater separation in phoneme boundaries and shallower slopes than adults, which was taken as evidence that they weighted the F2 transition more heavily than adults and weighted the fricative noise less heavily in making decisions of fricative identity. One objective of the present study was to see if a similar pattern of results could be obtained. If the patterns of phoneme boundaries and slopes obtained from the present study were similar to those obtained by Nittrouer, it could be concluded that those earlier results are reliable, lending further support for the conclusions reached regarding children's versus adults' relative weighting of the acoustic parameters of the speech signal for the purpose of making phonemic decisions.

As described previously, several procedural aspects of the present study differed from Nittrouer's (1992) study. In spite of these procedural differences, results from the two studies (for stimuli with natural vocalic portions) are remarkably similar. As in Nittrouer's study, children in the present study demonstrated greater separation in phoneme boundaries and shallower slopes than those demonstrated by the adults. In both studies, the 3 & 1/2-year-olds displayed lower phoneme boundaries than adults for the /(s)u/ syllables, which reflects their greater weighting of the F2 transition relative to adults. As shown in Table I, children in the present study needed the fricative noise to be 1129 Hz higher in the /ʃ/ transition condition than in the /s/ transition condition in order to arrive at an /s/ response. In other words, they needed 1129 Hz of fricative noise to counteract the /ʃ/ transition and identify the syllables as starting with /s/. Table I also shows that adults needed the fricative noise to be only 617 Hz higher in the /ʃ/ than in the /s/ condition to arrive at an /s/ response. The data from Nittrouer's 1992 study reflect similar separations in phoneme boundaries for the two transition conditions; specifically, a 1334-Hz separation for children's functions and a 646-Hz separation for adults' functions. Slopes obtained from Nittrouer's study and from the present study also reveal similar patterns of results: children's slopes are shallower than those of adults. This finding reflects the children's lesser weighting of the fricative noise during the identification task, relative to the adults' weighting of that parameter. (However, both the 3 & 1/2-year-olds' and adults' slopes from the current study are steeper than those of the 1992 study, due to the presentation of fewer stimuli.) These similarities in phoneme boundaries and slopes provide evidence that Nittrouer's data are reliable, and once again indicates that, relative to adults, children weighted the F2 transition more heavily, and the fricative noise less heavily, when making phonemic decisions. In the identification tasks with natural vocalic portions, then, children seemed to use the dynamic information pertaining to overall syllable production to a relatively greater extent than the adults, while using the static information specifically associated with the shape of the fricative constriction to a lesser extent.

The possibility that the phoneme boundaries from the present study might be generally lower than those obtained in 1992, due to the use of a truncated fricative-noise continuum, was not confirmed. The similarity in location of phoneme boundaries from both studies reveals the robustness of the boundary, in spite of the differences in manner of stimulus presentation, number of stimuli presented, and the stimuli themselves (i.e., step-size and range of the fricative-noise continua).

As part of her 1992 study, Nittrouer attempted to determine whether this greater weighting of the dynamic information by children indicated that the organizational unit in children's speech perception more closely approximates the syllable than the phoneme. Children's and adults' weighting of both intersyllabic and intrasyllabic formant transitions was measured. If the major organizational unit within children's speech perception is the syllable, it was speculated that their weighting of cross-syllabic formant transitions would be less than that of the adults. The children did demonstrate smaller effects of the formant transitions that crossed a syllable boundary than adults did. This, along with the children's greater weighting of the intrasyllabic formant transitions, was taken as evidence that the organizational unit in children's speech perception is probably more appropriately described as a syllabic unit, rather than as a phonemic unit. The findings from this study provide support for those earlier conclusions, namely, that children's weighting schemes during speech identification tasks differ from adults', and that those weighting schemes highlight the dynamic components of syllable production.

Jusczyk (1993) also asserts that children's early perceptual units are syllable-sized, and that their weighting schemes differ from those of adults. He proposes that this "global representation" of the components of the speech stream is adequate during the early stages of language acquisition. This is because, at an early age, the number of words in a child's lexicon is limited by his or her limited linguistic experience, and so the storage of learned words in memory can be based upon a less-detailed representation than is required by adults (i.e., the syllable). As linguistic experience (and thus, the lexicon) grows, representations based upon finer details (i.e., details that provide information about intrasyllabic components) are required in order to accomodate the increasing number of words to be stored. As greater attention to the details of the speech signal is required, a child's weighting scheme will adjust.

The second objective of this study was to determine whether 3 & 1/2-year-olds and adults would display similar phoneme boundaries and slopes for stimuli with natural and synthetic vocalic portions. One possibility was that both groups, or just the children, might not bring the speech-processing strategies used in identification of natural stimuli to bear on the identification of synthetic stimuli. Conversely, the possibility existed that they would use the processing strategies used in speech perception with these synthetic stimuli, and so would display even greater effects of the F2 transition in making phonemic judgments, due to the more extreme values of the onset of F2 in the synthetic vocalic portions.

Results for stimuli with synthetic vocalic portions suggest that the children, but not the adults, used a different processing strategy for the synthetic than for the natural stimuli. While children's phoneme boundaries for the natural stimuli reflect their greater weighting of the F2 transition relative to the adults', the phoneme boundaries obtained from their responses to the synthetic stimuli are very similar to those of the adults. It might have been predicted that all listeners would have weighted F2 transitions more for the synthetic than for the natural stimuli. However, children weighted the F2 transition for the synthetic stimuli in this study in a manner that corresponds to the adults' weighting of this acoustic cue. At the same time, the finding of shallower slopes for children's identification functions, compared to adults, for synthetic stimuli indicates that children were not weighting the fricative-noise spectrum as greatly as adults did in making decisions of fricative identity.

It is tempting to speculate that children's processing of the synthetic stimuli might have attained mature patterns sooner than their processing of natural stimuli, given the greater similarity in children's and adults' results for the synthetic than for the natural stimuli. On the other hand, it does not seem reasonable to suggest that the less-natural speech stimuli would evoke more advanced speech-processing strategies. Therefore, another possibility is that the children had difficulty processing these synthetic stimuli using the weighting schemes they normally reserve for the processing of speech, but that their somewhat 'nonspeech' strategies led to results that resembled those of adults. A similar suggestion is offered by Jusczyk and Derrah (1987) as a possible explanation for why infants' discrimination of speech stimuli might display similar patterns to those of adults. Those authors caution that " ... although the experimenter may be able to provide a description of the results in terms of phonetic features, this does not mandate that the infant has a phonetic feature-based representation. Instead, there may be other descriptions, compatible with the data, that better characterize the infant's responses" (p. 653).

Although still speculative, it is possible that the source of the seemingly contradictory results for speech perception tasks with infants and young children is related to the weighting schemes for acoustic parameters used by the two groups. Perhaps infants under the age of 6 months use a sort of default weighting scheme, not specific to speech stimuli. Young children, on the other hand, might use a weighting scheme specific to speech, but one that has not yet achieved mature dimensions. Regardless of whether or not the weighting schemes used by children for speech stimuli turn out to be specific to those stimuli, one conclusion does seem clear: the patterns of results observed for the natural stimuli in this study and in others do not simply reflect immature psychoacoustic capacities. That is, in studies with natural vocalic portions, children have not simply been weighting the fricative noise less heavily than adults because they are unable to resolve spectral differences in noise as well as adults. Rather, it is clear that children must have been directing their attention away from those components of the syllables.

All results presented here are consistent with a view of speech perception suggesting that listeners treat speech stimuli differently from how they treat other auditory stimuli. The specific suggestion being made is that the weighting schemes for the acoustic components of speech and non-speech signals differ. In listening to speech, competent language users appear to weight the various components of the signal in ways that they have learned will lead to correct judgments about phonemic identity in their native language. For children at the earliest stages of linguistic development, the weighting scheme used in speech perception may be one that facilitates the parsing of the acoustic signal into syllabic units: namely, intrasyllabic formant transitions. This scheme might be the most efficient for an individual trying to discover the word boundaries in an unfamiliar and ambiguous signal. Gradually then, the child becomes more skilled at recognizing word boundaries. At the same time, the need also increases for more detailed representations of the internal structure of those words. Consequently, the perceptual weighting scheme for speech shifts to emphasize those more static components of the signal that will support phoneme recognition in the child's native language.

# REFERENCES

Berg, B.G. (1989). Analysis of weights in multiple observation tasks. Journal of the Acoustical Society of America, 86, 1743-1746.

Cooper, F.S., Delattre, P.C., Liberman, A.M., Borst, J.M., & Gerstman, L.J. (1952). Some experiments on the perception of synthetic speech sounds. Journal of the Acoustical Society of America, 24, 597-606. also in Lehiste, I. (Ed.), Readings in Acoustic Phonetics (pp. 273-282). Cambridge, MA: M.I.T. Press.

Daniloff, R.G., & Hammarberg, R.E. (1973). On defining coarticulation. Journal of Phonetics, 1, 239-248.

Eilers, R.E., Wilson, W.R., & Moore, J.M. (1977). Developmental changes in speech discrimination in infants. Journal of Speech and Hearing Research, 20, 766-780.

Eimas, P.D., & Miller, J.L. (1980). Contextual effects in infant speech perception. Science, 209, 1140-1141.

Finney, D.J. (1964). Probit analysis. Cambridge: Cambridge University Press.

Fodor, J.A., Garrett, M.F., & Brill, S.L. (1975). Pi ka pu: The perception of speech sounds by prelinguistic infants. Perception and Psychophysics, 18, 74-78.

Fox, B., & Routh, D.K. (1975). Analyzing spoken language into words, syllables, and phonemes: A developmental study. Journal of Psycholinguistic Research, 4, 331-342.

Greenlee, M. (1980). Learning the phonetic cues to the voiced-voiceless distinction: A comparison of child and adult speech perception. Journal of Child Language, 7, 459-468.

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

House, A.S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. Journal of the Acoustical Society of America, 25, 105-113. also in Lehiste, I. (Ed.), Readings in Acoustic Phonetics (pp. 128-136). Cambridge, MA: M.I.T. Press.

Jastak, S., & Wilkinson, G.S. (1984). The Wide Range Achievement Test-Revised. Wilmington, DE: Jastak Associates

Jusczyk, P.W. (1986). Toward a model of the development of speech perception. In J.S. Perkell & D.H. Klatt (Eds.), Invariance and variability in the speech processes (pp. 1-19). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jusczyk, P.W. (1992). Developing phonological categories from the speech signal. In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), Phonological Development: Models, Research, Implications (pp. 17-64). Timonium, MD: York Press.

Jusczyk, P.W. (1993). From general to language-specific capabilities: The WRAPSA model of how speech perception develops. Journal of Phonetics, 21, 3-28.

Jusczyk, P.W., & Derrah, C. (1987). Representation of speech sounds by young infants. Developmental Psychology, 23, 648-654.

Jusczyk, P.W., Pisoni, D.B., Reed, M.A., Fernald, A., & Myers, M. (1983). Infants' discrimination of the duration of a rapid spectrum change in nonspeech signals. Science, 222, 175-177.

Jusczyk, P.W., Rosner, B.S., Cutting, J.E., Foard, C.F., & Smith, L.B. (1977). Categorical perception of nonspeech sounds by 2-month-old infants. Perception and Psychophysics, 21, 50-54.

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. Journal of the Acoustical Society of America, 73, 322-335.

Krause, S.E. (1982). Vowel duration as a perceptual cue to postvocalic consonant voicing in young children and adults. Journal of the Acoustical Society of America, 71, 990-995.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of the speech code. Psychological Review, 74, 431-461.

Liberman, A.M., Delattre, P.C., & Cooper, F.S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. Language and Speech, 1, 153-167.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? Journal of the Acoustical Society of America, 57, 1547-1551.

Mack, M., & Blumstein, S.E. (1983). Further evidence of acoustic invariance in speech production: The stop-glide contrast. Journal of the Acoustical Society of America, 73, 1739-1750.

Miller, J.L., & Eimas, P.D. (1983). Studies on the categorization of speech by infants. Cognition, 13, 135-165.

Morrongiello, B.A., Robson, R.C., Best, C.T., & Clifton, R.K. (1984). Trading relations in the perception of speech by five-year-old children. Journal of Experimental Child Psychology, 37, 231-250.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. Journal of Phonetics, 20, 1-32.

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.

Parnell, M.M., & Amerman, J.D. (1978). Maturational influences on perception of coarticulatory effects. Journal of Speech and Hearing Research, 21, 682-701.

Pisoni, D.B. (1985). Speech perception: some new directions in research and theory. Journal of the Acoustical Society of America, 78, 381-388.

Repp, B.H. (1981). On levels of description in speech research. Journal of the Acoustical Society of America, 69, 1462-1464.

Simon, C., & Fourcin, A.J. (1978). Cross-language study of speech-pattern learning. Journal of the Acoustical Society of America, 63, 925-935.

Stevens, K.N., & Blumstein, S.E. (1978). Invariant cues for place of articulation in stop consonants. Journal of the Acoustical Society of America, 64, 1358-1368.

Tallal, P. (1980). Perceptual requisites for language. In R.L. Schiefelbusch (Ed.), Nonspeech language and communication: Analysis and Intervention (pp. 450-467). Baltimore: University Park Press.

Wardrip-Fruin, C., & Peach, S. (1984). Developmental aspects of the perception of acoustic cues in determining the voicing feature of final stop consonants. Language and Speech, 27, 367-379.