
Developmental Effects of Multiple Looks in Speech Sound Discrimination

Rachael Frush Holt
Arlene Earley Carney
University of Minnesota

Purpose: The change/no-change procedure (J. E. Sussman & A. E. Carney, 1989), which assesses speech discrimination, has been used under the assumption that the number of stimulus presentations does not influence performance. Motivated by the tenets of the multiple looks hypothesis (N. F. Viemeister & G. H. Wakefield, 1991), work by R. F. Holt and A. E. Carney (2005) called this assumption into question (at least for adults): Nonsense syllable discrimination improved with more stimulus presentations. This study investigates the nature of developmental differences and the effects of multiple stimulus presentations in the change/no-change procedure.

Method: Thirty normal-hearing children, ages 4.0–5.9 years, were tested on 3 consonant–vowel contrasts at various signal-to-noise ratios using combinations of 2 and 4 standard and comparison stimulus repetitions.

Results: Although performance fell below that which is predicted by the multiple looks hypothesis in most conditions, discrimination was enhanced with more stimulus repetitions for 1 speech contrast. The relative influence of standard and comparison stimulus repetitions varied across the speech contrasts in a manner different from that of adults.

Conclusion: Despite providing no additional sensory information, multiple stimulus repetitions enhanced children's discrimination of 1 set of nonsense syllables. The results have implications for models of developmental speech perception and assessing speech discrimination in children.

KEY WORDS: multiple looks, children, speech perception, discrimination

Developmental changes in speech perception appear to parallel basic changes in psychoacoustic properties, such as temporal, frequency, and intensity resolution. Temporal and frequency resolution appear to mature by about 6 months of age—earlier than intensity processing, which develops into the school-age years (Werner, 2002). For example, young infants are comparable to adults at following amplitude modulations (Werner, 1996) and infants' psychophysical tuning-curve widths are adultlike by 6 months of age (Olsho, 1985; Spetner & Olsho, 1990). Although these data suggest mature frequency and temporal resolution by the middle of the first year of life, other investigators, using a different psychophysical procedure than Olsho and colleagues, have reported immature auditory filter widths in 3- and 4-year-olds (e.g., Allen, Wightman, Kistler, & Dolan, 1989; Hall & Gross, 1991). Hall and Gross suggested that the apparent discrepancy seems to be due to poor processing efficiency rather than immature frequency selectivity on the part of 4-year-old children. *Processing efficiency* refers to an ability of the central auditory system to extract signals from noise. The *processing efficiency hypothesis* (Hill, Hartley, Glasberg, Moore, & Moore, 2004) suggests that aspects of young children's peripheral auditory systems are mature but that their central auditory systems are still developing the ability to

extract signals from noise efficiently such that they need more advantageous signal-to-noise ratios (SNRs) than adults. Therefore, it appears that frequency resolution is mature by 6 months of age but that nonsensory processes (such as processing efficiency) can influence the outcome of the psychophysical procedure and ultimately what children perceive.

In contrast to temporal and frequency resolution, absolute sensitivity is not mature at all audiometric frequencies until 10 years of age (Werner, 2002). The largest gains are made in the first 6 months of life, with high-frequency thresholds approaching adult values before those at low frequencies (Werner & Marean, 1996). Changes in absolute sensitivity are likely due to a combination of factors: primary neural maturation of the auditory pathway before 6 months of age (Werner, Folsom, & Mancl, 1993, 1994); external ear maturation, which improves middle ear transmission efficiency in early infancy (Keefe, Bulen, Campbell, & Burns, 1994); and middle ear maturation itself, which develops into the school-age years (Keefe, Bulen, Arehart, & Burns, 1993). In fact, developmental changes in the middle ear have been reported to occur until approximately 8 years of age (Werner, 1996), whereas developmental changes in speech perception occur until at least 10 years of age (Elliott, 1986; Elliott, Longinotti, Meyer, Raz, & Zucker, 1981; Sussman & Carney, 1989) and possibly through the teenage years, especially in the presence of background noise (Johnson, 2000). Further, there is evidence to suggest that the auditory cortex develops up to 12 years of age (Moore, 2002; Moore & Guan, 2001). Thus, to some extent, changes in speech perception parallel developmental changes in the auditory system.

Development of basic neural processing in the early postnatal period likely is responsible for the early maturation of some aspects of auditory processing, such as frequency and temporal resolution. Other aspects of auditory development, such as intensity resolution, absolute sensitivity, and the processing of complex sounds (such as speech) mature later in development and depend, in part, on higher-level processing considered to be central in nature (Werner, 2002). Note that higher-level processing is required for all auditory stimuli but might play a more prominent role in complex sound processing. Higher-level auditory processing encompasses nonsensory factors that can influence speech perception, particularly in children, such as attention, motivation, memory, and response bias. These factors have long been a part of models of signal detection (Green & Swets, 1966) and more recently have been applied to children's perception of speech.

The influence of selective attention in speech perception has primarily centered on the issue of what children attend to in the speech signal. From this work,

considerable support has emerged for the notion that children and adults divide their perceptual attention differently for acoustic components of the speech signal. In her *developmental weighting shift (DWS) hypothesis*, Nittrouer (1992) proposed that the linguistic decision required in a speech perception task likely influences where a listener's perceptual attention is directed and that this can change with development: Children ages 3–7 years showed smaller effects than did adults for formant transitions that crossed syllable boundaries in a phoneme identification task. Thus, selective attention to acoustic cues in the ambient language appears to play a role in the maturation of speech perception.

The second nonsensory factor is that of *motivation*, which gives the listener incentive to respond and pay attention (Werner & Marean, 1996). In infant testing, this usually takes the form of a lighted, moving toy that testers presume is interesting to the infant (e.g., Eilers, Wilson, & Moore, 1977). With children, this usually takes the form of verbal praise. Whatever the form motivational force takes, there must be a payoff for listening.

Much of the interest in memory's influence on speech perception has come from attempts to account for the wide variability in speech and language outcomes of young children with cochlear implants. Pisoni and Geers (2000) identified one factor that accounts for some of the variability: working memory capacity. Quantifying working memory by measuring auditory digit span capacity, Pisoni and Geers reported a moderately strong relationship between measures of forward digit span and open-set word recognition scores ($r = .64$). More recently, Pisoni and Cleary (2003) partialled out other factors, statistically, that are known to influence open-set word recognition ability and reported that nearly 20% of the variance in pediatric cochlear implant word recognition performance could be explained by individual differences in cognitive factors underlying the speed and efficiency in which phonological and lexical representations of words are stored and retrieved from working memory. These types of findings were used as evidence that "processing variables" (Pisoni & Geers, 2000, p. 93) that are related to memory such as "coding, rehearsal, storage, retrieval, and manipulation of phonological representations of spoken words" (Pisoni & Geers, 2000, p. 93) influence open-set word recognition abilities. Thus, memory plays a role in the perception of speech—at least at the level of word recognition.

Response bias, or one's tendency to respond, is the final nonsensory influence on perception performance discussed in the literature. Psychophysicists often cite response bias as a potential contributor to the difference in detection thresholds between adults and children. However, Werner (1992) argued that even if procedures are used to control response bias, threshold differences

are still found among infants, children, and adults. Further, response bias can be controlled by using a performance measure of sensitivity not influenced by bias, such as d' . Numerous investigators have demonstrated age-related changes in threshold, even when using bias-free estimates of sensitivity (Trehub, Schneider, Thorpe, & Judge, 1991; Werner & Mancl, 1993).

The purpose of this discussion was to highlight the fact that both sensory and nonsensory factors influence speech perception testing in children (and infants and adults, as well). Further, as processing becomes more complex for stimuli such as speech, higher-level nonsensory factors likely play an even greater role in perception (Werner, 2002). In this investigation, we attempted to limit the effects of some of these nonsensory factors while examining the effects of one nonsensory factor that we recently identified—the salience or strength of the internal representation of speech sounds—on speech discrimination.

Examination of the strength of the internal representation of speech sounds grew out of previous work in which we attempted to bridge the gap between findings from psychoacoustic research and those from speech perception experiments, an effort that often has been met with limited success. We investigated the effects of multiple looks on speech sound discrimination in normal-hearing adults (Holt & Carney, 2005). The *multiple looks hypothesis*, proposed by Viemeister and Wakefield (1991), is a promising theoretical approach from the psychoacoustic literature for speech perception research. The hypothesis was developed to explain the contradictory findings that there are two time constants involved in temporal integration or summation (the observation that detection threshold decreases with longer stimulus durations). Specifically, the hypothesis proposes that rather than working as long-term integrators, the auditory system's filters sample the incoming signal at a relatively fast rate. Each sample provides the system a 3- to 5-ms "look" at the signal. "These samples or 'looks' are stored in memory and can be accessed and processed selectively" (Viemeister & Wakefield, 1991, p. 864). Long-duration signals provide the system more opportunities to be sampled. Therefore, the longer a stimulus, the more "looks" there will be at the stimulus and the more likely that at least one look will be above the listener's detection threshold. This results in lower detection thresholds for longer stimuli relative to shorter stimuli. The multiple looks hypothesis predicts that if the looks are independent and the information is combined optimally, d' for two pulses should be 1.4 (square root of 2) times larger than d' for a single pulse.

Holt and Carney (2005) extended the multiple looks hypothesis to longer stimuli in a speech discrimination task with 14 normal-hearing adults. We used the

change/no-change procedure (Sussman & Carney, 1989) in which a series of standard and comparison nonsense speech stimuli are presented to a listener who is to detect a change in the stimulus array. During a change trial, the stimuli change midway through the string; during a no-change trial, all of the stimuli are identical. For example, if the number of standard and comparison presentations is two each, a change trial might consist of the stimuli /pa/ /pa/ /ta/ /ta/ and a no-change trial might consist of /pa/ /pa/ /pa/ /pa/. Adults' discrimination of three pairs of synthetic nonsense syllables (/pa/ vs. /ta/, /ra/ vs. /la/, and /sa/ vs. /ja/) was assessed at SNRs of -10 , -8 , -6 , and -4 dB using one, two, and four repetitions of standard and comparison stimuli. The 108 conditions were pseudo-randomly presented to each listener: Conditions were blocked by syllable pair, and within each syllable pair, repetition conditions were blocked by SNR. The syllable pair, SNR, and repetition conditions were randomly presented. Within each syllable comparison, half of the listeners received one stimulus in the pair as the standard; the other half of the participants received the alternate stimulus as the standard.

Each condition consisted of 50 trials—25 change trials and 25 no-change trials. Listeners were tested in a double-walled sound booth with the overall level of the speech presented at 68 dB SPL from loudspeakers placed at $\pm 45^\circ$ azimuth relative to the listener. At the beginning of each visit, participants completed ten 2:2 (in which the numbers preceding and following the colon indicate the number of standard and comparison repetitions, respectively) trials in quiet to familiarize them with the stimuli and to ensure task understanding. No feedback was given during either the familiarization trials or the test phase. Participants were told how many repetitions they would hear of the identified standard and comparison stimuli in each condition to reduce any auditory uncertainty. They were instructed to touch one side of a touch screen monitor labeled "change" if they perceived a change in the string of syllables and to touch the side labeled "no-change" if they did not perceive a change in the string.

We reported that discrimination performance improved with more stimulus presentations. Performance varied across syllable pair, with the best performance achieved with the fricative pair, followed by the liquid and stop-consonant pairs, respectively. Although the multiple looks hypothesis' prediction of an increase in d' by a factor of 1.4 for every doubling in the number of looks at the stimuli was not demonstrated under every condition, the prediction was more consistent with performance under conditions using lower numbers of presentations (e.g., 1 and 2) near d' scores of 1.0 than for higher numbers of presentations (e.g., 4) and d' scores substantially above or below scores of 1.0. For example,

for the fricative contrast at -8 dB SNR (at which d' estimates were near 1.0), average d' estimates improved by a factor of 1.46, 1.19, 1.41, and 1.14 when the number of repetitions increased from 1:1 to 2:1, 1:1 to 1:2, 1:2 to 2:2, and 2:1 to 2:2, respectively. In contrast, d' estimates improved by a factor of 0.877, 0.40, 0.70, and 1.53 when the number of repetitions increased from 2:2 to 2:4, 2:2 to 4:2, 2:4 to 4:4, and 4:2 to 4:4, respectively. These low-repetition number conditions near threshold (d' values of approximately 1.0) are similar to those used by Viemeister and Wakefield (1991), in which they compared threshold for detection (near $d' = 1.0$) for 1 pulse versus 2 pulses. When similar conditions were used in our speech discrimination procedure, the multiple looks hypothesis was supported.

We are not the only investigators to suggest that the multiple looks hypothesis might play a role in higher-level auditory tasks such as discrimination and recognition. In his spectrotemporal excitation pattern (STEP) work, Moore (2003) proposed that the internal representation of an auditory stimulus in working memory can be considered a vector of samples or looks that can be analyzed, compared, and weighted intelligently within the system. Further, the internal representation might vary depending on whether the task is one of detection, discrimination, or recognition. For example, the template or internal representation of stimuli for a discrimination task might be formed by a combination of the incoming looks at the stimuli in working memory and previous experience with similar stimuli stored in long-term memory. Therefore, it is possible that more presentations of a stimulus could enhance or strengthen the stimulus representation in working memory, allowing for improved discrimination with other stimuli. Our work with adults showing improved discrimination in the absence of additional sensory information suggests that the internal representation of the stimuli is enhanced for adults with increased presentations of the nonsense speech stimuli.

Viemeister and Wakefield's (1991) hypothesis explains sensory data from adults, and our work on discrimination of speech sounds used adult listeners. Because children's auditory and speech perception skills develop through age 10 years to adult values (Elliott, 1986; Elliott et al., 1981; Sussman & Carney, 1989) and potentially into the teenage years in noise (Johnson, 2000), it is likely that the application of the multiple looks hypothesis might be different for children than for adults. The current investigation addressed the following question: Can children's discrimination be enhanced with multiple presentations of nonsense speech stimuli as it is in adults? If pediatric performance indeed improves with more looks at the stimuli, then it is prudent to examine the nature of the influence. In other words, do the looks at the standard and comparison stimuli contribute equally to discrimination performance, or is one more important

than the other? Our work with adults (Holt & Carney, 2005) suggests that for low-repetition numbers, the number of standard and comparison stimulus repetitions is equally important. However, for higher-repetition numbers, the number of comparison stimuli was more important. A proposed explanation for this finding was that memory loads increase with higher numbers of stimulus repetitions. This might force listeners to rely more on the percept formed from the most recently heard comparison stimuli than on the "older" standard stimuli (which were heard first).

Much of the work on developmental speech perception has concluded that children focus on different aspects of the acoustic signal and, thus, might approach the current discrimination task differently from adults. Early work by Parnell and Amerman (1978) and Morrongiello, Robson, Best, and Clifton (1984) suggested that there are developmental changes in the importance of acoustic cues, such as formant transitions. And, more recently, in her work on the DWS, Nittrouer and colleagues (Nittrouer, 1992; Nittrouer, Manning, & Meyer, 1993) proposed that the linguistic decision required by children in an auditory task likely influences where their perceptual attention is directed. Specifically, children preferentially weight the dynamic spectral properties of the speech signal, whereas adults weight the static properties more. Focusing on the dynamic properties of the speech signal might help the listener identify the syllabic structure of the language, whereas attention to static properties might assist in identifying the phonetic structure. In Jusczyk's (1993) word recognition and phonetic structure acquisition model of developmental speech perception, he proposed that children's attention to larger units is an adequate strategy when children's lexicons are small, but as the lexicon grows, the need to attend to finer differences at the phoneme level increases in order to accommodate the increasing number of words learned and stored. The DWS attempts to account for how children's attention to acoustic cues might change as their lexicon grows to accommodate the need for learning an increasing number of words. Because children's memory spans increase with age (Dempster, 1981), under the current procedure, children might weight the comparisons more than the standards (as the adults did when the total number of presentations was high), even in the low-repetition conditions. Alternatively, they might use a different strategy altogether than the adults, as has been found by other investigators on various kinds of speech perception tasks.

The purpose of this investigation was threefold. First, we set out to determine if the multiple looks hypothesis applies to discrimination of speech sounds with children as it does with adults. We hypothesized that discrimination ability would improve with more looks at the stimuli. Second, we aimed to examine the nature of the influence—are looks at the standard and comparison

stimuli equally important? We predicted that the number of comparison stimulus presentations would be weighted more than the number of standard stimulus presentations. Third, we aimed to identify any developmental differences in performance between children and adults. We hypothesized that children would be able to benefit from multiple stimulus presentations but to a lesser extent than the adults tested in Holt and Carney (2005) and that they would weight the standards and comparisons differently than the adults, even when other nonsensory factors are carefully controlled.

Method

General Design

Typically developing children with normal hearing were tested using the change/no-change procedure in a combined repeated measures and factorial design. Performance was measured for repetition conditions—4:4, 4:2, 2:4, and 2:2—at two SNRs that varied with the nonsense syllable comparison. Children were separated into three groups based on the syllable comparison (stimuli varied by manner of articulation across the groups): 10 were assigned to /pa/ versus /ta/, 10 to /ra/ versus /la/, and 10 to /sa/ versus /ʃa/. Within each group, the syllable comparison differed in place of consonant articulation. In addition, all four repetition conditions were repeated at a single SNR to assess test–retest reliability. This resulted in each participant completing a total of 12 conditions (4 repetition comparisons \times 1 syllable contrast \times 2 SNRs + 4 repetition conditions repeated at a single SNR). The SNR order was counterbalanced, and the repetition comparisons were presented in random order within each SNR. The design was similar to that used with the adults in Holt and Carney (2005), with the following exceptions: The adults were run in a completely repeated measures design in which all participants completed all conditions, whereas the children were separated into syllable contrast groups; the children completed a subset of the repetition comparison conditions tested in adults; and children completed fewer SNR conditions than did the adults. These differences were due to the fact that children take longer to complete the testing than adults and, thus, are unable to complete as many conditions.

Power Analysis

To determine how many trials and the number of children that would be needed to attain sufficient power, a simulation was performed using R (Ihaka & Gentleman, 1996), in which power was calculated for various combinations of number of change trials, number of participants, and nominal values of changes in d' . A nominal power value of 0.80 was selected as the desired level of power

because this makes a Type II error (accepting a false null hypothesis) 4 times as likely as a Type I error (rejecting a true null hypothesis). This is generally considered a reflection of their relative importance (Howell, 1997). Based on the simulation, this level of power (or higher) could be achieved with at least 10 participants and 25 change trials per condition, assuming a minimum d' change of 1.3 for a doubling in the number of stimulus repetitions. This allowed us to achieve the desired level of power with a slightly lowered change in d' (by a factor of 1.3) than is expected based on the multiple looks hypothesis.

Participants

Forty-five children were recruited to participate. Five children were eliminated for the following reasons: Two children served as pilot participants to determine the feasibility of the protocol; one child could not return to complete the entire protocol; another child did not pass the hearing screening because of middle ear problems and was referred to a physician for evaluation; and a fifth child was unable to understand the task in quiet and subsequently did not complete the protocol. This was the only child encountered who was unable to learn the task. After analyzing the data from the first set of 10 children assigned to the /ra/ versus /la/ contrast condition, we realized that the SNRs selected were too easy—the average d' achieved even in the easiest SNR was above 2.5. This left a total of 30 participants who were randomly assigned to one of three syllable contrast conditions, resulting in 10 children completing the protocol for each syllable pair. These children ranged in age from 4;0 to 5;11 (years;months), with a mean age of 5;2. The average age of the children in the /pa/ versus /ta/ group, /ra/ versus /la/ group, and /sa/ versus /ʃa/ group was 5;1, 5;2, and 5;3, respectively. Each child passed a bilateral hearing screening at 20 dB HL (see American National Standards Institute [ANSI], 1989) at 0.5, 1.0, 2.0, 4.0, and 8.0 kHz and 25 dB HL at 0.25 kHz. In addition, each child passed the articulation/intelligibility, language, information, sentence completion, and repeating sentences portions of the Minneapolis Public Schools' Early Childhood Screening Instrument (Minneapolis Public Schools, 1980) and the supplemental articulation screener of the Preschool Language Scale-3 (Zimmerman, Steiner, & Pond, 1992). All of the children were native English speakers.

Stimuli

The same stimuli used in Holt and Carney (2005) were used in the current investigation. These stimuli were synthesized prototypical exemplars for four of the six syllables and for the vowels of the remaining two stimuli. The synthesized syllables were /pa/, /ta/, /ra/, and

/la/, whereas /sa/ and /ʃa/ were created by splicing a synthetic /a/ to digitized natural tokens of /s/ and /ʃ/. This was done because it was difficult to synthesize fricative-vowel stimuli that sounded natural. A detailed description of the creation of the stimuli and their respective waveforms and spectrograms are provided in Holt and Carney. These stimuli were selected for several reasons. First, nonsense syllables were used because we did not want lexical knowledge to influence performance on the discrimination task. Second, /sa/ and /ʃa/ were selected because these fricatives have been used extensively in Nittrouer's work on the DWS (Nittrouer, 1992, 1996; Nittrouer et al., 1993; Nittrouer & Miller, 1997). Nittrouer and colleagues have demonstrated that these fricatives are particularly sensitive to developmental changes in speech perception. Third, all of the syllable contrasts vary in place of articulation, which causes the most errors in perception for individuals with normal hearing (Miller & Nicely, 1955) and those with hearing loss (Bilger & Wang, 1976) while varying in manner of articulation (stop consonants, liquids, and fricatives). These tokens will be used in the future with listeners with hearing loss. Therefore, it was important to have baseline data from normal-hearing individuals.

The noise used in each syllable pair condition consisted of white noise shaped to match the long-term average spectrum of each syllable pair. For example, the noise used in the /ra/ versus /la/ condition was shaped to have the long-term spectrum of /ra/ and /la/. Therefore, three slightly differently shaped noises were used—one for each syllable pair.

Equipment and Calibration

E-prime software Version 1.1 (Psychology Software Tools, 2002) on a Pentium III computer ran the procedure and recorded responses. The signal from the computer's sound card was routed to a Crown D-75 amplifier and then to two GSI speakers placed at $\pm 45^\circ$ azimuth relative to the listener in a double-walled sound booth. Responses were made on a touch screen monitor placed directly in front of the child. The overall level of the speech was 68 dB SPL, with the noise level varying with SNR. Calibration was checked each day of testing. The same equipment and setup was used with the adults in our previous investigation.

Procedure

Within each participant group, the SNR order was counterbalanced, and the repetition condition order (4:4, 4:2, 2:4, and 2:2) was randomized within SNR. All repetition conditions were completed in a single SNR before moving on to the next SNR. Half of the children assigned to each syllable-pair comparison were retested

at all repetition conditions at the lower of the two tested SNRs, whereas the other half were retested at the higher of the two tested SNRs. This was done to assess test-retest reliability without completing the entire protocol twice. Half of the children in a given syllable-pair comparison condition heard one stimulus as the standard, whereas the other half heard the other stimulus in the pair as the standard. Within each tested repetition condition, children listened to 50 trials, half of which were change trials and half of which were no-change trials. These procedures mimicked those used with the adults in our previous investigation.

SNRs of -4 and $+2$ dB were chosen for the /sa/ versus /ʃa/ comparison because they were successfully used with children in a pilot study (Frush, 1999). Examination of the adults' performance-intensity functions in Holt and Carney (2005) for /ra/ versus /la/ and /pa/ versus /ta/ indicated that less adverse SNRs needed to be used for the children for these comparisons than were used for /sa/ versus /ʃa/ in order for them to have any success with the task. Two older children (8 and 9 years of age, respectively) were pilot tested at a wide range of SNRs in both the /pa/ versus /ta/ and /ra/ versus /la/ conditions. Based on their results, $+2$ and $+8$ dB SNR were selected for use in the /pa/ versus /ta/ condition, and 0 and $+6$ dB SNR were initially selected for /ra/ versus /la/. As was discussed previously, the SNRs initially selected for the liquid pair were too easy, so a different group of 10 children completed the protocol at more difficult SNRs of -2 and $+4$ dB. Even though the absolute values of the SNRs were different across syllable comparison, the difference between the two selected SNRs within each comparison was held constant at 6 dB.

Prior to testing in noise, children listened to 10 trials in quiet that corresponded to the first repetition condition that they were assigned to that day. For example, if a child's randomization called for starting with the 2:4 repetition condition and he or she was in the /ra/ versus /la/ group and /la/ was his or her assigned standard, the five change trials in quiet would be "/la/, /la/, /ra/, /ra/, /ra/, /ra/" and the five no-change trials would be "/la/, /la/, /la/, /la/, /la/." This was done to familiarize children with the stimuli and to ensure that they understood the task. No feedback was given during either the familiarization or testing phases. A similar familiarization procedure was used with adults in our previous investigation with one exception: The adults' familiarization trials were always presented in a 2:2 format, regardless of the first repetition condition to which they were assigned that day.

During each trial, two rows of pictures appeared on the touch screen monitor. The top row corresponded to the change condition, and the bottom row corresponded to the no-change condition. All of the pictures in the

bottom row were identical, whereas those in the top row changed midway through the row, depending on the repetition condition. The number of pictures corresponded directly to the repetition condition. For example, for the 2:4 condition, the top row might have two soccer balls pictured, followed by four bunches of bananas, whereas the bottom row would have six soccer balls pictured. Five potential combinations of response pictures were available, although only one set was used at a time within each condition. This served two purposes. First, by keeping the pictures the same within a condition, the child experienced no uncertainty about which was the change row and which was the no-change row. Second, having five combinations of response pictures from which to choose helped keep interest high. Children were instructed to point to the top row in which the pictures changed if they heard a change in the string of syllables. If the children heard no change in the string of syllables, they were instructed to point to the bottom row, in which the pictures did not change. In this way, no verbal response was necessary, nor was it required that children be able to read the labels on the touch screen monitor, as it was for the adults in the previous study.

After the child touched the screen, the tester was required to indicate if the answer was accepted or rejected and to record the other option (either “change” or “no-change”) as the response. This was important because children tend to be more impulsive than adults and occasionally, they would accidentally touch one response row when they meant to press the alternate row; therefore, a mechanism was needed for dealing with unintended touch screen presses. The only occasion in which the tester rejected a child’s response was if, independently, the child verbally indicated that he or she had meant to press the alternate row of pictures. This verbal indication occurred an average of 4.1 times per child over the entire course of the study, which encompassed 600 total trials.

Citing a pilot study (Frush, 1999), we demonstrated that keeping track of the number of trials remaining was important in keeping children on task and maintaining high motivation levels. To implement this in the current investigation, we modified a procedure used by Allen and Wightman (1994) in which 1 piece of a 10-piece puzzle appeared successively on the touch screen monitor after each trial. Unlike the method used by Allen and Wightman, the presentation of each puzzle piece was not used as feedback; rather, a puzzle piece appeared after each response regardless of whether it was correct. After every one or two puzzles (corresponding to 10 or 20 trials) were completed, depending on the attention level of the individual child, he or she was given a short break. During the breaks, children could play a game with the tester or could color or talk for a few minutes. Five

different puzzles (all of which were pictures of animals) were randomly used, such that within a 50-trial condition, each puzzle only appeared once. The purpose of varying the puzzles randomly was to keep interest as high as possible. Along with verbal praise after each trial, the puzzle also served as a reinforcer. Neither verbal praise nor the puzzle itself was used as feedback as to the correctness of a child’s response.

All testing took place in a double-walled sound booth. Children completed the testing in two or three visits of approximately 2 hr each. In addition to short play breaks, children had a long snack break midway through each visit.

Nonsensory factors such as motivation, attention, response bias (Werner & Marean, 1996), and memory (Pisoni, 2000; Pisoni & Cleary, 2003; Pisoni & Geers, 2000) influence child perceptual development research. In the current investigation, we motivated children with effusive praise after each response (whether correct or not) by giving them breaks after every 10–20 trials and by filling in a puzzle piece after each trial. Frush (1999) demonstrated that the first two of these types of reinforcement helped maintain children’s interest in the task (by maintaining the operant). The nature of the change/no-change procedure allows one to minimize the variable of sustained attention through cueing by either telling the listener where in the array a change will take place within a block of trials or varying where the change can take place within the block. The latter approach requires the listener to pay much more attention during each trial (as he or she would not be able to predict where the change might take place in the array), whereas the former approach leaves no ambiguity as to where the change will occur and thus, requires less attention of the individual. The first approach was used in this investigation to reduce the influence of sustained attention on the task. By providing multiple opportunities to hear a stimulus, a clear and straightforward procedure with no auditory stimulus uncertainty, and a pictorial “answer sheet” (pictures on a touch screen monitor) with the number of pictures matching the number of auditory stimuli, memory loads should be reduced (although not eliminated) using the change/no-change procedure in this investigation.

Finally, d' was used to measure performance. By taking into account hits and false alarms, d' is unbiased. It was calculated by subtracting the z score for a listener’s false alarm rate from the z score for his or her hit rate. A listener can have 25 of 25 hits and some false alarms or 25 of 25 correct rejections with many hits. This type of response pattern requires corrections in the calculation of d' . In this study (as was used in our previous study with adults), a value of 0.01 was added to all hit and false alarm rates of 0.0, and a value of 0.01

was subtracted from all hit and false alarm rates of 1.0 (similar to a procedure described by Hautus, 1995, and Miller, 1996). The maximum d' that could be achieved in this study was 4.65, which is considered by many experimenters to be an effective ceiling (Macmillan & Creelman, 2005). With these procedural controls, the influence of these nonsensory factors—motivation, attention, memory, and response bias—was expected to be reduced, leaving the nonsensory factor of multiple looks more prominent.

Results

Individual d' estimates for speech discrimination performance across test conditions for each child are displayed in Table 1. Consistent with the pediatric speech perception literature, performance was variable across

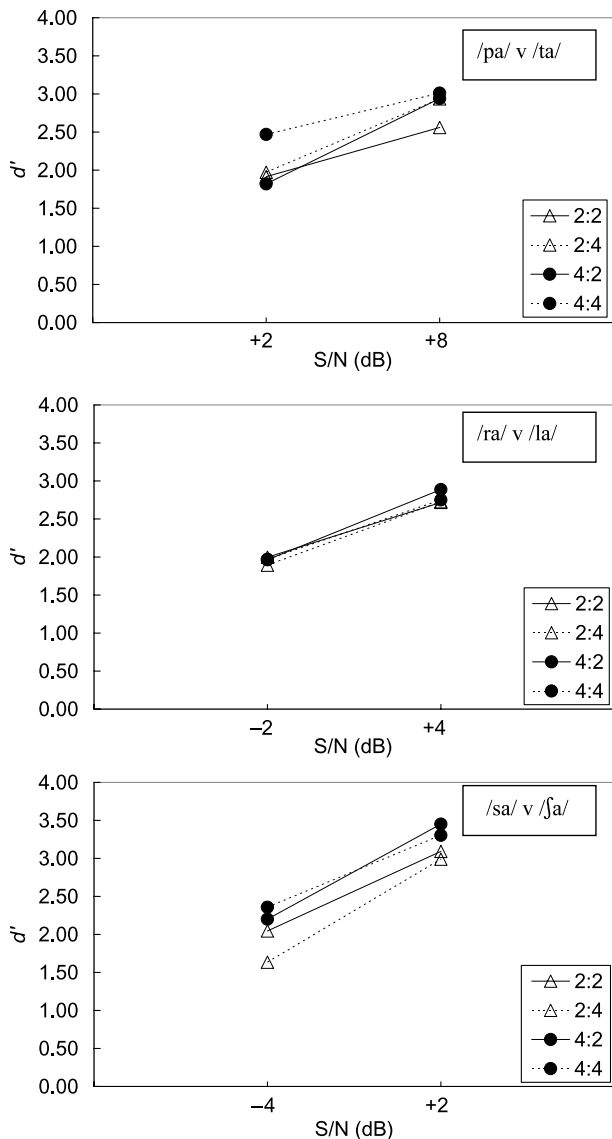
children. Figure 1 displays mean group performance by SNR. The top panel displays data for the stop-consonant contrast, the middle panel displays those for the liquid contrast, and the bottom panel displays data for the fricative contrast. The parameter in all three panels is repetition number: The open triangles and filled circles each represent a standard number of 2 and 4 repetitions, respectively, whereas the solid and dotted lines indicate a comparison number of 2 and 4 repetitions, respectively. For the stop-consonant and fricative pairs, there was a trend for increased discriminability with more looks at the stimuli. For example, for both /sa/ versus /ja/ and /pa/ versus /ta/, either the 4:4 or the 4:2 conditions resulted in the highest performance, especially at the more difficult SNRs, whereas the 2:2 and 2:4 conditions resulted in the poorest performance. For /ra/ versus /la/, performance was similar across repetition number within a given SNR.

Table 1. Individual d' estimates for the test conditions.

| Participant ID | Syllable contrast | Low SNR | | | | High SNR | | | |
|----------------|-------------------|---------|-------|------|------|----------|------|------|------|
| | | 2:2 | 2:4 | 4:2 | 4:4 | 2:2 | 2:4 | 4:2 | 4:4 |
| 1 | Stop | 4.08 | 4.08 | 3.73 | 3.50 | 4.65 | 4.65 | 4.65 | 4.65 |
| 2 | Stop | 1.99 | 1.76 | 2.33 | 4.08 | 2.33 | 2.58 | 2.68 | 2.93 |
| 3 | Stop | 4.65 | 3.16 | 2.28 | 3.73 | 3.73 | 3.73 | 4.65 | 4.65 |
| 4 | Stop | 1.99 | 2.68 | 2.25 | 2.40 | 2.75 | 3.50 | 2.25 | 3.16 |
| 5 | Stop | 1.74 | 1.33 | 0.76 | 1.02 | 1.80 | 2.68 | 2.79 | 2.38 |
| 6 | Stop | 1.46 | 1.66 | 1.50 | 2.68 | 2.58 | 3.50 | 4.08 | 3.16 |
| 7 | Stop | 1.39 | 1.99 | 0.79 | 2.18 | 2.46 | 2.91 | 3.17 | 2.33 |
| 8 | Stop | 1.23 | 1.25 | 1.35 | 1.35 | 1.70 | 1.66 | 1.46 | 2.40 |
| 9 | Stop | 0.41 | 2.02 | 2.46 | 3.50 | 2.25 | 3.50 | 2.91 | 3.73 |
| 10 | Stop | 0.20 | -0.20 | 0.76 | 0.24 | 1.35 | 0.64 | 0.76 | 0.73 |
| 31 | Liquid | 0.00 | -0.41 | 1.70 | 0.86 | 2.79 | 3.50 | 1.84 | 1.99 |
| 32 | Liquid | 3.32 | 3.73 | 2.38 | 2.28 | 3.03 | 1.70 | 2.18 | 1.74 |
| 33 | Liquid | 4.65 | 3.73 | 2.59 | 3.50 | 4.08 | 4.65 | 4.65 | 4.08 |
| 34 | Liquid | 4.08 | 4.08 | 4.08 | 3.17 | 4.08 | 4.08 | 4.08 | 4.65 |
| 35 | Liquid | 0.59 | 0.10 | 1.46 | 1.23 | 0.21 | 1.05 | 3.73 | 2.00 |
| 36 | Liquid | 3.50 | 4.65 | 2.35 | 2.02 | 3.50 | 2.79 | 3.50 | 3.50 |
| 37 | Liquid | 1.66 | 1.15 | 1.88 | 3.73 | 4.08 | 3.50 | 3.16 | 3.16 |
| 38 | Liquid | 0.14 | 0.59 | 0.82 | 0.94 | 1.25 | 1.56 | 0.66 | 0.21 |
| 39 | Liquid | 0.71 | 1.41 | 1.56 | 1.04 | 2.59 | 2.46 | 2.75 | 2.46 |
| 40 | Liquid | 1.33 | -0.10 | 0.82 | 0.94 | 1.56 | 1.99 | 2.33 | 3.73 |
| 21 | Fricative | 0.69 | 0.63 | 1.15 | 0.84 | 3.50 | 3.16 | 3.73 | 3.16 |
| 22 | Fricative | 4.08 | 4.08 | 4.08 | 4.65 | 3.50 | 3.73 | 4.08 | 3.50 |
| 23 | Fricative | 1.15 | 0.00 | 1.50 | 0.45 | 4.65 | 2.58 | 3.73 | 3.50 |
| 24 | Fricative | 1.43 | 0.72 | 1.29 | 2.81 | 1.88 | 2.58 | 2.17 | 2.17 |
| 25 | Fricative | 3.73 | 2.93 | 3.50 | 3.73 | 4.65 | 4.65 | 4.65 | 3.16 |
| 26 | Fricative | 1.33 | 1.76 | 2.58 | 2.17 | 1.76 | 1.84 | 1.99 | 3.50 |
| 27 | Fricative | 1.41 | 1.42 | 1.35 | 1.41 | 1.99 | 2.58 | 3.32 | 2.75 |
| 28 | Fricative | 4.08 | 2.17 | 3.16 | 4.08 | 4.65 | 4.08 | 4.65 | 4.65 |
| 29 | Fricative | 1.76 | 1.76 | 1.87 | 2.91 | 2.46 | 3.17 | 3.03 | 4.08 |
| 30 | Fricative | 0.79 | 0.86 | 1.53 | 0.52 | 1.87 | 1.55 | 3.16 | 2.58 |

Note. SNR = signal-to-noise ratio.

Figure 1. Mean d' performance across signal-to-noise ratio (S/N). The top panel displays mean performance for the stop-consonant contrast, /pa/ versus /ta/ ($n = 10$), the middle panel displays that for the liquid contrast, /ra/ versus /la/ ($n = 10$), and the bottom panel displays mean performance for the fricative contrast, /sa/ versus /ʃa/ ($n = 10$). Triangles and circles indicate a standard number of 2 and 4 repetitions, respectively. Solid and dotted lines indicate a comparison number of 2 and 4 repetitions, respectively.



Figures 2 through 4 display the same data shown in Figure 1 but do so using box plots for each syllable pair to show the variability in performance. Figure 2 shows data for the stop-consonant contrast, Figure 3 shows those for the liquid contrast, and Figure 4 displays data for the fricative contrast. The panels in each figure reflect each repetition condition tested—2:2, 2:4, 4:2, and 4:4. The variability across participants tended to decrease as

the SNR improved for all syllable pairs. This effect was less for /pa/ versus /ta/ than it was for other two pairs of syllables. Inspection of the box plots also assists in demonstrating the trend for the advantage of multiple presentations of the stimuli. For example, the top and bottom panels of Figure 2 display data for the /pa/ versus /ta/ comparison in the 2:2 and 4:4 conditions, respectively. Median d' in the most advantageous SNR (+8 dB) for the 2:2 condition was equal to 2.4; this value was nearly identical to the median performance ($d' = 2.5$) at the more difficult SNR (+2 dB) in the 4:4 condition. The addition of more looks (e.g., the 4:4 condition) allowed listeners to detect the difference between /pa/ and /ta/ at nearly the same rate as in the 2:2 condition, when the SNR dropped by 6 dB. Similar results can be seen for /sa/ versus /ʃa/. Performance for the liquid pair was too similar across the repetition conditions to draw these types of conclusions.

The data were entered into three separate three-way repeated measures analyses of variance (ANOVAs)—one for each syllable contrast, because the variance in performance across syllable pairs differed (evaluated by testing for equality of group variances using the Levene statistic). The factors were standard repetition number (2, 4), comparison repetition number (2, 4), and SNR (low, high). For all three syllable contrasts, the effect of SNR was significant: stop-consonants, $F(1, 9) = 40.835, p < .001$; fricatives, $F(1, 9) = 13.579, p = .005$; and liquids, $F(1, 9) = 9.815, p = .012$. For the fricative contrast, the effect of standard repetition number was significant, $F(1, 9) = 15.115, p = .004$, whereas the effect for comparison repetition number was not significant, $F(1, 9) = 0.535, p = .483$. Participants' fricative discrimination significantly improved when the number of standard repetitions increased from 2 to 4 (2:2 to 4:2 and 2:4 to 4:4). For the stop-consonant contrast, standard and comparison repetition number were not significant, $F(1, 9) = 2.129, p = .179$ and $F(1, 9) = 4.345, p = .067$, respectively, although there was a trend for participants' stop-consonant discrimination to be enhanced when the number of comparison repetitions increased from 2 to 4 (2:2 to 2:4 and 4:2 to 4:4). For the liquid contrast, neither the number of standard repetitions nor the number of comparison repetitions significantly influenced discrimination performance, $F(1, 9) = 0.065, p = .805$ and $F(1, 9) = 0.800, p = .394$, respectively. No interactions were significant for any syllable contrast.

Figure 5 shows the effect of doubling the number of comparison stimuli while keeping the number of standards constant by displaying the actual mean d' values (filled circles connected by lines) and predicted values (open circles) in each repetition condition grouped by number of standard stimuli. The predicted values were calculated by multiplying the actual performance in the 2:2 and 4:2 conditions by a factor of 1.4 (as is predicted by the multiple looks hypothesis) to predict performance in the 2:4 and 4:4 conditions, respectively.

Figure 2. Group performance for the stop-consonant contrast, /pa/ versus /ta/, in the 2:2 condition (top panel), the 2:4 condition (second panel), the 4:2 condition (third panel), and the 4:4 condition (bottom panel; $n = 10$).

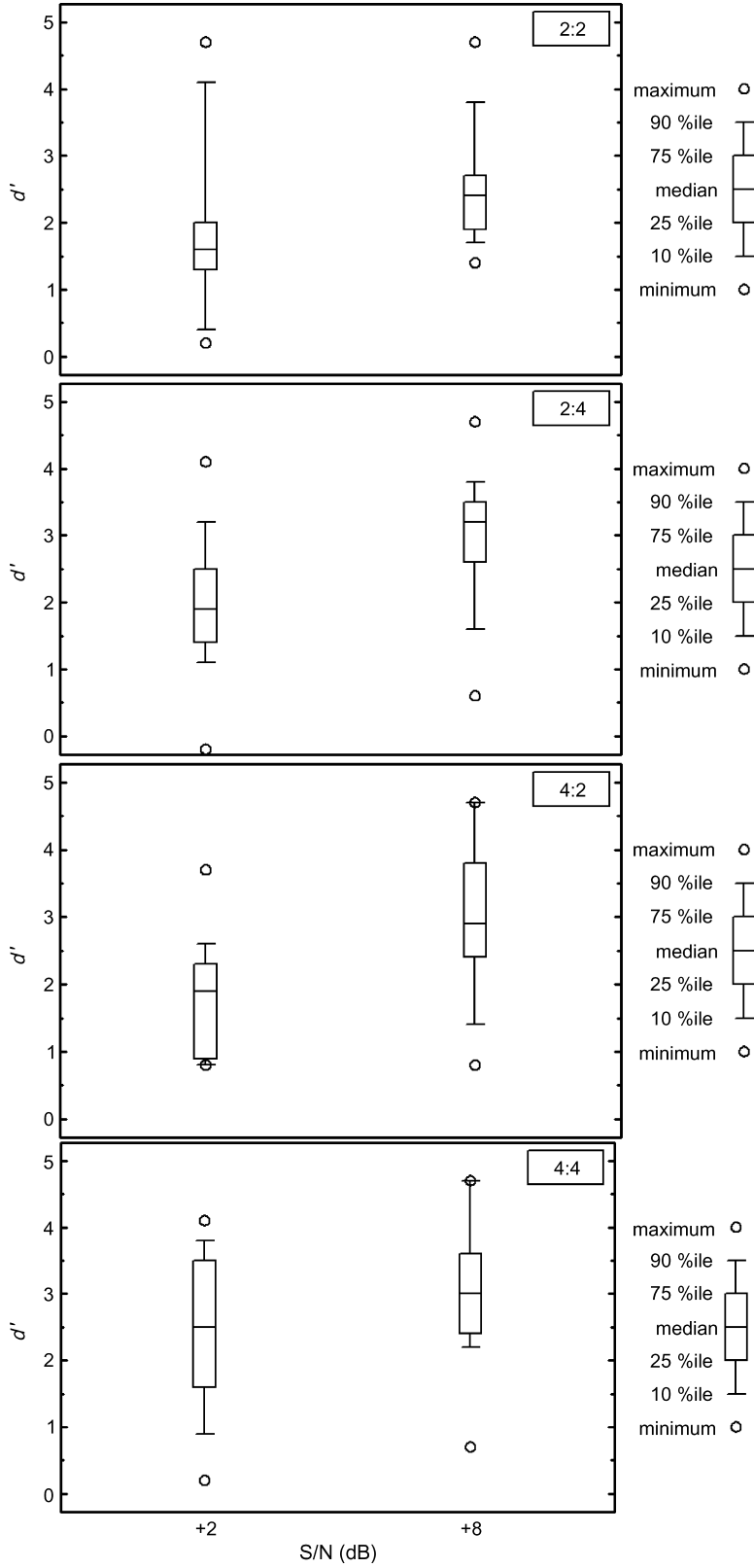


Figure 3. Group performance for the liquid contrast, /ra/ versus /la/, in the 2:2 condition (top panel), the 2:4 condition (second panel), the 4:2 condition (third panel), and the 4:4 condition (bottom panel); $n = 10$).

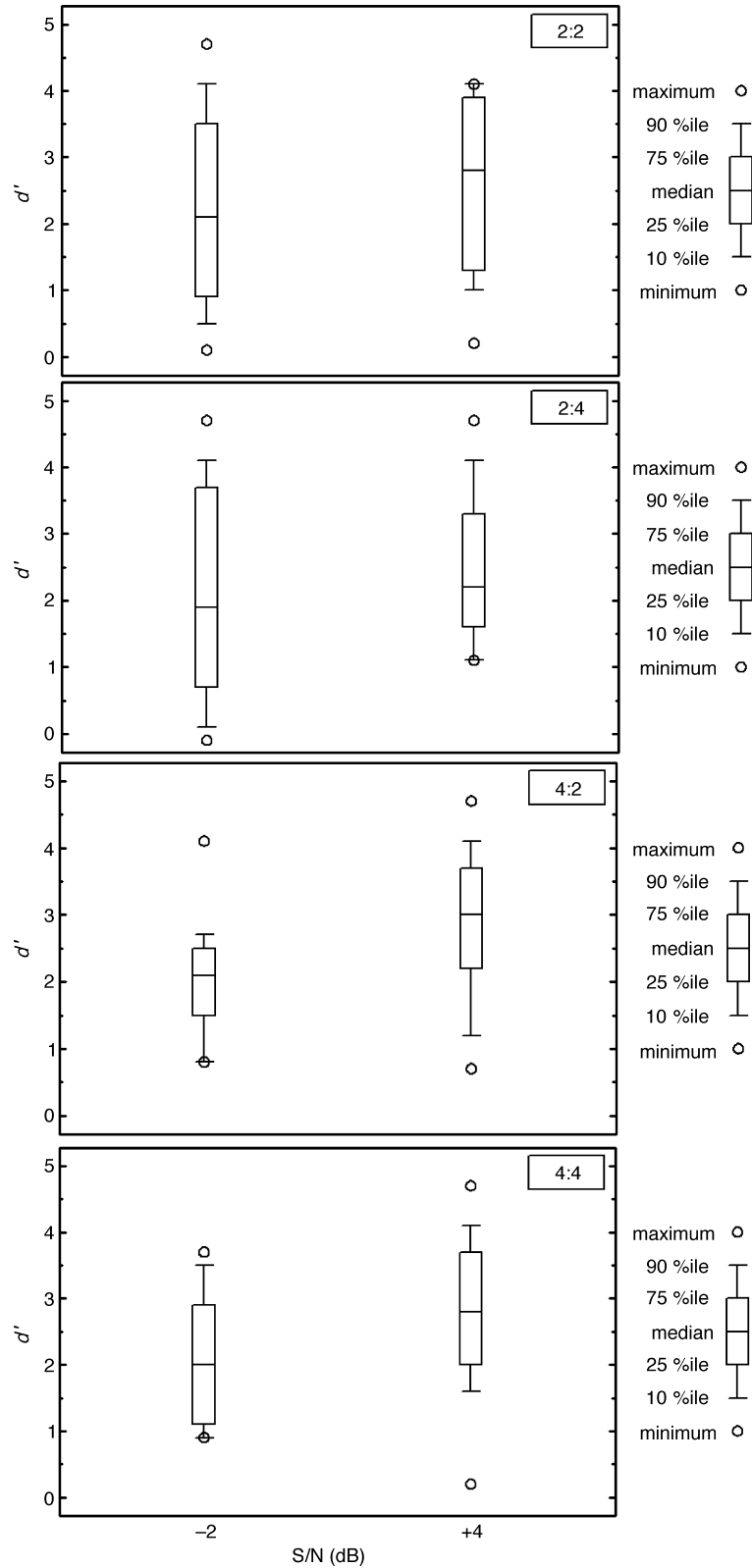


Figure 4. Group performance for the fricative contrast, /sa/ versus /ʃa/, in the 2:2 condition (top panel), the 2:4 condition (second panel), the 4:2 condition (third panel), and the 4:4 condition (bottom panel); $n = 10$).

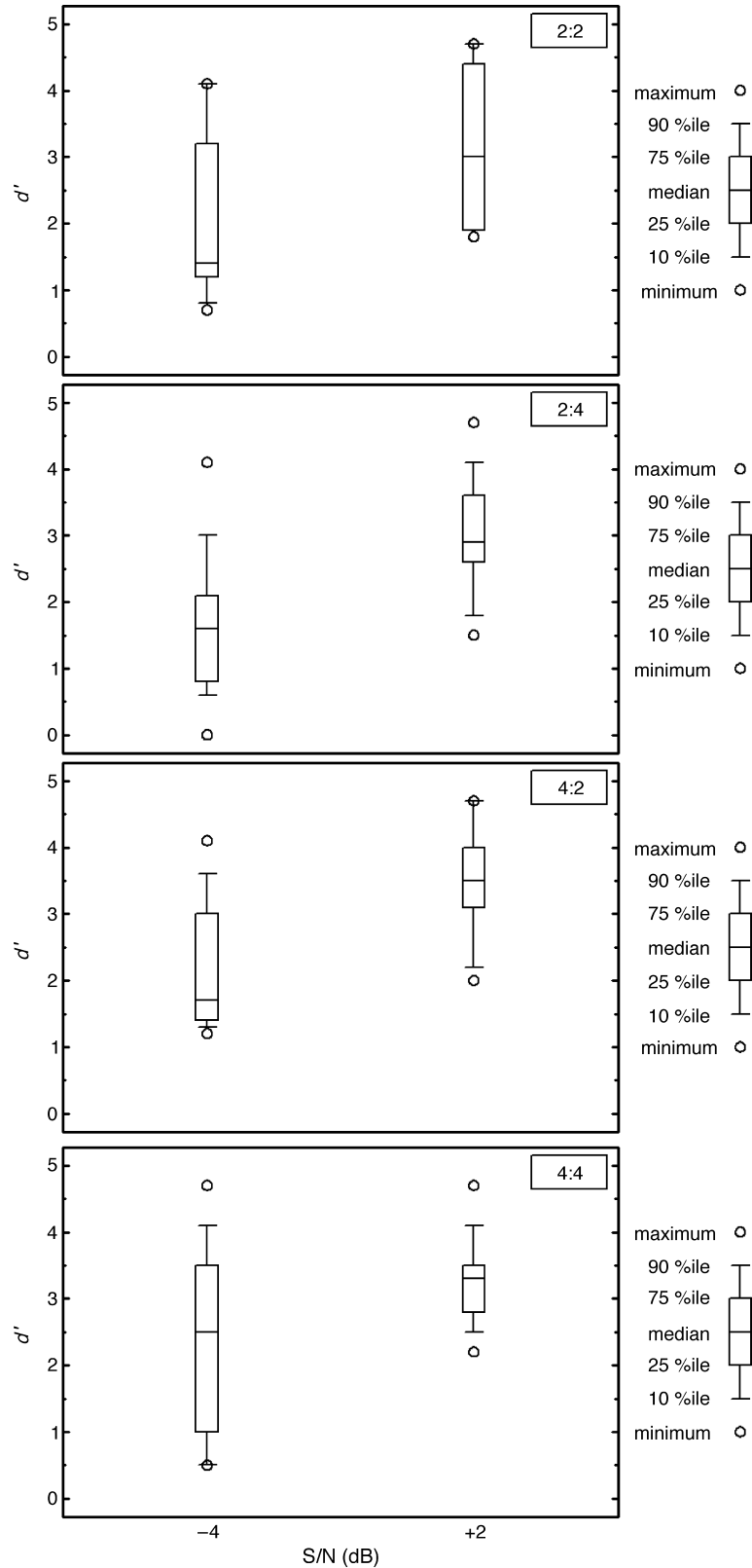
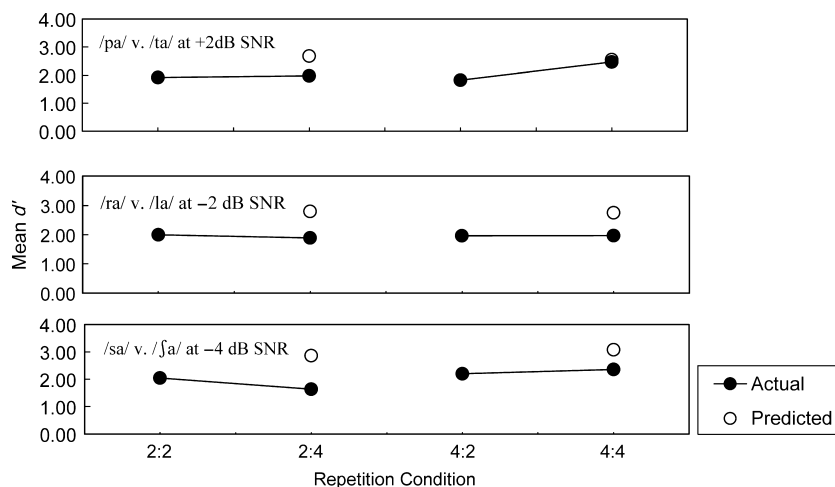


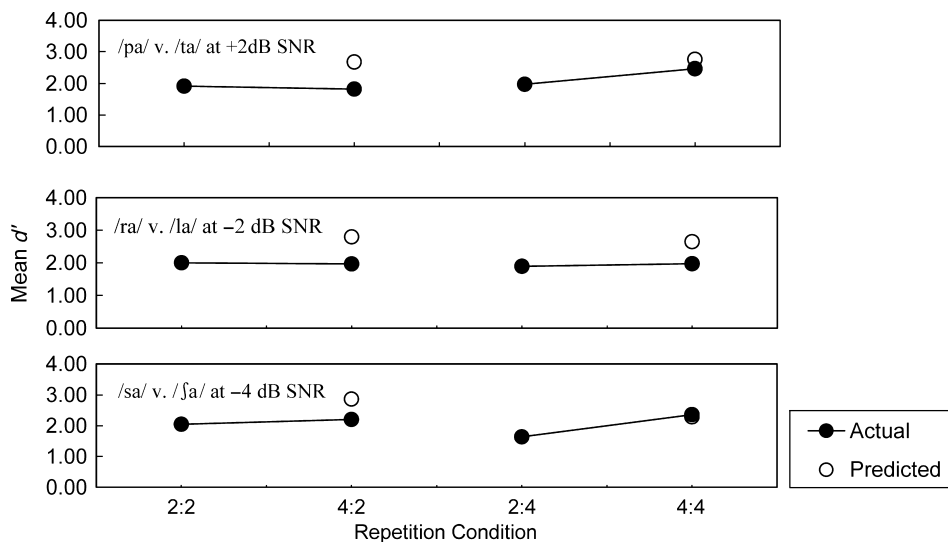
Figure 5. Actual (filled circles) versus predicted (open circles) performance (based on multiple looks), grouped by number of standard stimulus repetitions. The top panel shows performance for /pa/ versus /ta/ at +2 dB signal-to-noise ratio (SNR; $n = 10$), the middle panel shows that for /ra/ versus /la/ at -2 dB SNR ($n = 10$), and the bottom panel shows performance for /sa/ versus /ʃa/ at -4 dB SNR ($n = 10$).



Note that there are no data points for the predicted d' at 2:2 and 4:2 because these are the data points from which the predictions for 2:4 and 4:4 were made. The top panel displays data for the stop-consonant contrast, the middle panel displays data for the liquid contrast, and the bottom panel shows the data from the fricative contrast for the lowest SNR tested in each syllable pair. Each pair of data points along the abscissa represents a serial

doubling in the number of comparison repetitions, while keeping the number of standard repetitions constant. Figure 6 is set up similarly to Figure 5, except that it shows the effects of doubling the number of standard repetitions while keeping the number of comparisons constant by grouping the actual and predicted data by number of comparison stimuli. In general, actual mean performance fell below that which was predicted by the

Figure 6. Actual (filled circles) versus predicted (open circles) performance (based on multiple looks), grouped by number of comparison stimulus repetitions. The top panel shows performance for /pa/ versus /ta/ at +2 dB SNR ($n = 10$), the middle panel shows that for /ra/ versus /la/ at -2 dB SNR ($n = 10$), and the bottom panel shows performance for /sa/ versus /ʃa/ at -4 dB SNR ($n = 10$).



multiple looks hypothesis in most of the conditions tested. However, when the number of comparisons increased from 2 to 4 with four standard stimuli (going from 4:2 to 4:4), the multiple looks hypothesis explained performance for the stop-consonant contrasts well (see Figure 5). Similarly, when the number of standards increased from 2 to 4 with four comparison stimuli (going from 2:4 to 4:4), the multiple looks hypothesis explained performance for the fricative contrasts well (see Figure 6), consistent with the results of the repeated measures ANOVA for the fricative group.

Test–Retest Reliability

The mean d' difference scores between test and retest (retest score – test score) in each syllable repetition condition ranged from -0.28 to 0.25 . To assess reliability across testing sessions, Pearson correlations between test and retest data were calculated for each syllable pair separately. As with any test procedure that is administered more than once to the same group of participants, it is expected that performance at the first session will not be independent from performance at the second session. However, if the test procedure is reliable, performance at test and retest should be highly correlated (Anastasi, 1961). Table 2 presents the r values for each syllable pair's test–retest comparison along with the corresponding alpha levels. For both the stop-consonant and liquid contrasts, significant ($p \leq .05$) test–retest correlations were found in all conditions except the 4:2 condition (in which the r values were $.562$ and $.458$, respectively). For the fricative contrast, test–retest correlations in all four repetition conditions were significant ($p \leq .05$). The significant r values ranged from $.657$ to $.941$ across syllable pairs. Correlation coefficients equal to or greater than $.80$ are considered strong for real variables, whereas those near $.63$ are considered moderate in strength (Howell, 1997). These results suggest that the change/no-change procedure has good test–retest reliability (particularly for the 4:4, 2:4, and 2:2 conditions), thereby extending the findings of Carney et al. (1991) to normal-hearing children.

Table 2. Pearson correlation and alpha values for d' estimates at test and retest.

| Syllable contrast | Repetition condition | | | |
|-------------------|----------------------|--------------|--------------|---------------|
| | 2:2 | 2:4 | 4:2 | 4:4 |
| Stop | .912* (<.001) | .891* (.001) | .562 (.091) | .657* (.039) |
| Liquid | .941* (<.001) | .771* (.009) | .458 (.183) | .750* (.013) |
| Fricative | .916* (<.001) | .845* (.002) | .847* (.002) | .894* (<.001) |

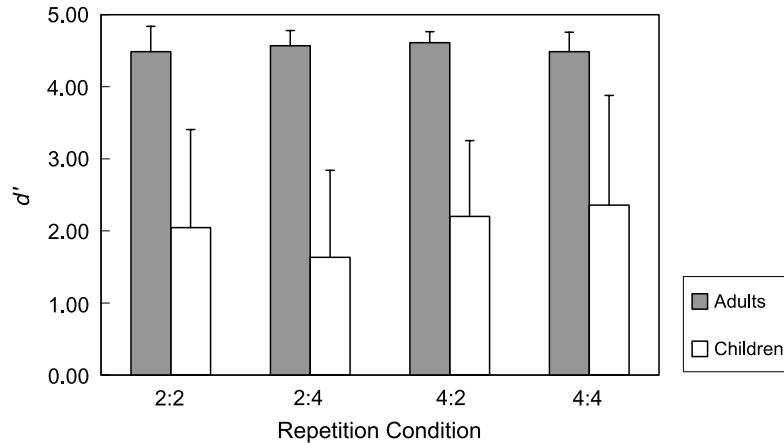
* $p \leq .05$.

Developmental Effects

There were four conditions in common between this investigation with children and our previous study with adults: /sa/ versus /ʃa/ at -4 dB SNR in the 4:4, 4:2, 2:4, and 2:2 conditions. Because the children were separated into groups on the basis of syllable pair, 10 children completed these four conditions, whereas 14 adults were tested in these conditions in our previous investigation. Figure 7 displays mean d' performance by repetition condition for the adults from Holt and Carney (2005; filled bars) and the children in the current investigation (unfilled bars) for the repetition conditions common to both studies. The error bars indicate $+1$ standard deviation. The data were entered into a one-way ANOVA to examine the effect of age in each repetition condition. Children performed significantly more poorly than adults in all four conditions—4:4 condition, $F(1, 23) = 25.570, p < .001$; 4:2 condition, $F(1, 23) = 72.313, p < .001$; 2:4 condition, $F(1, 23) = 80.743, p < .001$; and 2:2 condition, $F(1, 23) = 41.788, p < .001$.

Recall that adults from our previous investigation demonstrated improvements in d' near the predicted factor of 1.4 under low stimulus number repetition conditions and when performance was near $d' = 1.0$, but with higher stimulus repetition numbers and d' scores substantially above or below scores of 1.0, performance fell below that predicted by the multiple looks hypothesis. A direct comparison of degree of benefit from multiple stimulus presentations between the current data and those from our previous investigation is limited by at least two factors. First, because of time constraints and children's shorter attention spans, children were tested at fewer SNRs than were adults, thereby limiting the number of sampling points along the psychometric function. Further, the SNRs selected were advantageous enough that on average, children were performing around d' values of 2.0, even at the more difficult SNR. This occurred despite pilot testing the SNRs on older children, who presumably would have even less difficulty discriminating speech sounds at these SNRs. Second, children were tested at fewer stimulus repetition combinations than were adults. If children's performance were like that of the adults, we would expect that increasing the number of stimulus presentations from two to four at SNRs at which they were performing above threshold would result in d' improvements of less than a factor of 1.4. Table 3 displays the children's factorial increase in d' with a doubling in the number of standard and comparison stimulus repetitions at both the lower and higher SNRs for each syllable comparison. The mean factorial increase in d' did not always undershoot the multiple looks hypothesis' prediction of a factor of 1.4. In fact, 6 of the 24 values were at or above the multiple looks prediction, and another 5 were within $.21$ of it.

Figure 7. Mean adult performance (shaded bars; reprinted from Holt & Carney, 2005) and mean child performance (unfilled bars) at conditions in which they had in common: /sa/ versus /ʃa/ at -4 dB SNR in the 2:2, 2:4, 4:2, and 4:4 repetition conditions. Error bars indicate +1 SD.



The separate syllable pair analyses on the child data indicated that the effects of multiple stimulus repetitions varied across syllables with different manners of articulation. In our previous investigation, we entered the adult data into one overall repeated measures ANOVA because all adults completed all of the testing conditions. All of the main effects (standard and comparison repetition number, SNR, and syllable pair) were significant, as were a number of interactions ($p \leq .05$). To compare the current results with the adult data in our previous investigation, we reanalyzed our adult data by entering the data for each syllable pair into a three-way repeated measures ANOVA. The factors were standard repetition number, comparison repetition number, and SNR. The adult results for the stop-consonant contrast were similar to the children's results: the effect of SNR was significant, $F(3, 39) = 37.711, p < .001$, whereas standard and comparison repetition numbers were not significant, $F(2, 26) = 1.269, p = .298$ and $F(2, 26) = 2.648, p = .090$, respectively, although as with the children, there was a trend for adults' stop-consonant discrimination to be enhanced when the number of comparison repetitions increased. Recall that although SNR was significant for the liquid contrast,

neither the number of standard repetitions nor the number of comparison repetitions significantly influenced children's discrimination of /ra/ and /la/. In contrast, both the number of standard repetitions, $F(2, 26) = 3.663, p = .040$, and the number of comparison repetitions, $F(2, 26) = 10.599, p < .001$, significantly influenced adults' discrimination of the liquid syllable contrast, as did SNR, $F(3, 39) = 213.625, p < .001$. Finally, as with the liquid contrast, there were developmental differences in discrimination of the fricative contrast. Children's performance was significantly influenced by the number of standard repetitions and SNR, whereas adults' discrimination was significantly influenced by the number of comparison repetitions, $F(2, 26) = 4.405, p = .023$, and SNR, $F(3, 39) = 147.344, p < .001$. Although there was a trend for adults' fricative discrimination to be enhanced when the number of standard repetitions increased, the effect was not significant, $F(2, 26) = 3.020, p = .066$.

Discussion

Despite providing no additional sensory information, multiple stimulus repetitions enhanced children's

Table 3. Mean factorial improvement in d' .

| Syllable contrast | Doubling in standard repetition number | | | | Doubling in comparison repetition number | | | |
|-------------------|--|---------|----------|---------|--|---------|----------|---------|
| | Low SNR | | High SNR | | Low SNR | | High SNR | |
| | 2:2-4:2 | 2:4-4:4 | 2:2-4:2 | 2:4-4:4 | 2:2-2:4 | 4:2-4:4 | 2:2-2:4 | 4:2-4:4 |
| Stop | 1.65 | 1.04 | 1.13 | 1.05 | 1.22 | 1.40 | 1.14 | 1.07 |
| Liquid | 1.71 | 0.90 | 2.58 | 1.07 | 1.27 | 1.02 | 1.40 | 0.92 |
| Fricative | 1.25 | 1.55 | 1.19 | 1.19 | 0.82 | 1.05 | 1.02 | 1.01 |

discrimination of a single set of nonsense syllables. Specifically, increasing the number of opportunities to perceive the standard fricative stimuli significantly improved children's discrimination. No significant effects of stimulus repetition were found for the liquid or stop-consonant contrasts. Further, even though the SNRs selected might not have been optimal for stringently examining the multiple looks hypothesis, gains of the magnitude predicted by the multiple looks hypothesis were still made in some conditions, whereas other conditions resulted in smaller gains than the hypothesis would suggest. Therefore, our first hypothesis—that discrimination ability would improve with more looks at the stimuli—was supported for the fricative contrast.

The stimuli used in each syllable contrast in this investigation vary linguistically in their places of articulation. However, the contrasts are quite different from one another acoustically. The primary acoustic cues available for discriminating /pa/ and /ta/ are the second and third formant transitions: Both transitions increase for /pa/, whereas they decrease for /ta/ due to the relative frequency content of their respective bursts. The primary acoustic cue for discriminating /ra/ and /la/ is the third formant transition: /ra/ has a much lower third formant than /la/. Finally, the primary acoustic cue for discriminating /sa/ and /ʃa/ is the bandwidth of the fricatives' relative noise spectra: The spectrum for /ʃ/ is much wider than that for /s/, which has energy restricted primarily to the high frequencies. The fricative contrast is the only one of the three contrasts that does not rely primarily on a formant transition cue for discrimination, and it is the only one in which children benefited from multiple stimulus repetitions.

This contrast also happens to be the same one used extensively by Nittrouer in developing the DWS hypothesis. This hypothesis suggests that one important developmental speech perception difference between adults and children is that children tend to weight formant transition cues more than do adults, whereas adults weight spectral noise cues more than do children, at least for /ʃ/ and /s/ (Nittrouer, 1996; Nittrouer & Miller, 1997; Nittrouer & Studdert-Kennedy, 1987). However, Harris (1958) demonstrated that adults change their listening strategy based on context: for the /s/ versus /ʃ/ contrast, adults weight the spectral noise cues, whereas for /f/ versus /θ/, perceptual decisions are based on formant transition cues. Based on these results, Nittrouer (2002) predicted that children's weighting functions for /f/ and /θ/ would be similar to those of adults because adults weight the formant transition cues for this contrast just as children have been shown to do for /s/ and /ʃ/. The results largely supported the hypothesis: Children and adults' weighting functions were similar for /f/ and /θ/ but were different for /s/ and /ʃ/ (as had been shown previously). Therefore, developmental differences

in acoustic cue weighting strategies appear to vary with phonetic context.

Our data suggest that the advantage of multiple stimulus presentations also varies with context. Multiple opportunities to perceive the fricative stimuli enhanced children's discrimination sensitivity, whereas it did not have the same effect for the other contrasts. Perhaps repetition of the fricative stimuli helps children attend to the acoustic cue that they generally do not weight—the noise spectra—thereby improving their discrimination. Multiple repetitions of the stop-consonant and liquid stimuli allow children access to the acoustic cues that they already heavily weight—formant transitions—which might not be as useful to them as highlighting the cues to which they do not generally attend.

Our second hypothesis—that comparison stimuli would be weighted more than the standards—was based on the performance of adults in our previous work who showed a stronger weighting for the comparison stimuli with more stimulus repetitions. We interpreted these results to mean that listeners weighted the most recently heard stimuli when the memory load increased. Because children's memory spans are shorter than adults', we predicted that children would weight the comparisons more than the standards regardless of the number of stimulus presentations. However, we did not find this to be true: For the fricative contrast, the standards were more important than the comparisons. This outcome, however, is consistent with our original hypothesis with adults regarding the relative importance of standard and comparison stimuli. At the time, we proposed that the standards would be weighted more than the comparisons because once the change was detected, more presentations of the comparisons were superfluous to the task. However, adults weighted the standards and comparisons equally at low-repetition numbers, whereas with higher-repetition numbers, comparisons were weighted more heavily. Perhaps this experimental outcome is not a true developmental difference but, rather, a reflection of methodology. Both adults and children were told which stimulus was the standard, which was the comparison, and where the change would occur during change trials. It is possible that adults, who are more experienced listeners than children, focused most of their attention on the comparison stimuli with high numbers of stimulus repetitions, a strategy that is sufficient for completing the task with relatively good accuracy. Children, with less listening experience, might have used everything offered, not knowing that paying more attention to the second set of stimuli over the first was sufficient for completing the task. Thus, the result that children attend more to the standards than the comparison stimuli for the fricative contrast would suggest that once they perceived the change, further presentations of the comparisons were, indeed, no longer necessary for completing

the task. Further work is needed to determine if these different weighting strategies are true reflections of development or of methodology.

Addressing our third hypothesis—that children would benefit from multiple stimulus presentations but to a lesser extent than adults, and that they would weight standard and comparison stimulus repetitions differently than adults—is limited, to some extent, by the SNRs used with the children. Despite this limitation, some conclusions regarding developmental differences between adults and children can be drawn. First, the children performed significantly more poorly than did the adults in our previous investigation in the conditions common to both experiments. Even in this task, in which there is virtually no auditory stimulus uncertainty (meaning that the listener was told what speech stimuli would be presented, how many would be presented, and when the change would occur on a change trial), these 4- and 5-year-old children had much more difficulty discriminating /sa/ and /ʃa/ than did adults at a SNR commonly found in educational environments, -4 dB SNR (Knecht, Nelson, Whitelaw, & Feth, 2002). Although no direct developmental comparisons can be made between the other conditions tested in this investigation, it is clear that children need more advantageous SNRs to perform like adults. This finding supports the results of many investigators on the developmental changes in speech perception in noise (e.g., Elliott et al., 1979; Johnson, 2000; Nabelek & Nabelek, 1994; Neuman & Hochberg, 1983). Further, the data here also show that children, like adults, are differentially susceptible to varying amounts of noise. It is not that noise in general interferes with their perceptual processing but, rather, that differing noise levels differentially influence their ability to discriminate between stimuli.

Second, despite higher average d' estimates than adults (because of the SNRs used), the magnitude of the increase in d' was still in line with that predicted by the multiple looks hypothesis in 25% of the conditions and undershot the prediction by less than .22 in another 25% of the conditions. These results suggest that even in conditions in which adults show a benefit from multiple looks that is consistently less than that predicted by the hypothesis, children still can benefit from multiple stimulus repetitions in a way that is consistent with the hypothesis some of the time. Future work with children using more difficult SNRs would help clarify if children actually benefit more from multiple looks than do adults when the difficulty levels of the listening conditions between the two age groups are more comparable.

The relative weighting of standard and comparison stimuli was different between adults and children for the liquid and fricative stimulus pairs. For the liquid contrast, multiple stimulus repetitions did not affect

discrimination performance in children, whereas both multiple standard and comparison stimulus repetitions enhanced adults' discrimination. For the fricative contrast, multiple repetitions of the standard stimuli enhanced children's discrimination, whereas the comparison repetitions significantly enhanced discrimination in adults. In contrast, both adults' and children's discrimination was not influenced by either multiple repetitions of the standard or comparison stimuli for the stop-consonant contrast. These findings partially support our third hypothesis—that children would weight the standards and comparisons differently than the adults. As previous investigators have found, children often use different listening strategies than adults under similar listening conditions (Morrongiello et al., 1984; Nitttrouer, 1992, 1996; Parnell & Amerman, 1978). Further work is needed to determine if the developmental differences are due to the way in which the task was performed or the relative locations of the adults and children on the psychometric function, or if they are accurate reflections of true developmental differences.

Our extension of the multiple looks hypothesis suggests that increased opportunities to hear stimuli also improve discrimination, perhaps by enhancing the stimuli's representation in working memory, thereby making each individual stimulus more salient so that differences between them are more easily detected. Therefore, the increased number of opportunities not only enhances the individual internal representations but also might increase the contrast between the two stimuli. A somewhat analogous argument has been proposed in vision research on edge detection (Palmer, 1999) and in the psychology literature in research on category formation. In vision, object edges are believed to be signaled by changes in luminance. Uniformly colored, smooth surfaces tend to have subtle changes in luminance, whereas visual fields with multiple surfaces tend to have abrupt changes in luminance. These abrupt changes in luminance are important in visual processing because they tend to signify the edge of one object or a change in the surface orientation of an object. Admittedly, the edge detection process is quite complex, but it appears that edge operators, which calculate the values of adjacent pixels in an image, will not respond when two adjacent visual regions have the same (or nearly the same) luminance values. However, if two adjacent visual regions are very different in their luminance values, the edge operators' output is great. This increase in output occurs at the edges of images where the contrast between one image and another is enhanced. In auditory discrimination terms, it appears that increasing the internal representation of at least one of the stimuli in a pair helps enhance the contrast between it and a novel stimulus, thereby facilitating better discrimination. We are now proposing that the same neural mechanisms involved in

edge detection are necessarily at work in our speech discrimination task but that, functionally, a similar effect appears to take place.

In category formation, the term *concept* refers to mental representations of classes of objects, and the term *category* refers to the classes themselves (Murphy, 2002). A relatively new debate has arisen over whether the formation of categories is primarily knowledge driven, primarily perception driven, or some combination of the two (Jones & Smith, 1993; Murphy, 2002). The knowledge-driven approach suggests that children use prior knowledge of a particular relevant domain to determine category membership of an object, whereas the perception-driven approach argues that it is our perceptual representation of objects that allows us to make similarity comparisons between objects. If, indeed, category formation has a perceptual component, then multiple encounters with a particular object could enhance the representation and the formation of categories. For children, the change/no-change procedure could be a useful tool for assessing children's abstract knowledge of categories. In this study, we used the change/no-change procedure to examine slightly more concrete knowledge regarding discrimination of speech sounds. But the same paradigm could be adapted to sample more abstract knowledge of categories involved in language acquisition across the developmental span by examining a potential basis for category formation—perception of speech units. This kind of methodology might be mutually beneficial for speech and hearing sciences and psychology.

These results have both theoretical and clinical implications. From a theoretical perspective, the results highlight the importance of a newly identified factor in speech perception—that of multiple looks at the stimuli—and extend the findings to children's speech discrimination. The data suggest that for at least one fricative speech contrast, additional repetitions of the stimuli in a pair allow children more opportunities to form a stronger perceptual representation of the stimulus in working memory. This leads to an improved ability to discriminate that stimulus from a novel one. Even when context and vocabulary are removed from the task, discrimination ability is not solely driven by the acoustic–phonetic characteristics of the input. The acoustic–phonetic cues were constant across each fricative token, and yet there were significant performance differences.

The results have implications for models of developmental speech perception in that models will need to account for the strength of the internal perceptual representation, particularly under adverse listening conditions such as background noise and/or hearing loss. Accounting for the strength of the internal representation of the stimulus might be important under these

conditions because multiple looks appear to be particularly useful when perceptual errors are more likely. Further, it is possible that as speech categories take form during development and as children gain experience with the ambient language, the influence of multiple repetitions strengthening the internal representation might change over time. For example, in his STEP model, Moore (2003) proposed that the internal representation of stimuli for a discrimination task might be formed by a combination of (a) the incoming looks at the stimuli in working memory and (b) previous experience with similar stimuli stored in long-term memory. Perhaps early in development, when experience with similar stimuli is limited, the incoming looks are more important than later in development, when the representations are better formed in long-term memory because of more experience with the ambient language. This might partially explain why children in this investigation still tended to demonstrate enhanced discrimination under more advantageous listening conditions, whereas adults' enhancement was generally limited to more difficult listening conditions. Perhaps adults had stronger templates formed for the speech stimuli processed in long-term memory than children and did not need to rely on the incoming looks at the stimuli when there was less stimulus uncertainty. In this way, the strength of the internal representation of speech sounds could be an important component of developmental speech perception theories.

Another potential theoretical application is to the body of research on the formation of categories. The results might suggest experimental design considerations and the general idea that the perceptual salience of objects (whether concrete, such as “cat,” or relatively abstract, such as speech) might contribute to how children form categories for objects. It is possible that salience could be enhanced with multiple presentations of stimuli or objects.

Entire bodies of literature are dedicated to infant detection and discrimination abilities through about 18 months of age, beginning with the seminal work of Eimas, Siqueland, Jusczyk, and Vigorito (1971). However, between 18 and 36 months of age, there are few data on speech perception development, even at this critical age when children's vocabularies are growing at a rate of about two to four new words per day (Smith, 1926). Children at this age tend to be difficult to test, and there are few techniques available for assessing toddlers. Consistent with the paucity of data on toddler and young children's speech perception development, there is little theoretical discussion on speech perception development at these ages, even though we know that speech perception develops at least through 10 years of age (Elliott, 1986; Elliott et al., 1981; Sussman & Carney, 1989). The change/no-change procedure is a tool that can be used to evaluate changes in speech perception that

occur during this critical time in development and, as shown here, can be used to identify developmental differences in listening strategies, suggested by Nittrouer and colleagues (Nittrouer, 1992; Nittrouer et al., 1993) to be important in speech perception development. Further, the procedure was shown in this investigation to have good test–retest reliability across all syllable contrasts with the children. This lends further support for this procedure’s use with children to track performance over time.

From a clinical perspective, there is a need for speech perception procedures that can assess perceptual development at the level of discrimination without the confounding influence of vocabulary skill. Increasing numbers of children need objective evaluation of sensory aid benefit. At this time, nearly 10,000 children have cochlear implants in the United States. Because many implanted children have not developed sufficient speech and auditory language skills to use word recognition measures due to their age and/or degree of hearing loss, it is critical to have alternative tools that can measure the development of speech perception skills both pre- and postoperatively. Another growing group of children who will need this same type of monitoring is the group of infants identified with hearing loss through universal newborn hearing screening (UNHS) programs. As of 2006, 37 states (plus the District of Columbia and Puerto Rico) have instituted UNHS programs, resulting in more very young children being fitted with amplification than ever before. This group of children, regardless of their degree of hearing loss, will have limited spoken language skills because of their developmental ages. Therefore, measuring the development of their speech perception skills will be critical in determining device benefit and potentially could influence the choice of sensory aid. At least two other populations might benefit from this line of work on non-language-biased assessment tools—children with phonological impairments and children with specific language impairment. To assess perception in these populations, it is important to have a tool that does not confound basic perception with expressive language production knowledge.

The final potential clinical application relates to the finding that multiple repetitions enhance discrimination for stimuli that differ in the very acoustic cue to which children do not generally attend. This finding might suggest that multiple presentations in a speech discrimination task could be used to help train listeners on difficult contrasts—contrasts with cues that are limited by hearing loss or those that are nonnative to second-language learners.

There are some limitations of the current investigation that require further work. First, the procedures used in this study were developed to reduce the effects of nonsensory factors in order to focus on the effects of

multiple looks. In doing so, we told participants where the change would occur in the string of nonsense syllables. This cueing might affect how listeners completed the task. Rather than listening to the entire stimulus array, listeners might have primarily focused their attention on the stimuli surrounding the transition between the standards and comparisons rather than during the entire stimulus array. Although this cannot be ascertained from the current procedure, it could be tested by varying where the change occurs without telling participants where the change will take place in the stimulus array. This might uncover further developmental differences between adults and children because there is no guarantee that either group, but especially children, will approach the task in the most efficient way. In other words, even though cueing was used, we cannot be sure that children, in particular, focused their attention at the critical time during each trial when, on a change trial, the stimuli would change. Second, the SNRs were not difficult enough to force the children to make enough errors that would place them near d' values of about 1.0—the best location on the psychometric function to test the multiple looks hypothesis on the basis of what we know of adults’ performance. The SNRs used were sufficient for demonstrating the effects of multiple stimulus repetitions for the fricative contrast and for showing gains of the magnitude predicted by the multiple looks hypothesis in 25% of the conditions but not for stringently testing the multiple looks hypothesis. Therefore, future work could involve more difficult listening conditions to determine if children’s performance consistently follows the predictions of the multiple looks hypothesis.

The change/no-change procedure is sensitive to developmental changes in auditory discrimination and is easy enough to be used with young children. The procedure also fills the need for a non-language-biased, flexible, reliable, and developmentally appropriate technique for assessing speech perception in children with hearing loss, very young children, children with specific language impairment, children with phonological impairments, and children who use American Sign Language or Signed English. Finally, it has been used with limited success in toddlers (e.g., Dawson, Nott, Clark & Cowan, 1998), a population on which we have little data about speech perception development. If the procedure could be refined in future work to be used more successfully with toddlers, it could be useful in beginning to develop theories on speech perception between late infancy and early childhood and for examining how speech perception abilities serve as a basis from which children form more abstract categories for the acquisition of language. With a better understanding of typical speech perception development during the early stages of word learning in toddlers, we might be able to provide better intervention for young children using sensory devices.

These children may, for the first time, be experiencing sensory input or, at least, less degraded sensory input. However, almost immediately, demands are placed on them to organize this new perceptual input and to attach linguistic meaning to it. The change/no-change procedure is a tool that can be used to better understand how children organize the perceptual representations of speech into a higher-level linguistic code.

Acknowledgments

This research was supported by National Research Service Award F31 DC05919 (predoctoral fellowship) from the National Institute on Deafness and Other Communication Disorders and by the 2002 Student Research Grant in Audiology from the American Speech-Language-Hearing Foundation. Preliminary findings were presented at the 2002 American Speech-Language-Hearing Association Annual Convention (Atlanta, GA) and at the 2004 Acoustical Society of America Annual Meeting (New York, NY). We thank Edward Carney for his expertise in computer programming and data analysis; Benjamin Munson for his assistance in synthesizing the stimuli; Kimberly Krambeck for assistance in portions of the data collection; Karlind Moller, Peggy Nelson, Robert Schlauch, and Neal Viemeister for their valuable insights on this project; and Judith Gierut for her valuable comments on earlier versions of this article.

References

- Allen, P., & Wightman, F.** (1994). Psychometric functions for children's detection of tones in noise. *Journal of Speech and Hearing Research, 37*, 205–215.
- Allen, P., Wightman, F., Kistler, D., & Dolan, T.** (1989). Frequency resolution in children. *Journal of Speech and Hearing Research, 32*, 317–322.
- American National Standards Institute.** (1989). *Specification for audiometers (ANSI S3.6-1989)*. New York: Author.
- Anastasi, A.** (1961). Test reliability. In *Psychological Testing* (2nd ed.; pp. 105–134). New York: Macmillan.
- Bilger, R. C., & Wang, M. D.** (1976). Consonant confusions in patients with sensorineural hearing loss. *Journal of Speech and Hearing Research, 19*, 718–748.
- Carney, A. E., Osberger, M. J., Miyamoto, R. T., Karasek, A., Dettman, D. L., & Johnson, D. L.** (1991). Speech perception along the sensory aid continuum: From hearing aids to cochlear implants. In J. Feigin & P. Stelmachowicz (Eds.), *Proceedings of the 1991 National Conference on Pediatric Amplification* (pp. 93–113). Omaha, NE: Boys Town National Research Hospital.
- Dawson, P. W., Nott, P. E., Clark, G. M., & Cowan, R. S. C.** (1998). A modification of play audiometry to assess speech discrimination ability in severe-profoundly deaf 2- to 4-year-old children. *Ear and Hearing, 19*, 371–384.
- Dempster, F. N.** (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*, 63–100.
- 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., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J.** (1971, January 22). Speech perception in infants. *Science, 171*, 303–306.
- Elliott, L. L.** (1986). Discrimination and response bias for CV syllables differing in voice onset time among children and adults. *The Journal of the Acoustical Society of America, 80*, 1250–1255.
- Elliott, L. L., Connors, S., Kille, E., Levin, S., Ball, K., & Katz, D.** (1979). Children's understanding of monosyllabic nouns in quiet and noise. *The Journal of the Acoustical Society of America, 66*, 12–21.
- Elliott, L. L., Longinotti, C., Meyer, D., Raz, I., & Zucker, K.** (1981). Developmental differences in identifying and discriminating CV syllables. *The Journal of the Acoustical Society of America, 70*, 669–677.
- Frush, R. A.** (1999). *The effects of multiple looks on speech perception in children*. Unpublished master's thesis, University of Minnesota, Minneapolis.
- Green, D. M., & Swets, J. A.** (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Hall, J. W., III, & Gross, J. H.** (1991). Notched-noise measures of frequency selectivity in adults and children using fixed-masker-level and fixed-signal-level presentation. *Journal of Speech and Hearing Research, 34*, 651–660.
- Harris, K. S.** (1958). Cues for the discrimination of American English fricatives in spoken syllables. *Language and Speech, 1*, 1–7.
- Hautus, M. J.** (1995). Corrections for extreme proportions and their biasing effects on estimated values of d' . *Behavioral Research Methods, Instruments, & Computers, 27*, 46–51.
- Hill, P. R., Hartley, D. E. H., Glasberg, B. R., Moore, B. C. J., & Moore, D. R.** (2004). Auditory processing efficiency and temporal resolution in children and adults. *Journal of Speech, Language, and Hearing Research, 47*, 1022–1029.
- Holt, R. F., & Carney, A. E.** (2005). Multiple looks in speech sound discrimination in adults. *Journal of Speech, Language, and Hearing Research, 48*, 922–943.
- Howell, D. C.** (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Wadsworth Publishing Company.
- Ihaka, R., & Gentleman, R.** (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics, 5*, 299–314.
- Johnson, C. E.** (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research, 43*, 144–157.
- Jones, S. S., & Smith, L. B.** (1993). The place of perception in children's concepts. *Cognitive Development, 8*, 113–139.
- Jusczyk, P. W.** (1993). From general to language-specific capabilities: The WRAPSA model of how speech perception develops. *Journal of Phonetics, 21*, 3–28.
- Keefe, D. H., Bulen, J. C., Arehart, K. H., & Burns, E. M.** (1993). Ear-canal impedance and reflection coefficient in human infants and adults. *The Journal of the Acoustical Society of America, 94*, 2617–2638.
- Keefe, D. H., Bulen, J. C., Campbell, S. L., & Burns, E. M.** (1994). Pressure transfer function from the diffuse field to the human infant ear canal. *The Journal of the Acoustical Society of America, 95*, 355–371.
- Knecht, H. A., Nelson, P. B., Whitelaw, G. M., & Feth, L. L.** (2002). Background noise levels and reverberation times in

- unoccupied classrooms: Predictions and measurements. *American Journal of Audiology*, 11, 65–71.
- Macmillan, N. A., & Creelman, C. D.** (2005). *Detection theory*. Mahwah, NJ: Erlbaum.
- Miller, G. A., & Nicely, P. E.** (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27, 338–352.
- Miller, J. O.** (1996). The sampling distribution of d' . *Perception & Psychophysics*, 58, 65–72.
- Minneapolis Public Schools.** (1980). *Minneapolis Preschool Screening Instrument*. Minneapolis, MN: Author.
- Moore, B. C. J.** (2003). Temporal integration and context effects in hearing. *Journal of Phonetics*, 31, 563–574.
- Moore, J. K.** (2002). Maturation of human auditory cortex: implications for speech perception. *Annals of Otolaryngology, Rhinology, and Laryngology*, 189(Suppl.), 7–10.
- Moore, J. K., & Guan, Y. L.** (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, 2, 297–311.
- Morrongiello, B. A., Robson, R. C., Best, C. T., & Clifton, R. K.** (1984). Trading relations in perception of speech by five-year-old children. *Journal of Experimental Child Psychology*, 37, 231–250.
- Murphy, G. L.** (2002). *The big book of concepts*. Cambridge, MA: The MIT Press.
- Nabelek, A. K., & Nabelek, I. V.** (1994). Room acoustics and speech perception. In J. Katz (Ed.), *Handbook of clinical audiology* (4th ed., pp. 624–637). Baltimore: Williams & Wilkins.
- Neuman, A. C., & Hochberg, I.** (1983). Children's perception of speech in reverberation. *The Journal of the Acoustical Society of America*, 73, 2145–2149.
- 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.** (1996). Discriminability and perceptual weighting of some perceptual cues to speech perception by 3-year-olds. *Journal of Speech and Hearing Research*, 39, 278–287.
- Nittrouer, S.** (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *The Journal of the Acoustical Society of America*, 112, 711–719.
- Nittrouer, S., Manning, C., & Meyer, G.** (1993). The perceptual weighting of acoustic cues changes with linguistic experience. *The Journal of the Acoustical Society of America*, 94, S1865.
- Nittrouer, S., & Miller, M. E.** (1997). Predicting developmental shifts in perceptual weighting schemes. *The Journal of the Acoustical Society of America*, 101, 2253–2266.
- 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.
- Olsho, L. W.** (1985). Infant auditory perception: Tonal masking. *Infant Behavior and Development*, 7, 27–35.
- Palmer, S. E.** (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: The MIT Press.
- Parnell, M. M., & Amerman, J. D.** (1978). Maturation influences on perception of coarticulatory effects. *Journal of Speech and Hearing Research*, 21, 682–701.
- Pisoni, D. B.** (2000). Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear and Hearing*, 21, 70–78.
- Pisoni, D. B., & Cleary, M.** (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing*, 24(Suppl.), 106S–120S.
- Pisoni, D. B., & Geers, A.** (2000). Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language processing. *Annals of Otolaryngology, Rhinology, & Laryngology*, 109, 92–93.
- Psychology Software Tools.** (2002). *E-Prime 1.1*. Pittsburgh, PA: Author.
- Smith, M. E.** (1926). An investigation of the development of the sentence and the extent of vocabulary in young children. *Studies in Child Welfare*, Vol. 3. Iowa City, IA: University of Iowa.
- Spetner, N. B., & Olsho, L. W.** (1990). Auditory frequency resolution in human infancy. *Child Development*, 61, 632–652.
- Sussman, J. E., & Carney, A. E.** (1989). Effects of transition length on the perception of stop consonants by children and adults. *Journal of Speech and Hearing Research*, 32, 151–160.
- Trehub, S. E., Schneider, B. A., Thorpe, L. A., & Judge, P.** (1991). Observational measures of auditory sensitivity in early infancy. *Developmental Psychology*, 27, 40–49.
- Viemeister, N. F., & Wakefield, G. H.** (1991). Temporal integration and multiple looks. *The Journal of the Acoustical Society of America*, 90, 858–865.
- Werner, L. A.** (1992). Interpreting developmental psychoacoustics. In L. A. Werner & E. W. Rubel (Eds.), *Developmental psychoacoustics* (pp. 47–88). Washington, DC: American Psychological Association.
- Werner, L. A.** (1996). The development of auditory behavior (or what anatomists and physiologists have to explain). *Ear and Hearing*, 17, 438–446.
- Werner, L. A.** (2002). Infant auditory capabilities. *Current Opinion in Otolaryngology and Head and Neck Surgery*, 10, 398–402.
- Werner, L. A., Folsom, R. C., & Mancl, L. R.** (1993). The relationship between auditory brainstem response and behavioral thresholds in normal hearing infants and adults. *Hearing Research*, 68, 131–141.
- Werner, L. A., Folsom, R. C., & Mancl, L. R.** (1994). The relationship between auditory brainstem response latency and behavioral thresholds in normal hearing infants and adults. *Hearing Research*, 77, 88–98.
- Werner, L. A., & Mancl, L. R.** (1993). Pure-tone thresholds of 1-month-old human infants. *The Journal of the Acoustical Society of America*, 93, 2367.
- Werner, L. A., & Marean, G. C.** (1996). *Human auditory development*. Boulder, CO: Westview Press.
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E.** (1992). *Preschool Language Scale-3*. San Antonio, TX: Psychological Corporation.

Received March 30, 2006

Revision received November 17, 2006

Accepted May 1, 2007

DOI: 10.1044/1092-4388(2007)098

Contact author: Rachael Frush Holt, who is now with the Department of Speech and Hearing Sciences, Indiana University, 200 South Jordan Avenue, Bloomington, IN 47405. E-mail: raholt@indiana.edu.

Copyright of *Journal of Speech, Language & Hearing Research* is the property of American Speech-Language-Hearing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.