



HHS Public Access

Author manuscript

J Speech Lang Hear Res. Author manuscript; available in PMC 2017 August 01.

Published in final edited form as:

J Speech Lang Hear Res. 2007 February ; 50(1): 41–57. doi:10.1044/1092-4388(2007/004).

Children With Specific Language Impairments Perceive Speech Most Categorically When Tokens Are Natural and Meaningful

Jeffry A. Coady,

Boston University

Julia L. Evans,

University of Wisconsin–Madison

Elina Mainela-Arnold, and

University of Wisconsin–Madison

Keith R. Kluender

University of Wisconsin–Madison

Abstract

Purpose—To examine perceptual deficits as a potential underlying cause of specific language impairments (SLI).

Method—Twenty-one children with SLI (8;7–11;11 [years;months]) and 21 age-matched controls participated in categorical perception tasks using four series of syllables for which perceived syllable-initial voicing varied. Series were either words or abstract nonword syllables and either synthesized or high-quality edited natural utterances. Children identified and discriminated (a) digitally edited tokens of naturally spoken “bowl”–“pole”, (b) synthesized renditions of “bowl”–“pole”, (c) natural “ba”–“pa”, and (d) synthetic “ba”–“pa”.

Results—Identification crossover locations were the same for both groups of children, but there was modestly less accuracy on unambiguous endpoints for children with SLI. Planned comparisons revealed these effects to be limited to synthesized speech. Children with SLI showed overall reduced discrimination, but these effects were limited to abstract nonword syllables.

Conclusion—Overall, children with SLI perceived naturally spoken real words comparably to age-matched peers but showed impaired identification and discrimination of synthetic speech and of abstract syllables. Poor performance on speech perception tasks may result from task demands and stimulus properties, not perceptual deficits.

Keywords

specific language impairment; categorical perception; speech perception

(C) The American-Speech-Language-Hearing Association is the publisher of this article and holder of the copyright. Further reproduction of this article in violation of copyright is prohibited without the consent of the publisher. To contact the publisher: <http://www.asha.org/>

Contact author: Jeffry A. Coady, Sargent College, Boston University, 635 Commonwealth Avenue, Boston, MA 02215. coady@bu.edu.

Julia L. Evans is now at the School of Speech, Language, and Hearing Sciences, San Diego State University, and Elina Mainela-Arnold is now at the Department of Communication Sciences and Disorders, Pennsylvania State University.

The present investigation was conducted to examine how speech perception abilities of children with specific language impairments (SLI) change as a function of the nature of the tokens with which they are tested. These children have normal hearing and nonverbal intelligence, with no obvious oral-motor or neurological deficits. Nevertheless, they experience difficulty learning language in spite of having all of the requisite cognitive abilities generally considered to support normal language acquisition (see Leonard, 1998, for a review). Although these children do acquire language, they experience deficits in many linguistic areas, including phonology, morphology, and syntax. Among the possible underlying causes that have been considered are deficits in , auditory processing (e.g., Eisenson, 1972). Explanations of these potential processing difficulties range from slower processing across all perceptual modalities (Leonard, 1998; Leonard, McGregor, & Allen, 1992; Miller, Kail, Leonard, & Tomblin, 2001) to a deficit in processing rapidly changing auditory information characteristic of speech (Tallal & Piercy, 1973,1974,1975; Tallal, Stark, Kallman, & Mellits, 1981). These children are hypothesized to have inadequately specified phonological representations in long-term memory (Stark & Heinz, 1996a, 1996b), which would make extracting linguistic generalizations more difficult (Joanisse & Seidenberg, 1998) or make encoding information in phonological working memory less efficient (Gathercole, 1999).

Coady, Kluender, and Evans (2005) hypothesized that these putative perceptual deficits do not result from a deficit in auditory processing per se, but rather from increased sensitivity to task demands used to examine auditory processing in children with SLI. Previous studies that have provided evidence taken to be indicative of an auditory processing deficit showed that children with SLI perform relatively poorer when tested with synthetic versions of nonsense syllables in tasks with high memory demands (Elliott & Hammer, 1988; Elliott, Hammer, & Scholl, 1989; Evans, Viele, Kass, & Tang, 2002; Joanisse, Manis, Keating, & Seidenberg, 2000; Leonard et al., 1992; McReynolds, 1966; Stark & Heinz, 1996a, 1996b; Sussman, 1993, 2001; Tallal & Piercy, 1974, 1975; Tallal & Stark, 1981; Tallal et al., 1981; Thibodeau & Sussman, 1979). Among these are experiments that have examined categorical perception of speech in children with SLI (Joanisse et al., 2000; Sussman, 1993; Thibodeau & Sussman, 1979).

In categorical perception tasks, listeners hear speech tokens that have been created to vary incrementally along one or more relevant acoustic dimensions, ranging perceptually from one phoneme to another. This results in a series of ambiguous speech tokens between two unambiguous endpoints. Listeners are asked to identify tokens drawn from such a series, and they readily partition the series into two groups. Listeners then discriminate tokens drawn from the series, either by identifying two tokens as being the same or different or by identifying two of three tokens that are the same. As outlined by Wood (1976), categorical perception entails three components: (a) a sharp labeling or identification function; (b) discontinuous discrimination performance, with near perfect discrimination for pairs of tokens drawn from across the labeling boundary and near chance discrimination for pairs drawn from the same side of the labeling boundary; and (c) the ability to predict discrimination performance from labeling. When children with SLI participate in a categorical perception task, their crossover point for identification data is at the same

location as typically developing children, but they often show less consistent labeling of unambiguous endpoint tokens with a consequent shallower identification function (Joanisse et al., 2000; Sussman, 1993; Thibodeau & Sussman, 1979). These results are cited as evidence that children with SLI perceive speech less categorically than children who develop language normally.

Coady and colleagues (2005) observed that categorical perception appears to be a basic perceptual property of any general learning system. Infants (Eimas, Siqueland, Jusczyk, & Vigorito, 1971), chinchillas (Kuhl, 1981; Kuhl & Miller, 1975, 1978), and a variety of simple neural network models covering a range of architectures (see Damper & Hamad, 2000, for a review) all provide evidence of relatively simple processes underlying categorical perception. Further, categorical perception has been demonstrated for a host of nonspeech and nonauditory stimuli (Kluender, Lotto, & Holt, 2006). Owing to the ubiquity of categorical perception across age, species, and stimulus domain, deficits in categorical perception, if real, could be taken to indicate a very general and profound pathology. Given the broad perceptual and cognitive successes of children with SLI, a true deficit in categorical perception should be unexpected. In light of these considerations, Coady and colleagues examined categorical perception of speech in a group of children with SLI, aged 7;3–11;5 (years; months), using a digitally edited naturally spoken series of real-word tokens, varying acoustically in voice-onset time (VOT) and ranging perceptually from “bowl” to “pole.” They found that children with SLI did not differ from an age-matched control group in their identification functions but did show lower peak discrimination values. Nevertheless, children with SLI discriminated pairs of tokens drawn from across the identification crossover relatively well. Further, when group identification functions were used to predict group discrimination functions, both groups of children showed similar levels of predictability. Coady and colleagues interpreted these results as evidence against auditory or speech perception deficits as an underlying cause of SLI. Instead, they suggested that poorer performance on speech perception tasks by children with SLI may be attributed to memory, processing, or representational limitations.

While these results suggest that children with SLI can perceive speech categorically under optimal conditions, they fail to reveal specific underlying causes of differences between performance by children with SLI and typically developing children. Coady and colleagues only examined perception of a single digitally edited, naturally spoken, real-word series. They did not systematically manipulate processing demands associated with synthetic speech or representational demands associated with abstract nonsense syllables.

Naturally produced speech contrasts are specified by multiple redundant acoustic properties, and listeners are quite sensitive to these when making phonetic judgments (see, e.g., Kluender, 1994; Pisoni & Luce, 1987). However, these multiple acoustic cues are typically under-represented in less acoustically rich computer-generated speech. Luce, Feustel, and Pisoni (1983) reported that listeners show better perception and comprehension of natural speech tokens compared with synthesized speech tokens (see also Nusbaum & Pisoni, 1985). Evans et al. (2002) hypothesized that children with SLI would likewise perceive naturally spoken tokens better than synthetic tokens. They asked children with SLI to identify naturally spoken and synthetic versions of the vowels [i] and [u] in isolation and in a [dab-V-

ba] context, and the consonants [s] and [ʃ] in a [daC] context. The children could identify both naturally produced and synthetic vowels in isolation but only naturally produced vowels when embedded in the phonetic context. They also had difficulty identifying both natural and synthetic versions of [s]–[ʃ] embedded in the [daC] context. Reynolds and Fucci (1998) have also reported that children with SLI had more difficulty comprehending synthesized sentences than did age-matched control children.

In a recent study examining speech perception of children with developmental dyslexia, Blomert and Mitterer (2004) reported similar results for categorical perception of a more natural series varying in place of articulation ([ta]–[ka]) and a more synthetic place series ([ba]–[da]). They found that children with dyslexia showed typical patterns of impaired categorical perception when tested with synthesized versions of the speech series but not for more natural versions. There were no group differences in placement of the labeling boundary or in slopes of the identification functions. Children with dyslexia were less consistent than a control group when labeling unambiguous endpoint tokens from the synthetic series but not from the more natural series. Thus, direct comparison of perception of natural and synthetic speech might reveal that children with SLI perceive a natural speech series more categorically than a synthetic speech series.

Stark and Heinz (1996a, 1996b), Bishop (2000), and Evans (2002) have hypothesized that children with SLI have underspecified or fragile underlying linguistic representations. This would suggest that these children would be less facile with abstract nonsense-syllable test items compared with real-word test items for which lexical representations could be drawn upon. In order to perceive series of abstract speech sounds such as those used in categorical perception tasks, listeners may benefit from lexical structure gained through experience hearing a large corpus of speech. After hearing many instances of a particular phonemic class in a variety of phonetic contexts, children are more likely to create a single underlying representation of the individual sounds (Eimas, 1994). That is, phonemes may emerge as dimensions of the developing lexical space rather than being lexical primitives themselves (Kluender & Lotto, 1999).

Chiat (2001) has argued that SLI might be explained as a mapping deficit at multiple linguistic levels. She suggested that children with SLI have particular difficulty mapping abstract syllables with dimensions of lexical organization. However, she did not consider a potential deficit in mapping from acoustic input to phonetic or phonological representations because evidence available at the time suggested that children with SLI have basic speech perception deficits. More recent evidence suggests that children with SLI perceive speech much like their typically developing peers (Coady et al., 2005). If Chiat's mapping account were extended to include a mapping process such as that proposed by Eimas (1994), then these children's difficulty in processing nonword stimulus items can be easily explained. Children with SLI may be perfectly capable of perceiving minimal phonetic pairs (varying along a single acoustic dimension) in the words that they hear but may not have established an underlying representation sufficiently robust to facilitate perception of nonword syllables with only limited resemblance to lexical representations.

The current study was designed to examine the effects of speech quality (natural vs. synthetic speech) and lexical status (real word vs. nonsense syllable) on categorical perception of speech by children with SLI. Children with SLI and age-matched control children identified and discriminated tokens from four different speech series: (a) digitally edited tokens of naturally spoken “bowl”– “pole,” (b) synthesized renditions of “bowl”– “pole,” (c) natural “ba”– “pa,” and (d) synthetic “ba”– “pa.” If children with SLI are sensitive to speech quality, then they should perceive synthetic speech series less categorically than naturally spoken series. If they are sensitive to the lexical status of test items, then they should perceive nonsense-syllable test series less categorically than real-word test series. If their performance suffers from both of these manipulations, then only their perception of naturally spoken real words should match that of age-matched control children.

Method

Participants

Participants included 21 monolingual English-speaking children (12 females, 9 males) with SLI (mean age 10;3; range: 8;7–11;11) and 21 typically developing children (11 females, 10 males with a mean age of 10;1; range: 8;0–12;3) matched for chronological age. The age difference between groups was not significant, $t(40) = 0.49$, $p = .62$, $\omega^2 < 0$. Children were drawn from a larger sample of children in local schools. Children with SLI met exclusion criteria (Leonard, 1998), having no frank neurological impairments, no evidence of oral-motor disabilities, normal hearing sensitivity, and no social or emotional difficulties (based on parent report). Nonverbal IQs were at or above 85 (1 *SD* below the mean or higher) as measured by the Leiter International Performance Scale– Revised (Leiter-R; Roid & Miller, 1997). To control for possible confounding effects of articulation impairments, only children without articulation deficits were included. Speech intelligibility, as measured during spontaneous production, was at or above 98% for all children. All children also had normal range hearing sensitivity on the day of testing as indexed by audiometric pure-tone screening at 25 dB for 500 Hz tones and at 20 dB for 1000-, 2000-, and 4000-Hz tones.

Language assessment measures included (a) the Clinical Evaluation of Language Fundamentals–Third Edition (CELF-III; Semel, Wiig, & Secord, 1995), (b) the Peabody Picture Vocabulary Test–Third Edition (PPVT-III; Dunn & Dunn, 1997), (c) the Expressive Vocabulary Test (EVT; Williams, 1997), (d) the Nonword Repetition Task (NWR; Dollaghan & Campbell, 1998), and (e) the Competing Language Processing Task (CLPT; Gaulin & Campbell, 1994). Children with SLI received the full Expressive and Receptive language batteries of the CELF-III, and composite Expressive (ELS) and Receptive (RLS) language scores were calculated. Typically developing children received the full expressive language battery of the CELF-III, while their receptive language was screened with the Concepts and Directions subtest of the receptive language battery.

The group of children with SLI included 2 children with only expressive language impairments (E-SLI) and 19 with both expressive and receptive language impairments (ER-SLI). The language criteria for E-SLI were ELS at least 1 *SD* below the mean (<85) and RLS greater than 1 *SD* below the mean (>85). Criteria for ER-SLI were both ELS and RLS at

least 1 *SD* below the mean (<85). Language criteria for the age-matched control group were ELS above 85 and standard score on the Concepts and Directions subtest at or above 8. Group summary statistics are provided in Table 1. Children with SLI scored significantly below typically developing children on all diagnostic measures: Leiter-R, $t(40) = 2.755$, $p < .005$, $\omega^2 = .136$, power = .70; CELF-III ELS, $t(40) = 10.64$, $p < .0001$, $\omega^2 = .728$, power = .99; PPVT-III, $t(40) = 4.83$, $p < .0001$, $\omega^2 = .347$, power = .99; EVT, $t(40) = 5.06$, $p < .0001$, $\omega^2 = .369$, power = .99; NWR, $t(40) = 3.94$, $p = .0002$, $\omega^2 = .257$, power = .96; CLPT, $t(40) = 4.96$, $p < .0001$, $\omega^2 = .360$, power = .99. Individual scores for the children with SLI are provided in the Appendix.

Stimuli

Four different test series were constructed: (a) digitally edited tokens of naturally spoken “bowl”–“pole” and (b) natural “ba”–“pa,” (c) synthesized renditions of “bowl”–“pole,” and (d) synthetic “ba”–“pa.” The naturally spoken real-word series was the “bowl”–“pole” series used previously by Coady et al. (2005). These words were chosen because they differ along the [voice] dimension on which children with SLI have been repeatedly tested. Further, the words have very similar (adult) ratings of word familiarity, concreteness, imageability, meaningfulness (Coltheart, 1981), and word frequency (Kucera & Francis, 1967).

An adult female with an upper Midwestern accent produced several versions of the words in pairs, as similar to one another as possible, thereby minimizing any potential pitch, duration, or amplitude differences. The speaker produced the words in a soundproof chamber, and utterances were recorded directly into a Windows-based waveform analysis program. They were digitized at a 44.1-kHz sampling rate with 16-bit resolution. One pair of tokens was chosen for further processing. The pair of tokens was used to construct a six-member “bowl”–“pole” series, in which the perceptual change from [b] to [p] was accomplished by manipulating the delay between syllable onset and onset of periodicity for voicing. Delay of periodicity was 0 ms for “bowl” and 52 ms for “pole.” As is typical of bilabial stop consonants, the noise burst at release was weak for both tokens. The naturally spoken token of “bowl” was the first endpoint stimulus. Intermediate steps in the series were created by deleting the burst and successively larger acoustic segments of periodic energy from “bowl” and replacing these acoustic segments with the burst and equally long acoustic segments of aperiodic energy taken from “pole.” Each step represented two pitch pulses of voicing, or approximately 10 ms, being replaced by an equal amount of aspiration. The token with the 50 ms VOT served as the “pole” endpoint. All cuts were made at the end of a pitch pulse at a zero crossing. Thus, there was no audible indication (e.g., presence of a click) that stimuli had been made by digitally splicing two different words. Naive listeners consistently perceived the stimuli to be unedited tokens of natural speech. The total duration of each stimulus was 391 ms.

Stimuli were transferred to compact discs for presentation. For the identification task, each stimulus was presented twice in a single trial separated by a 1-s inter-stimulus interval (ISI). For the discrimination portion, stimuli were modeled after those used by Sussman (1993). An identical pair of tokens was followed by another identical pair of tokens. In some cases, the first pair did not match the second pair (change), while in other cases, all four tokens

were identical (no change). Each token was paired with itself and with tokens two or more steps away in the series. So, for example, the third token in the series (20 ms VOT) was paired with the first, third, fifth, and sixth tokens (0, 20, 40, and 50 ms VOTs, respectively). Test items were pairs of tokens differing in VOT by 20 ms, or two steps in the series (VOT0–20, VOT10–30, VOT20–40, and VOT30–50). Pairs consisting of identical tokens served as catch trials, while pairs consisting of tokens with more than a 20 ms VOT difference served as filler trials. The four tokens within each trial were separated by 500 ms of silence. Children could respond “same” or “different” based on whether they heard a change in stimulus identity. Each pair of tokens (in both orders) was presented twice in a fixed random order. That is, each test trial was presented two times in two different orders for a total of four presentations.

Abstract syllables “ba” and “pa” were chosen because they contain the same voicing contrast being tested in the real-word series and because they are meaningless syllables. The syllables “bo” and “po” could not be used because “bow” is a real word that appears in children’s speech. The naturally spoken abstract-syllable series was created from utterances of the same speaker, and digital editing and stimulus creation followed the same procedures described above. The total duration of each stimulus was 364 ms.

The six-step synthetic real-word (“bowl”–“pole”) series was modeled after the natural series, with delay of onset of periodicity varying incrementally in 10-ms steps. Tokens were synthesized using the parallel branch of the Klatt synthesizer (Matt, 1980; Matt & Matt, 1990). Syllables were 350 ms long including a 10-ms burst, aspiration noise, and periodic vocalic portions. The burst was created by setting the amplitude of frication (AF) to 65 dB, the amplitude of the frication-excited parallel second formant (A2F) to 65 dB, and the amplitude of the frication-excited parallel bypass path (AB) to 55 dB over the first 10 ms. For every incremental step in the series, 10 ms of voicing after the burst was replaced with 10 ms of aspiration at 70 dB. The fundamental frequency was set to 200 Hz at the beginning of the token, falling to 180 Hz over the first 315 ms and remaining constant thereafter. The amplitude of voicing was set to 60 dB at the onset of voicing, remaining constant until the 315 ms mark, then falling to 40 dB over the next 25 ms, and falling to 0 dB over the last 10 ms. The frequency of the first formant was set to 460 Hz at the beginning of each token, rising to 555 Hz over the first 50 ms, remaining constant for the next 200 ms, then falling to 470 Hz over the last 100 ms. The amplitude of the first formant varied with the amplitude of voicing, starting at 0 dB but rising abruptly at the onset of voicing to 60 dB and remaining constant thereafter. The frequency of the second formant was set to 810 Hz over the first 250 ms and rose to 950 Hz over the last 100 ms. The frequency of the third formant rose from 2000 Hz to 2800 Hz over the first 50 ms, remained constant for the next 200 ms, and rose to 3800 Hz over the last 100 ms. The bandwidths of the first three formants remained constant at 90 Hz, 90 Hz, and 130 Hz, respectively. The amplitudes of the second and third formants remained constant at 54 dB and 48 dB, respectively.

The six-step synthetic “ba”–“pa” series was created in the same way as the synthetic word series. Voicing, amplitude, and timing parameters were identical to the synthetic “bowl”–“pole” series. The only difference was in the formant frequencies for the vowel portion of the syllables. The first formant rose from 460 Hz to 1000 Hz over the first 50 ms and remained

constant thereafter. The second formant started at 810 Hz and rose to 1600 Hz over the first 50 ms, remaining constant over the last 200 ms. The third formant rose from 2000 Hz to 2800 Hz over the first 50 ms and stayed at 2800 Hz thereafter.

Procedure

Children participated in both identification and discrimination tasks with each of the four stimulus series during a single session as part of a larger experimental test battery. These tasks were interspersed with an unrelated perceptual task, also presented in four blocks. The order of presentation was as follows: two blocks of the unrelated perceptual task, break, categorical perception tasks with two of the series, break, last two blocks of the unrelated perceptual task, break, two categorical perception tasks with each of the remaining two series. For the categorical perception tasks, the order was that children always identified tokens from one of the test series then discriminated tokens from that same series. Once one series had been completed, children identified and discriminated tokens from a different series. The order in which the children received the four test series was randomized across subjects.

Listeners were tested individually in a large soundproof chamber (Acoustic Systems). Test items were presented over a single speaker (Realistic Minimus 7) that had been calibrated to 75 dB SPL at the beginning of each session. The frequency response (100–10000 Hz) was measured and found to be acceptably flat. The speaker was positioned approximately 2 ft in front of the listener. For the identification task, children were told that they would be hearing a woman (or a computer voice) saying the words “bowl” or “pole” (or “ba” or “pa”). Their job was to listen to and identify each word by pointing to a picture (i.e., a two-alternative forced choice [2AFC] task). For the real-word series, colored pictures of bowls and poles were contained in a notebook, with item positions and positions of correct answers randomized (right–left). For the abstract-syllable series, colored versions of printed letters *B* and *P* were contained in a similar notebook, with item positions randomized. The experimenter provided two practice trials, in which he said each word aloud and prompted the listener to point to the appropriate picture. None of the children had any difficulty with the identification portion of the task. Once testing began, children heard two presentations of each token within a single trial and indicated their response by pointing to one of the pictures. The experimenter recorded their responses. Over 60 trials, listeners heard presentations of all six stimuli in 10 randomized blocks. Immediately following the identification task, listeners participated in the discrimination task with tokens drawn from the same series. In a variant of an AX discrimination task, listeners were told that they would be hearing two words, then another two words (AAXX). Their job was to determine if all of the words were the same or different. Again, the experimenter produced one of each stimulus type to ensure that the listener understood the task. All children heard 48 pairs of tokens in a fixed random order, for a total of four presentations of each test trial, two presentations of each catch trial, and 20 filler trials. Children responded verbally, and the experimenter recorded their responses.

Results

In the analyses that follow, the alpha level was set at .05. Results from analyses of variance were used to estimate population effect sizes (ω^2), or the amount of total population variance accounted for by variance due to effects of interest. According to Cohen (1977), small effects result in an ω^2 of approximately .01, medium effects in an ω^2 of approximately .06, and large effects in an ω^2 of approximately .15. Effect sizes, sample sizes, and alpha level were then used to calculate statistical power, or the probability of finding a significant effect if one truly exists. Based on the way ω^2 was calculated, *F* or *t* values less than 1 result in ω^2 values less than 0, and power could not be calculated.

Identification

Identification functions for each series were obtained by calculating the percentage of voiceless [p] responses (out of 10) for each stimulus along each of the six-step series. The results for all series are shown in Figure 1. Tokens with lower values were consistently identified as [+voice] tokens (“bowl” or “ba”), while tokens with higher values were identified as [–voice] tokens (“pole” or “pa”). For both groups and for all speech series, there was a sharp shift in identification. To examine potential perceptual deficits in children with SLI, group performance was compared to age-matched children’s group identification performance. Potential group differences in identification were first examined by comparing the number of [p] responses as a function of stimulus. They were next examined by comparing the results of probit analyses for both groups. Probit analysis fits a cumulative normal curve to probability estimates as a function of stimulus level by the method of least squares (Finney, 1971), estimating mean (identification crossover) and standard deviation for each distribution.

Overall, children with SLI made an average of 30.3 [p] responses, while the typically developing children made an average of 30.5 [p] responses. The percentage of [p] responses was entered into a 2 (group) \times 6 (VOT) \times 2 (speech source) \times 2 (lexical status) mixed design analysis of variance (ANOVA). There was a significant main effect of stimulus level, $F(5, 800) = 1478.07$, $P < .0001$, $\omega^2 = .97$, power = .99, indicating that tokens with lower VOT values were consistently given the [+voice] label, while tokens with higher VOT values were consistently given the [–voice] label. Identification functions did not differ significantly by group, $F(1, 5) = 0.017$, $p = .901$, $\omega^2 < 0$; speech source, $F(1, 5) = 0.233$, $p = .650$, $\omega^2 < 0$; or lexical status of test stimuli, $F(1, 5) = 3.237$, $p = .132$, $\omega^2 = .05$, power = .25. None of the two-way interactions was significant, but the three-way interaction approached significance, $F(1, 5) = 5.603$, $p = .064$, $\omega^2 = .10$, power = .45. A strict interpretation is that the two groups do not differ in their perceptual abilities. However, this statistic approached the absolute α -level cutoff of .05, with a medium to large effect size. Therefore, this interaction was investigated further.

Subsequent analysis revealed that none of the two-way interactions was mediated by a third factor. The Group \times Speech Source interaction did not vary by lexical status of test items: real words, $F(1, 5) = 2.472$, $p = .177$, $\omega^2 = .03$, power = .20; abstract syllables, $F(1, 5) = 0.707$, $p = .439$, $\omega^2 < 0$. The Group \times Lexical Status interaction did not vary by speech source: natural speech, $F(1, 5) = 1.128$, $p = .337$, $\omega^2 = .003$, power = .18; synthetic speech,

$F(1, 5) = 3.605, p = .116, \omega^2 = .06, \text{power} = .24$. Finally, the Speech Source \times Lexical Status interaction did not vary by group: children with SLI, $F(1, 5) = 0.030, p = .869, \omega^2 < 0$; typical children, $F(1, 5) = 4.255, p = .094, \omega^2 = .07, \text{power} = .33$. Because the three-way interaction was not a function of any of the two-way interactions, group effects were examined separately for each stimulus series.

As is evident from a visual inspection of the functions, children with SLI and typically developing children performed almost identically with naturally spoken versions of real words, but there were visible group differences for the other three series. To quantify these differences, each speech series was considered in turn. For the naturally spoken real words, the percentage of “pole” responses was entered into a 2 (group) \times 6 (VOT) mixed design ANOVA, which revealed no significant differences in the two groups’ identification functions, $F(5, 200) = 0.183, p = .969, \omega^2 < 0$. An ANOVA on the percentage of “pa” responses for the naturally spoken abstract syllables revealed no group differences in identification functions, $F(5, 200) = 1.385, p = .231, \omega^2 = .01, \text{power} = .20$. For synthetic renditions of real words, there were no significant differences in group identification functions, $F(5, 200) = 1.485, p = .196, \omega^2 = .01, \text{power} = .32$. Finally, for synthetic versions of abstract syllables, the two groups differed significantly in their identification functions, $F(5, 200) = 5.10, p = .0002, \omega^2 = .09, \text{power} = .80$.

Next, accuracy on only endpoint stimuli was examined. Overall, children with SLI were at 93.63% accuracy on unambiguous endpoints, while typically developing children were at 97.74% accuracy, a significant difference, $F(1, 160) = 5.708, p = .018, \omega^2 = .10, \text{power} = .57$. There were no overall differences in endpoint accuracy due to the lexical status of test items, $F(1, 160) = 0.030, p = .86, \omega^2 < 0$. However, there was a significant main effect of speech source, with higher accuracy for natural as opposed to synthetic speech, $F(1, 160) = 6.744, p = .01, \omega^2 = .12, \text{power} = .66$. Further, the Group \times Speech Source interaction approached significance, $F(1, 160) = 3.627, p = .059, \omega^2 = .06, \text{power} = .37$. Examination of this interaction revealed that the children with SLI were more accurate on endpoints in the natural speech conditions than in the synthetic speech conditions, $F(1, 80) = 6.546, p = .01, \omega^2 = .12, \text{power} = .65$, while the typically developing children showed no such accuracy difference, $F(1, 80) = 0.763, p = .385, \omega^2 < 0$.

As for the overall identification functions, endpoint accuracy was examined for each speech series separately. Groups did not differ on endpoint accuracy for real words or for natural speech: natural “bowl”–“pole,” $F(1, 40) = 0.623, p = .435, \omega^2 < 0$; natural “ba”–“pa,” $F(1, 40) = 2.107, p = .154, \omega^2 = .03, \text{power} = .28$; and synthetic “bowl”–“pole,” $F(1, 40) = 2.06, p = .159, \omega^2 = .03, \text{power} = .25$. Children with SLI were slightly less accurate identifying unambiguous endpoint tokens for the synthetic “ba”–“pa” series, $F(1, 40) = 3.52, p = .068, \omega^2 = .06, \text{power} = .36$, with this effect approaching significance.

Identification data for each series for each listener were entered into probit analysis. As explained above, probit analysis fits identification functions to a cumulative normal curve, providing a mean (50% point) and standard deviation (Finney, 1971). The mean 50% crossover point for children with SLI was at 24.2 ms, while that for typically developing children was at 25.8 ms, a nonsignificant difference, $F(1, 160) = 2.59, p = .109, \omega^2 = .04$,

power = .30. Crossover points did not differ significantly for natural versus synthetic speech, $F(1, 160) = .001, p = .974, \omega^2 < 0$. However, crossover points did vary as a function of the lexical status of the test syllable, with the boundary shifted toward [p] (more “bowl” responses) for the real-word series, $F(1, 160) = 5.687, p = .018, \omega^2 = .10, \text{power} = .58$. Recall, however, that the real words and abstract syllables provided different phonetic contexts for the test segments. The differences in crossover points may be a consequence of the different phonetic context ([Ca] vs. [Co]) or of effects of familiarity, concreteness, imageability, meaningfulness, or word frequency for which adult data were used as estimates. Both groups showed a similar shift in category boundary as a function of this phonetic context, $F(1, 160) = 1.105, p = .295, \omega^2 = .002, \text{power} = .25$. No other interactions were significant.

As a measure of the sharpness of the labeling boundaries, standard deviations from the probit analyses were used to calculate boundary widths (Kuhl & Miller, 1978; Zlatin & Koenigsnecht, 1975). Boundary width was defined as the linear distance between the 25th and 75th percentiles as determined by the mean and standard deviation obtained from probit analysis. Larger boundary widths have sometimes been hypothesized to indicate more overlap between adjacent phonemic classes. Objectively, wider widths are a direct consequence of lower slopes consequent to less extreme performance on extreme tokens. Children with SLI had a mean boundary width of 22.57 ms, while that for typically developing children was 9.88 ms, a significant difference, $F(1, 160) = 4.096, p = .045, \omega^2 = .07, \text{power} = .42$. For all children, boundary widths were larger for synthetic speech than for natural speech, $F(1, 160) = 8.49, p = .004, \omega^2 = .15, \text{power} = .78$, but there was no difference in boundary width for real words versus abstract syllables, $F(1, 160) = 0.244, p = .622, \omega^2 < 0$. No interactions were significant.

Finally, boundary widths were compared separately for each speech series. Groups did not differ for naturally spoken versions of real words, $F(1, 40) = 0.010, p = .923, \omega^2 < 0$. However, children with SLI exhibited greater boundary width for naturally spoken versions of meaningless syllables, and this difference approached significance, $F(1, 40) = 3.606, p = .06, \omega^2 = .06, \text{power} = .36$. For synthetic renditions of real words, there were no significant differences in boundary width, $F(1, 40) = .768, p = .386, \omega^2 < 0$. However, children with SLI had slightly larger boundary widths for synthetic renditions of abstract syllables, $F(1, 40) = 3.977, p = .053, \omega^2 = .07, \text{power} = .37$, with this effect also approaching significance.

Discrimination

Discrimination functions for each series were obtained by calculating the percentage of “different” responses (out of four) for each stimulus pair. As explained above, stimulus pairs consisted of tokens differing by 20 ms VOT. The four stimulus pairs were 00–20, 10–30, 20–40, and 30–50. The results are shown in Figures 2,3,4, and 5. To compare discrimination functions, the percentage of “different” responses was entered into a 2 (group) \times 4 (stimulus pair) \times 2 (speech source) \times 2 (lexical status) mixed design ANOVA. Overall, there was a significant effect of stimulus level, $F(3, 480) = 162.368, p < .0001, \omega^2 = .79, \text{power} = .99$, indicating better discrimination of cross-boundary pairs and poorer discrimination of pairs drawn from the same side of the labeling boundary. The two groups did not differ in overall

discrimination functions, $F(1, 3) = 0.004$, $p = .954$, $\omega^2 < 0$. Further, there were no differences in discrimination due to speech source, $F(1, 3) = 0.246$, $p = .654$, $\omega^2 < 0$, or lexical nature of test items, $F(1, 3) = 0.023$, $p = .889$, $\omega^2 < 0$. None of the two-way interactions approached significance, but the three-way interaction was significant, $F(1, 3) = 15.565$, $p = .029$, $\omega^2 = .26$, power = .35. Analysis of this interaction revealed that none of the two-way interactions was mediated by the third variable. Discrimination results were therefore considered separately for each speech series.

For naturally spoken real words, the percentage of “different” responses was entered into a 2 (group) \times 4 (stimulus level) mixed design ANOVA, which revealed that the two groups differed in their discrimination functions, $F(3, 120) = 4.521$, $p = .005$, $\omega^2 = .08$, power = .57. Analysis of this difference revealed that the groups did not differ on discriminating the 00-ms “bowl” from the 20-ms “bowl,” $F(1, 40) = 0.045$, $p = .832$, $\omega^2 < 0$, nor in discriminating the 20-ms “bowl” from the 40-ms “pole,” $F(1, 40) = 0.031$, $p = .862$, $\omega^2 < 0$. However, children with SLI were significantly less likely to discriminate the 10-ms “bowl” from the 30-ms “pole,” $F(1, 40) = 4.971$, $p = .032$, $\omega^2 = .09$, power = .37, and more likely to discriminate the 30-ms “pole” from the 50-ms “pole,” $F(1, 40) = 5.488$, $p = .024$, $\omega^2 = .10$, power = .39.

For the naturally spoken “ba”–“pa” series, the two groups again differed in their discrimination functions, $F(3, 120) = 4.11$, $p = .008$, $\omega^2 = .07$, power = .55. The only difference in this function was for the two “pa” tokens. Children with SLI were significantly more likely to call them “different,” $F(1, 40) = 10.554$, $p = .002$, $\omega^2 = .19$, power = .83. The two groups did not differ when discriminating any of the other tokens, 00–20, $F(1, 40) = 0.971$, $p = .330$, $\omega^2 < 0$; 10–30, $F(1, 40) = 2.119$, $p = .153$, $\omega^2 = .02$, power = .30; 20–40, $F(1, 40) = 1.98$, $p = .167$, $\omega^2 = .02$, power = .29.

For the synthetic real words, the two groups again differed in their discrimination functions, $F(3, 120) = 2.691$, $p = .049$, $\omega^2 = .04$, power = .33. The children with SLI were more likely to call the two “bowl” tokens “different,” $F(1, 40) = 5.792$, $p = .021$, $\omega^2 = .10$, power = .57, with no difference for any other pairs, 10–30, $F(1, 40) = 1.486$, $p = .230$, $\omega^2 = .01$, power = .28; 20–40, $F(1, 40) = 1.25$, $p = .270$, $\omega^2 = .01$, power = .27; 30–50, $F(1, 40) = 2.738$, $p = .106$, $\omega^2 = .04$, power = .30.

For synthetic abstract syllables, the groups differed in their discrimination functions, $F(1, 40) = 3.867$, $p = .011$, $\omega^2 = .06$, power = .36. The groups did not differ in their discrimination of the 00-ms “ba” and the 20-ms “ba,” $F(1, 40) = 0.870$, $p = .357$, $\omega^2 < 0$; in their discrimination of the 20-ms “ba” from the 40-ms “pa,” $F(1, 40) = 1.243$, $p = .272$, $\omega^2 = .01$, power = .27; or in their discrimination of the 30-ms “pa” from the 50-ms “pa,” $F(1, 40) = 3.018$, $p = .09$, $\omega^2 = .05$, power = .33. Children with SLI were less likely to discriminate the 10-ms “ba” from the 30-ms “pa,” $F(1, 40) = 11.173$, $p = .002$, $\omega^2 = .20$, power = .85.

As a check on attention, group differences in discrimination of “no change” (catch) tokens was examined. Over all conditions, children with SLI identified 15.77% of catch tokens as “different,” while typically developing children identified 10.17% of catch trials as

“different.” This difference between groups was significant, $F(1, 160) = 4.50, p = .035, \omega^2 = .08, \text{power} = .38$. There was no effect of lexical status on catch trial accuracy, $F(1, 160) = 0.846, p = .359, \omega^2 < 0$, nor of speech source, $F(1, 160) = 1.227, p = .270, \omega^2 = .01, \text{power} = .20$. None of the interactions was significant. When each speech series was considered separately, group differences only approached significance for the naturally spoken “ba”–“pa” series. Children with SLI were slightly more likely to call catch trials “different,” $F(1, 40) = 3.102, p = .086, \omega^2 = .05, \text{power} = .34$. There were no group differences for any of the other speech series: natural “bowl”–“pole,” $F(1, 40) = 0.437, p = .512, \omega^2 < 0$; synthetic “bowl”–“pole,” $F(1, 40) = 0.704, p = .406, \omega^2 < 0$; and synthetic “ba”–“pa,” $F(1, 40) = 0.752, p = .391, \omega^2 < 0$.

Predictability

Individual identification functions were used to predict discrimination functions based on simple probabilities. For example, if two tokens were both perceived as [–voice] or [p] 20% of the time, then they should be perceived as “different” 32% of the time [Token 1 “p” probability multiplied by Token 2 “b” probability ($.20 \times .80 = .16$) plus Token 1 “b” probability multiplied by Token 2 “p” probability ($.80 \times .20 = .16$)] and the “same” 68% of the time [Token 1 “p” probability multiplied by Token 2 “p” probability ($.20 \times .20 = .04$) plus Token 1 “b” probability multiplied by Token 2 “b” probability ($.80 \times .80 = .64$)]. The correlation between predicted and obtained discrimination values was then calculated for both groups.

Over all four speech series, the correlation between observed and predicted values for children with SLI was $r = .559$, while that for typically developing children was $r = .704$. The correlation for typically developing children (TYP) was significantly higher than that for children with SLI ($z = 4.45, p < .0001, \omega^2 = .31, \text{power} = .99$). For only naturally spoken real words, the correlations between observed and predicted values did not differ between groups ($r = .684$ for SLI, $r = .712$ for TYP; $z = .49, p = .62, \omega^2 < 0$). For the other three series, however, the correlations for children with SLI were significantly lower than those for typically developing children. For naturally spoken “ba”–“pa,” SLI $r = .531$, TYP $r = .747$ ($z = 3.37, p = .0008, \omega^2 = .20, \text{power} = .89$). For synthetic “bowl”–“pole,” the correlation for children with SLI was $r = .467$, while that for typically developing children was $r = .641$ ($z = 2.29, p = .02, \omega^2 = .09, \text{power} = .54$). For synthetic “ba”–“pa,” the correlation for children with SLI was $r = .554$, while the correlation for typically developing children was $.716$ ($z = 2.48, p = .013, \omega^2 = .11, \text{power} = .62$).

For the naturally spoken real words, children with SLI and typically developing children had significantly different discrimination functions but remarkably similar correlations between observed and predicted discrimination functions. Therefore, it is worthwhile to examine how the two groups deviated from predicted discrimination values. Recall that the two groups differed in their discrimination of one cross-boundary pair, the 10-ms “bowl” and the 30-ms “pole,” and one pair labeled the same, the 30-ms “pole” and the 50-ms “pole.” Comparisons of observed and predicted values reveal that these two comparisons were indeed different for the two groups. Observed discrimination was slightly lower than predicted discrimination for the 10 -ms and 30-ms VOT to-kens for the children with SLI, $F(1, 40) = 3.987, p = .053$,

$\omega^2 = .07$, power = .38, but not for the typically developing children, $F(1, 40) = 0.507$, $p = .480$, $\omega^2 < 0$. Conversely, the difference between observed and predicted values for discriminating the 30- and 50-ms tokens was greater for typically developing children, $F(1, 40) = 3.183$, $p = .082$, $\omega^2 = .05$, power = .35, than for children with SLI, $F(1, 40) = 0.995$, $p = .325$, $\omega^2 < 0$. No other observed values were significantly different from predicted values for either group. Children with SLI discriminated the 10-ms and 30-ms tokens less well than would be predicted by their identification functions, while typically developing children discriminated the 30- and 50-ms tokens less well than would be predicted by their identification functions. These results suggest that although there were significant differences between the two groups' discrimination functions, both groups differed similarly from their predicted values.

IQ-Matched Groups

All children in the current study had nonverbal IQ scores greater than 1SD below the mean (>85). While all of the children scored in the normal range, children with SLI scored significantly lower on the Leiter-R, a measure of nonverbal intelligence (Roid & Miller, 1997). A reasonable conclusion of the current results is that children with lower language and/or nonverbal intelligence scores perform more poorly than unaffected children on speech perception measures. To rule out nonverbal intelligence as a mediating factor, the results were reanalyzed for a subgroup of children matched on nonverbal intelligence. The IQ-matched groups were created by removing 4 children from both groups—the 4 children with SLI with the lowest nonverbal IQ scores (4 at 87), and the 4 typically developing children with the highest nonverbal IQ scores (range: 119–129). The remaining 17 children with SLI had a mean nonverbal IQ of 99.2, while the remaining 17 typically developing children had a mean nonverbal IQ of 103.0, a nonsignificant difference, $t(32) = 1.331$, $p = .193$, $\omega^2 = .012$, power = .22. However, the two groups still differed on all diagnostic measures: CELF-III ELS, $t(30) = 9.11$, $p < .0001$, $\omega^2 = .66$, power = .99; PPVT-III, $t(32) = 3.54$, $p < .001$, $\omega^2 = .22$, power = .93; EVT, $t(32) = 3.57$, $p = .001$, $\omega^2 = .22$, power = .93; NWR, $t(32) = 3.52$, $p = .001$, $\omega^2 = .21$, power = .92; CLPT, $t(32) = 3.59$, $p = .001$, $\omega^2 = .22$, power = .93.

Generally speaking, the results for groups matched on nonverbal intelligence replicated the results for the complete groups. For the naturally spoken real-word series, there were no group differences in overall identification functions, $F(5, 160) = 0.114$, $p = .989$, $\omega^2 < 0$; boundary widths, $F(1, 32) = 0.106$, $p = .747$, $\omega^2 < 0$; or endpoint accuracy, $F(1, 32) = 0.634$, $p = .432$, $\omega^2 < 0$. As in the complete groups, there were group differences in discrimination functions, $F(3, 96) = 3.278$, $p = .02$, $\omega^2 = .05$, power = .42, carried by almost-significant differences in the ability to discriminate the 10-ms and 30-ms tokens, $F(1, 32) = 3.777$, $p = .06$, $\omega^2 = .06$, power = .40, and the 30-ms and 50-ms tokens, $F(1, 32) = 3.796$, $p = .06$, $\omega^2 = .06$, power = .40. Further, the two groups did not differ in how accurately they responded to catch trials, $F(1, 32) = 0.037$, $p = .849$, $\omega^2 < 0$. Finally, both groups showed similar correlations between observed and predicted discrimination values: SLI $r = .718$, TYP $r = .684$ ($z = .537$, $p = .592$, $\omega^2 < 0$). For the naturally spoken abstract-syllable series, there were no group differences in identification functions, $F(5, 160) = 0.843$, $p = .521$, $\omega^2 < 0$; boundary widths, $F(1, 32) = 2.076$, $p = .159$, $\omega^2 = .03$, power = .26; or accuracy on

unambiguous endpoints, $F(1, 32) = 1.354, p = .253, \omega^2 = .01, \text{power} = .20$. There were still group differences in discrimination functions, $F(3, 96) = 3.497, p = .019, \omega^2 = .06, \text{power} = .41$, but the difference in accuracy on catch trials was no longer significant, $F(1, 32) = 2.497, p = .124, \omega^2 = .03, \text{power} = .28$. However, identification functions were still less predictive of discrimination functions for children with SLI: SLI $r = .556$, TYP $r = .716$ ($z = 2.2, p = .03, \omega^2 = .08, \text{power} = .52$).

For the synthetic real-word series, there were no group differences in identification functions, $F(5, 160) = 0.675, p = .643, \omega^2 < 0$; boundary width measurements, $F(1, 32) = 0.347, p = .560, \omega^2 < 0$; or endpoint accuracy, $F(1, 32) = 0.672, p = .418, \omega^2 < 0$. Group differences in discrimination functions approached significance, $F(3, 96) = 2.341, p = .078, \omega^2 = .03, \text{power} = .30$, but there were no significant differences in accuracy on catch trials, $F(1, 32) = 1.017, p = .321, \omega^2 = .0004, \text{power} = .20$. Identification functions were still less predictive of discrimination functions for children with SLI: SLI $r = .457$, TYP $r = .679$ ($z = 2.69, p = .016, \omega^2 = .13, \text{power} = .72$). For the synthetic abstract-syllable series, there were significant group differences in identification functions, $F(5, 160) = 3.822, p = .003, \omega^2 = .06, \text{power} = .60$, but no differences in boundary width, $F(1, 32) = 2.074, p = .160, \omega^2 = .03, \text{power} = .30$, or endpoint accuracy, $F(1, 32) = 1.551, p = .222, \omega^2 = .01, \text{power} = .22$. There were still group differences in discrimination, $F(3, 96) = 3.417, p = .02, \omega^2 = .05, \text{power} = .40$, but no group differences on catch trials, $F(1, 32) = 0.895, p = .351, \omega^2 < 0$. However, discrimination functions of children with SLI were less predicted by identification functions: SLI $r = .535$, TYP $r = .772$ ($z = 3.45, p = .003, \omega^2 = .21, \text{power} = .91$).

Discussion

The purpose of the present study was to examine how differences in the quality of speech tokens affect how children with SLI perceive speech. To this end, four different speech series were created, differing in their source (natural vs. synthetic speech) and in their lexical status (real words vs. abstract nonsense syllables). The hypothesis was that children with SLI would perceive speech most categorically with naturally spoken versions of real words and that perception would be less categorical for synthetic speech and for nonword test tokens. This hypothesis was directly supported. When children with SLI and typically developing children participated in a categorical perception task with digitally edited versions of naturally spoken real words, there were no group differences in performance. Crossover points, boundary widths, and accuracy of unambiguous endpoint tokens were the same for the two groups. There were differences in group discrimination functions, but similar correlations between observed and predicted discrimination values revealed that both groups' observed discrimination functions differed similarly from predicted functions. When children with SLI hear naturally spoken versions of real words, their perception matches that of typically developing age-matched children. This replicates previous work suggesting that children with SLI do not have an underlying auditory or phonological perceptual deficit, but rather that their performance is limited by memory, processing, and representational demands (Coady et al., 2005).

This study also replicates previous research showing that children with SLI perceive synthetic versions of abstract syllables less accurately than typically developing children

(Elliott & Hammer, 1988; Elliott et al., 1989; Evans et al., 2002; Joanisse et al., 2000; Leonard et al., 1992; McReynolds, 1966; Stark & Heinz, 1996a, 1996b; Sussman, 1993, 2001; Tallal & Piercy, 1974, 1975; Tallal & Stark, 1981; Tallal et al., 1981; Thibodeau & Sussman, 1979). Overall, in categorical perception tasks, children with SLI placed their labeling boundaries at the same locations as typically developing children. However, they were less accurate on unambiguous endpoint tokens, their labeling boundary widths were larger (suggesting shallower identification functions), their peak discrimination values were lower, and their identification functions were less predictive of their discrimination functions, all relative to age-matched control children (Joanisse et al., 2000; Sussman, 1993; Thibodeau & Sussman, 1979). This pattern of results was independent of group differences in nonverbal intelligence. The notable exception to this pattern was for naturally spoken real-word series.

In several instances, group differences in perception patterns approached, but did not reach, an absolute cutoff value of $\alpha = .05$. This is due to greater response variability in children with SLI, replicating previous studies. Thibodeau and Sussman (1979), Sussman (1993), and Joanisse and colleagues (2000) all reported greater variability in the response patterns of children with language impairments, relative to children with normal language abilities. This increased variability in the performance of children with language impairments has traditionally been viewed as “noise” in the data. However, more recent models view it as the actual phenomenon to be explained (see, e.g., Evans, 2002). Whatever its nature, increased variability has the effect of increasing error variance, reducing the likelihood of finding a significant effect. In terms of the current experiments, children with SLI responded less consistently to synthetic speech and to meaningless syllables. Again, the notable exception to this pattern was the naturally spoken real-word series. As a group, children with SLI responded as consistently as their age-matched peers to naturally spoken real words.

In some past studies, authors have focused on these differences in endpoint performance. Slope differences, typically shallower slopes for children with SLI, are a direct consequence of performance on endpoints in as much as slopes always are shallower when performance on extreme tokens is less likely to approach minima and maxima (0 and 100%). Here, the largest differences between endpoint performances for children with SLI and typically developing children are only about 5%. It is worthwhile to consider exactly what this entails. Children respond to every stimulus 10 times. If a child fails to attend to any single presentation of a stimulus, performance for that single presentation would be at chance (50–50). Thus, if one assumes that a typically developing child always hears the endpoint stimulus the same when he or she attends to it, 95% identification corresponds to paying attention 9 out of 10 times and randomly providing one or the other responses on the 10th presentation. Now consider children with SLI. Performance of 90% on endpoint stimuli may correspond to paying close attention to 8, instead of 9, out of 10 trials. Note that, for intermediate stimuli for which ideal identification may be 50%, neglect of some trials has no effect on the data.

Wightman and Allen (1992) provided multiple simulations demonstrating how guessing on unattended trials may account for most, if not all, of the observed differences between typical children and adults on a variety of listening tasks, and that these data could

erroneously be taken to indicate rather dramatic differences in auditory function for the two groups. They further concluded that “apparent differences between adults and children in auditory skills may reflect nothing more than the influence of nonsensory factors such as memory and attention” (p. 133). Similar conclusions can be drawn from Viemeister and Schlauch’s (1992) detailed simulations of attentional effects in data from listening tasks with infants. Given that differences found here are very subtle relative to much more dramatic differences between children and adults, it may be unproductive to perseverate upon small group differences for endpoint stimuli and consequent changes in slope. Caution should be taken, as these small differences could lead to erroneous conclusions of perceptual impairments in children with SLI.

Word-nonword effects of the present study are consistent with most contemporary models of auditory lexical access for which the lexicon is not viewed as a passive recipient of preliminary extraction of phonetic information. Instead, acoustic/auditory information is viewed as interacting with lexical structure and processes. McClelland and Elman’s (1986) TRACE model is one of the earliest formalizations of such models. Lexical effects on speech perception have been reported in experiments examining adults’ perception of word-nonword series. If listeners identify tokens from a VOT series ranging from a real word to a nonword (e.g., *beace-peace* or *beef-peef*), the labeling boundary is shifted toward the nonword because listeners are more likely to attach the real-word label to ambiguous tokens from the middle of the series (Ganong, 1980). These effects hold for perception of a VOT series in which both endpoints are real words that differ in word frequency (e.g., *best-pest*; Connine, Titone, & Wang, 1993) and for VOT series in which end-points differ in neighborhood density (Newman, Sawusch, & Luce, 1997). In both of these cases, listeners were more likely to label ambiguous middle tokens as the more frequently occurring option—the more frequently occurring word or the nonword with higher neighborhood density.

Besides the effects for adult listeners, researchers have long recognized that speech perception abilities and lexical knowledge are linked very early in development. Ferguson and Farwell (1975) were among the first to point out that children learn to discriminate minimal phonetic pairs by learning to discriminate words in their language. They referred to this as “the primacy of lexical learning in phonological development” (p. 437). Recent models of developmental speech processing (e.g., Fennell & Werker, 2003; Pater, Stager, & Werker, 2004; Walley, Metsala, & Garlock, 2003; Werker & Curtin, 2005) place particular emphasis on growth and organization of the lexicon as a driving force for phonemic distinctions as emergent properties of lexical development. According to these models, children’s earliest words may be organized in a holistic fashion, without regard to phonological structure. However, as more words are added to the lexicon, children are forced to pay attention to phonetic differences that distinguish words. Because children with SLI have smaller lexicons (e.g., Leonard, 1998), they will have a smaller corpus over which regularities such as phonemes can be extracted.

Thus, there is evidence that adults and children rely on lexical knowledge during speech perception. If auditory information is of lower quality, such as speech embedded in noise or synthetic speech, one would expect greater influence of lexical structure on performance. If lexical representations also are compromised, explicitly so for nonword stimuli, speech

perception suffers. In the case of children developing language typically, this will not be a problem. As Kingston and Diehl (1994) have suggested, most people are such skilled phonetic perceivers that the entire process appears automatic. However, children with SLI have known lexical difficulties (see Leonard, 1998) and are less able to compensate for compromised input. In the present study, performance by children with SLI suffered with synthetic speech, and it also suffered when nonword stimuli were used. When the same children were presented the opportunity to exploit both of these sources of information, their perception was within the normal range.

While the findings of this study have been interpreted not as a failure by children with SLI to perceive speech categorically but rather as a result of task demands and stimulus properties, two alternative interpretations are possible. The first is that less categorical performance by children with SLI could be due to deficits in phonological working memory. The children with SLI in the current study repeated nonwords less accurately than age-matched controls, replicating a long history of such findings (for reviews, see Coady & Evans, in press; Graf Estes, Evans, & Else-Quest, 2007). While reduced accuracy in nonword repetition tasks typically has been interpreted as resulting from a deficit in phonological working memory, Coady and Evans (in press) questioned whether phonological working memory is actually a viable construct, distinct from the lexicon. They pointed out that phonological working memory is tightly linked to lexical abilities in both children with SLI and children developing language normally. If accuracy on nonword repetition tasks can be linked directly to lexical knowledge, then phonological working memory becomes an extraneous construct. This supports the preferred interpretation in which perceptual abilities in children with SLI are tightly tied to lexical knowledge, even for syllable-level contrasts.

An alternate explanation is that the findings from this study suggest that speech perception deficits in children with SLI are secondary to processing capacity deficits. There is a growing body of work showing that processing demands exceed processing capacity more quickly in children with SLI than in children acquiring language typically (e.g., Leonard, 1998). While this study sought to minimize processing demands by reducing the memory requirements of the task, the stimuli themselves provided multiple redundant cues to signal the VOT contrast. For the naturally spoken real-word stimuli, children with SLI could exploit the stacking of multiple redundant cues, both acoustic and semantic, to facilitate identification and discrimination. In the synthetic non-word condition, however, the only cue available was the relative timing between the burst and the onset of voicing. Any supplementary acoustic or semantic cues were unavailable, and performance suffered in the group of children with SLI. These findings suggest that children with SLI may need multiple redundant cues to perceive speech stimuli in a categorical fashion.

A previous study on categorical perception by children with SLI (Coady et al., 2005) included an AXAX discrimination task, in which stimuli were separated by a 100-ms ISI. In that condition, children with SLI showed a lower peak discrimination value than children developing language typically. In the current study, an AAXX discrimination task with a 500-ms ISI was modeled after that used by Sussman (1993). She found that children with SLI discriminated synthetic “ba”–“da” much like typically developing children. In the present study, children with SLI and typically developing children discriminated natural

“bowl”–“pole” in very similar ways (observed values for both groups differed equally from predicted values). However, unlike the Sussman (1993) study, children with SLI were less likely to discriminate tokens than were typically developing children for synthetic and nonword speech series. This can be explained in terms of differences in the test items. In the Sussman study, the first token in all discrimination pairs was the unambiguous “ba” endpoint token. The second token was then one of the ambiguous tokens from the synthetic “ba”–“da” series. In this anchored discrimination task, Sussman only measured discriminability of test tokens relative to the “ba” endpoint. In the current study, discrimination was measured for all stimulus pairs having a 20-ms difference. Consistent with the present authors’ hypothesis, performance was best when children with SLI discriminated tokens from the natural, real-word series. For the natural nonword series, the synthetic real-word series, and the synthetic nonword series, children with SLI showed lower peak discrimination values and lower correlations between observed and predicted discrimination values.

The results of the current study show that children with SLI perceive speech categorically when it is both natural and meaningful. Further, performance on these tasks is compromised in children with SLI when they are presented with either synthetic speech or nonsense syllables. These findings provide further evidence that auditory or phonetic perception abilities are intact in children with SLI and should be excluded as an underlying source of impairment in this population. This has clear implications for future research. Because perception is compromised for synthetic speech, experimental outcomes for any measure that uses synthetic speech instead of naturally produced speech can be misinterpreted (see also Reynolds & Fucci, 1998). Similarly, if abstract nonsense syllables are used instead of real words in studies with children with SLI, any performance deficits must be attributed, at least in part, to the language deficit itself, and not to putative auditory deficits.

Acknowledgments

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders: DC-05263 to the first author, DC-04072 to the second author, and DC-005650 to the fourth author. We are grateful to the children and their parents for participating. We thank Lisbeth Simon and Kristin Ryan for help with standardized testing, and Ariel Young Shibilski for recording the stimuli.

References

- Bishop DVM. How does the brain learn language? Insights from the study of children with and without language impairment. *Developmental Medicine and Child Neurology*. 2000; 42:133–142. [PubMed: 10698331]
- Blomert L, Mitterer H. The fragile nature of the speech-perception deficit in dyslexia: Natural vs. synthetic speech. *Brain and Language*. 2004; 89:21–26. [PubMed: 15010233]
- Chiat S. Mapping theories of developmental language impairment: Premises, predictions, and evidence. *Language and Cognitive Processes*. 2001; 16:113–142.
- Coady JA, Evans JL. The uses and interpretations of nonword repetition tasks in children with and without specific language impairments. *International Journal of Language and Communication Disorders*. in press.
- Coady JA, Kluender KR, Evans JL. Categorical perception of speech by children with specific language impairments. *Journal of Speech, Language, and Hearing Research*. 2005; 48:944–959.

- Cohen, J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press; 1977. Rev. ed.
- Coltheart M. The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*. 1981; 33(A):497–505.
- Connine CM, Titone D, Wang J. Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1993; 19:81–94.
- Damper RI, Harnad SR. Neural network models of categorical perception. *Perception & Psychophysics*. 2000; 62:843–867. [PubMed: 10883589]
- Dollaghan CA, Campbell TF. Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*. 1998; 41:1136–1146.
- Dunn, LM., Dunn, LM. *Peabody Picture Vocabulary Test*. 3. Circle Pines, MN: American Guidance Service; 1997.
- Eimas PD. Categorization in early infancy and the continuity of development. *Cognition*. 1994; 50:83–93. [PubMed: 8039377]
- Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. Speech perception in infants. *Science*. 1971 Jan 22; 171:303–306. [PubMed: 5538846]
- Eisenson, J. *Aphasia in children*. New York: Harper & Row; 1972.
- Elliott LL, Hammer MA. Longitudinal changes in auditory discrimination in normal children and children with language-learning problems. *Journal of Speech and Hearing Disorders*. 1988; 53:467–474. [PubMed: 3263544]
- Elliott LL, Hammer MA, Scholl ME. Fine-grained auditory discrimination in normal children and children with language-learning problems. *Journal of Speech and Hearing Research*. 1989; 32:112–119. [PubMed: 2704185]
- Evans JL. Variability in comprehension strategy use in children with SLI: A dynamical systems account. *International Journal of Language and Communication Disorders*. 2002; 37:95–116. [PubMed: 12012619]
- Evans JL, Viele K, Kass RE, Tang F. Grammatical morphology and perception of synthetic and natural speech in children with specific language impairments. *Journal of Speech, Language, and Hearing Research*. 2002; 45:494–504.
- Fennell CT, Werker JF. Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*. 2003; 46:245–264. [PubMed: 14748446]
- Ferguson CA, Farwell CB. Words and sounds in early language acquisition. *Language*. 1975; 51:419–439.
- Finney, DJ. *Probit analysis*. New York: Cambridge University Press; 1971.
- Ganong WF. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*. 1980; 6:110–125. [PubMed: 6444985]
- Gathercole SE. Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*. 1999; 3:410–419. [PubMed: 10529796]
- Gaulin CA, Campbell TF. Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Perceptual and Motor Skills*. 1994; 79:55–64. [PubMed: 7991333]
- Graf Estes K, Evans JL, Else-Quest N. Differences in nonword repetition performance for children with and without specific language impairment: A metaanalysis. *Journal of Speech, Language, and Hearing Research*. 2007; 50:177–195.
- Joanisse MF, Manis FR, Keating P, Seidenberg MS. Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology*. 2000; 77:30–60. [PubMed: 10964458]
- Joanisse MF, Seidenberg MS. Specific language impairment: A deficit in grammar or processing? *Trends in Cognitive Sciences*. 1998; 2:240–247. [PubMed: 21244922]
- Kingston J, Diehl RL. *Phonetic knowledge*. *Language*. 1994; 70:419–454.
- Klatt DH. Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*. 1980; 67:971–995.

- Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality parameters among female and male talkers. *Journal of the Acoustical Society of America*. 1990; 87:820–857. [PubMed: 2137837]
- Kluender, KR. Speech perception as a tractable problem in cognitive science. In: Gernsbacher, MA., editor. *Handbook of psycholinguistics*. San Diego, CA: Academic Press; 1994. p. 173-217.
- Kluender KR, Lotto AJ. Virtues and perils of an empiricist approach to speech perception. *Journal of the Acoustical Society of America*. 1999; 105:503–511.
- Kluender, KR., Lotto, AJ., Holt, LL. Contributions of nonhuman animal models to understanding human speech perception. In: Greenberg, S., Ainsworth, W., editors. *Listening to speech: An auditory perspective*. Mahwah, NJ: Erlbaum; 2006. p. 203-220.
- Kucera, F., Francis, W. *Computational analysis of present-day American English*. Providence, RI: Brown University Press; 1967.
- Kuhl PK. Discrimination of speech by nonhuman animals: Basic sensitivities conducive to the perception of speech-sound categories. *Journal of the Acoustical Society of America*. 1981; 70:340–349.
- Kuhl PK, Miller JD. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar-plosive consonants. *Science*. 1975 Oct 3.190:69–72. [PubMed: 1166301]
- Kuhl PK, Miller JD. Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America*. 1978; 63:905–917. [PubMed: 670558]
- Leonard, LB. *Children with specific language impairment*. Cambridge, MA: MIT Press; 1998.
- Leonard L, McGregor K, Allen G. Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech and Hearing Research*. 1992; 35:1076–1085. [PubMed: 1447919]
- Luce PA, Feustel TC, Pisoni DB. Capacity demands in short-term memory for synthetic and natural speech. *Human Factors*. 1983; 25:17–32. [PubMed: 6840769]
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology*. 1986; 18:1–86. [PubMed: 3753912]
- McReynolds LV. Operant conditioning for investigating speech sound discrimination in aphasic children. *Journal of Speech and Hearing Research*. 1966; 9:519–528.
- Miller CA, Kail R, Leonard LB, Tomblin JB. Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. 2001; 44:416–433.
- Newman RS, Sawusch JR, Luce PA. Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception and Performance*. 1997; 23:873–889. [PubMed: 9180048]
- Nusbaum HC, Pisoni DB. Constraints on the perception of synthetic speech generated by rule. *Behavior Research Methods, Instruments, and Computers*. 1985; 17:235–242.
- Pater J, Stager C, Werker J. The perceptual acquisition of phonological contrasts. *Language*. 2004; 80:384–402.
- Pisoni, DB., Luce, PA. Trading relations, acoustic cue integration, and context effects in speech perception. In: Schouten, MEH., editor. *The psychophysics of speech perception*. Dordrecht, The Netherlands: Martinus Nijhoff; 1987. p. 155-172.
- Reynolds ME, Fucci D. Synthetic speech comprehension: A comparison of children with normal and impaired language skills. *Journal of Speech, Language, and Hearing Research*. 1998; 41:458–466.
- Roid, GH., Miller, LJ. *Leiter International Performance Scale—Revised*. Wood Dale, IL: Stoelting; 1997.
- Semel, E., Wiig, E., Secord, W. *Clinical Evaluation of Language Fundamentals*. 3. San Antonio, TX: The Psychological Corporation; 1995.
- Stark RE, Heinz JM. Perception of stop consonants in children with expressive and receptive-expressive language impairments. *Journal of Speech and Hearing Research*. 1996a; 39:676–686. [PubMed: 8844549]
- Stark RE, Heinz JM. Vowel perception in children with and without language impairment. *Journal of Speech and Hearing Research*. 1996b; 39:860–869. [PubMed: 8844565]

- Sussman JE. Perception of formant transition cues to place of articulation in children with language impairments. *Journal of Speech and Hearing Research*. 1993; 36:1286–1299. [PubMed: 8114495]
- Sussman JE. Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions? *Journal of the Acoustical Society of America*. 2001; 109:1173–1180. [PubMed: 11303931]
- Tallal P, Piercy M. Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologist*. 1973; 11:389–398.
- Tallal P, Piercy M. Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*. 1974; 12:83–93. [PubMed: 4821193]
- Tallal P, Piercy M. Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia*. 1975; 13:69–74. [PubMed: 1109463]
- Tallal P, Stark RE. Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America*. 1981; 69:568–574. [PubMed: 7462479]
- Tallal P, Stark R, Kallman C, Mellits D. A reexamination of some nonverbal perceptual abilities of language-impaired and normal children as a function of age and sensory modality. *Journal of Speech and Hearing Research*. 1981; 24:351–357. [PubMed: 7300276]
- Thibodeau LM, Sussman HM. Performance on a test of categorical perception of speech in normal and communication disordered children. *Journal of Phonetics*. 1979; 7:375–391.
- Viemeister, NF., Schlauch, RS. Issues in infant psychoacoustics. In: Werner, LA., Rubel, EW., editors. *Developmental psychoacoustics*. Washington, DC: American Psychological Association; 1992. p. 191-209.
- Walley AC, Metsala JL, Garlock VM. Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading & Writing*. 2003; 16:5–20.
- Werker JF, Curtin S. PRIMIR: A developmental model of speech processing. *Language Learning and Development*. 2005; 1:197–234.
- Wightman, F., Allen, P. Individual differences in auditory capability among preschool children. In: Werner, LA., Rubel, EW., editors. *Developmental psychoacoustics*. Washington, DC: American Psychological Association; 1992. p. 113-134.
- Williams, KT. *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Service; 1997.
- Wood CC. Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *Journal of the Acoustical Society of America*. 1976; 60:1381–1389. [PubMed: 1010890]
- Zlatin MA, Koenigsknecht RA. Development of the voicing contrast: Perception of stop consonants. *Journal of Speech and Hearing Research*. 1975; 18:541–553. [PubMed: 1186163]

Appendix

Chronological age (years;months.days), Composite Expressive (ELS) and Receptive (RLS) Language Scores on the CELF-III, standard receptive and expressive vocabulary scores, percent phonemes correct on the NWR, percent final words recalled on CLPT, and standard scores on nonverbal IQ measure (Leiter-R) for children with SLI.

Participant	Age	ELS	RLS	PPVT-III	EVT	NWR	CLPT	Nonverbal IQ
E-SLI1	9;0.14	84	86	112	80	66.67	26.19	102
E-SLI2	10;4.13	72	90	97	89	81.25	52.38	102
ER-SLI1	8;7.8	80	65	91	88	83.33	40.48	95
ER-SLI2	8;11.1	78	75	96	87	81.25	26.19	97
ER-SLI3	9;2.1	78	69	92	86	84.93	52.38	111
ER-SLI4	9;4.1	80	80	88	88	81.25	38.1	89
ER-SLI5	9;4.11	84	82	103	109	67.71	42.86	107

Participant	Age	ELS	RLS	PPVT-III	EVT	NWR	CLPT	Nonverbal IQ
ER-SLI6	9;7.13	84	75	93	81	86.46	2.38	111
ER-SLI7	9;8.21	69	53	90	74	63.54	30.95	87
ER-SLI8	9;11.29	84	78	91	93	81.25	45.24	98
ER-SLI9	10;2.29	53	57	95	80	65.63	28.57	89
ER-SLI10	10;4.9	57	50	88	74	79.17	16.67	87
ER-SLI11	10;6.26	53	50	102	82	81.25	33.33	87
ER-SLI12	10;6.28	57	50	105	80	61.46	26.19	119
ER-SLI13	10;8.3	53	53	66	74	75	28.57	87
ER-SLI14	10;8.29	75	72	87	81	85.42	54.76	102
ER-SLI15	11;2.12	69	50	69	81	77.08	23.81	91
ER-SLI16	11;3.28	65	80	94	70	75	42.86	98
ER-SLI17	11;7.8	72	50	82	77	81.25	47.62	89
ER-SLI 18	11;8.8	78	53	93	81	91.67	40.48	100
ER-SLI19	11;10.18	53	53	66	74	89.58	26.19	87

Note. CELF-III = Clinical Evaluation of Language Fundamentals–Third Edition; PPVT-III = Peabody Picture Vocabulary Test–Third Edition; EVT = Expressive Vocabulary Test; NWR = Nonword Repetition Task; CLPT = Competing Language Processing Task; Leiter-R = Leiter International Performance Scale–Revised; SLI = specific language impairments.

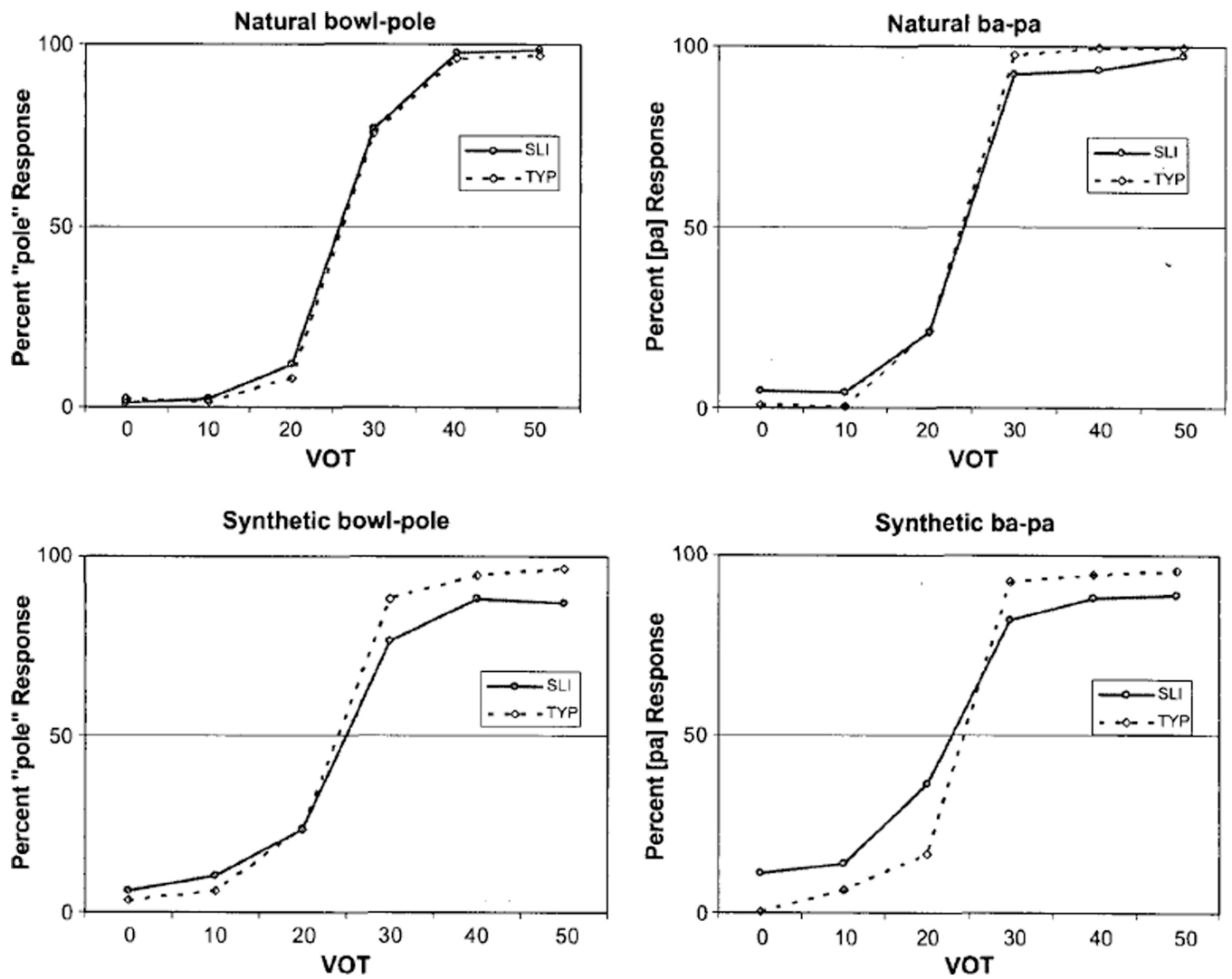


Figure 1. Identification functions for each of the four speech series for both children with specific language impairment (SLI; solid lines) and typically developing children (TYP; dotted lines). Percentage of “pole” responses is plotted as a function of voice-onset time (VOT). Identification of the naturally spoken real-word series is shown in the upper left, naturally spoken abstract syllables in the upper right, synthesized renditions of real words in the lower left, and synthesized abstract syllables in the lower right.

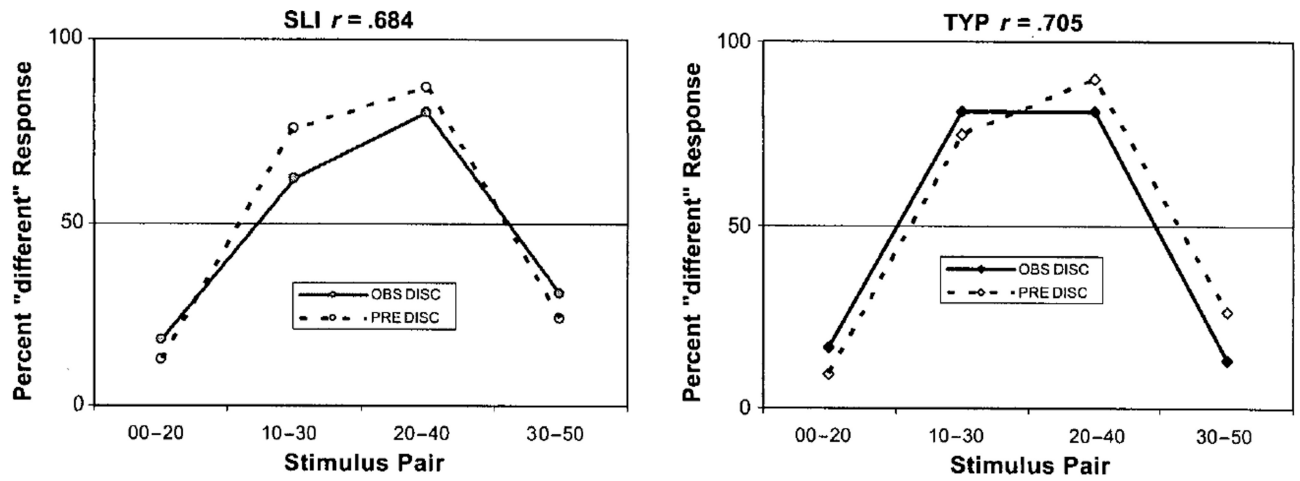


Figure 2. Discrimination functions for children with SLI (left plot) and typically developing children (TYP; right plot) for the naturally spoken real-word series. Percentage of “different” responses is plotted as a function of stimulus pair. OBS DISC = observed discrimination; PRE DISC = predicted discrimination.

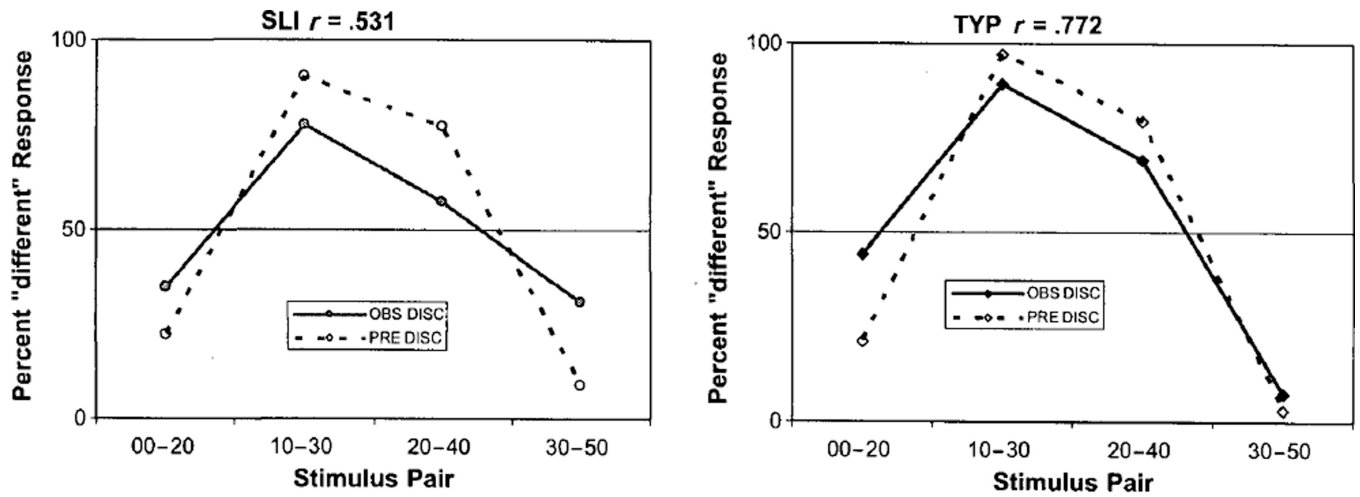


Figure 3. Discrimination functions for children with SLI (left plot) and typically developing children (TYP; right plot) for the naturally spoken nonword series. Percentage of “different” responses is plotted as a function of stimulus pair. OBS DISC = observed discrimination; PRE DISC = predicted discrimination.

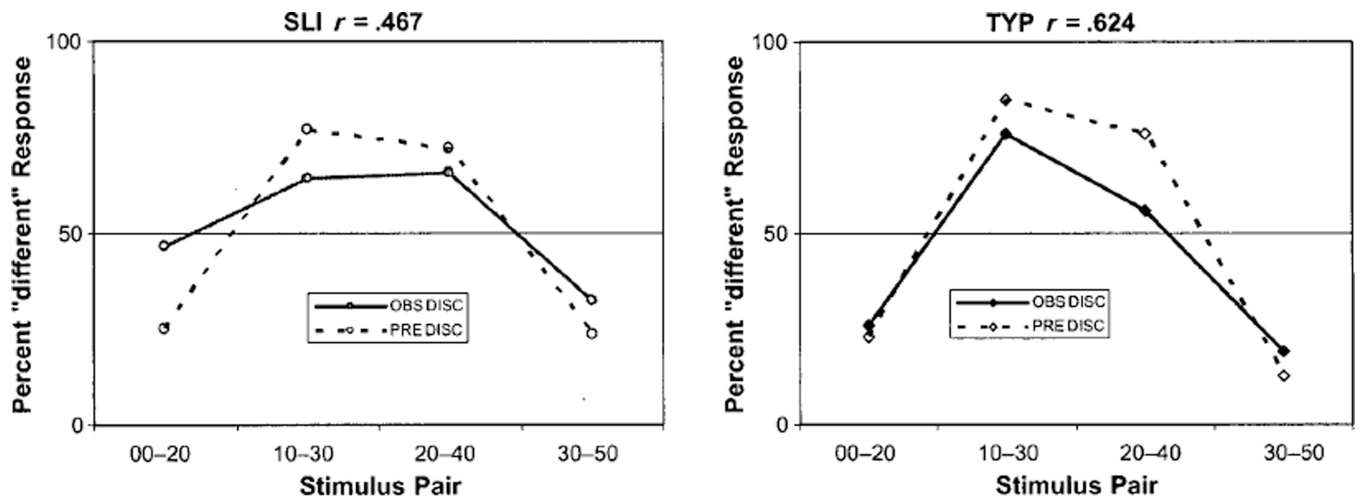


Figure 4.

Discrimination functions for children with SLI (left plot) and typically developing children (TYP; right plot) for synthesized real word series. Percentage of “different” responses is plotted as a function of stimulus pair. OBS DISC = observed discrimination; PRE DISC = predicted discrimination.

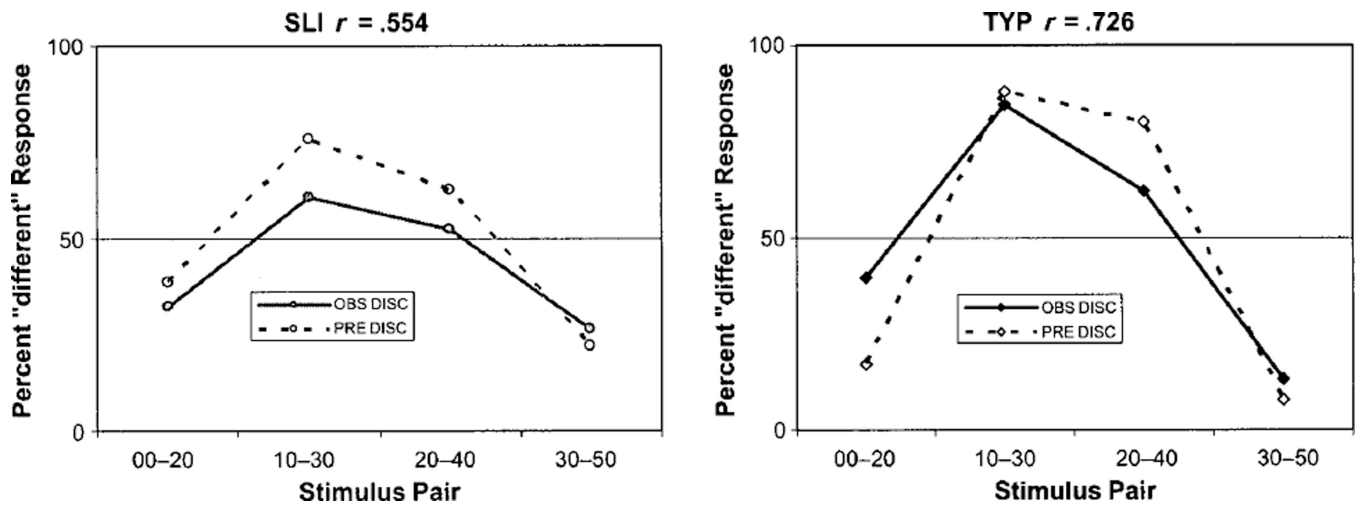


Figure 5. Discrimination functions for children with SLI (left plot) and typically developing children (TYP; right plot) for the synthetic nonword series. Percentage of “different” responses is plotted as a function of stimulus pair. OBS DISC = observed discrimination; PRE DISC = predicted discrimination.

Table 1

Group summary statistics for children with specific language impairments (SLI) and for typically developing children.

	Children with SU	Typically developing children
Age	10;3(1;0)	10;1 (1;2)
CELF-ELS	70.38(11.7)	110.7(12.3)
CELF-RLS	65.28(14.1)	—
PPVT-R	90.48(11.9)	108.33(11.5)
EVT	82.33 (8.5)	99.5(12.6)
NWR	78.10(8.6)	87.69 (6.7)
CLPT	34.58(13.0)	55.22(13.3)
Leiter-R	96.90 (9.5)	107.14(10.8)

Note. Means (with standard deviations in parentheses) are presented for chronological age (years;months), composite Expressive (ELS) and Receptive (RLS) Language Scores on the Clinical Evaluation of Language Fundamentals–Third Edition (CELF-III), standard scores on the Peabody Picture Vocabulary Test–Third Edition (PPVT-III) and the Expressive Vocabulary Test (EVT), percent phonemes correct on the Nonword Repetition Task (NWR), percent final words recalled on the Competing Language Processing Task (CLPT), and standard score on the Leiter International Performance Scale–Revised (Leiter-R).