

## Research Article

# Cognitive and Linguistic Sources of Variance in 2-Year-Olds' Speech-Sound Discrimination: A Preliminary Investigation

Kaylah Lalonde<sup>a</sup> and Rachael Frush Holt<sup>a</sup>

**Purpose:** This preliminary investigation explored potential cognitive and linguistic sources of variance in 2-year-olds' speech-sound discrimination by using the toddler change/no-change procedure and examined whether modifications would result in a procedure that can be used consistently with younger 2-year-olds.

**Method:** Twenty typically developing 2-year-olds completed the newly modified toddler change/no-change procedure. Behavioral tests and parent report questionnaires were used to measure several cognitive and linguistic constructs. Stepwise linear regression was used to relate discrimination sensitivity to the cognitive and linguistic measures. In addition, discrimination results from the current experiment were compared with those from 2-year-old children tested in a previous experiment.

**Results:** Receptive vocabulary and working memory explained 56.6% of variance in discrimination performance. Performance

was not different on the modified toddler change/no-change procedure used in the current experiment from in a previous investigation, which used the original version of the procedure.

**Conclusions:** The relationship between speech discrimination and receptive vocabulary and working memory provides further evidence that the procedure is sensitive to the strength of perceptual representations. The role for working memory might also suggest that there are specific subject-related, nonsensory factors limiting the applicability of the procedure to children who have not reached the necessary levels of cognitive and linguistic development.

**Key Words:** speech discrimination, children, individual differences, speech perception, development

Early identification of hearing loss and early intervention require developmentally appropriate tests and procedures that can track the growth of young children's speech perception skills, assess the benefit of sensory aids, optimize adjustments to those aids, and guide decisions about (re)habilitative intervention (Boothroyd, 2004). Speech discrimination testing is a valuable tool for determining whether children of varying ages have the auditory capacity to perceptually discriminate speech contrasts that are important for the development of phonetic representations and thus speech and language (Eisenberg, Martinez, & Boothroyd, 2007; Holt & Carney 2007; Sussman & Carney, 1989; Tyler, 1993). When young children and toddlers complete a speech

discrimination procedure, they do so with incomplete and highly variable cognitive and linguistic skills, which constrains the methods that can be used, complicates interpretation of results, and introduces high levels of individual variability that are caused by both sensory and nonsensory factors. Nonsensory factors include child factors, such as cognitive and linguistic skills; and task factors, such as how much the task demands cognition and language (Boothroyd, 2004). Accounting for nonsensory variables is the primary obstacle to creating valid and reliable tests and procedures for assessing toddlers' speech discrimination (Boothroyd, 1991; Eisenberg et al., 2007; Tyler, 1993).

Whereas investigators and clinicians reasonably try to minimize the influences of nonsensory factors, they often do not examine their inevitable contribution to the variability in performance of toddlers and young children. Recently, we modified the change/no-change procedure (Sussman & Carney, 1989), a speech discrimination task wherein listeners indicate whether they detect a change in the string of speech sounds by using a developmentally appropriate motor response. These

<sup>a</sup>Indiana University, Bloomington

Correspondence to Kaylah Lalonde: klalonde@indiana.edu

Editor: Craig Champlin

Associate Editor: Emily Tobey

Received July 13, 2012

Revision received April 15, 2013

Accepted May 23, 2013

DOI: 10.1044/1092-4388(2013)12-0227

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

modifications made the procedure suitable for toddlers age 2.5 years and older (Holt & Lalonde, 2012). As is typical in toddler speech perception, performance on the task was widely variable, and the youngest 2-year-olds struggled with the procedure. The purpose of the current investigation was to identify nonsensory (e.g., cognitive and linguistic) sources of variance in performance on this toddler speech-sound discrimination task, the toddler change/no-change procedure. By investigating the relationship between individual differences in nonsensory child factors (in this case, cognitive and linguistic development) and individual differences in discrimination sensitivity, we can make conclusions about nonsensory task requirements and get a better sense of what the task is measuring. A secondary aim was to examine whether further modifications to the procedure would lead to reliable testing of 24- to 30-month-old toddlers.

### ***Role of Nonsensory Factors in Toddler Speech Perception Testing***

In an ideal behavioral test, the response to a stimulus reflects only the listener's perceptual abilities. In reality, it is impossible to eliminate all nonsensory factors from behavioral testing, especially in pediatric populations but particularly in toddlers (Allen & Wightman, 1992; Boothroyd, 1991; Eisenberg et al., 2007; Tyler, 1993). As Boothroyd (1971) pointed out, child factors such as attention, orientation, linguistic skills, motor development, and phonological knowledge constrain methods that can be used to test toddlers. These factors specifically create test-related effects that may influence interpretation of test results and introduce high levels of variability, even among typically developing children. Further, the parallel development of cognition, language, and speech perception causes difficulty in determining how much developmental change and individual differences in discrimination (particularly on the poor end of the spectrum) are related to sensory-perceptual factors versus task-related factors (Eisenberg et al., 2007; Holt & Carney, 2007).

### ***Toddler Cognition and Language***

Two-year-olds are a particularly difficult age group to evaluate because of their developmental status in many domains. First, on average only 50% of their speech is intelligible (Coplan & Gleason, 1988; Weiss & Lillywhite, 1976). Second, there is evidence that 2-year-olds' phonetic representations are more holistic (less specified) than those of adults (Charles-Luce & Luce, 1990, 1995; Jusczyk, 1986; Walley, 1993), as less detail is required to distinguish between items in their limited lexicon (Metsala & Walley, 1998). For example, toddlers discriminate minimally different word pairs only when they know the words well, despite being able to discriminate phonological contrasts soon after birth (Barton, 1980; Eimas, Siqueland, Jusczyk, & Vigorito, 1971). Third, 2-year-olds have not yet reached the general executive function growth spurt that occurs between 3 and 6 years of age (Carlson, 2005; Diamond 2001; Garon, Bryson, & Smith, 2008; Rothbart & Posner, 2001), likely because the slow development of the prefrontal cortex, which is related to the

development of executive function, is far from complete (Benes, 2001; Madsen et al., 2010; Scheibel & Levin, 1997). Consequently, 2-year-olds are limited in their ability to hold information in sensory and short-term memory (Glass, Sachse, & von Suchodoletz, 2008; Pelphrey & Reznick, 2002), inhibit automatic responses (Garon et al., 2008), sustain focused attention (Ruff & Capozzoli, 2003), and delay gratification (Carlson, 2005; Kochanska, Murray, & Harlan, 2000; Kochanska, Murray, Jacques, Koenig, & Vandegest, 1996). Fourth, they are able to hold fewer items in short-term memory than older peers (Pelphrey & Reznick, 2002) and are only beginning to exhibit use of working memory (Call, 2001; Collier-Baker & Suddendorf, 2006; Corrigan, 1981; Ross, Boatright, Auld, & Nass, 1996; Triana & Pasnak, 1986). Fifth, their language is less complex than that of older peers (Bates, Dale, & Thal, 1995). Sixth, although they can learn rules on the basis of environmental feedback (Diamond, 1990), they encounter difficulty in associating an abstract rule with a reward (Diamond, 2006). Finally, combining demands from these various domains further taxes the developing system (Berger, 2004; Garon et al., 2008).

There are individual differences in the time course of maturation of each of these cognitive and linguistic domains, making 2-year-olds particularly susceptible to task-related variability. For example, when asked to carry out a sorting task in a counterintuitive manner (similar to a Stroop task) after sorting them in an intuitive way, some 24-month-old children fail to inhibit the initial prepotent response whereas others carry out the new task without difficulty (Carlson, Mandell, & Williams, 2004). The ability to delay gratification (e.g., delay eating a smaller snack to be rewarded with a larger one) at 24 months of age varies in duration from 0.9 to 12.5 s (Carlson et al., 2004). Further, focused attention during play with toys averages 22.8 s ( $SD = 22.6$  s) for 26-month-old children (Ruff & Capozzoli, 2003). Considerable variability in performance was observed on Hughes and Ensor's (2007) test of 2-year-olds' working memory using a multilocation search task called *spin-the-pots*, with performance varying from perfect to up to 10 of 16 possible incorrect responses. Finally, there also is large, consistent variability in the rate of early development of language (Bates et al., 1995). For example, expressive vocabulary ranges from 89 to 534 words among typically developing 24-month-old children (Bates et al., 1995). Others have proposed that this large variation in lexicon size reflects underlying variability in specificity of phonetic representations (Walley, 2004), which likely has important implications for speech discrimination assessments.

### ***Constraints on Methodology***

Limitations and variability in cognitive and linguistic skills make toddlers especially susceptible to nonsensory task factors. To minimize the influence of nonsensory task factors, researchers and clinicians attempt to work within certain methodological constraints. Specifically, the task must be interesting and short enough to sustain the child's motivation and attention. The response task must not exceed the child's

motor abilities (including gross, fine, and speech motor skills). The level of reasoning involved in understanding task requirements and the relationship between the response and the reward or reinforcement must be considered. The amount of information that must be stored and the duration of the delay between stimulus and response cannot exceed the child's short-term memory span. Effective verbal instructions must have limited linguistic complexity. Finally, stimuli with linguistic content may be included only at the expense of introducing confounds related to lexical development (Holt & Lalonde, 2012).

### ***Variability in Pediatric Speech Perception***

In addition to constraining methodology, cognitive and linguistic factors also contribute to the inevitable variability in speech perception performance. With the focus in pediatric speech perception research on group averages, variability is sometimes not reported (e.g., Dawson, Nott, Clark, & Cowan, 1998; Eilers, Wilson, & Moore, 1977; L. L. Elliott, 1986; Erber, 1971; Sussman & Carney, 1989). However, the literature is replete with examples of variability in performance, even among typically developing school-age children with normal hearing. For example, in one study, the performance of 19 normal-hearing 5- to 7-year-old children on a three-interval forced-choice speech feature discrimination paradigm ranged from approximately 30% to 87% relative to chance (Hnath-Chisolm, Laipply, & Boothroyd, 1998). Using a visually reinforced head turn procedure similar to the visually reinforced infant speech discrimination procedure (VRISD; Eilers et al., 1977), Eisenberg and colleagues (2007) reported that children 6 to 30 months of age could discriminate particular speech features on the basis of probability theory. Scores for consonant place, manner, and voicing discrimination varied from 0% to 90%. Using a different response task—a play (motor) act—in response to a change in the stimulus array, the authors reported that percent confidence varied between 70% and 95% in children 2;9 (years;months) to 3;6. Finally, in a study that used an imitative response, accuracy scores ranged from 58% to 96% in 2- and 3-year-olds (Boothroyd, Eisenberg, & Martinez, 2010). These results highlight the consistent findings of enormous individual differences in children's speech perception, even among typically developing children with normal hearing.

The change/no-change procedure (Sussman & Carney, 1989) is similarly sensitive to a wide range of children's discrimination performance. Sussman (1993) tested nine children between 4;9 and 6;6. Discrimination of the end points of a /ba/ to /ga/ continuum varied from chance performance to perfect accuracy. Holt and Carney (2007) tested thirty 4- and 5-year-olds using the same procedure. In any one of the more difficult listening conditions (e.g., -4 dB signal-to-noise ratio [SNR]), performance also varied from chance to ceiling levels of performance. Finally, we recently tested 2- and 3-year-old typically developing children on a toddler-based version of Sussman and Carney's (1989) change/no-change procedure. Discrimination sensitivity in quiet again ranged from chance levels to perfect discrimination (Holt & Lalonde, 2012).

This variability in performance is not limited to the perceptual level of discrimination. Speech recognition testing (particularly in noise) shows a wide range of scores among normal-hearing children. For instance, Sanderson-Leepa and Rintelman (1976) administered the Word Intelligibility by Picture Identification (WIPI; Ross & Lerman, 1970) test and the Phonetically Balanced Kindergarten Word Lists (PBK-50; Haskins, 1949) to twelve 3.5-year-old typically developing children. Mean accuracy was 88.3% and 71.7% with standard deviations of 10.85 and 15.75 on the WIPI and PBK-50, respectively. In addition, typically developing 3- to 5-year-old children identified key words in sentences with 55%–100% accuracy at 7 dB SNR in one study (Holt, Kirk, & Hay-McCutcheon, 2011). In another, typically developing children between 4.5 and 6.5 years identified words in sentences presented at 0 dB SNR with 43.8%–85% accuracy (Nittrouer & Boothroyd, 1990).

Together these data suggest that regardless of the procedure or task, individual differences in pediatric speech perception are the norm, even among typically developing populations. Identifying factors that contribute to individual variability in pediatric speech perception is important for understanding both typical and atypical speech perception. A reasonable place to begin examining nonsensory contributions to speech perception is at a level where complex neural encoding of the stimulus is required, but lexical knowledge is not. The change/no-change procedure or, in the case of young children, the toddler change/no-change procedure (Holt & Lalonde, 2012) allows assessment at this neural encoding level (Holt, 2011).

### ***Toddler Change/No-Change Procedure***

The change/no-change procedure (Sussman & Carney, 1989) is a forced-choice procedure that involves presenting standard and comparison auditory speech stimuli and requiring the listener to indicate, with a developmentally appropriate motor response, whether the stimuli are the same (no change) or different (change). The procedure has been used successfully with normal-hearing and hearing-impaired children and adults (Carney et al., 1991, 1993; Dawson et al., 1998; Holt, 2011; Holt & Carney, 2005, 2007; Osberger et al., 1991; Sussman, 1991, 1993) but with only limited success in children under 4 years of age (Dawson et al., 1998). The toddler version of this task (Holt & Lalonde, 2012) uses procedural modifications from the original version that address toddlers' developmental constraints. These targeted modifications include reducing test time and using a teaching session, a developmentally appropriate gross motor response, and multiple forms of reinforcement. In the modified version, toddlers stood on a star on a mat in a sound booth facing a computer monitor and two sets of pictures on white fabric placed on a mat on the floor. When a no-change stimulus array was presented (e.g., /ba ba ba ba/), the child was instructed to jump or step to a set of four identical pictures; when a change stimulus array was presented (e.g., /ba ba bu bu/), the child was instructed to jump or step to two sets of two different pictures. Each trial consisted of two presentations of the standard

stimulus followed by two presentations of the comparison stimulus.

Children were taught the response task using live-voice presentation of animal sounds (“moo moo moo moo” and “moo moo ribbit ribbit”) and semantically related pictures of animals (e.g., four cows or two cows followed by two frogs). Thirty of the thirty-four 2- and 3-year-old toddlers who were taught the procedure reached the response criterion of five consecutive correct responses (consistent with the criterion used in Trehub, Schneider, & Henderson, 1995). These 30 toddlers were then tested on 36 recorded trials of maximally contrastive training stimuli (long /u/ vs. short /ga/), followed by 36 recorded trials each of perceptually easy (acoustically distinct: /ba/ vs. /bu/) and perceptually hard (acoustically similar: /sa/ vs. /ja/) contrasts. Each child also was retested on either the easy or the hard contrast. During training, test, and retest phases, trials began with a picture of a woman cupping her hand around her ear while the auditory stimulus was presented in the sound field. After the child responded, a puzzle piece appeared on the monitor, and the child was verbally praised. After correct responses, the child was reinforced with a 3-s animated video. After every 12 trials, the child was given tangible reinforcement (M&M’s or Cheerios).

The results indicated that performance relied on perceptual processing and development, in that the procedure was sensitive to both the perceptual difficulty (acoustic distinctiveness) of the speech contrasts and the age of the listener. Results were also highly reliable from test to retest. Although these results were promising, 4 of the 12 children between the ages of 24 and 30 months did not learn the response task, one other child (27 months of age) who learned the response task failed to clearly demonstrate the ability to complete the procedure, and many younger 2-year-olds had low discrimination sensitivity scores. Further, as noted above, intersubject variability was high, with scores ranging from chance performance to perfect accuracy. Because the toddler change/no-change procedure (a) was shown to be valid and reliable for assessing speech discrimination in 3-year-olds and older 2-year olds, (b) has high intersubject variability, and (c) can be used to examine the early sensory–neural encoding of speech (Holt & Lalonde, 2012), it is an appropriate procedure to begin investigating sources of individual differences in speech discrimination. Furthermore, because the youngest 2-year-olds struggled to master the procedure, further modifications need to be examined to determine whether the procedure can be used with these youngest toddlers.

The purpose of this investigation was twofold: (a) to identify likely nonsensory sources of variance in toddler speech-sound discrimination by examining cognitive and linguistic development; and (b) to examine whether further modifications would lead to a procedure that can be used with 24- to 30-month-old toddlers. This investigation introduces three developmentally based modifications to the toddler change/no-change procedure—an orientation cue; immediate, selective reinforcement; and fewer trials per condition—to evaluate whether they are effective at allowing younger 2-year-olds to overcome task-related effects and thus extend the

youngest age with which the procedure can be used. Although modifications based on known developmental constraints increase the validity of a task, it is impossible to eliminate nonsensory factors from pediatric psychoacoustic methods (Allen & Wightman, 1992). Further, given individual differences among toddlers and their notoriously variable speech discrimination performance, even the most theory- and developmentally based procedure will not yield uniform performance in this age group. Therefore, it is important to understand how cognitive and linguistic factors might influence speech discrimination, particularly if the testing procedure is intended for use with clinical populations (Boothroyd, 1991; Tyler, 1993). This study takes the approach of acknowledging and investigating individual differences in cognitive and linguistic variables that are likely to contribute to individual differences in speech discrimination performance, including attention, working memory, reasoning, executive functioning, and receptive and expressive vocabulary.

## Materials and Method

### Participants

Of the thirty 2-year-olds recruited to participate in the study, 28 met the inclusion criteria of native English background and normal hearing, speech, and language development. One was excluded because of bilingual status; another failed the speech screening. Two additional participants cried and refused to perform the experimental tasks, and one parent never rescheduled follow-up appointments. The remaining 25 children completed all of the cognitive and linguistic tests. Three of these children were unable to learn the toddler change/no-change procedure (ages 25, 25, and 27 months), and two refused to complete all speech discrimination testing conditions (ages 27 and 28 months). The remaining 20 participants completed the entire protocol and ranged in age between 27 and 36 months ( $M = 31.76$ ,  $SD = 3.02$ ). All children passed a bilateral hearing screening using a four-frequency distortion product otoacoustic emission (DPOAE) test, conditioned play, or modified visual reinforcement audiometry at 0.25, 0.5, 1, 2, and 4 kHz. Conditioned play and modified visual reinforcement audiometry were used when children did not pass the DPOAE screening or would not allow the researcher to place the DPOAE probe in the ear canal. Behavioral screening was completed at 20 dB HL for 0.5 through 4 kHz and at 25 dB HL at 0.25 kHz (American National Standards Institute, 2004). All children also passed the Early Language Milestone Scale—2 (Coplan, 1993).

### Linguistic Measures

Standardized tests and questionnaires widely used to measure linguistic development in toddlers were used in the current study to assess receptive and expressive vocabulary.

*Receptive One-Word Picture Vocabulary Test.* The Receptive One-Word Picture Vocabulary Test—Fourth Edition (ROWPVT-4; Martin & Brownell, 2011) is a norm-referenced measure of receptive vocabulary. This test was administered in accordance with the instruction manual.

During test administration, an easel displaying a row of four pictures was placed in the child's view. The examiner verbally presented test items, and the participant touched the full-color picture corresponding to the item presented. Test items are presented in a developmental sequence, based on the age when examinees are likely to encounter the concepts. Testing ends when the participant misses 6 of 8 consecutive items. Raw scores and standard scores based on normative data were used. Based on test norms, the mean standard score for any age is 100.

*MacArthur-Bates Communicative Development Inventory.* The MacArthur-Bates Communicative Development Inventory (MCDI; Fenson et al., 1994) was used to measure expressive vocabulary. The Words and Sentences subtest was used for participants younger than 30 months of age. This subtest is based on normative data from 1,789 children ages 8 to 30 months (Dale & Fenson, 1996). A newly published version, the CDI-III subtest (Fenson et al., 2007), was used for participants age 30 months and older. The CDI-III is based on normative data from 356 children ages 30 to 37 months. One of each child's parents read a list of words (680 on the Words and Sentences subtest, 100 on the CDI-III) and marked a circle next to words that the child was known to say (regardless of articulation accuracy).

### **Cognitive Measures**

Standardized tests and questionnaires widely used to measure cognitive development in toddlers were used in the present study. The areas assessed included executive functioning, short-term and working memory, reasoning, nonverbal intelligence, and attention.

*Behavioral Rating Inventory of Executive Function—Preschool Version.* The Behavioral Rating Inventory of Executive Function—Preschool Version (BRIEF-P; Gioia, Epsy, & Isquith 2001) was used to assess executive functions. The parent read a list of behaviors and indicated the frequency with which his or her child's use of that behavior had caused a problem in the past 6 months (never, sometimes, or often). This standardized rating scale yields scores for inhibitory control, shifting, emotional control, working memory, and planning and organization indexes. Those scores are combined to generate the General Executive Composite reported using *t* scores in the current study, because all of the children tested fit into one normative group. Based on the test norms, the mean *t* score is 50. Higher scores on the BRIEF-P are associated with parent-reported executive function problems.

*Leiter International Performance Scales—Revised.* The Leiter International Performance Scales—Revised (Leiter-R; Roid & Miller, 1997a) were used to assess reasoning, nonverbal intelligence, attention, and memory because this completely nonverbal test is not confounded with the child's language development, uses no verbal instruction, and requires no verbal response. The Leiter-R consists of two batteries: Visualization and Reasoning; and Attention and Memory. For 2-year-olds, the Visualization and Reasoning battery consists of seven subtests that are combined to

generate composite scores of fluid reasoning, fundamental visualization, and nonverbal IQ. The fluid reasoning score is based on performance on a Sequential Order subtest and the Repeated Patterns subtest, both of which require the child to observe a pattern and generate rules. The fundamental visualization score is based on a Picture Context subtest and a Classification subtest, both of which require the child to perceive conceptual similarity and match pictures on the basis of classes of information. In addition to the four subtests already described, the full nonverbal IQ is based on three additional subtests: the Figure Ground, the Form Completion, and the Matching subtests. The Figure Ground subtest is a visual recognition task with distracters in the form of complex backgrounds. It requires visual scanning, inhibition, and freedom from distractibility. The Form Completion subtest is a visual organization task that requires the child to mentally organize fragmented pieces so as to perceive the fragments as a whole. It also requires perceptual scanning and visual recognition. The Matching subtest measures visual awareness, scanning, and spatial orientation as well as patience and freedom from impulsivity.

The Attention and Memory battery consists of three subtests for 2-year-olds: Associated Pairs, Forward Memory, and Attention Sustained. The Associated Pairs subtest is a paired-associates learning task with familiar and unfamiliar pairs that measures short-term retention. The test calls on echoic memory and possibly rehearsal skills. The Forward Memory subtest is a measure of sequential memory span that also requires sustained attention and inhibition of proactive interference from previous trials. The Attention Sustained subtest is a cancellation task that measures sustained visual attention. It requires vigilance, focused attention, motoric inhibition, and visual scanning as the child crosses out stimuli of one class without marking another class. Some of the children tested struggled with this task, scribbling on the paper rather than following the nonverbal instructions. According to the test guidelines, these children were assigned scores of "negative," a score that was also assigned if children marked more items from the wrong class than from the right class. Seven of the 20 children tested received a score of negative, suggesting that this test was particularly difficult for nearly half of the children and thus was not particularly sensitive to individual differences in this group's selective attention.

The order of testing and methods for teaching the task followed the instructions outlined in the Leiter-R examiner's manual. The validity and reliability of the Leiter-R is well established, and normative data are based on 1,719 typically developing children (Roid & Miller, 1997b). Scaled scores for each individual subtest and the composite measures are based on normative data. Raw and scaled scores were used for individual subtests and composite measures.

*Spin-the-pots.* The only working memory measure available for 2-year-olds is Hughes and Ensor's (2005) spin-the-pots task, a multilocation search task in which the child and experimenter place stickers in six of eight boxes on a lazy Susan. After covering the boxes and spinning the lazy Susan, the child is given 16 opportunities to find all of the stickers,

choosing one box at a time and returning it to its position on the tray after opening it. Children earned a score between 0 and 16, representing the difference between the number of opportunities (16) and the number of errors. Scores for the 122 2-year-old children tested by Hughes and Ensor (2005) varied from 6 to 16, suggesting that performance on this task is sufficiently variable to correlate spin-the-pots scores with speech discrimination sensitivity. Because the reliability of this task has not previously been established, this task was administered twice. A Pearson correlation revealed that scores at test and retest were significantly related, but the correlation was not strong ( $r = .59, p = .0019$ ).

### ***Strategic Modifications to the Toddler Change/No-Change Procedure***

In an effort to test younger 2-year-olds using the toddler change/no-change procedure (Holt & Lalonde, 2012), we implemented three additional developmentally based modifications in the current investigation: an orientation cue, immediate and selective reinforcement, and fewer trials per condition. In the previous version of this task (Holt & Lalonde, 2012), animated reinforcement was provided for correct responses. However, the substantial delay between the correct response and the reinforcement was filled with noncontingent verbal praise, a noncontingent puzzle piece appearing on the screen to keep the task interesting, and discussion about the puzzle. In the version used in the current investigation, the 3-s animated video was replaced with lighted, animated toys inside smoked, Plexiglas boxes (toys commonly used in visual reinforcement audiometry) placed on the floor directly in front of each response space. After stepping on the correct response space, children were immediately reinforced by the activation of the visual reinforcement toy just in front of the correct response space. Only clear, correct responses were rewarded. This reinforcement method was also used during the teaching portion to allow maximal transfer and to motivate younger children to learn the task.

The implementation of immediate, selective reinforcement was based on findings that infants are better able to infer the relationship between response and reward when the two are physically attached (Diamond, Churchland, Cruess, & Kirkham, 1999). Specifically, 9- and 12-month-old infants can perform a delayed match-to-sample task above chance when the reward is attached to the novel stimulus (but not visible) but not when the reward sits in a basin below the novel stimulus. It is reasonable to expect that toddlers will also benefit from the tighter coupling of the response and reward (in time and space). Tight coupling between the response and reward should alleviate some task-related difficulty regarding toddlers' trouble with associating the abstract rule with rewards (Diamond, 2006). Only clear, correct responses were rewarded because research with infants (Eilers et al., 1977; Primus & Thompson, 1985) and older children (R. Elliott, 1970) has shown that contingent rewards lead to better, faster responses.

Before each stimulus was presented, an orientation cue consisting of a 2-s video of a baby laughing silently was

used to get the child into a "ready" state. This video was developed in Hirsh-Pasek's lab (e.g., Roseberry, Hirsh-Pasek, Parish-Morris, & Golinkoff, 2009) and has been used in research with 6- to 30-month-old children (e.g., Houston, Ying, Pisoni, & Kirk, 2003; Roseberry et al., 2009). This change was meant to ameliorate constraints related to focused attention (Ruff & Capozzoli, 2003) and is based on adult detection literature showing that orientation cues decrease detection thresholds (Watson & Nichols, 1976). It is difficult to know what a child is attending to prior to the onset of a trial, and pediatric research has demonstrated that the degree to which a listener is involved in some other task can influence the degree to which he or she reacts to a stimulus (Tellinghuisen & Oakes, 1997). By directing each child's attention to the same cue prior to beginning the trial, we attempted to limit any deficits in discrimination that may have arisen because of distractions.

To further reduce the effects of limited focused attention, we further reduced the number of trials per condition relative to our previous study. The previous implementation of this procedure included 36 trials in each condition (Holt & Lalonde, 2012). Twenty trials were used in each test and retest condition in the current investigation. However, 36 trials were used in training because learning effects were observed in the previous study. Additional analysis of the data justified these modifications. Specifically, average discrimination scores using only the first 10 change trials and first 10 no-change trials were greater than or equal to scores using all 36 trials. Furthermore, analyzing only a subset of trials did not compromise test validity and reliability. The effects of age and difficulty of speech contrast and test-retest reliability remained significant. Finally, interpretation of individual results is largely unaffected by this change. Children who had poor discrimination sensitivity with 36 trials continued to discriminate poorly when 20 trials were analyzed; children who discriminated well continued to discriminate well. The increased number of training trials (relative to test-retest trials) was used to help ensure that participants had sufficient practice with the task.

### ***Discrimination Stimuli***

The stimuli used in the discrimination experiment consisted of two standard syllables followed by two comparison syllables that were either the same as the standard syllables (no-change stimuli) or different from the standard syllables (change stimuli). The syllables were separated by 100-ms silent intervals. The same three contrasts were used in the current experiment as were used in the previous one (i.e., Holt & Lalonde, 2012): a maximally contrastive training contrast of long /u/ versus short /ga/, an acoustically distinct (perceptually easy) vowel height contrast of /bu/ versus /ba/, and an acoustically similar (perceptually hard) place contrast of /sa/ versus /ja/. The training contrast consisted of 18 change and 18 no-change trials, randomly presented. The easy and hard contrasts consisted of 10 change and 10 no-change trials each, randomly presented. The syllables were digitally recorded from a female speaker and equalized in total root-mean-square

amplitude. Multiple tokens of each syllable were used to eliminate nonphonemic differences between syllables. The stimulus recording, editing, and selection procedures, as well as statistics related to duration, amplitude, accuracy of identification, and rating of goodness, have been described elsewhere (Holt & Lalonde, 2012). Adult listeners identified the individual syllables with 100% accuracy and rated the syllables as very good examples of the target production (Holt & Lalonde, 2012).

### Discrimination Test Setup

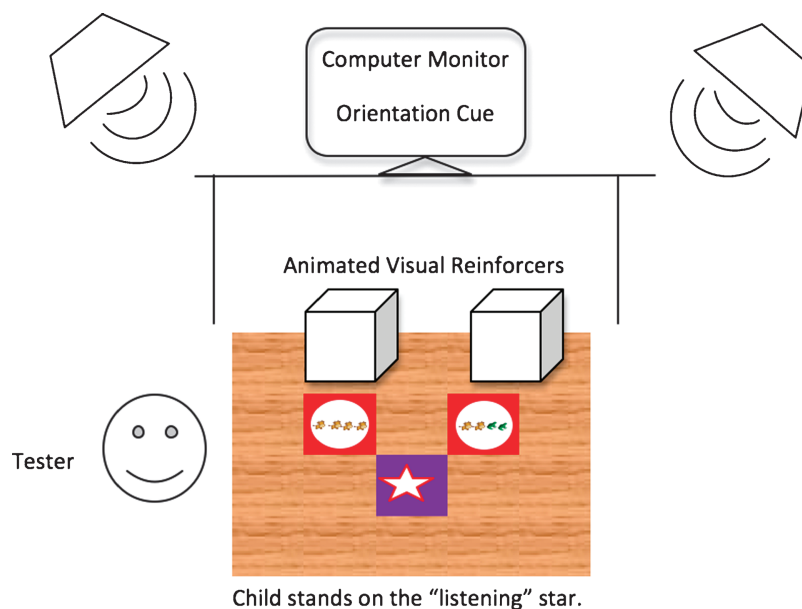
Figure 1 shows a diagram of the test setup for the current experiment. The discrimination stimuli were presented and data were recorded using E-Prime software (Version 2.0; Psychology Software Tools, 2007) and an Intel desktop computer. As in previous implementations of the change/no-change procedure (Holt & Carney, 2005, 2007; Holt & Lalonde, 2012), the stimuli were routed through an audiometer to two wall-mounted speakers in a double-walled sound booth, at  $\pm 45^\circ$  relative to the listener. The stimuli were presented at 65 dBA at the location of the listener's head. Sound-field presentation allows the participants to remain mobile while performing the gross motor response described below. In addition, the small variations in sound pressure level should not affect results, because 65 dBA is well above participants' detection thresholds. Calibration was checked at the start of each testing day.

A 4-ft  $\times$  5-ft SoftTile interlocking foam mat, set in the center of the floor of the testing booth, served as a response mat. The mat has a wood grain pattern, except for a purple piece with a star cutout placed in the center of the mat and two red pieces with circle cutouts placed directly in front of the star cutout—one to the left and one to the right. The child stood on the star facing the two circles. The circle on the left contained a picture of four identical animals in a row (four cows), representing a no-change response; the circle on the right contained a picture of two cows and two frogs in a row, representing a change response. The pictures were clip-art images from Microsoft Office ironed onto white fabric using Printworks white t-shirt transfers. Unlike in the previous study, the cows and frogs were used for teaching, training, and testing. A 19-in. monitor used to present an orientation cue was placed on a table approximately 1 m from the child at his or her eye level. A keyboard was placed to the left of the response mat and next to the examiner, who used the keyboard to record the child's response. Lighted, animated toys inside smoked, Plexiglas boxes were set on the floor, 1 ft in front of each response space. These toys were used to immediately reinforce correct responses.

### Discrimination Testing Protocol

The participants were run in a combined factorial and repeated measures design. Following live-voice teaching trials, participants completed the recorded training and easy

**Figure 1.** Schematic of the test setup, including the 4-ft  $\times$  5-ft response mat placed on the floor of the sound booth. The child stood on the star-shaped cutout facing the two red circles to the front-left and front-right of the star, which displayed pictures of cows and frogs corresponding to the no-change and change responses. Lighted, animated toys in smoked, Plexiglas boxes sat in front of each response space for immediate, selective reinforcement. The reinforcement controller and a keyboard for entering responses were placed to the left of the experimenter. A monitor placed on a table at the child's eye level displayed the orientation cue. Speakers were placed at  $\pm 45^\circ$  azimuth. Figure is not to scale.



and hard test contrasts. The 36 trials of the training contrast were completed first; order of easy and hard contrasts (20 trials each) was counterbalanced across participants. Each participant was also retested using 20 trials of either the easy or hard contrast.

During the teaching phase, the child was instructed to stand on the listening star, facing the response spaces, the reinforcement toys, and the monitor while the experimenter explained the task and introduced the reinforcement. Highly contrastive, live-voice, familiar animal sounds were used as teaching stimuli. The experimenter taught the child to jump to the no-change response space (picture of four cows) when the child heard “moo moo moo moo” and to the change space (picture of two cows and two frogs) when hearing “moo moo ribbit ribbit.” When the child jumped to or touched the correct response picture, the examiner activated the animated reinforcer directly in front of that response and explained that the toy “was dancing” because the child chose the right answer. The training phase continued until the child responded correctly to five consecutive trials (consistent with the criterion used in Trehub et al., 1995).

Following the teaching session, participants completed the training trials, recorded long /u/ versus /ga/. On each training trial, the 2-s orientation video of a laughing baby was followed by a 200-ms silent interval during which the baby’s still image remained on the screen, and then the stimulus was presented. If the child did not respond independently, the experimenter asked, “What did the baby say? Which picture?” If the child jumped to or touched the correct answer, the reinforcement was immediately presented and the participant praised. If the child jumped to or touched the wrong picture, the experimenter directed her or his attention to the correct response picture, demonstrating that the sounds presented matched that image. The experimenter recorded the child’s response when the participant was ready for the next trial (standing on the star quietly) and directed the child’s attention to the orientation cue and upcoming stimulus. Although these procedures remained exciting for many of the children, some needed breaks and other reinforcement to maintain attention to the task. Reinforcement methods were adapted to each individual participant, including a TossAcross game (beanbag throwing game) and tangible “listening tickets.” Test and retest conditions followed the same protocol as the teaching session.

### **Experimental Procedure**

This study was approved by the Indiana University institutional review board. All experimental tasks were completed in three or four 1-hr sessions. Following parental informed consent, all children underwent speech, language, and hearing screenings to determine eligibility to participate in the study. Once a child was deemed eligible to participate, the psychometrically rigorous parent questionnaires, standardized tests, and toddler change/no-change discrimination test were administered. Standardized testing always began with the Leiter–R, because this completely nonverbal test allowed shy toddlers to begin the experiment without needing to talk with

the experimenter. The order of the remaining tasks was counterbalanced, with the exception that the second administration of the spin-the-pots task always occurred during the last session. To avoid attrition, we tailored the order of testing to the temperament of the child whenever necessary. For instance, when children stopped responding to nonverbal instructions and encouragement during administration of the Leiter–R, the examiner administered the ROWPVT–4 (Martin & Brownell, 2011) and spin-the-pots task (Hughes & Ensor, 2005) or began testing speech discrimination, all of which allowed verbal encouragement. The Leiter–R was always administered over multiple sessions. To get the best possible data, we typically conducted speech perception testing over multiple sessions. The child led how much speech perception testing occurred on a given day, completing as much testing as possible in a given session before the experimenter judged that the child’s attention had waned. All cognitive and linguistic testing occurred in the quiet laboratory at a child-sized table, with the parent seated behind the child for comfort. Parents were instructed not to help their children with any of the tasks, especially not to provide verbal directions on the Leiter–R. For motivation to continue responding, children stamped a sheet of paper after each subtest was administered.

## **Results**

### **Speech Discrimination**

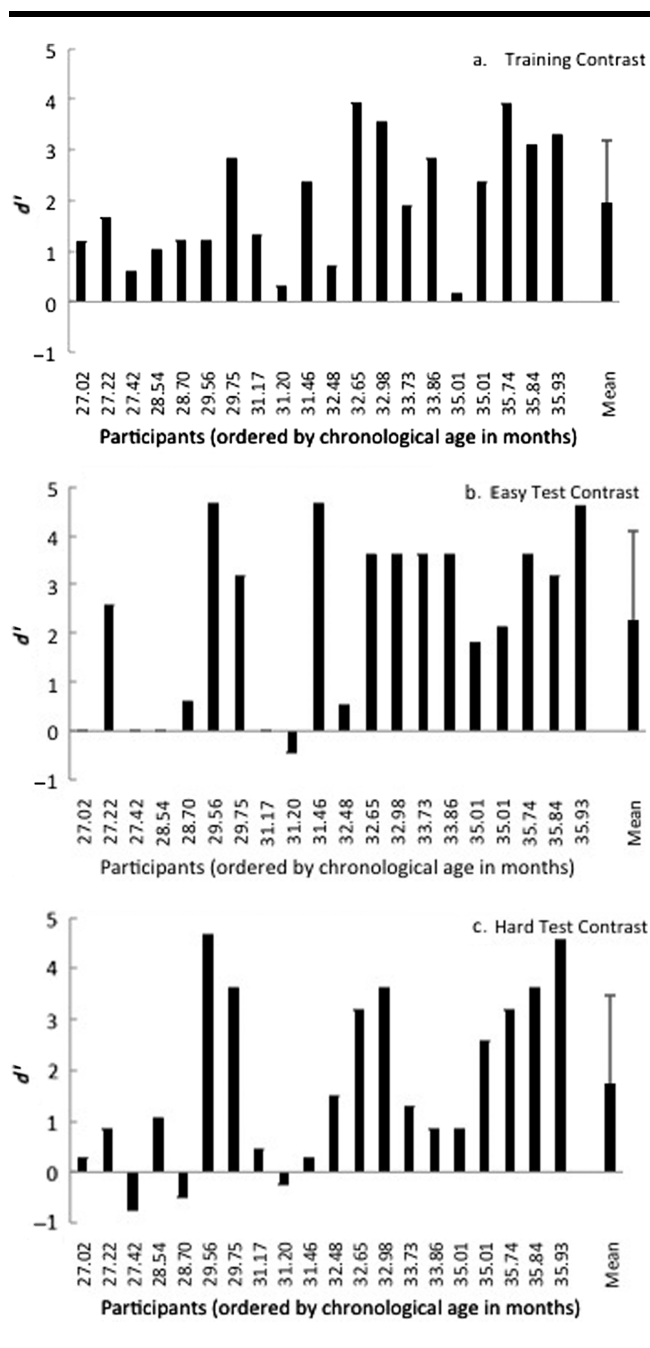
The dependent measure in this study was  $d'$ , a bias-free measure of sensitivity to the speech contrast (Macmillan & Creelman, 2005);  $d'$  is calculated by subtracting the  $z$  score for the false alarm rate from the  $z$  score for the hit rate. As in other studies (Holt, 2011; Holt & Carney, 2005, 2007; Holt & Lalonde, 2012), the hit and false alarm rates were limited to the range of 0.01 and 0.99, so that perfect performance corresponds to a  $d'$  of 4.65.

As previously noted, three of the children recruited were unable to learn the task (ages 25.4, 25.8, and 27.2 months), and two refused to complete all conditions (ages 27.8 and 28.7 months). Individual results and mean data on the toddler change/no-change procedure for the remaining participants are shown in Figure 2. Participants are ordered by chronological age on the  $x$ -axis, and sensitivity to the contrast ( $d'$ ) is shown on the  $y$ -axis for the training contrast (Figure 2a), easy contrast (Figure 2b), and hard contrast (Figure 2c). As a group, the children performed well above chance levels, and large individual variability was observed for each contrast: training (mean  $d' = 1.96$ ,  $SD = 1.21$ , range = 0.14 to 3.92), easy contrast (mean  $d' = 2.27$ ,  $SD = 1.80$ , range =  $-0.44$  to 4.65), and hard contrast (mean  $d' = 1.74$ ,  $SD = 1.73$ , range =  $-0.76$  to 4.65). All but one of the children (age 31.2 months) demonstrated task understanding, performing well above chance on at least one contrast ( $d' = 0.59$  or greater).

As in the previous investigation of the toddler change/no-change procedure (i.e., Holt & Lalonde, 2012), a Pearson correlation relating performance at test to performance on the same contrast at retest was conducted to assess test–retest



**Figure 2.** Individual performance on the (a) training, (b) easy, and (c) hard contrasts. The children are ordered by chronological age at first testing session. Group means (+1 standard deviation) are shown at the far right of each panel.



reliability. This correlation was significant ( $r = .671, p = .001$ ), supporting and extending previous reliability results to a slightly younger group of toddlers.

### Cognitive and Linguistic Measures

Group means, standard deviations, and ranges for each of the cognitive and linguistic measures are reported in Table 1.

Raw and standard scores are reported, where available. Standard scores are provided to allow comparisons with normative data. Raw scores are not reported for the MCDI, because different subscales were used for the younger and older 2-year-old children. Further, because many children received a score of “negative” on the attention sustained task, mean and standard deviation could not be calculated. Finally, there is no standardized measure for the spin-the-pots task. Here, the raw scores are compared with the mean raw score from another study (Hughes & Ensor, 2006).

One-sample  $t$  tests were conducted to assess whether the sample of children included in the study differed significantly from the mean normative data on each measure and thus the general population. As a group, the 2-year-olds were significantly above the average of the normative data on many of the measures, including receptive vocabulary,  $t(19) = 12.92, p < .001$ ; and several of the subtests from the nonverbal intelligence measure that contribute to the fundamental visualization composite,  $t(19) = 12.55, p < .001$ ; nonverbal IQ,  $t(19) = 13.96, p < .001$ ; and memory screener,  $t(19) = 3.89, p < .001$ . Although these relatively gifted children may not be representative of the general 2-year-old population, Table 1 shows that children varied considerably in their performance on each of the cognitive and linguistic measures. This is important given that a range of scores is necessary for investigating the possible relations between cognitive and linguistic development and discrimination sensitivity. As discussed below, it is very likely that other studies conducted in a university setting rely on similar, nonrepresentative samples.

### Relation Between Cognitive–Linguistic Measures and Speech Discrimination

To reduce the data set, the independent variables (easy and hard test contrast discrimination scores) were submitted to a principal component analysis. The two test contrasts loaded strongly (.922 each) on one component that accounted for 85% of variance. This component represents discrimination sensitivity and is perfectly correlated with the mean of the easy and hard contrast scores ( $r = 1, p < .0001$ ). Discrimination sensitivity was used as an outcome measure to investigate cognitive and linguistic factors in pediatric speech discrimination. Although the units are not meaningful, individual scores on this aggregate measure are displayed in Figure 3.

First-order Pearson correlations were performed to preliminarily test the relationship between this measure of discrimination sensitivity and each of the cognitive and linguistic variables. These are displayed in the rightmost column of Table 1. When possible, raw scores were used for each cognitive and linguistic variable to better investigate the strength of relationship between discrimination and the linguistic constructs tested. Raw scores were preferred over standard scores because, for example, if a 26-month-old child and a 34-month-old child knew the same number of words, the 26-month-old would have a higher scaled score than the 34-month-old despite the fact that their lexicons are likely of similar size.

As shown in Table 1, discrimination sensitivity significantly correlated with receptive vocabulary ( $r = .614, p = .004$ ),

**Table 1.** Results of cognitive–linguistic and neurocognitive measures.

Variable	Mean raw scores (SD)	Range of raw scores	Mean standard scores (SD)	Normative mean standard score	Correlation with discrimination ( <i>r</i> )
Linguistic					
Receptive vocabulary (ROWPVT–4)	50.45 (9.023)	35–69	119.6 (6.79)***	100	.614**
Expressive vocabulary (MCDI)		19.29–99	61.39 (24.87)	50	.292
Cognitive					
General executive composite (BRIEF–P)	91.21 (15.69)	65–129	50.26 (9.44)	50	–.118
Associated pairs (Leiter–R)	8.70 (3.95)	3–17	13.15 (3.13)***	10	.061
Forward memory (Leiter–R)	4.05 (2.28)	1–10	10.55 (2.11)	10	.532*
Attention sustained (Leiter–R)		“negative”–45	8.89 (4.04)	10	
Fundamental visualization (Leiter–R)	39.45 (6.46)	26–50	126.4 (9.41)***	100	.317
Fluid reasoning (Leiter–R)	10.45 (6.30)	3–26	100.1 (12.39)	100	.505*
Nonverbal intelligence (Leiter–R)	70.90 (8.25)	53–85	123.8 (7.63)***	100	.337
Spin-the-pots		5–16	12.7 (2.90)	12	.117

Note. ROWPVT–4 = Receptive One-Word Picture Vocabulary Test—Fourth Edition; MCDI = MacArthur-Bates Communicative Development Inventory; BRIEF–P = Behavioral Rating Inventory of Executive Function—Preschool Version; Leiter–R = Leiter International Performance Scales—Revised. \* $p < .025$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

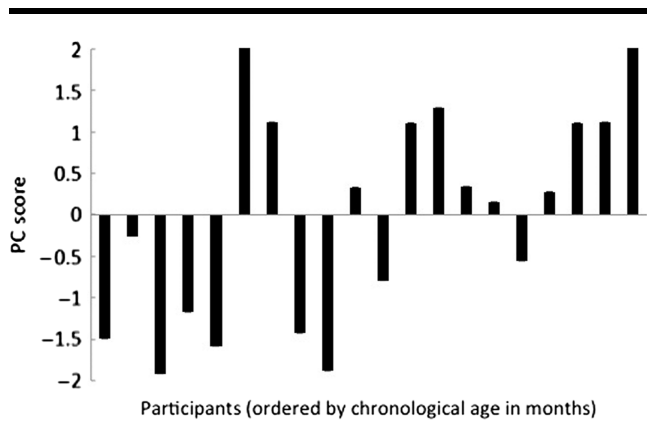
the nonverbal reasoning measure; the fluid reasoning score from the Leiter–R ( $r = .505, p = .023$ ); and the raw score for the short-term sequential memory measure, forward memory, from the Leiter–R ( $r = .532, p = .016$ ). In accordance with our previous data (Holt & Lalonde, 2012), the correlation between discrimination sensitivity and chronological age was also significant ( $r = .534, p = .015$ ). The same pattern of correlations is observed when performance on the individual contrasts (easy and hard) is used as the outcome variable. The correlations between individual contrast scores and cognitive–linguistic measures tend to be slightly weaker than those between the discrimination sensitivity measure and cognitive–linguistic measures. This suggests that the measure of overall discrimination sensitivity (across the contrasts of different perceptual difficulties) is more robust and supports the use of this variable as the outcome measure in further analyses.

The cognitive and linguistic variables in Table 1 (excluding the sustained attention measure, for which seven

children received scores of “negative”) and age were submitted to another principal component analysis. Four uncorrelated components emerged, accounting for 73.89% of variance in discrimination sensitivity. The loadings onto each component are shown in Appendix A. Most of the variables (receptive vocabulary, nonverbal intelligence, age, fundamental visualization, fluid reasoning, associated pairs, and forward memory) loaded positively onto Component 1, so the component could not be used to determine which of the specific cognitive or linguistic variables accounts for discrimination sensitivity. Executive function, short-term memory, and associated pairs all loaded onto Component 2. However, executive function loaded in the opposite direction from short-term memory and paired-associates learning. Thus, a high score on this component represents poor executive function but good short-term memory and paired-associates learning. This is not meaningful, because short-term memory is a subcomponent of executive function. Expressive vocabulary and paired-associates learning loaded strongly but in opposite directions onto Component 3. Finally, fluid reasoning and fundamental visualization loaded strongly but in opposite directions onto Component 4. In summary, the analysis did not provide meaningful components for exploring the relationship between discrimination sensitivity and cognitive–linguistic development. Therefore, the original cognitive and linguistic variables from Table 1 were used as the predictor variables in a stepwise regression analysis to further examine the relationship between the cognitive and linguistic variables and discrimination sensitivity.

Because there were significant correlations among some of the measures of cognitive and linguistic development (see Appendix B), we evaluated the predictor data set for potential collinearity issues by using variance inflation factors (VIFs). A common conservative criterion is that no variable in the predictor set should have  $VIF \geq 5$  (Belsley, Kuh, & Welsch, 1980). When VIFs were calculated for the cognitive and linguistic variables in Table 1 and chronological age, some were greater than 5 (VIF for age = 6.27, VIF for

**Figure 3.** Individual principal component (PC) scores for the measure of discrimination sensitivity. The children are ordered by chronological age at first testing session.



nonverbal IQ = 5.69), indicating that one or more predictor variables needed to be removed. The correlation matrix in Appendix B displays relatively large correlations between nonverbal intelligence and fluid reasoning ( $r = .851$ ) and age ( $r = .834$ ). Removing either age or nonverbal IQ from the predictor data set resulted in acceptable VIFs (VIF  $\leq 4.72$  without age, VIF  $\leq 4.22$  without nonverbal IQ). Therefore, age was not included as a predictor variable in the stepwise regression, because excluding age resulted in better VIF values than removing nonverbal intelligence and because nonverbal intelligence is more theoretically meaningful than age.

The nine cognitive and linguistic variables in Table 1 (excluding attention sustained) were entered as predictor variables in a stepwise linear regression with discrimination sensitivity as the outcome variable. Adjusted  $R^2$  values were used to estimate the proportion of independent variance in discrimination explained by each variable that emerged from the stepwise regression. Table 2 shows that 48.9% of variance in discrimination sensitivity could be accounted for by receptive vocabulary (34.2%),  $F(1, 18) = 10.871, p = .004$ ; and executive function (14.7%),  $F(1, 17) = 6.155, p = .001$ . No other variables were significant.

*Executive function* is an umbrella term for the cognitive processes used to regulate one's own actions and behaviors, such as inhibition, planning, and working memory. To further examine the relationship between executive function and discrimination sensitivity, a second stepwise regression was conducted. The predictor variables included in the analysis were receptive vocabulary and the five subscales of the BRIEF-P measure of executive function: Inhibition, Task Shifting, Emotional Control, Working Memory, and Planning/Organization. We chose to use this subset of variables, rather than adding the BRIEF-P subscales to the set of predictors from the previous regression, because using the full set led to unacceptable variance inflation factors (VIF  $\leq 12.32$ ). The VIFs for the subset were acceptable (VIF  $\leq 4.41$ ). As shown in Table 3, 56.6% of variance in discrimination sensitivity could be accounted for by receptive vocabulary (34.2%),  $F(1, 18) = 10.871, p = .004$ ; and working memory (22.4%),  $F(1, 17) = 10.286, p = .005$ . No other variables were significant.

### Comparison With Previous Data

A secondary goal of the current study was to examine whether evidence-based modifications to the toddler change/no-change procedure resulted in the procedure being appropriate for evaluating the discrimination skills of the youngest

2-year-olds. Data from the current study were compared with those from the first 20 trials per condition for 2-year-olds tested using the previous implementation of the toddler change/no-change procedure (Holt & Lalonde, 2012). The two groups of children were similar in age ( $p = .74$ ): current experiment ( $M = 31.76$  months,  $SD = 3.02$ ) and previous experiment ( $M = 31.59$  months,  $SD = 2.95$ ). Box plots comparing the two groups on each speech-sound contrast are displayed in Figure 4. To test for differences in performance between the group that participated in the previous experiment and the group that participated in the current experiment, we performed a mixed analysis of variance (ANOVA; variables: experiment [old or new] and contrast [training, easy, hard]). The effect of contrast was significant,  $F(2, 37) = 5.84, p = .004$ . Post hoc comparisons with Bonferroni corrections for multiple comparisons indicated that, as in the previous study (Holt & Lalonde, 2012), the effect of contrast was due to significantly better performance on the perceptually easy than the perceptually hard contrast. There were no significant differences between performance on the training contrast and the easy contrast ( $p = .141$ ) or hard contrast ( $p = .487$ ). There was no effect of experiment,  $F(1, 37) = 0.12, p = .729$ ; or interaction between the two variables,  $F(2, 37) = 0.236, p = .748$ ; suggesting that as a group, children in the current experiment with further procedural modifications did not perform differently than those in the previous experiment. However, we were specifically interested in the youngest 2-year-olds who struggled with the task in the previous investigation. Therefore, further analyses were carried out.

Seven of the children tested in each investigation were between 24 and 30 month of age at the time of testing. The ages of the two subgroups were well matched: current investigation ( $M = 28.30$  months,  $SD = 1.01$ ) and previous investigation ( $M = 28.32$  months,  $SD = 1.12$ ). On average, there was a trend for younger 2-year-old children in the current experiment to perform better than the younger 2-year-old children in the previous experiment on the training contrast ( $d' = 2.25$  in the current experiment,  $d' = 0.93$  in the previous experiment), easy contrast ( $d' = 2.22$  in the current experiment,  $d' = 1.60$  in the previous experiment), and hard contrast ( $d' = 1.92$  in the current experiment,  $d' = 0.54$  in the previous experiment). However, a mixed ANOVA with only the 24- to 30-month-old children revealed that these group differences were not significant,  $F(1, 12) = 0.354, p = .563$ . Further, the ages of the children who could not learn the task in the previous and current experiment were similar: two were 25 months old and one was 27 months old in the current experiment; two were 24 months old, one was 26 months old,

**Table 2.** Summary of significant results of stepwise linear regression for the discrimination sensitivity measure.

Dependent variable	Predictor variable	Standardized beta	% variance	F	p
Discrimination sensitivity	Receptive vocabulary	0.797	34.2	$F(1, 18) = 10.892$	.004
	Executive function <sup>a</sup>	-0.446	14.7	$F(1, 17) = 6.155$	.024

Note. Probability of  $F$  to enter  $\leq .05$ ; probability of  $F$  to remove  $\leq .1$ .

<sup>a</sup>Lower executive function scores mean fewer problems with executive function.

**Table 3.** Summary of significant results of stepwise linear regression for the discrimination sensitivity measure with Receptive Vocabulary and the Executive Function subscales as predictor variables.

Dependent variable	Predictor variable	Standardized beta	% variance	F	p
Discrimination sensitivity	Receptive Vocabulary	0.748	34.2	$F(1, 18) = 10.871$	.004
	Working Memory <sup>a</sup>	-0.503	22.4	$F(1, 17) = 10.286$	.005

Note. Probability of  $F$  to enter  $\leq .05$ ; probability of  $F$  to remove  $\leq .1$ .

<sup>a</sup>Lower executive function scores mean fewer problems with executive function.

and one was 30 months old in the previous experiment. The children who refused to finish testing in the current experiment were 27 and 28 months of age; those in the previous experiment were 28 and 29 months of age. Finally, Pearson correlations relating performance at test to performance at retest were approximately the same for 2-year-old children tested in the previous experiment ( $r = .691, p = .001$ ) and those tested in the current experiment ( $r = .776, p = .04$ ).

## Discussion

The primary purpose of the current investigation was to identify sources of variance in 2-year-old children's performance on a speech-sound discrimination task, the toddler change/no-change procedure, by investigating individual differences in cognitive and linguistic development across toddlers. The secondary purpose was to examine whether

modifications to the toddler change/no-change procedure would allow testing of younger 2-year-olds.

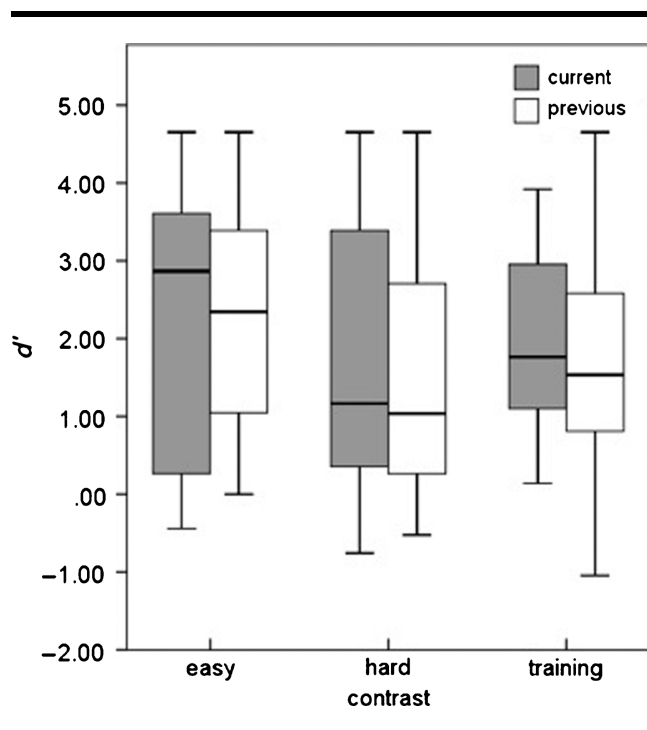
### Sources of Variance in Toddlers' Speech-Sound Discrimination

Two cognitive and linguistic variables—receptive vocabulary and working memory—emerged from regression analyses, explaining 56.6% of the variance in discrimination sensitivity. Receptive vocabulary accounted for 34.2% of variance in speech discrimination sensitivity. As discussed below, there are several possible interpretations of the relation between discrimination and receptive vocabulary: (a) Language development might strengthen phonetic representations and improve discrimination; (b) children with better discrimination abilities might have an advantage for word learning; or (c) children with larger vocabularies might have understood the oral instructions about the discrimination procedure better than those with smaller vocabularies.

The first interpretation, that language development might strengthen phonetic representations and improve discrimination, is supported by the lexical restructuring model (Metsala & Walley, 1998), which predicts that children with larger vocabularies will demonstrate better speech discrimination sensitivity, because vocabulary development prompts restructuring of representations from initially holistic to more detailed representations, as more detail is required to distinguish between items in our lexicon. Although it is clear that language and speech processing develop in parallel, much of the evidence for the link between the two has been indirect (Walley, 1993). In addition, children's receptive vocabulary is related to measures of their phonological sensitivity and/or phonological short-term memory (Baddeley, Gathercole, & Papagno, 1998; Metsala, 1999; Nazzi & Bertoncini, 2003; Storkel & Morissette, 2002), including 30- to 36-month old toddlers (Schwartz, Burnham, & Bowey, 2006; Smith, McGregor, & Demille, 2006). The current investigation extends this work by demonstrating a relationship between receptive vocabulary and more basic speech perception skills. Specifically, speech discrimination, which requires complex neural encoding and analysis of the auditory stimulus without necessarily invoking phonological awareness skills or the lexicon, was shown to relate to receptive vocabulary.

The second interpretation of this relation is that children who have better entry-level speech discrimination (those who potentially have better phonetic representations) might have an advantage for word learning. This is supported by

**Figure 4.** Mean performance ( $\pm 1$  standard deviation) on training and perceptually easy and hard test contrasts by children in the current experiment (filled bars) and 2-year-olds from Holt and Lalonde (2012; unfilled bars).



Kuhl and colleagues' research demonstrating that infants' native speech-sound discrimination at 6 months of age is correlated with parent reports of word and phrase understanding and word production at 13, 16, and 24 months of age (Tsao, Liu, & Kuhl, 2004). Differences in general auditory capacity and cognitive abilities could not explain the relationship between speech discrimination and later language skills, because discrimination of nonnative contrasts—a task with the same auditory and cognitive load—was negatively correlated with later language development (Kuhl, Conboy, Padden, Nelson, & Pruitt, 2005). The current investigation extends these earlier findings and suggests that entry-level speech discrimination, even at 2 years of age, continues to be related to children's language development. These findings are consistent with the literature on clinical populations, such as children with language and reading impairments, showing that they do not discriminate speech sounds as well as normal-hearing, typically developing controls (Bradlow et al., 1999; Godfrey, Syrdal-Lasky, Millay, & Knox, 1981; Kraus et al., 1996; Leonard, McGregor, & Allen, 1992; Manis et al., 1997; Reed, 1989; Stark & Heinz, 1996a, 1996b; Sussman, 1993, 2001; Tallal & Piercy, 1974, 1975; Werker & Tees, 1987).

These interpretations are not mutually exclusive, and it is likely that both are valid. Children with better discrimination abilities may have had an advantage for word learning and thus developed larger vocabularies. In turn, these larger vocabularies may lead to stronger phonetic representations and improved discrimination. Current models of developmental speech perception typically include bidirectional interactions (e.g., Kuhl et al., 2008; Werker & Curtin, 2005). Werker and Curtin's processing rich information from multidimensional interactive representations model includes three multidimensional spaces—a general perceptual space, a word form space, and a phoneme space—that mutually influence one another. Phase 3 of the expanded native language magnet theory includes bidirectional effects, wherein phonetic learning improves detection of word patterns and learning phonetically similar words improves awareness of phonetic distinctions (Kuhl et al., 2008).

Regardless of the direction of interpretation, these results provide further evidence that the toddler change/no-change procedure is sensitive to the strength of the child's phonetic representations, the development of which either results from vocabulary development, aids in vocabulary development, or both. This is consistent with studies that have used the change/no-change procedure with older children and adults. Holt and Carney (2005, 2007) and Holt (2011) demonstrated that adults' and older children's discrimination sensitivity, measured using the change/no-change procedure, improves when the number of stimulus repetitions increases. This was interpreted as a demonstration that repetition strengthens the early perceptual representation of the speech stimulus at the initial stage of speech-sound processing (Holt, 2011; Holt & Carney, 2005, 2007). The current study extends this finding to a younger age range and preliminarily suggests that the procedure is also sensitive to individual differences in the strength of phonetic representations. This procedure may serve as a means to further investigate the

development of phonetic representations. Unfortunately, this may also mean that the procedure is measuring more than whether children have the auditory capacity to perceptually discriminate speech contrasts that are important for the development of phonetic representations and thus speech and language (Eisenberg et al., 2007; Holt & Carney, 2007; Sussman & Carney, 1989; Tyler, 1993). If the development of speech discrimination, phonetic representations, and vocabulary are as intricately linked as the current findings, the literature, and models of developmental speech perception suggest, it may not be possible to separate the auditory capacity for speech discrimination from the phonetic and lexical knowledge that support the process.

The third interpretation of these results is that better vocabulary could facilitate better task understanding or oral instruction, leading to better discrimination performance. However, this seems unlikely because some of the children who were able to learn the task had poorer vocabularies than those who could not learn to perform the discrimination task.

Executive function also accounted for a sizable portion of the variance in speech discrimination sensitivity, after controlling for receptive vocabulary. Further analysis revealed that this variance could be attributed to a specific component of executive function—working memory—that accounted for an additional 22.4% of variance in speech discrimination sensitivity. The results suggest that speech discrimination, at least as tested by the procedure, relies in part on working memory.

The creators of the BRIEF-P describe the Working Memory subscale as measuring both working memory and sustained attention, processes they consider to be closely linked and behaviorally indistinguishable (Gioia et al., 2001). More specifically, the subscale measures the ability to “hold information in mind for the purpose of completing a task or making a response” (Gioia et al., 2001, p. 18), which is important for following directions and carrying out multistep activities. Children with high scores on the Working Memory subscale are described as having short attention spans and frequently forgetting things, including rules or direction, even for very short durations, such as a few seconds, and even while currently involved with a task (Gioia et al., 2001).

It is easy to see that these results might reflect demands specific to the toddler change/no-change procedure. The speech information must be held in memory during the interval between stimulus presentation and response, while it is associated with the appropriate sequence of pictures on the response mat, and a motor response is executed. Although the toddlers were encouraged to respond immediately after the stimulus was presented, they sometimes had to be prompted to respond, creating a substantial delay between the stimulus and response. Further, the child had to simultaneously remember the rules of the task over the duration of testing. After an incorrect response, the experimenter always directed the child's attention to the response picture and demonstrated that the sounds presented matched the image. Thus, there were frequent, systematic reminders of the rules of the task, which should have lessened the effects of working memory requirements.

Yoshida and colleagues highlighted the contributions of task demands (particularly memory demands) to 14-month-olds'

performance on a word-learning task by administering two slightly different testing protocols (Yoshida, Fennell, Swingley, & Werker, 2009). As in previous studies (Fennell & Werker, 2003; Stager & Werker, 1997; Werker, Fennell, Corcoran, & Stager, 2002), 14-month-old children were presented with two novel object-label mappings. The infants demonstrated the ability to discriminate the minimally contrastive object labels (“bin” and “din”) in a preferential looking paradigm, but not in the switch task commonly used by Werker and colleagues (e.g., Fennell & Werker, 2003; Stager & Werker, 1997; Werker et al., 2002). This was interpreted in terms of the amount of resource demands. Specifically, the switch task measures differences in looking time between test trials where the object is paired with the mapped label and test trials where the mismatched label is presented; this requires the infant to remember the object-label combinations and compare them with the combination presented during the trial. The preferential looking procedure likely involves a reduced memory load, as both objects are presented simultaneously (Fennell & Werker, 2003). This allows the infant to devote limited attentional resources to attending to the fine phonetic detail necessary to discriminate between the minimally contrastive labels and demonstrate the ability to use that detail to process recently learned words (Fennell & Werker, 2003; Stager & Werker, 1997; Werker et al., 2002; Yoshida et al., 2009). Just as children in Werker and colleagues’ set of word-learning experiments were required to link the label with the object, toddlers’ ability to link the speech sounds with the appropriate response space likely plays a role in determining whether a given toddler can meet the training criterion and how well he or she will discriminate the stimuli. The change/no-change procedure is more cognitively demanding than the head-turn procedure used with younger infants, in that it requires a similar linking of the speech sound with the appropriate response. However, performance on less demanding tasks, such as head-turn procedures, declines beyond the age of 12 months (Eilers et al., 1977; Martinez, Eisenberg, Boothroyd, & Visser-Dumont, 2008), making these procedures inappropriate for the 2-year-old age range.

The observed relation between discrimination sensitivity and working memory may reflect the role of working memory in speech perception, generally. In fact, Baddeley and colleagues have theorized that verbal working memory, in the form of the phonological loop, mediates the development of phonological representations, which in turn aid in acquiring new words (Baddeley et al., 1998). In fact, the relation between the capacity of the phonological loop and receptive vocabulary development is well documented (e.g., Baddeley et al., 1998). Some researchers have suggested that the phonological loop contributes to word learning, especially for younger children (age 5–6 years) (Baddeley et al., 1998; Gathercole & Baddeley, 1989; Gathercole, Hitch, Service, & Martin, 1997). Alternatively and consistent with the lexical restructuring model, others have suggested that the detailed phonological representations associated with larger vocabularies lead to better representation of the items in the phonological store, resulting in greater storage capacity

(Fowler, 1991; Metsala, 1999; Metsala & Walley, 1998; Walley, 1993). The phonological loop seems less important for older children who have acquired enough language to use their vocabulary knowledge to aid in learning new words (Baddeley et al., 1998; Gathercole & Baddeley, 1989). Consistent with this line of reasoning, there is evidence that the phonological loop is especially relevant when the listener cannot rely on language knowledge (such as vocabulary), when, for example, the stimulus is a nonword (Gathercole & Baddeley, 1989, 1990). This is clearly the case in the current investigation, in which the stimuli are meaningless strings of syllables. Although speech discrimination requires only complex neural encoding and analysis of the auditory stimulus without necessarily invoking phonological awareness skills or the lexicon, the relationship between working memory and discrimination sensitivity might suggest that the 2-year-old children tested were encoding the stimulus as a sequence of syllables. They seemed to be attempting to use their developing phonological knowledge to perform the task.

It is curious that the behavioral working memory measure, the spin-the-pots task (Hughes & Ensor, 2005), did not also emerge from the regression. Performance on the spin-the-pots task also was not significantly correlated with the Working Memory subscale of the BRIEF-P ( $r = -.335$ ,  $p = .149$ ). These results may question the validity of the spin-the-pots task as a measure of working memory. It may be the case that the two measures are assessing different aspects of working memory, that the spin-the-pots task is a pure visuospatial working memory measure whereas the BRIEF-P is also assessing verbal working memory. Finally, the BRIEF-P is based on parental report, whereas spin-the-pots is a performance measure.

### *Evaluating Speech-Sound Discrimination in Toddlers 24 to 30 Months of Age*

Three developmentally based modifications to the toddler change/no-change procedure were incorporated into the current experiment in an attempt to assess younger 2-year-olds: an orientation cue, immediate and selective reinforcement, and fewer trials per condition. Children in the current experiment did not perform significantly better than those in the previous experiment, and results were not more reliable from test to retest, suggesting that these modifications were not successful in making the task more appropriate for young toddlers. Of the 20 participants who completed the speech discrimination procedure, only two did not demonstrate task understanding during the experiment (27.42 and 31.20 months). However, three others were unable to learn the task and two chose to end testing.

Other testing procedures have proven successful with young children. However, some of these tests have upper limits to the age at which the test can be used, because older infants (age 22–26 months) habituate to a visual reinforcer faster than do younger infants (11–13 months), despite similar conditioning rates (Primus & Thompson, 1985). The VRISD procedure (Eilers et al., 1977) and the Visual Reinforcement Assessment of the Perception of Speech Pattern Contrasts

(Eisenberg, Martinez, & Boothroyd, 2004), a test based on the VRISD procedure, are two examples wherein performance is best for children under 12 months of age and declines for older infants (Martinez et al., 2008). Other tests have been successful with children 3 years of age and older, including the Speech Feature Test (SFT; Dawson et al., 1998); the Play Assessment of Speech Pattern Contrasts (Eisenberg et al., 2007), which is based on the SFT; the Imitative Test of Speech Pattern Contrasts (Boothroyd, 1985) as well as its online version (Eisenberg, Martinez, & Boothroyd, 2003); and the change/no-change procedure (Holt & Carney, 2007; Sussman & Carney, 1989). However, these tests and procedures are limited in their applicability to the 2-year-old range (Holt & Lalonde, 2012). Although progress has been made, no test or procedure has been designed that completely and consistently covers the entire 2-year-old age range. As in our previous investigation (Holt & Lalonde, 2012), some of the younger 2-year-old children (generally those less than 30 months of age) could not learn the task, did not complete the entire procedure, or performed poorly in the current investigation. This suggests that the procedure's application is limited with children below 2.5 years of age but does fill a gap by providing an appropriate test for the 2.5- to 3-year-old range. However, given that the sample of children used in the current study was not representative of the typical population, in that no children with low to average vocabularies and intelligence were recruited, it remains to be demonstrated that the procedure can be used with the average (or below average) 2.5- to 3-year-old who is more commonly seen in a clinical setting. It is also unclear whether the same variables would be most strongly related to speech discrimination if a more varied and representative sample were recruited.

It should be noted that although this sample was gifted in certain cognitive and linguistic domains, it is very likely that other studies conducted in a university setting rely on similar nonrepresentative samples. These participants are typical of the children recruited for research purposes at this university campus (e.g., recruited from our large child database of families who are willing to participate in research). It is likely that the families that voluntarily enroll their children in our research are less reflective of the general population than researchers desire. Consequently, the procedure's lower age limits for use with the general population might be higher than our research indicates, a fact we would not have discovered if we had not measured cognitive and linguistic function. These results support the notion proposed by others (e.g., Champion, Hyter, McCabe, & Bland-Stewart, 2003; Kresheck & Nicolosi, 1973; Washington, 1996; Washington & Craig, 1992) that pediatric scientists in university settings must be mindful of potentially exaggerated differences between the typical and disordered sample resulting from self-selected sampling.

### ***Limitations and Future Directions***

We identified two cognitive and linguistic sources of variance in toddler speech discrimination sensitivity that accounted for just over half of the variability in performance.

Obviously, there was remaining variance that was not captured by the cognitive and linguistic measures included in the study. It may be the case that vocabulary and working memory become less important when other variables are included. However, these preliminary results of our theoretically driven investigation suggest that the outcomes of speech discrimination tasks represent the complex interaction of many speech perception processes, including working memory and phonological processing. The findings also serve as a caution for clinicians and researchers about the complexity of what is asked of a child in a speech perception task (and perhaps in speech perception in daily life) and about the need to consider the development of specific cognitive and linguistic constructs when assessing speech perception in young children. This issue may be particularly relevant for clinical populations with poorer receptive vocabularies and/or executive function, such as those with hearing impairment, language impairment, or attention deficit disorders.

It is clear that task requirements have an effect on the degree to which 2-year-old children succeed on the toddler change/no-change procedure. However, it is unclear whether the cognitive and linguistic variables investigated in this study are related to the processes necessary to discriminate speech or the processes necessary to perform the task. This could be investigated by performing the same experiment using other discrimination tasks and/or nonspeech stimuli. In addition, this study was designed to investigate only between-subject variance. Certainly there are nonsensory factors affecting within-subject variation in performance, which should be considered when making choices about procedures and interpreting results in both clinical and research settings. It will be important for future studies to assess whether the same variables affect within-subject, trial-to-trial variability in performance.

Because of the limited scope of the current study, it was not possible to test every cognitive and linguistic construct that might influence performance. Each of the toddlers in the experiment was tested during approximately four 1-hr sessions. The need to limit test time precluded the inclusion of more measures. Given the known relation between receptive vocabulary and phonemic awareness and the potential role of general phonological knowledge in performing the change/no-change procedure or reducing task-related demands (Yoshida et al., 2009), future research should include measures of phonemic awareness as a potential mediating process. In addition, different types of measures were used to assess each of the constructs (behavioral and parent report), and at least one of our measures proved inappropriate for the experiment (attention), despite the use of standardized measures with normative data. Given that the two variables that emerged from the regression analysis were tested behaviorally (receptive vocabulary) and by parent report (executive function), this may not have been an issue for the current study. Finally, it is impossible to determine causation by using correlations and regression analyses. However, as a first attempt to account for individual differences in 2-year-olds' speech-sound discrimination, these results represent a promising area for further investigations.

## Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grant T32 DC00012, Indiana University's Faculty Research Support Program, and the Ronald E. McNair Research Foundation. Portions of this work were presented at the annual meeting of the American Auditory Society in Scottsdale, AZ (March 2012), and the Newborn Hearing Screening 2012 Conference in Cernobbio, Italy (June 2012). We are grateful for the comments of Tonya Bergeson on previous versions of this article.

## References

- Allen, P., & Wightman, F. L. (1992). Spectral pattern discrimination by children. *Journal of Speech and Hearing Research, 35*, 222–235.
- American National Standards Institute. (2004). *Specification for audiometers* (ANSI S3.6-2004). New York, NY: Author.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.
- Barton, D. (1980). Phonemic perception in children. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology: Volume 2. Perception*. New York, NY: Academic Press.
- Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language* (pp. 96–151). Oxford, England: Basil Blackwell.
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York, NY: Wiley.
- Benes, F. M. (2001). The development of prefrontal cortex: The maturation of neurotransmitter systems and their interactions. In C. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 79–92). Cambridge, MA: MIT Press.
- Berger, S. E. (2004). Demands on finite cognitive capacity cause infants' perseverative errors. *Infancy, 5*, 217–238.
- Boothroyd, A. (1971). Developmental factors in speech perception. *International Journal of Audiology, 9*, 30–38.
- Boothroyd, A. (1985). Evaluation of speech production of the hearing impaired: Some benefits of forced-choice testing. *Journal of Speech and Hearing Research, 28*, 185–196.
- Boothroyd, A. (1991). Speech perception measures and their role in the evaluation on hearing aid performance in a pediatric population. In J. A. Feigin & J. P. Stelmachowicz (Eds.), *Pediatric amplification* (pp. 77–91). Omaha, NE: Boys Town National Research Hospital.
- Boothroyd, A. (2004). Measuring auditory speech perception capacity in very young children. In R. Miyamoto (Ed.), *Proceedings of the 8th International Cochlear Implant Conference* (International Congress Series 1273, pp. 292–295). Amsterdam, the Netherlands: Elsevier.
- Boothroyd, A., Eisenberg, L. S., & Martinez, A. S. (2010). An on-line imitative test of speech-pattern perception (OlimSpace): Developmental effects in normally hearing children. *Journal of Speech, Language, and Hearing Research, 53*, 531–542.
- Bradlow, A. R., Kraus, N., Nicol, T. G., McGee, T. J., Cunningham, J., Zecker, S. G., & Carrell, T. D. (1999). Effects of lengthened formant transition duration on discrimination and neural representation of synthetic CV syllables by normal and learning-disabled children. *The Journal of the Acoustical Society of America, 106*, 2086–2096.
- Call, J. (2001). Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and children (*Homo sapiens*). *Journal of Comparative Psychology, 115*, 150–171.
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool. *Developmental Neuropsychology, 28*, 595–616.
- Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology, 40*, 1105–1122.
- Carney, A. E., Osberger, M. J., Carney, E., Robbins, A. M., Renshaw, J., & Miyamoto, R. T. (1993). A comparison of speech discrimination with cochlear implants and tactile aids. *The Journal of the Acoustical Society of America, 94*, 2036–2049.
- 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.
- Champion, T. B., Hyter, Y. D., McCabe, A., & Bland-Stewart, L. M. (2003). A matter of vocabulary: Performance of low-income African American Head Start children on the Peabody Picture Vocabulary Test—III. *Communication Disorders Quarterly, 24*, 121–128.
- Charles-Luce, J., & Luce, P. A. (1990). Similarity neighborhoods of words in young children's lexicons. *Journal of Child Language, 17*, 205–215.
- Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language, 22*, 727–735.
- Collier-Baker, E., & Suddendorf, T. (2006). Do chimpanzees (*Pan troglodytes*) and 2-year-old children (*Homo sapiens*) understand double invisible displacement? *Journal of Comparative Psychology, 120*, 89–97.
- Coplan, J. (1993). *Early Language Milestone Scale—2*. Austin, TX: Pro-Ed.
- Coplan, J., & Gleason, J. R. (1988). Unclear speech: Recognition and significance of unintelligible speech in preschool children. *Pediatrics, 82*, 447–452.
- Corrigan, R. (1981). The effects of task and practice on search for invisibly displaced objects. *Developmental Review, 1*, 1–17.
- Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28*, 125–127.
- Dawson, P. W., Nott, P. E., Clark, G. M., & Cowan, R. S. (1998). A modification of play audiometry to assess speech discrimination ability in severe-profound deaf 2- to 4-year-old children. *Ear and Hearing, 19*, 371–384.
- Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural bases of inhibitory control in reaching. *Annals of the New York Academy of Sciences, 608*, 637–676.
- Diamond, A. (2001). A model system for studying the role of dopamine in the prefrontal cortex during early development in humans: Early and continuously treated phenylketonuria. In C. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 433–472). Cambridge, MA: MIT Press.
- Diamond, A. (2006). Bootstrapping conceptual deduction using physical connection: Rethinking frontal cortex. *Trends in Cognitive Science, 10*, 212–218.
- Diamond, A., Churchland, A., Cruess, L., & Kirkham, N. Z. (1999). Early developments in the ability to understand the relation between stimulus and reward. *Developmental Psychology, 35*, 1507–1517.
- 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.
- Eisenberg, L. S., Martinez, A. S., & Boothroyd, A. (2003). Auditory-visual and auditory-only perception of phonetic contrasts in children. *Volta Review*, *103*, 327–346.
- Eisenberg, L. S., Martinez, A. S., & Boothroyd, A. (2004). Perception of phonetic contrasts in infants: Development of the VRASPAC. In R. Miyamoto (Ed.), *Proceedings of the 8th International Cochlear Implant Conference* (International Congress Series 1273, pp. 364–367). Amsterdam, the Netherlands: Elsevier.
- Eisenberg, L. S., Martinez, A. S., & Boothroyd, A. (2007). Assessing auditory capabilities in young children. *International Journal of Pediatric Otorhinolaryngology*, *71*, 1339–1350.
- 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, R. (1970). Simple reaction time: Effects associated with age, preparatory interval, incentive shift, and mode of presentation. *Journal of Experimental Child Psychology*, *9*, 86–104.
- Erber, N. P. (1971). Auditory and audiovisual reception of words in low-frequency noise by children with normal hearing and by children with impaired hearing. *Journal of Speech and Hearing Research*, *14*, 496–512.
- Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, *46*, 245–264.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. (1994). *MacArthur Communicative Development Inventory*. San Diego, CA: Singular/Thomson Learning.
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User guide and technical manual* (2nd ed.). Baltimore, MD: Brookes.
- Fowler, A. E. (1991). How early phonological development might set the stage of phonemic awareness. In S. A. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 97–117). Hillsdale, NJ: Erlbaum.
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*, 31–60.
- Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*, 200–213.
- Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, *81*, 439–454.
- Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, *33*, 966–979.
- Gioia, G. A., Epsy, K. A., & Isquith, P. K. (2001). *Behavioral Rating Inventory of Executive Function—Preschool Version (BRIEF-P)*. Lutz, FL: Psychological Assessment Resources.
- Glass, E., Sachse, S., & von Suchodoletz, W. (2008). Development of auditory sensory memory from 2 to 6 years: An MMN study. *Journal of Neural Transmission*, *115*, 1221–1229.
- Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology*, *32*, 401–424.
- Haskins, H. L. (1949). *A phonetically balanced test of speech discrimination for children* (Unpublished master's thesis). Northwestern University, Evanston, IL.
- Hnath-Chisolm, T. E., Laipply, E., & Boothroyd, A. (1998). Age-related changes on a children's test of sensory-level speech perception capacity. *Journal of Speech, Language, and Hearing Research*, *53*, 94–106.
- Holt, R. F. (2011). Enhancing speech discrimination through stimulus repetition. *Journal of Speech, Language, and Hearing Research*, *54*, 1431–1447.
- 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.
- Holt, R. F., & Carney, A. E. (2007). Developmental effects of multiple looks in speech sound discrimination. *Journal of Speech, Language, and Hearing Research*, *50*, 1404–1424.
- Holt, R. F., Kirk, K. I., & Hay-McCutcheon, M. (2011). Assessing multimodal spoken word-in-sentence recognition in children with normal hearing and children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *54*, 632–657.
- Holt, R. F., & Lalonde, K. L. (2012). Assessing toddlers' speech-sound discrimination. *International Journal of Pediatric Otorhinolaryngology*, *76*, 680–692.
- Houston, D. M., Ying, E., Pisoni, D. B., & Kirk, K. I. (2003). Development of preword-learning skills in infants with cochlear implants. *Volta Review*, *103*, 303–326.
- Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair? *Developmental Neuropsychology*, *28*, 645–668.
- Hughes, C., & Ensor, R. (2006). Behavioral problems in two-year-olds: Links with individual differences in theory of mind, executive function and negative parenting. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *47*, 488–497.
- Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from age 2 to 4. *Developmental Psychology*, *43*, 1447–1459.
- Jusczyk, P. W. (1986). Toward a model of the development of speech perception. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 1–35). Hillsdale, NJ: Erlbaum.
- Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*, *36*, 220–232.
- Kochanska, G., Murray, K. T., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development*, *67*, 490–507.
- Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., & Koch, D. B. (1996, August 16). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science*, *273*, 971–973.
- Kresheck, J. D., & Nicolosi, L. (1973). A comparison of Black and White children's scores on the Peabody Picture Vocabulary Test. *Language, Speech, and Hearing Services in Schools*, *4*, 37–40.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions: Biological Sciences*, *363*, 979–1000.
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the “critical period.” *Language Learning and Development*, *1*, 237–264.
- Leonard, L. B., McGregor, K. K., & Allen, G. D. (1992). Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech and Hearing Research*, *35*, 1076–1085.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Erlbaum.

- Madsen, K. S., Baare, W. F. C., Bestergaard, M., Skimminge, A., Ejersbo, L. R., Ramsøy, T. Z., ... Jernigan, T. L. (2010). Response inhibition is associated with white matter microstructure in children. *Neuropsychologia*, *48*, 854–862.
- Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., & Peterson, A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology*, *66*, 211–235.
- Martin, N. A., & Brownell, R. (2011). *Receptive One-Word Picture Vocabulary Test* (4th ed.). Los Angeles, CA: Western Psychological Services.
- Martinez, A., Eisenberg, L., Boothroyd, A., & Visser-Dumont, L. (2008). Assessing speech pattern contrast perception in infants: Early results on VRASPAC. *Otology and Neurotology*, *29*, 183–188.
- Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *16*, 1–27.
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth in prompting changes from initially holistic to gradually more specified representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Hillsdale, NJ: Erlbaum.
- Nazzi, T., & Bertoni, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition. *Developmental Science*, *6*, 136–142.
- Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America*, *87*, 2705–2715.
- Osberger, M. J., Miyamoto, R. T., Zimmerman-Philips, S., Kemink, J., Stroer, B. S., Firszt, J., & Novak, M. A. (1991). Independent evaluation of the speech perception abilities of children with the Nucleus 22-Channel cochlear implant system. *Ear and Hearing*, *12*(Suppl.), 66S–80S.
- Pelphrey, K., & Reznick, J. (2002). Working memory in infancy. *Advances in Child Development and Behavior*, *31*, 173–227.
- Primus, M. A., & Thompson, G. (1985). Response strength of young children in operant audiometry. *Journal of Speech and Hearing Research*, *28*, 539–547.
- Psychology Software Tools. (2007). E-Prime 2.0 [Computer program]. Pittsburgh, PA: Author.
- Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in dyslexic children. *Journal of Experimental Child Psychology*, *48*, 270–292.
- Roid, G. H., & Miller, L. J. (1997a). *Leiter International Performance Scale—Revised*. Wood Dale, IL: Stoelting.
- Roid, G. H., & Miller, L. J. (1997b). *Leiter International Performance Scale—Revised: Examiner's manual*. In *Leiter International Performance Scale—Revised*. Wood Dale, IL: Stoelting.
- Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development*, *80*, 1360–1375.
- Ross, G., Boatright, S., Auld, P. A. M., & Nass, R. (1996). Specific cognitive abilities in 2-year-old children with subependymal and mild intraventricular hemorrhage. *Brain and Cognition*, *32*, 1–13.
- Ross, M., & Lerman, J. (1970). A picture identification test for hearing-impaired children. *Journal of Speech and Hearing Research*, *13*, 44–53.
- Rothbart, M., & Posner, M. (2001). Mechanisms and variation in the development of attentional networks. In C. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 353–363). Cambridge, MA: MIT Press.
- Ruff, H., & Capozzoli, M. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology*, *39*, 877–890.
- Sanderson-Leepa, M. E., & Rintelman, W. F. (1976). Articulation functions and test-retest performance of normal-hearing children on three speech discrimination tests: WIPI, PBK-50, and NU Auditory Test No. 6. *Journal of Speech and Hearing Disorders*, *41*, 503–519.
- Scheibel, R., & Levin, H. (1997). Frontal lobe dysfunction following closed head injury in children: Findings from neuropsychology and brain imaging. In N. Rasneger, G. Lyon, & P. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 241–260). Baltimore, MD: Brookes.
- Schwartz, I., Burnham, D., & Bowey, J. A. (2006). Phoneme sensitivity and vocabulary size in 2 1/2- to 3-year-olds. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australasian International Conference on Speech Science and Technology* (pp. 142–147). Canberra, Australian Capital Territory, Australia: Australasian Speech Science and Technology Association.
- Smith, B. L., McGregor, K. K., & Demille, D. (2006). Phonological development in lexically precocious 2-year-olds. *Applied Psycholinguistics*, *27*, 355–375.
- Stager, C. L., & Werker, J. F. (1997, July 24). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382.
- Stark, R., & Heinz, J. M. (1996a). Perception of stop consonants in children with expressive and receptive-expressive language impairments. *Journal of Speech and Hearing Research*, *39*, 676–686.
- Stark, R., & Heinz, J. M. (1996b). Vowel perception in language-impaired children and language normal children. *Journal of Speech and Hearing Research*, *39*, 860–869.
- Storkel, H., & Morrisette, M. L. (2002). The lexicon and phonology: Interactions in language acquisition. *Language, Speech, and Hearing Services in Schools*, *33*, 23–37.
- Sussman, J. E. (1991). Stimulus ratio effects on speech discrimination by children and adults. *Journal of Speech and Hearing Research*, *34*, 671–678.
- Sussman, J. E. (1993). Auditory processing in children's speech perception: Results of selective adaptation and discrimination tasks. *Journal of Speech and Hearing Research*, *36*, 380–395.
- Sussman, J. E. (2001). Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions? *The Journal of the Acoustical Society of America*, *109*, 1173–1180.
- 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*, *36*, 380–395.
- Tallal, P., & Piercey, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, *12*, 83–93.
- Tallal, P., & Piercey, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia*, *13*, 69–74.
- Tellinghuisen, D. J., & Oakes, L. M. (1997). Distractibility in infancy: The effects of distractor characteristics and type of attention. *Journal of Experimental Child Psychology*, *64*, 232–254.
- Trehub, S. E., Schneider, B. A., & Henderson, J. L. (1995). Gap detection in infants, children, and adults. *The Journal of the Acoustical Society of America*, *98*, 2532–2541.
- Triana, E., & Pasnak, R. (1986). A distraction task as an assessment of object permanence. *Infant Behavior & Development*, *9*, 151–165.
- Tsao, F., Liu, H., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*, 1067–1084.
- Tyler, R. S. (1993). Speech perception by children. In R. S. Tyler (Ed.), *Cochlear implants: Audiological foundations*. San Diego, CA: Singular.
- Walley, A. C. (1993). The role of vocabulary development in spoken word recognition and segmentation ability. *Developmental Review*, *13*, 286–350.

- Walley, A. C. (2004). Speech perception in childhood. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 449–468). Malden, MA: Blackwell.
- Washington, J. (1996). Issues in assessing language skills in African American children. In A. G. Kamhi, K. E. Pollock, & J. L. Harris (Eds.), *Communication development and disorders in African American children: Research assessment and intervention* (pp. 35–54). Baltimore, MD: Brookes.
- Washington, J. A., & Craig, H. K. (1992). Performance of low-income African American preschool and kindergarten children on the Peabody Picture Vocabulary Test—Revised. *Language, Speech, and Hearing Services in Schools, 23*, 329–333.
- Watson, C. S., & Nichols, T. L. (1976). Detectability of auditory signals presented without defined observation intervals. *The Journal of the Acoustical Society of America, 59*, 655–668.
- Weiss, C. E., & Lillywhite, H. S. (1976). *Communicative disorders*. St. Louis, MO: Mosby.
- Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1*, 197–234.
- Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy, 3*, 1–30.
- Werker, J. F., & Tees, R. C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology, 41*, 48–61.
- Yoshida, K. A., Fennell, C. T., Swingle, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar sounding words. *Developmental Science, 12*, 412–418.

## Appendix A

Results of Principal Component Analysis, Showing Loading of Predictor Variables on Principal Components

Variable	Principle Component 1	Principle Component 2	Principle Component 3	Principle Component 4
Linguistic				
Receptive vocabulary (ROWPVT-4)	0.734	0.405	0.360	-0.005
Expressive vocabulary (MCDI)	-0.136	-0.086	0.783	0.237
Cognitive				
General executive composite (BRIEF-P)	0.151	0.836	-0.072	0.301
Associated pairs (Leiter-R)	0.423	0.327	-0.508	-0.222
Forward memory (Leiter-R)	0.403	0.578	0.299	-0.204
Fundamental visualization (Leiter-R)	0.580	-0.242	0.181	0.644
Fluid reasoning (Leiter-R)	0.539	-0.328	0.413	-0.544
Nonverbal intelligence (Leiter-R)	0.880	-0.282	-0.192	0.045
Spin-the-pots	0.294	-0.208	-0.378	0.268
Other				
Chronological age	0.909	-0.199	-0.110	-0.035

Note. ROWPVT-4 = Receptive One-Word Picture Vocabulary Test—Fourth Edition; MCDI = MacArthur-Bates Communicative Development Inventory; BRIEF-P = Behavioral Rating Inventory of Executive Function—Preschool Version; Leiter-R = Leiter International Performance Scales—Revised.

## Appendix B

First-Order Pearson Correlations Among Predictor Variables

Variable	Linguistic		Cognitive						
	Receptive vocabulary	Expressive vocabulary	Fluid reasoning	Fundamental visualization	Nonverbal intelligence	Paired-associates learning	Short-term memory	Working memory	Executive function
Chronological age	.542*	-.256	.449*	.540*	.834***	.315	.274	.261	-.075
Executive function	.491*	-.079	-.213	.069	.222	.408	.350	-.032	
Working memory	.101	-.160	-.110	.199	-.006	.134	.042		
Short-term memory	.574**	.072	.183	-.060	.042	.154			
Paired-associates learning	.147	-.268	-.026	.095	.216				
Nonverbal intelligence	.416	-.042	.489*	.851***					
Fundamental visualization	.266	.042	.249						
Fluid reasoning	.401	.146							
Expressive vocabulary	.079								

Note. Receptive vocabulary was measured using the ROWPVT-4. Expressive vocabulary was measured using the MCDI. Working memory refers to the spin-the-pots task. Short-term memory refers to the forward memory task. Fundamental visualization measures visual abilities at a basic level. Fluid reasoning refers to the ability to solve novel problems.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .