Research Article

Nonword Repetition and Language Outcomes in Young Children Born Preterm

Lisa D. Gresch,^a Virginia A. Marchman,^b Elizabeth C. Loi,^a
Anne Fernald,^b and Heidi M. Feldman^a

Purpose: The aims of this study were to examine phonological short-term memory in children born preterm (PT) and to explore relations between this neuropsychological process and later language skills.

Method: Children born PT ($n = 74$) and full term (FT; $n = 60$) participated in a nonword repetition (NWR) task at 36 months old. Standardized measures of language skills were administered at 36 and 54 months old. Group differences in NWR task completion and NWR scores were analyzed. Hierarchical multiple regression analyses examined the extent to which NWR ability predicted later performance on language measures.

Results: More children born PT than FT did not complete the NWR task. Among children who completed the task, the performance of children born PT and FT was not statistically different. NWR scores at 36 months old accounted for significant unique variance in language scores at 54 months old in both groups. Birth group did not moderate the relation between NWR and later language performance.

Conclusions: These findings suggest that phonological short-term memory is an important skill underlying language development in both children born PT and FT. These findings have relevance to clinical practice in assessing children born PT.

In 2015, approximately one in 10 infants in the United States was born preterm (PT; Centers for Disease Control and Prevention, 2016). Although medical advancements have increased survival rates and improved health outcomes, children born PT continue to be at an increased risk for neurodevelopmental disorders compared with children born full term (FT; Aylward, 2005; Vieira & Linhares, 2011; Williams et al., 2013). Previous studies have demonstrated that approximately one third of children born PT can be classified as having a language disorder (Briscoe, Gathercole, & Marlow, 1998; Sansavini et al., 2010; Woodward et al., 2009), and children born PT consistently perform more poorly on language tasks than their age-matched, FT peers (Sansavini et al., 2007; Wolke, Samara, Bracewell, & Marlow, 2008; Woodward et al., 2009). A meta-analysis integrating findings in children from preschool to adolescence revealed that, as a group, children born PT scored between 0.38 and 0.77 *SD* below children born FT on receptive and expressive language

tasks (Barre, Morgan, Doyle, & Anderson, 2011). These group differences have been found across development (Guarini et al., 2016; Luu et al., 2009; Sansavini et al., 2010), irrespective of the presence of major disabilities, such as cerebral palsy, intellectual disabilities, and sensory impairments and independent of socioeconomic status (SES; van Noort-van der Spek, Franken, & Weisglas-Kuperus, 2012). Later in the school years, children born PT have been found to demonstrate greater deficits than children born FT in decoding and reading comprehension, skills that require a strong foundation in oral language (Kovachy, Adams, Tamareis, & Feldman, 2015).

A critical gap in the literature is understanding whether children born PT show patterns of language development that are similar to those of children born FT or whether they show distinctive cognitive-linguistic profiles. An important domain to consider is phonological short-term memory because of its putative critical role in language development (Bishop, North, & Donlan, 1996; Gathercole & Baddeley, 1990; Parra, Hoff, & Core, 2011; Sahlén, Reuterskiöld-Wagner, Nettelbladt, & Radeborg, 1999). To date, this domain has been minimally studied in children born PT. Children born PT have been shown to have deficits in spatial and object working memory (Vicari, Caravale, Carlesimo, Casadei, & Allemand, 2004; Woodward, Edgin, Thompson, & Inder, 2005). They also have other cognitive deficits, including depressed intelligence scores, weak

^aDepartment of Pediatrics, Stanford University School of Medicine, CA

^bDepartment of Psychology, Stanford University, CA

Correspondence to Virginia A. Marchman: marchman@stanford.edu

Editor-in-Chief: Sean Redmond

Editor: Lizbeth Finestack

Received June 5, 2017

Revision received September 12, 2017

Accepted January 22, 2018

https://doi.org/10.1044/2018_JSLHR-L-17-0217

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

executive function skills, and poor attention, which may negatively affect phonological short-term memory or may uncouple the association between phonological short-term memory and subsequent language performance (Aylward, 2002; Bhutta, Cleves, Casey, Craddock, & Anand, 2002; Frye, Landry, Swank, & Smith, 2009). We undertook this study to compare the performance of 36-month-old children born PT with the performance of an age-matched group of children born FT on a task of phonological short-term memory, specifically nonword repetition (NWR). We further explored how performance on this task would be related to language ability concurrently and a year and a half later to determine whether the degree of association would be similar in both groups of children.

Nonword Repetition

Tasks of NWR have been suggested as a measure of phonological short-term memory (Gathercole & Baddeley, 1990) and have gained prominence in the study and assessment of children with known or suspected language deficits (Bishop et al., 1996; Boerma et al., 2015; Roy & Chiat, 2004; Weismer et al., 2000). Typical NWR tasks require children to listen to audio recordings of novel phonetic sequences of varying lengths and to verbally repeat each item. Responses are scored based on the child's ability to accurately reproduce the entire word or on the proportion of individual phonemes produced correctly. Previous studies have shown that children with language impairments demonstrate poorer performance on tasks of NWR than children with typical language skills (Bishop et al., 1996; Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990; Gray, 2003; Leonard et al., 2007; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998; Weismer et al., 2000). Moreover, children with language deficits, especially specific language impairment (SLI), demonstrate disproportionate difficulty with nonwords of increasing length compared with their typically developing peers (Dollaghan & Campbell, 1998; Weismer et al., 2000).

Difficulty with NWR has been associated with deficits in multiple language domains but has been most widely studied in relation to semantic and morphosyntactic skills (Adams & Gathercole, 1995; Bishop et al., 1996; Chiat & Roy, 2007). Several studies note a relation between vocabulary and NWR (Adams & Gathercole, 1995; Edwards, Beckman, & Munson, 2004; Gathercole & Adams, 1993; Gathercole & Baddeley, 1990; Guarini et al., 2016; Hoff, Core, & Bridges, 2008). This association suggests an empirical connection between the processes required for repeating nonwords and for learning new vocabulary, consistent with the theory that children need to recognize, remember, and recall phonological sequences when learning new words (Adlof & Patten, 2017; Gathercole, 2006). In addition, studies have suggested connections between NWR and general language ability, including morphosyntactic skills (Adams & Gathercole, 1995; Chiat & Roy, 2008; Sahlén et al., 1999). This finding provides evidence that phonological short-term memory plays a role in a child's ability to remember and

recall phonological information, allowing them not only to learn individual lexical items but also to form generalizations regarding morphosyntactic structures.

NWR tasks have traditionally been used in the study of school-age children aged 5 years and older (e.g., Dollaghan & Campbell, 1998; Gathercole, Willis, Baddeley, & Emslie, 1994; Wagner, Torgesen, Rashotte, & Pearson, 2013). The nature of the traditional task is difficult for children under the age of 5 years who are unfamiliar with decontextualized repetition tasks (Chiat & Roy, 2007). Researchers have altered the procedure of the NWR test to reduce task demands for younger children in two ways (Chiat & Roy, 2007; Hoff et al., 2008). First, children are asked to repeat nonwords produced live by an experimenter, rather than via an audio recording. In addition, the nonwords are used in a meaningful context for the child, such as presenting the stimuli as the "name" of a toy figure (e.g., "This is my friend *Jot!* Can you say *Jot!*"). These procedural changes have created a naturalistic NWR task in which young children are more likely to participate.

Another difficulty with using NWR tasks with young children is that these children present with immature phonological systems that may impede their ability to accurately repeat presented items. Although the impact of speech sound production ability has not been systematically considered in relation to NWR in children born PT, it has been shown that speech sound production ability affects NWR in children born FT (Adams & Gathercole, 1995; Krishnan et al., 2013) and suggested that children born PT have poorer articulatory and phonological skills than children born FT (Guarini et al., 2016; Wolke & Meyer, 1999). Therefore, using scoring methods that determine repetition accuracy based on adultlike productions may negatively impact NWR scores and obscure errors that specifically result from deficits in phonological short-term memory in either or both populations. To differentiate between phonological errors that reflect immature speech sound production and errors that reflect limitations in short-term memory, researchers have modified scoring procedures to allow for developmentally appropriate phonological substitutions and phoneme distortions (Deevy, Weil, Leonard, & Goffman, 2010; Roy & Chiat, 2004; Stokes & Klee, 2009).

NWR and Children Born PT

To date, a limited number of studies have explored NWR in children born PT (Briscoe et al., 1998; De Hirsch, Jansky, & Langford, 1964; Fraello et al., 2011; Guarini et al., 2016; Jansson-Verkasalo et al., 2004; Odd, Emond, & Whitelaw, 2012; Sansavini et al., 2007). Two of these studies lack relevance to this study because of critical sample differences. The first evaluated 5- to 6-year-old children born between 1955 and 1965 (De Hirsch et al., 1964); such children tended to be born at older gestational ages and faced very different medical challenges than children born after the advent of mechanical ventilation and other medical advances. The second compared cognitive outcomes in 8- to 11-year-old children born at 32–36 weeks of gestation with

children born at 37–42 weeks of gestation (Odd et al., 2012); these children were born at older gestational ages than PT samples in other studies, including our study.

Two other previous studies that included evaluating NWR in children born PT are pertinent to this study. The first evaluated 12-year-old adolescents on a variety of language measures and found no differences between PT and FT groups on NWR, although the PT group performed below FT peers on recalling sentences, another task related to short-term memory (Fraello et al., 2011). By contrast, Guarini and colleagues (2016) found that 5-year-old children born PT at less than 32 weeks of gestation scored significantly below their FT peers on an NWR task. These inconsistent findings suggest that age at the time of testing may be a relevant variable. NWR may be insensitive to PT–FT group differences by the time children have mastered language skills and reading, so studies with younger children who are at the beginning stages of learning are warranted.

Two studies have explored NWR in preschool-aged children born PT. Both studies used a whole-word, correct/incorrect scoring procedure. Sansavini and colleagues (2007) explored group differences in task participation and non-word item length and compared NWR performance with grammatical abilities in 3.5-year-old Italian children born PT. Overall performance on the NWR task did not significantly differ between children born PT and FT. However, children born PT demonstrated significantly more item omissions and performed significantly more poorly when repeating four-syllable words compared with their FT peers. The researchers also found that significantly more children born PT than FT did not complete the NWR task, and these children demonstrated poorer grammatical skills as measured by mean length of utterance (MLU) on an Italian test of sentence repetition (Test di Ripetizione di Frasi) (Devescovi & Caselli, 2001). Correlations between NWR performance and Test di Ripetizione di Frasi MLU were found for both groups, suggesting that the relation between phonological short-term memory and MLU is characteristic of children regardless of birth status. These findings align with those of Briscoe and colleagues (1998) who also found no significant group differences between 3- to 4-year-old children born PT and FT on an NWR task. However, when these researchers divided the children into groups based on risk for language impairment, as determined by a test of narrative language, the mean NWR total score for at-risk PT children was significantly lower than that of no-risk PT and FT children (Briscoe et al., 1998). It should be noted, however, that the at-risk PT, no-risk PT, and FT samples included only eight, 18, and 26 children, respectively. In addition, these researchers only used two- and three-syllable nonwords. It is possible that the small sample size and lack of variability in scores due to limited items obscured significant results. Taken together, the findings of Sansavini and colleagues (2007) and Briscoe and colleagues (1998) suggest that deficits in phonological short-term memory are associated with weaker language skills in children born PT, similar to the association in children born FT. However, the cross-sectional nature of these studies prevents conclusions regarding the

developmental impact of early phonological short-term memory skills on later language skills in children born PT.

The Present Study

In this study, we sought to determine whether phonological short-term memory is affected by PT birth and whether this skill is related to language development in children born PT, as it is in children born FT. We chose to study this skill in children at 36 months old when they are still rapidly developing their linguistic skills in order to inform understanding of the developmental profile of children born PT and to maximize the likelihood of finding significant group differences. The following hypotheses and questions were tested. (a) We hypothesized that, compared with children born FT, 3-year-old children born PT would have more difficulty completing an NWR task and would achieve lower overall scores on the NWR task (Sansavini et al., 2007). (b) To date, the relation between speech sound production ability and NWR has not been explicitly explored in children born PT, but some studies have suggested that children born PT have poorer articulatory and phonological skills than children born FT (Guarini et al., 2016; Wolke & Meyer, 1999). We predicted that differences between PT and FT groups in NWR performance would persist or emerge when articulation and phonological skills for age were taken into account in scoring. (c) Given that nonword length has been shown to be related to language deficits and children born PT have been shown to be at a greater risk for language deficits, we hypothesized that the gap in performance in NWR between PT and FT 3-year-olds would increase as nonword length increased (Briscoe et al., 1998; Dollaghan & Campbell, 1998; Sansavini et al., 2007; Weismer et al., 2000). (d) Although previous studies of children born PT have examined concurrent associations of NWR and language ability, they have not evaluated the children longitudinally (Briscoe et al., 1998; Sansavini et al., 2007). We sought to understand the relation between phonological short-term memory ability and subsequent language skills by examining associations between NWR at the age of 36 months and general language skills a year and a half later. We predicted that NWR performance at the age of 3 years would be associated with language skills at the age of 4.5 years in children born PT and FT, and we explored whether birth group would moderate the strength of this association.

Method

Participants

Participants were 134 children (64 girls, 70 boys) born PT ($n = 74$) or FT ($n = 60$). All participants were part of a longitudinal study tracking language development in PT and FT children from 16 to 54 months old. Children in the PT group were born ≤ 32 weeks of gestational age. PT participants were primarily recruited from a local children's hospital via the neonatal intensive care unit and the hospital's High-Risk Infant Follow-up Clinic. Children in the FT group

were born ≥ 37 weeks of gestational age. FT participants were primarily recruited through birth records. PT and FT families were also recruited via a research registry, local parent groups, birth records, and an early intervention service provider. Exclusionary criteria for all children included a history of meningitis, seizure disorder requiring medication, ventriculoperitoneal shunt, genetic disorder, or a visual or hearing impairment. All children were primary English speakers (i.e., heard less than 25% of another language at home).

The current NWR analyses used data collected at 36 months old. There was no significant difference in chronological age between children born PT ($M = 36.3$, $SD = 0.64$) and FT ($M = 36.4$, $SD = 0.65$) at the 36-month visit, $t(132) = -0.60$, $p = .55$. As seen in Table 1, at 36 months old, the two birth groups did not significantly differ in percentage of male participants, maternal years of education, or SES, as calculated from an updated version of the Hollingshead Four Factor Index of Socioeconomic Status (Hollingshead, 1975; range = 8–66). On average, both samples of children were from college-educated and high-SES families, although there was variation within each group.

Follow-up language data were available on a subset of children born PT ($n = 35$) or FT ($n = 20$) who had reached 54 months old at the time of data analysis. At the 54-month visit, the PT group ($M = 54.5$, $SD = 1.6$) was, on average, about one month older than the FT group ($M = 53.7$, $SD = 1.0$), $t(53) = 2.11$, $p = .04$. Children in the PT and FT groups at 54 months old did not significantly differ in gender, $\chi^2(1) = 0.003$, $p = .96$, or maternal years of education (PT: $M = 16.4$, $SD = 1.94$; FT: $M = 17.1$, $SD = 0.97$), $t(52.3) = -1.85$, $p = .07$. Among those seen at the 54-month visit, children born PT ($M = 57.1$, $SD = 7.2$) were from significantly lower-SES households than children born FT ($M = 62.4$, $SD = 3.00$), $t(49.5) = -3.79$, $p < .001$; however, both groups were primarily from college-educated, higher-SES families. SES and age are controlled for in all analyses at this data point.

Procedure

This study includes data collected from an NWR task at 36 months old, a measure of receptive vocabulary

at 36 and 54 months old, and a measure of general language ability at 54 months old. The experimental tasks were administered by trained research assistants at a university laboratory. The child's caregiver(s) were allowed in the room during evaluation to provide emotional support for their child during the experimental tasks. Caregivers were asked to remain as unobtrusive as possible and to avoid comments related to the accuracy of their child's responses. All sessions were audio- and/or video-recorded to allow for offline transcription of the NWR data.

Phonological Short-Term Memory Task

Real Word Production

This task was administered to accustom the child to the experimental procedures. In the Real Word Production task, the experimenter asked the child to name or repeat the names of familiar real objects presented as photographs one at a time. Real word stimuli are listed in the Appendix.

Nonword Repetition

The NWR task was a modified version of the task used by Hoff et al. (2008) in which test stimuli were administered live, rather than through an audio recording. This procedure is more natural than when prerecorded stimuli are used because stimuli presentation can be easily adjusted to each child's behavioral and attention needs (Adams & Gathercole, 1995; Bishop et al., 1996; Chiat & Roy, 2007; Hoff et al., 2008). In addition, nonword stimuli were presented as names for toy animals and people (e.g., "This is my friend *Jot*. Can you say *Jot*?"). Nonword stimuli, found in the Appendix, ranged from one to five syllables in length, for a total of 20 items. The one-, two-, and three-syllable items were borrowed from Parra et al. (2011). Additional four- and five-syllable nonword items were created to increase opportunities for production attempts and to prevent a ceiling effect. Prior studies have found that children with deficits in phonological short-term memory demonstrate more difficulty repeating longer nonwords (Bishop et al., 1996; Briscoe, Bishop, & Norbury, 2001).

Trained research assistants administered the task. If the child did not respond to the initial presentation of the

Table 1. Descriptive statistics of birth history and sociodemographic factors.

Characteristic	PT, <i>M</i> (<i>SD</i> ; <i>n</i> = 74)	FT, <i>M</i> (<i>SD</i> ; <i>n</i> = 60)	<i>t</i> (132) or χ^2	<i>p</i> value
Birth history				
Gestational age at birth (weeks)	29.3 (1.8)	40.2 (1.0)	–40.7	< .001*
Birth weight (g)	1202.2 (275.4)	3549.4 (444.6)	–37.4	< .001*
Gender (%)				
Male/Female	52.7/47.3	51.7/48.3	0.01	.91
Family demographics				
Maternal years of education	16.3 (1.8)	16.5 (1.6)	–0.7	.48
SES	55.8 (9.8)	58.1 (9.0)	–1.4	.15

Note. PT = preterm; FT = full term; SES = Hollingshead Four Factor Index of Socioeconomic Status.

* $p \leq .001$.

stimulus, at least one repetition of the stimulus was attempted. When the child's failure to respond was caused by apparent distraction away from the task, the research assistant attempted to draw the child's attention back to the stimulus before the second presentation. When the failure to respond was due to active refusal, the research assistant paused and attempted to allow the child control of other aspects of the play situation to encourage cooperation with the NWR task. We attempted to present all of the items to all of the participants, unless the child actively refused to participate or disengaged from the task. We limited the number of repetitions of each stimulus item because more than two repetitions might give some children unfair advantage over other children.

Scoring and Reliability

A certified speech-language pathologist who was blind to birth group status transcribed and scored participant responses. Narrow transcription using the International Phonetic Alphabet was used to account for variations in speech sound production. Two scoring criteria were used. First, percentage of consonants correct (PCC) was calculated using procedures described by Shriberg and Kwiatkowski (1982). This method was judged to be the most conservative measure, not allowing for any deviation from adultlike production of the target words.

Second, because of the young age and diverse speech status of our participants, we also incorporated a relaxed scoring system, herein referred to as Relaxed PCC, based on Shriberg, Austin, Lewis, Mcsweeney, and Wilson's (1997) PCC-Revised. Consistent with the original PCC-Revised scoring method, all phoneme distortions were scored as correct. Our modified system also accepted developmentally appropriate substitutions for consonant clusters and the following phonemes, which are acquired by fewer than 75% of children at 36 months old: /l ɪ ŋ j v θ ð s z ʃ ʒ tʃ dʒ/ (Smit, Hand, Freilinger, Bernthal, & Bird, 1990). Previous studies of NWR with young participants have used similar scoring procedures that allow for articulatory and phonological differences (Roy & Chiat, 2004).

To assess reliability, NWR responses from 20% of the participants were independently transcribed and scored a second time by the same speech-language pathologist 6 months after the initial transcriptions. Intraclass correlation coefficients (ICC) were used to calculate intrarater reliability. Reliability was excellent for both PCC total scores (ICC = 0.92) and Relaxed PCC total scores (ICC = 0.98). In addition, NWR responses from 20% of the participants were randomly selected to be listened to and scored a second time by a trained research assistant who was unaware of birth group status. Reliability was excellent for both PCC total scores (ICC = 0.85) and Relaxed PCC total scores (ICC = 0.92).

Task Completion and Nonresponses

Of the 134 children, 16 attempted less than four of 20 possible items; the remaining 118 completed 10 or more

items. We excluded the 16 children who did not complete at least four items (25% of the task) because their NWR scores might not be reliable estimates of their phonological short-term memory. For example, a child who attempted only one item could receive a PCC score of 100% for correctly repeating only two of two phonemes, but this score is not comparable with the same score of a child who completes all of the items. The 118 children included in further analyses successfully completed $\geq 50\%$ of the items. Of this group, 82.2% completed all 20 items.

When calculating PCC and Relaxed PCC, only items for which the child attempted a response were scored. All other items were considered "missing" and did not receive a score. Phoneme counts were adjusted accordingly. Missing items included items with inaudible responses (e.g., due to environmental noise), items to which the child did not respond, and items that were not administered. The decision to not include missing items in scoring was based on procedures used by Hoff and colleagues (2008). Excluding items that were missing in computation of NWR scores reduced variability in the range of scores across the sample. However, we chose to exclude the items because the young age of the participants and possible concomitant attention and behavioral challenges made it difficult to determine whether nonresponses were due to the child's difficulty with recalling the phoneme sequences or to unrelated factors. In addition, some children attempted to repeat the nonwords regardless of whether they thought they could do so accurately, whereas others were reticent to respond if they thought they would repeat the word incorrectly. Therefore, not scoring missing items leads to a measure of phonological short-term memory that is not confounded with child temperament.

Language Outcome Measures

Measures of receptive vocabulary and general language ability were available for use during analyses exploring the relation between phonological short-term memory and later language skills. Receptive vocabulary was measured at 36 months old and again at 54 months old using the Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; Dunn & Dunn, 2007), a well-known measure with strong reliability and validity. The PPVT-4 was administered according to standardized procedures by trained research assistants. Standard scores were derived based on chronological age for both PT and FT children.

General language ability was measured at 54 months old using the Core Language Score from the Clinical Evaluation of Language Fundamentals–Preschool-2 (CELF-P2; Semel, Wiig, & Secord, 2004), a widely used measure of language ability with strong validity and reliability. The Core Language Score is derived from scores on the Sentence Structure, Word Structure, and Expressive Vocabulary subtests. Subtests were administered by trained research assistants according to standardized procedures, and standard scores were derived based on chronological age for both PT and FT children.

Results

Task Completion

We first sought to test the hypothesis that children born PT would have more difficulty completing the NWR task than children born FT. Of the 134 children who were tested at 36 months old, 16 children were excluded from NWR analyses because they completed < 25% of the items. Of this group, 13 were from the PT group (17.6% of PT children) and three were from the FT group (5% of FT children). A chi-square test demonstrated that children from the PT group were more likely to be excluded because of insufficient task completion, $\chi^2(1) = 4.98, p = .03$.

Given that more children born PT than FT had difficulty completing the NWR task, it was possible that even those children born PT who participated in the NWR task would complete fewer items than children born FT. To test this, we compared the total number of NWR items to which children born PT and FT responded. Children born PT ($M = 19.4, SD = 2.0$) did not significantly differ from children born FT ($M = 19.3, SD = 1.8$) in the number of items attempted, $t(116) = -0.42, p = .68, d = 0.05$.

Group Comparisons Based on Scoring Method

We next compared NWR total scores of children born PT and FT who completed the task. The final sample for this analysis included 61 children born PT and 57 children born FT. As with the full sample, the birth groups did not differ significantly in terms of gender, $\chi^2(1) = 0.28, p = .60$, maternal years of education (PT: $M = 16.4, SD = 1.8$; FT: $M = 16.5, SD = 1.5$), $t(116) = -0.44, p = .66$, or SES, (PT: $M = 55.8, SD = 9.3$; FT: $M = 58.1, SD = 9.2$), $t(116) = -1.32, p = .19$.

We used two different scoring methods because speech sound production ability can impact NWR performance (Adams & Gathercole, 1995; Krishnan et al., 2013), and some studies have suggested that children born PT have poorer articulatory and phonological skills than children born FT (Guarini et al., 2016; Wolke & Meyer, 1999). Therefore, speech sound production ability had the potential to complicate interpretation of group differences.

Table 2 presents mean scores on the NWR task for children born PT and FT using the traditional PCC and the Relaxed PCC scoring methods. Children born PT scored below children born FT on average; however, group differences in NWR performance using either scoring method did

not achieve statistical significance, and the effect sizes were small. As expected, NWR scores were significantly higher with the Relaxed PCC than with the PCC scoring, for both children born PT, $t(60) = 16.54, p < .001$, and FT, $t(56) = 17.62, p < .001$.

To determine if the Relaxed PCC scoring method resulted in higher NWR scores to the same extent in both groups, a mixed analysis of variance was conducted, with Group (PT vs. FT) as the between-subject factor and Scoring Method (PCC vs. Relaxed PCC) as the within-subject factor. The results indicated a main effect of Scoring Method, $F(1, 116) = 574.83, p < .001, \eta_p^2 = .83$, but no main effect of Group, $F(1, 116) = 2.40, p = .12, \eta_p^2 = .02$, and no Group $\times$ Scoring Method interaction, $F(1, 116) = 0.21, p = .65, \eta_p^2 = .002$ (Figure 1). The Relaxed PCC scoring method increased scores to the same extent for children in both groups. Because of the young age and diverse speech status of our participants, Relaxed PCC scores were used for all subsequent analyses to allow for developmental and individual variations in speech sound production.

It should be noted that, if we had scored nonadministered and nonattempted items as zero and included all children in statistical analyses, regardless of the number of items completed, then differences in performance as a function of birth group status would have reached statistical significance for both the PCC (PT: $M = 44.3, SD = 25.6$; FT: $M = 53.8, SD = 19.6$), $t(131.6) = -2.43, p = .02, d = 0.42$, and Relaxed PCC (PT: $M = 56.7, SD = 31.0$; FT: $M = 67.1, SD = 21.9$), $t(129.7) = -2.27, p = .03, d = 0.39$, scoring methods. However, effect sizes remained small. Statistics were corrected for unequal variances.

Effect of Nonword Length

To assess whether articulation and phonological skills influenced performance for either birth group, we compared the Relaxed PCC scores at each nonword length (number of syllables) for each birth group (Table 3). Because mean scores for one- and two-syllable items and for three- and four-syllable items were similar, scores were collapsed across these syllable lengths in further analyses to reduce the number of statistical comparisons. We conducted a repeated measures mixed analysis of variance with Nonword Length as a within-subject factor and Group as a between-subject factor. Three children from the PT group and two children from the FT group were excluded because of incomplete data across item lengths. The Huynh-Feldt

Table 2. Birth group differences in nonword repetition (NWR) scores using percentage of consonants correct (PCC) and Relaxed PCC scoring.

NWR scoring	PT ($n = 61$)		FT ($n = 57$)		$t(116)$	p	d
	M	SD	M	SD			
PCC	54.8	15.8	59.2	14.2	-1.61	.11	0.29
Relaxed PCC	70.1	16.4	74.0	13.5	-1.41	.16	0.26

Note. PT = preterm; FT = full term.

Figure 1. Mean nonword repetition scores using percentage of consonants correct (PCC) and Relaxed PCC scoring methods as a function of birth group.

correction was applied to correct for violations of sphericity in the sample (i.e., to correct for greater variance in the NWR scores for children born PT compared with children born FT).

As shown in Figure 2, there was a significant main effect of nonword length, $F(1.88, 208.59) = 247.29, p < .001, \eta_p^2 = .69$, with children from both groups having more difficulty repeating longer nonwords. There was no significant main effect of group, $F(1, 111) = 0.78, p = .38, \eta_p^2 = .007$, and no significant interaction between group and length, $F(1.88, 208.59) = 0.79, p = .45, \eta_p^2 = .007$. These results suggest that, regardless of birth group, children typically demonstrated greater difficulty repeating longer nonwords.

NWR Relations to Language Performance

Task Completion and Language Ability

We next sought to determine if performance on a task of NWR was related to concurrent and later language ability, as measured by standardized assessments. We first compared PPVT-4 standard scores at 36 months old for

Table 3. Mean and standard deviation of relaxed percentage of consonants correct scores at each nonword length for children born preterm and full term.

Number of syllables	Preterm			Full term		
	n	M	SD	n	M	SD
1	61	86.95	16.27	57	87.21	14.81
2	61	88.67	14.03	57	91.21	12.08
3	61	70.78	22.67	57	75.69	17.49
4	60	76.37	19.00	56	79.40	17.00
5	58	54.86	20.88	55	58.05	19.75

Note. Nonresponses were not scored. Smaller sample sizes reflect no response to any item at a given nonword length.

Figure 2. Effect of nonword length on Relaxed percentage of consonants correct (PCC) scores as a function of birth group.

children who did and did not sufficiently complete the task across both birth groups. Statistics were corrected for unequal variances. Children who did not complete the NWR task ($M = 94.5, SD = 23.6$) demonstrated significantly lower PPVT-4 scores, on average, than those who sufficiently completed the task ($M = 111.1, SD = 16.9$), $t(17.2) = -2.72, p = .01$. However, 45% ($n = 5$) of the 11 children who scored $-1 SD$ below the mean on the PPVT-4 sufficiently completed the NWR task.

We also compared language performance at 54 months old of those children with insufficient NWR task completion at 36 months old ($n = 10$) with that of the children with sufficient NWR task completion ($n = 55$; Table 4). Similar to our earlier results, children who had difficulty completing the NWR task at 36 months old demonstrated significantly poorer PPVT-4 and CELF-P2 Core Language scores at 54 months old in comparison with those children who completed the NWR task. Therefore, those children who had greater difficulty completing the NWR task at 36 months old demonstrated poorer later language skills.

Table 4. Comparison of mean standard scores on language outcomes at 54 months old for children who did not complete ($n = 10$) and those who did complete ($n = 55$) the nonword repetition (NWR) task at 36 months old.

Measure	NWR noncompleters		NWR completers		$t(63)$	p	d
	M	SD	M	SD			
PPVT-4	98.2	21.1	114.6	17.1	-2.69	.01**	0.85
CELF-P2	98.1	16.3	109.1	15.8	-2.01	.05*	0.69

Note. PPVT-4 = Peabody Picture Vocabulary Test—Fourth Edition; CELF-P2 = Clinical Evaluation of Language Fundamentals—Preschool-2 Core Language Score.

* $p \leq .05$. ** $p \leq .01$.

Prediction and Moderation Analyses

Our final analysis was conducted to determine the extent to which performance on an NWR task would be associated with later language scores for a subsample of 55 children for whom data were available for the PPVT-4 and the CELF-P2 at 54 months old. We also explored whether birth group would moderate the strength of the association between NWR scores and later language skills. Looking first at descriptive statistics, Table 5 shows that the mean scores for the children born PT were generally lower than those for children born FT on both measures. However, group differences achieved statistical significance only for the CELF-P2. Nevertheless, mean scores for children born PT fell within the average range on both language measures, with only four children scoring -1 *SD* below the mean on each measure.

Table 6 shows results from multiple regression analyses that explored whether NWR Relaxed PCC scores at 36 months old predict language outcomes at 54 months old, as measured by the PPVT-4 and CELF-P2 Core Language. SES and age at visit were covariates due to significant group differences. Because earlier language skills may explain variance in later language outcomes, we also included PPVT-4 scores at 36 months old. Correlations between each of the independent variables in the model were weak or moderate; therefore, the assumption of collinearity was not violated. Table 6 shows that SES, age at visit, earlier language skills, and birth group accounted for 61% of the variance in PPVT-4 scores, $p < .001$, with no significant contribution of birth group, $p = .41$. Relaxed PCC scores explained 12.6% unique variance, $p < .001$. We then applied the PROCESS macro in SPSS 23.0 (Hayes, 2013) to explore the contribution of the interaction of Group $\times$ NWR Relaxed PCC. The interaction term did not account for any additional variance, $F(1, 48) = 0.08$, $p = .78$, $\Delta R^2 = 0.09\%$, indicating that birth group did not moderate these relations.

We found analogous results for the CELF-P2 (Table 6). The covariates accounted for about 41% of the variance, $p < .001$, with only previous language skills accounting for unique variance. However, Relaxed PCC scores accounted for about 16% additional variance, $p < .001$, when added to the model. The relation between NWR Relaxed PCC scores and CELF-P2 was not moderated by birth group,

$F(1, 48) = 0.013$, $p = .91$, $\Delta R^2 = 0.04\%$. Taken together, these results suggest significant predictive validity between NWR performance at 36 months old and language outcomes at 54 months old for all children, regardless of birth group.

Discussion

The overall goal of this study was to explore the neuropsychological process of phonological short-term memory, as measured by an NWR task, and to investigate its relation to language skills in children born PT, a population known to demonstrate poorer language skills than their FT peers. Our findings revealed that children born PT had greater difficulty completing the NWR task at 36 months old than their FT peers. Among those children who completed the task, there were no significant differences in NWR performance, regardless of scoring measure and nonword length. Furthermore, there was a significant, positive association between NWR and later language skills for both children born PT and FT. We discuss each of these findings in turn.

NWR Task Completion

Our findings support the hypothesis that children born PT would be more likely to have difficulty completing the NWR task than children born FT. One possible reason for this finding is that children born PT have poorer receptive language skills, which may have compromised their ability to understand the task and/or remember the stimuli. Consistent with this hypothesis, across both groups, children who had difficulty completing the task demonstrated significantly poorer concurrent receptive vocabulary scores than children with sufficient completion. These children also demonstrated significantly lower scores on measures of receptive vocabulary and general language ability at 54 months old. These results suggest that noncompletion of the task may be a clinically relevant indicator of high risk for poor language outcomes, although further research is needed to determine the reason for the relation between noncompletion and poor language outcomes.

There may be other factors that contribute to difficulty completing the task. Previous studies have found that children born PT demonstrate poorer attention and executive functioning skills (Aarnoudse-Moens, Weisglas-Kuperus, van Goudoever, & Oosterlaan, 2009; Bhutta et al., 2002). In addition to having potentially low receptive language skills, it is possible that children who had more difficulty approaching and understanding the NWR task did so because of deficits in these other areas. These skills were not assessed as a part of this study. However, future studies should consider exploring these relations.

NWR Task Performance

Contrary to our hypothesis, we found that, as a group, the performance of those children born PT who

Table 5. Mean standard score for preterm (PT) and full-term (FT) groups on standardized language measures at 54 months old.

Measure	PT (<i>n</i> = 35)		FT (<i>n</i> = 20)		<i>t</i> (53)	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
PPVT-4	111.97	17.27	119.25	16.25	-1.54	.13	0.43
CELF-P2	105.31	14.69	115.70	15.86	-2.45	.02*	0.68

Note. PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition; CELF-P2 = Clinical Evaluation of Language Fundamentals–Preschool-2 Core Language Score.

* $p \leq .05$.

Table 6. Summary of hierarchical regression analyses predicting PPVT-4 and CELF-P2 standard scores at 54 months old ($N = 55$).

Variable	PPVT-4				CELF-P2	
SES	0.15 (0.26)	0.06 (0.22)	0.07 (0.30)	0.20 (0.30)	0.10 (0.26)	0.11 (0.35)
Age (months)	1.24 (1.08)	1.74 (0.91)	1.72 (1.07)	0.11 (1.23)	0.63 (1.07)	0.62 (1.42)
PPVT-4 at 36 months old	0.79 (0.10)***	0.60 (0.09)***	0.61 (0.12)***	0.54 (0.12)***	0.34 (0.11)**	0.34 (0.12)**
Birth group	3.03 (3.68)	2.55 (3.06)	8.26 (21.07)	-2.24 (4.16)	-2.74 (3.60)	0.87 (34.52)
NWR	—	0.46 (0.10)***	0.51 (0.22)*	—	0.48 (0.11)***	0.50 (0.35)**
NWR $\times$ Group	—	—	-0.07 (0.26)	—	—	-0.05 (0.42)
ΔR^2	—	12.6%***	0.09%	—	15.8%***	0.04%
R^2	61.0%***	73.6%***	73.7%***	41.3%***	57.1%***	57.1%***

Note. Values presented are unstandardized coefficients, B (SE). PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition; CELF-P2 = Clinical Evaluation of Language Fundamentals–Preschool-2 Core Language Score; SES = Hollingshead Four Factor Index of Socioeconomic Status; Age (months) = age in months at the 54-month-old visit; NWR = nonword repetition task Relaxed percentage of consonants correct.

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

completed the NWR task did not significantly differ from that of their FT peers. Thus, children born PT who had sufficient skills to engage in this task at this young age achieved scores that were comparable with those of children born FT. Note that, when we included all children in the analyses, both those who completed and those who did not complete the task, PT–FT group differences reached the level of statistical significance. Including the children who could not complete the task in the subsequent analyses would have resulted in different conclusions about the presence of group differences.

In general, results of this study add to the existing literature finding no significant group differences in phonological short-term memory skills between children born PT and FT if they complete the NWR task (Briscoe et al., 1998; De Hirsch et al., 1964; Fraello et al., 2011; Jansson-Verkasalo et al., 2004; Odd et al., 2012; Sansavini et al., 2007). These findings align with research showing that children born PT demonstrate generally lower cognitive and linguistic skills than their FT peers, but their overall performance is not necessarily discrepant (Aram, Hack, Hawkins, Weissman, & Borawski-Clark, 1991; Aylward, 2005). Our findings differ from those of Guarini and colleagues (2016), which may be due to differences in the test and scoring procedures.

We evaluated whether articulation and phonological skills would affect performance. Compared with children born FT, children born PT were not more significantly aided by an adjusted scoring measure that accounted for articulatory differences and developmentally appropriate sound substitutions (Relaxed PCC). This finding suggests that children born PT in this sample did not demonstrate articulatory or phonological challenges that interfered with their NWR performance to a greater extent than their FT peers. These findings do not support the suggestion that deficits in language in children born PT are attributable to limitations in articulatory skills (Guarini et al., 2016; Wolke & Meyer, 1999). However, we acknowledge that our findings may not align with those of previous studies because of sample characteristics, such as age of participants, as both Guarini et al. (2016) and Wolke and Meyer (1999) studied

older children, aged 5–6 years. In this study, speech sound production ability did not differentially affect NWR performance at 36 months old for the PT and FT groups, an age at which children are still developing their speech skills. It is possible that significant differences in speech sound development cannot be detected at this young age.

Consistent with our hypothesis that nonword length would affect performance, children born PT and FT both demonstrated increasingly poorer NWR scores for items of increasing length. However, contrary to the findings of Sansavini and colleagues (2007), children born PT in this study did not have significantly more difficulty repeating longer nonwords than did children born FT. Studies of clinical populations show that children with SLI perform significantly more poorly on nonword items with a greater number of syllables (Dollaghan & Campbell, 1998; Weismer et al., 2000). The lack of a significant interaction between item length and birth group suggests that children born PT, in this way, do not resemble children with SLI. These findings align with previous results suggesting that children born PT do not demonstrate language deficits that are characteristic of SLI (Smith, DeThorne, Logan, Channell, & Petrill, 2014).

Relations to Language Outcomes

It is generally accepted that performance on NWR tasks is strongly associated with language performance in children born FT (Adams & Gathercole, 1995; Chiat & Roy, 2008; Edwards et al., 2004; Gathercole & Adams, 1993; Gathercole & Baddeley, 1990; Hoff et al., 2008; Sahlén et al., 1999). The current study aligns with these findings, demonstrating that performance on an NWR task at 36 months old is significantly associated with concurrent language skills and is also a significant predictor of later language outcomes in both children born PT and FT, even after controlling for earlier language skills. Furthermore, the current longitudinal study contributes evidence that early phonological short-term memory and later language skills are significantly linked in parallel ways for both children born PT and FT. Thus, performance on

a test of NWR, even as young as 3 years old, provides a clinically useful tool for identifying an increased risk of subsequent language deficits in children born PT, as it has been shown in children born FT.

Limitations

Although the sample of 118 children was acceptable, sample sizes for subanalyses were relatively small. In particular, we only had follow-up language data for 55 participants. In addition, our sample generally consisted of children from higher SES backgrounds who had relatively strong language skills. Because environment can have a significant impact on language development, our participants' performance may not represent the general abilities of children born PT and FT from diverse populations. In addition, although live presentation has the advantage of increasing the likelihood that young children will remain engaged in the task, rate of item presentation, amount of prompting for attention and engagement, and other small differences in item presentation may have affected performance for some children. It should also be noted that most of our participants demonstrated language abilities within the average range. It is possible that a sample with a greater range of language abilities across both groups may reveal different results. Finally, we did not have access to information regarding underlying neuropsychological skills, such as attention or executive function, that might have differentiated performance on the NWR task in PT and FT children. Future studies should incorporate measures that would allow such comparisons.

Implications and Conclusions

The findings of this study demonstrate that children born PT were less likely to engage in a test of phonological short-term memory, and those children born PT and FT who did not complete the task exhibited poorer performance on standardized language measures both concurrently and later. Therefore, difficulty completing NWR tasks could be a marker of language deficit in and of itself, and lack of task compliance should not be dismissed as solely an attention or behavior problem. Alternative means of assessment should be used in such children to explore possible language weaknesses.

Among the children who completed the task, children born PT did not perform significantly below children born FT on a test of phonological short-term memory. Rather, they showed a comparable cognitive–linguistic profile. Despite the lack of group differences, individual differences in NWR performance were positively associated with variation in later language outcomes in both children born PT and FT. Although deficits in this skill alone may not account for the lower language skills seen in children born PT, phonological short-term memory remains an important neuropsychological process underlying language development in this population. Because early NWR skills were significantly predictive of later language skills, even after considering

age, SES, and early receptive vocabulary, we suggest that assessment of early phonological short-term memory skills may be one task that can assist clinicians in identifying those children born PT who are at the greatest risk for language deficits. Future studies should explore whether other skills, such as speed of processing, cognition, attention, or executive function, also contribute to the increased risk for language deficits in this population.

Acknowledgments

This study was funded by a research grant from the National Institute of Child Health and Human Development (R01 HD069150) awarded to Dr. Heidi Feldman and Dr. Anne Fernald. We would like to thank Katherine Adams, Melanie Ashland, the staff of the Language Learning Lab at Stanford University, and the Developmental Behavioral Pediatrics lab at the Stanford University School of Medicine for their support. We would also like to thank the children and families who participated in this study.

References

- Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics*, 124(2), 717–728. <https://doi.org/10.1542/peds.2008-2816>
- Adams, A., & Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research*, 38(2), 403–414. <https://doi.org/10.1044/jshr.3802.403>
- Adlof, S. M., & Patten, H. (2017). Nonword repetition and vocabulary knowledge as predictors of children's phonological and semantic word learning. *Journal of Speech, Language, and Hearing Research*, 60(3), 682–693. https://doi.org/10.1044/2016_JSLHR-L-15-0441
- Aram, D. M., Hack, M., Hawkins, S., Weissman, B. M., & Borawski-Clark, E. (1991). Very-low-birthweight children and speech and language development. *Journal of Speech and Hearing Research*, 34(5), 1169–1179. <https://doi.org/10.1056/NEJM199107253250403>
- Aylward, G. P. (2002). Cognitive and neuropsychological outcomes: More than IQ scores. *Mental Retardation and Developmental Disabilities Research Reviews*, 8, 234–240. <https://doi.org/10.1002/mrdd.10043>
- Aylward, G. P. (2005). Neurodevelopmental outcomes of infants born prematurely. *Journal of Developmental and Behavioral Pediatrics*, 26(6), 394–407. <https://doi.org/10.1097/01.DBP.0000452240.39511.d4>
- Barre, N., Morgan, A., Doyle, L. W., & Anderson, P. J. (2011). Language abilities in children who were very preterm and/or very low birth weight: A meta-analysis. *Journal of Pediatrics*, 158(5), 766.e1–774.e1. <https://doi.org/10.1016/j.jpeds.2010.10.032>
- Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm. *Journal of the American Medical Association*, 288(6), 728–737. <https://doi.org/10.1001/jama.288.6.728>
- Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a behavioral marker for inherited language impairment: Evidence from a twin study. *The Journal of Child Psychology*

- and *Psychiatry*, 37(4), 391–403. <https://doi.org/10.1111/j.1469-7610.1996.tb01420.x>
- Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A quasi-universal nonword repetition task as a diagnostic tool for bilingual children learning Dutch as a second language. *Journal of Speech, Language, and Hearing Research*, 58(6), 1747–1760. https://doi.org/10.1044/2015_JSLHR-L-15-0058
- Briscoe, J., Bishop, D. V. M., & Norbury, C. F. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *The Journal of Child Psychology and Psychiatry*, 42(3), 329–340. <https://doi.org/10.1111/1469-7610.00726>
- Briscoe, J., Gathercole, S. E., & Marlow, N. (1998). Short-term memory and language outcomes after extreme prematurity at birth. *Journal of Speech, Language, and Hearing Research*, 41(3), 654–666. <https://doi.org/10.1044/jslhr.4103.654>
- Centers for Disease Control and Prevention. (2016). *Premature birth*. <http://www.cdc.gov/features/prematurebirth/>
- Chiat, S., & Roy, P. (2007). The Preschool Repetition Test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research*, 50(2), 429–443. <https://doi.org/10.1092-4388/07/5002-0429>
- Chiat, S., & Roy, P. (2008). Early phonological and sociocognitive skills as predictors of later language and social communication outcomes. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49(6), 635–645. <https://doi.org/10.1111/j.1469-7610.2008.01881.x>
- De Hirsch, K., Jansky, J. J., & Langford, W. S. (1964). The oral language performance of premature children and controls. *Journal of Speech and Hearing Disorders*, 29(1), 60–69. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emcl1&NEWS=N&AN=0008862412>
- Deevy, P., Weil, L. W., Leonard, L. B., & Goffman, L. (2010). Extending use of the NRT to preschool-age children with and without specific language impairment. *Language, Speech, and Hearing Services in Schools*, 41(3), 277–288. [https://doi.org/10.1044/0161-1461\(2009/08-0096\)](https://doi.org/10.1044/0161-1461(2009/08-0096))
- Devescovi, A., & Caselli, M. C. (2001). Una prova di ripetizione di frasi per la valutazione del primo sviluppo grammaticale [A sentence repetition task for the evaluation of initial grammatical development]. *Psicologia Clinica Dello Sviluppo*, 5(3), 341–364. <https://doi.org/10.1449/633>
- Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, 41(5), 1136–1146. <https://doi.org/10.1044/jslhr.4105.1136>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test—Fourth Edition*. Minneapolis, MN: Pearson.
- Edwards, J., Beckman, M. E., & Munson, B. (2004). Vocabulary size and phonotactic production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, 47(2), 421–436. [https://doi.org/10.1044/1092-4388\(2004/034\)](https://doi.org/10.1044/1092-4388(2004/034))
- Fraello, D., Maller-Kesselman, J., Vohr, B., Katz, K. H., Kesler, S., Schneider, K., ... Spann, M. N. (2011). Consequence of preterm birth in early adolescence: the role of language on auditory short-term memory. *Journal of Child Neurology*, 26(6), 738–742. <https://doi.org/10.1177/0883073810391904>
- Frye, R. E., Landry, S. H., Swank, P. R., & Smith, K. E. (2009). Executive dysfunction in poor readers born prematurely at high risk. *Developmental Neuropsychology*, 34(3), 254–271. <https://doi.org/10.1080/87565640902805727>
- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(4), 513–543. <https://doi.org/10.1017/S0142716406060383>
- Gathercole, S. E., & Adams, A. M. (1993). Phonological working memory in very young children. *Developmental Psychology*, 29(4), 770–778. <https://doi.org/10.1037/0012-1649.29.4.770>
- Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, 29(3), 336–360. [https://doi.org/10.1016/0749-596X\(90\)90004-J](https://doi.org/10.1016/0749-596X(90)90004-J)
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory*, 2(2), 103–127. <https://doi.org/10.1080/09658219408258940>
- Gray, S. (2003). Diagnostic accuracy and test-retest reliability of nonword repetition and digit span tasks administered to pre-school children with specific language impairment. *Journal of Communication Disorders*, 36(2), 129–151. [https://doi.org/10.1016/S0021-9924\(03\)00003-0](https://doi.org/10.1016/S0021-9924(03)00003-0)
- Guarini, A., Marini, A., Savini, S., Alessandrini, R., Faldella, G., & Sansavini, A. (2016). Linguistic features in children born very preterm at preschool age. *Developmental Medicine & Child Neurology*, 58(6), 949–956. <https://doi.org/10.1111/dmcn.13118>
- Hayes, A. F. (2013). *Methodology in the Social Sciences. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: Guilford Press.
- Hoff, E., Core, C., & Bridges, K. (2008). Non-word repetition assesses phonological memory and is related to vocabulary development in 20- to 24-month-olds. *Journal of Child Language*, 35(4), 903–916. <https://doi.org/10.1017/S0305000908008751>
- Hollingshead, A. B. (1975). *Four factor index of social status*. (Unpublished manuscript). New Haven, CT: Yale University.
- Jansson-Verkasalo, E., Valkama, M., Vainionpää, L., Pääkkö, E., Ilkko, E., & Lehtihalmes, M. (2004). Language development in very low birth weight preterm children: A follow-up study. *Folia Phoniatrica et Logopaedica*, 56(2), 108–119. <https://doi.org/10.1159/000076062>
- Kovachy, V. N., Adams, J. N., Tamaresis, J. S., & Feldman, H. M. (2015). Reading abilities in school-aged preterm children: A review and meta-analysis. *Developmental Medicine and Child Neurology*, 57(5), 410–419. <https://doi.org/10.1111/dmcn.12652>
- Krishnan, S., Alcock, K. J., Mercure, E., Leech, R., Barker, E., Karmiloff-Smith, A., & Dick, F. (2013). Articulating novel words: Children's oromotor skills predict nonword repetition abilities. *Journal of Speech, Language, and Hearing Research*, 56(6), 1800–1812. [https://doi.org/10.1044/1092-4388\(2013\)12-0206](https://doi.org/10.1044/1092-4388(2013)12-0206)
- Leonard, L. B., Ellis Weismer, S., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech, Language, and Hearing Research*, 50(2), 408–428. [https://doi.org/10.1044/1092-4388\(2007\)029](https://doi.org/10.1044/1092-4388(2007)029)
- Luu, T. M., Ment, L. R., Schneider, K. C., Katz, K. H., Allan, W. C., & Vohr, B. R. (2009). Lasting effects of preterm birth and neonatal brain hemorrhage at 12 years of age. *Pediatrics*, 123(3), 1037–1044. <https://doi.org/10.1542/peds.2008.1162>
- Odd, D. E., Emond, A., & Whitelaw, A. (2012). Long-term cognitive outcomes of infants born moderately and late preterm. *Developmental Medicine & Child Neurology*, 54(8), 704–709. <https://doi.org/10.1111/j.1469-8749.2012.04315.x>
- Parra, M., Hoff, E., & Core, C. (2011). Relations among language exposure, phonological memory, and language

- development in Spanish-English bilingually developing 2-year-olds. *Journal of Experimental Child Psychology*, 108(1), 113–125. <https://doi.org/10.1016/j.jecp.2010.07.011>
- Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: Evidence from typically developing children. *Journal of Speech, Language, and Hearing Research*, 47, 223–234. <https://doi.org/10.1092-4388/04/4701-0223>
- Sahlén, B., Reuterskiöld-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Language comprehension and non-word repetition in children with language impairment—Pitfalls and possibilities. *Clinical Linguistics & Phonetics*, 34(3), 337–352. <https://login.iris.etsu.edu:3443/login?url=http://search.proquest.com/docview/85496148?accountid=10771>
- Sansavini, A., Guarini, A., Alessandrini, R., Faldella, G., Giovanelli, G., & Salvioli, G. (2007). Are early grammatical and phonological working memory abilities affected by preterm birth? *Journal of Communication Disorders*, 40(3), 239–256. <https://doi.org/10.1016/j.jcomdis.2006.06.009>
- Sansavini, A., Guarini, A., Justice, L. M., Savini, S., Broccoli, S., Alessandrini, R., & Faldella, G. (2010). Does preterm birth increase a child's risk for language impairment? *Early Human Development*, 86(12), 765–772. <https://doi.org/10.1016/j.earlhumdev.2010.08.014>
- Semel, E., Wiig, E. H., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals—Preschool-2*. San Antonio, TX: Pearson. <http://www.pearsonclinical.com/language/products/100000316/celf-preschool-2-celf-preschool-2.html>
- Shriberg, L. D., Austin, D., Lewis, B. A., Mcsweeney, J. L., & Wilson, D. L. (1997). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, 40(4), 708–722. <https://doi.org/10.1044/jslhr.4004.708>
- Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders*, 47(3), 256–270. <https://doi.org/10.1044/jshd.4703.256>
- Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, 55(4), 779–798. <https://doi.org/10.1044/jshd.5504.779>
- Smith, J. M., DeThorne, L. S., Logan, J. A. R., Channell, R. W., & Petrill, S. A. (2014). Impact of prematurity on language skills at school age. *Journal of Speech, Language, and Hearing Research*, 57(3), 901–916. [https://doi.org/10.1044/1092-4388\(2013\)12-0347](https://doi.org/10.1044/1092-4388(2013)12-0347)
- Stokes, S. F., & Klee, T. (2009). The diagnostic accuracy of a new Test of Early Nonword Repetition for differentiating late talking and typically developing children. *Journal of Speech, Language, and Hearing Research*, 52(4), 872–882. [https://doi.org/10.1044/1092-4388\(2009\)08-0030](https://doi.org/10.1044/1092-4388(2009)08-0030)
- Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*, 41(2), 407–418. <https://doi.org/10.1044/jslhr.4102.407>
- van Noort-van der Spek, I. L., Franken, M.-C. J. P., & Weisglas-Kuperus, N. (2012). Language functions in preterm-born children: A systematic review and meta-analysis. *Pediatrics*, 129(4), 745–754. <https://doi.org/10.1542/peds.2011-1728>
- Vicari, S., Caravale, B., Carlesimo, G. A., Casadei, A. M., & Allemand, F. (2004). Spatial working memory deficits in children at ages 3–4 who were low birth weight, preterm infants. *Neuropsychology*, 18(4), 673–678. <https://doi.org/10.1037/0894-4105.18.4.673>
- Vieira, M. E. B., & Linhares, M. B. M. (2011). Developmental outcomes and quality of life in children born preterm at pre-school- and school-age. *Jornal de Pediatria*, 87(4), 281–291. <https://doi.org/10.2223/JPED.2096>
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing—Second Edition*. Austin, TX: Pro-Ed.
- Weismer, S. E., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword repetition performance in school-age children with and without language impairment. *Journal of Speech, Language, and Hearing Research*, 43(4), 865–878. <https://doi.org/10.1044/jslhr.4304.865>
- Williams, B. L., Dunlop, A. L., Kramer, M., Dever, B. V., Hogue, C., & Jain, L. (2013). Perinatal origins of first-grade academic failure: Role of prematurity and maternal factors. *Pediatrics*, 131(4), 693–700. <https://doi.org/10.1542/peds.2012-1408>
- Wolke, D., & Meyer, R. (1999). Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: The Bavarian Longitudinal Study. *Developmental Medicine & Child Neurology*, 41(2), 94–109. <https://doi.org/10.1111/j.1469-8749.1999.tb00561.x>
- Wolke, D., Samara, M., Bracewell, M., & Marlow, N. (2008). Specific language difficulties and school achievement in children born at 25 weeks of gestation or less. *The Journal of Pediatrics*, 152(2), 256–262. <https://doi.org/10.1016/j.jpeds.2007.06.043>
- Woodward, L. J., Edgin, J. O., Thompson, D., & Inder, T. E. (2005). Object working memory deficits predicted by early brain injury and development in the preterm infant. *Brain*, 128(11), 2578–2587. <https://doi.org/10.1093/brain/awh618>
- Woodward, L. J., Moor, S., Hood, K. M., Champion, P. R., Foster-Cohen, S., Inder, T. E., & Austin, N. C. (2009). Very preterm children show impairments across multiple neurodevelopmental domains by age 4 years. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 94(5), F339–F344. <https://doi.org/10.1136/adc.2008.146282>

Appendix

Real Word and Nonword Stimuli

Real Word	Nonword	Nonword IPA Transcription
dog	<i>kog</i>	'kag
juice	<i>boose</i>	'bus
cat	<i>jot</i>	'dʒat
book	<i>dook</i>	'dʊk
balloon	<i>cha-LOON</i>	tʃə'lun
cookie	<i>POO-kie</i>	'puki
puppy	<i>KU-ppy</i>	'kəpi
chicken	<i>BI-ken</i>	'bɪkən
lollipop	<i>ba-JA-pop</i>	bə'dʒapap
banana	<i>ta-LEE-na</i>	tə'line
pajamas	<i>la-LE-mas</i>	lə'leməs
telephone	<i>pa-NA-fone</i>	pə'næfon
watermelon	<i>wa-nut-SA-plin</i>	wənət'sæplɪn
motorcycle	<i>MOW-lee-mu-kle</i>	'mouliməkəl
helicopter	<i>HE-der-be-ckle</i>	'hɛdə-bɛkəl
peanut butter	<i>pi-lo-KI-der</i>	pilo'kaɪdə
	<i>see-koo-DA-ja-mint</i>	siku'dadʒəmɪnt
	<i>fee-glo-ME-za-cheeb</i>	figlo'mɛzɛtʃɪb
	<i>la-po-VEE-doo-breen</i>	lapo'vidubrin
	<i>ze-ki-SPA-di-tog</i>	zɛki'spɒdɪtəg

Nonwords adapted from "Relations Among Language Exposure, Phonological Memory, and Language Development in Spanish–English Bilingually Developing 2-Year-Olds," by M. Parra, E. Hoff, and C. Core, 2011, *Journal of Experimental Child Psychology*, 108, pp. 113–125. Copyright © Elsevier Inc. Reprinted with permission.